Molecular Cloning: A Laboratory Manual, 3rd ed., Vols 1,2 and 3
J.F. Sambrook and D.W. Russell, ed., Cold Spring Harbor Laboratory Press, 2001, 2100 pp., soft cover

Appendices

Appendix 1

Preparation of Reagents and Buffers Used in Molecular Cloning

BUFFERS

Tris Buffers

Biological reactions work well only within a narrow concentration range of hydrogen ions. Paradoxically, however, many of these reactions themselves generate or consume protons. Buffers are substances that undergo reversible protonation within a particular pH range and therefore maintain the concentration of hydrogen ions within acceptable limits. Perfect buffers are, like the Holy Grail, always beyond reach. An ideal biological buffer should

- have a pK_a between pH 6.0 and pH 8.0
- be inert to a wide variety of chemicals and enzymes
- be highly polar, so that it is both exquisitely soluble in aqueous solutions and also unlikely to diffuse across biological membranes and thereby affect intracellular pH
- be nontoxic
- be cheap
- not be susceptible to salt or temperature effects
- not absorb visible or ultraviolet light

None of the buffers used in molecular biology fulfill all of these criteria. Very few weak acids are known that have dissociation constants between 10^{-7} and 10^{-9}. Among inorganic salts, only borates, bicarbonates, phosphates, and ammonium salts lie within this range. However, they are all incompatible in one way or another with physiological media.

In 1946, George Gomori (Gomori 1946) suggested that organic polyamines could be used to control pH in the range 6.5–9.7. One of the three compounds he investigated was Tris(2-amino-2-hydroxymethyl-1,3-propanediol), which had been first described in 1897 by Piloty and Ruff. Tris turned out to be an extremely satisfactory buffer for many biochemical purposes and today is the standard buffer used for most enzymatic reactions in molecular cloning.

TABLE A1-1 Preparation of Tris Buffers of Various Desired pH Values

DESIRED pH (25°C)	VOLUME OF 0.1 N HCl (ml)
7.10	45.7
7.20	44.7
7.30	43.4
7.40	42.0
7.50	40.3
7.60	38.5
7.70	36.6
7.80	34.5
7.90	32.0
8.00	29.2
8.10	26.2
8.20	22.9
8.30	19.9
8.40	17.2
8.50	14.7
8.60	12.4
8.70	10.3
8.80	8.5
8.90	7.0

Tris buffers (0.05 M) of the desired pH can be made by mixing 50 ml of 0.1 M Tris base with the indicated volume of 0.1 N HCl and then adjusting the volume of the mixture to 100 ml with water.

TRIS BUFFERS

One of Tris' first commercial successes, which received wide attention, was the reduction of mortality during handling and hauling of fish. In the 1940s, live fish were carried to market in tanks of seawater. Unfortunately, many of the fish died because of the decline in pH resulting from an accumulation of CO_2. This problem was only partially alleviated by including anesthetics in the water that minimized the fishes' metabolic activities. What these anesthetics did to the people who ate the fish is not recorded. Tris certainly reduced the mortality rate of the fish (McFarland and Norris 1958) by stabilizing the pH of the seawater and may also have kept the fish eaters more alert. Tris also turned out to be an extremely satisfactory buffer for many biochemical purposes and today is the standard buffer used for most enzymatic reactions in molecular cloning.

Tris [Tris(hydroxymethyl)aminomethane] has a very high buffering capacity, is highly soluble in water, and is inert in a wide variety of enzymatic reactions. However, Tris also has a number of deficiencies:

- ***The pK_a of Tris is pH 8.0 (at 20°C),*** which means that its buffering capacity is very low at pHs below 7.5 and above 9.0.
- ***Temperature has a significant effect on the dissociation of Tris.*** The pH of Tris solutions decreases ~0.03 pH units for each 1°C increase in temperature. For example, a 0.05 M solution has pH values of 9.5, 8.9, and 8.6 at 5°C, 25°C, and 37°C, respectively. By convention, the pH of Tris solutions given in the scientific literature refers to the pH measured at 25°C. When preparing stock solutions of Tris, it is best to bring the pH into the desired range and then allow the solution to cool to 25°C before making final adjustments to the pH.
- ***Tris reacts with many types of pH electrodes*** that contain linen-fiber junctions, apparently because Tris reacts with the linen fiber. This effect is manifested in large liquid-junction potentials, electromotive force (emf) drift, and long equilibration times. Electrodes with linen-fiber junctions, therefore, cannot accurately measure the pH of Tris solutions. Use only those electrodes with ceramic or glass junctions that are warranted by the manufacturer to be suitable for Tris.
- ***Concentration has a significant effect on the dissociation of Tris.*** For example, the pHs of solutions containing 10 mM and 100 mM Tris will differ by 0.1 of a pH unit, with the more concentrated solution having the higher pH.
- ***Tris is toxic to many types of mammalian cells.***
- ***Tris, a primary amine, cannot be used with fixatives*** such as glutaraldehyde and formaldehyde. Tris also reacts with glyoxal. Phosphate or MOPS buffer is generally used in place of Tris with these reagents.

Good Buffers

Tris is a poor buffer at pH values below 7.5. In the mid 1960s, Norman Good and his colleagues responded to the need for better buffers in this range by developing a series of *N*-substituted aminosulfonic acids that behave as strong zwitterions at biologically relevant pH values (Good et al. 1966; Ferguson et al. 1980). Without these buffers, several techniques central to molecular cloning either would not exist at all or would work at greatly reduced efficiency. These techniques include high-efficiency transfection of mammalian cells (HEPES, Tricine, and BES), gel electrophoresis of RNA (MOPS), and high-efficiency transformation of bacteria (MES).

Table A1-2 Properties of Good Buffers

Acronym	Chemical Name	FW	pK_a	Useful Range (in pH units)
MES	2-(*N*-morpholino)ethanesulfonic acid	195.2	6.1	5.5–6.7
Bis-Tris	*bis*(2-hydroxyethyl)imino*tris*(hydroxymethyl)methane	209.2	6.5	5.8–7.2
ADA	*N*-(2-acetamido)-2-iminodiacetic acid	190.2	6.6	6.0–7.2
ACES	2-[(2-amino-2-oxoethyl)amino]ethanesulfonic acid	182.2	6.8	6.1-7.5
PIPES	piperazine-*N*,*N*′-*bis*(2-ethanesulfonic acid)	302.4	6.8	6.1–7.5
MOPSO	3-(*N*-morpholino)-2-hydroxypropanesulfonic acid	225.3	6.9	6.2–7.6
Bis-Tris Propane	1,3-*bis*[*tris*(hydroxymethyl)methylamino]propane	282.3	6.8[a]	6.3–9.5
BES	*N*,*N*-*bis*(2-hydroxyethyl)-2-aminoethanesulfonic acid	213.2	7.1	6.4–7.8
MOPS	3-(*N*-morpholino)propanesulfonic acid	209.3	7.2	6.5–7.9
HEPES	*N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonic acid)	238.3	7.5	6.8–8.2
TES	*N*-*tris*(hydroxymethyl)methyl-2-aminoethanesulfonic acid	229.2	7.4	6.8–8.2
DIPSO	3-[*N*,*N*-*bis*(2-hydroxyethyl)amino]-2-hydroxypropanesulfonic acid	243.3	7.6	7.0–8.2
TAPSO	3-[*N*-*tris*(hydroxymethyl)methylamino]-2-hydroxypropanesulfonic acid	259.3	7.6	7.0–8.2
TRIZMA	*tris*(hydroxymethyl)aminomethane	121.1	8.1	7.0-9.1
HEPPSO	*N*-(2-hydroxyethyl)piperazine-*N*′-(2-hydroxypropanesulfonic acid)	268.3	7.8	7.1–8.5
POPSO	piperazine-*N*,*N*′-*bis*(2-hydroxypropanesulfonic acid)	362.4	7.8	7.2-8.5
EPPS	*N*-(2-hydroxyethyl)piperazine-*N*′-(3-propanesulfonic acid)	252.3	8.0	7.3–8.7
TEA	triethanolamine	149.2	7.8	7.3–8.3
Tricine	*N*-*tris*(hydroxymethyl)methylglycine	179.2	8.1	7.4–8.8
Bicine	*N*,*N*-*bis*(2-hydroxyethyl)glycine	163.2	8.3	7.6–9.0
TAPS	*N*-*tris*(hydroxymethyl)methyl-3-aminopropanesulfonic acid	243.3	8.4	7.7–9.1
AMPSO	3-[(1,1-dimethyl-2-hydroxyethyl)amino]-2-hydroxypropanesulfonic acid	227.3	9.0	8.3–9.7
CHES	2-(*N*-cyclohexylamino)ethanesulfonic acid	207.3	9.3	8.6–10.0
CAPSO	3-(cyclohexylamino)-2-hydroxy-1-propanesulfonic acid	237.3	9.6	8.9–10.3
AMP	2-amino-2-methyl-1-propanol	89.1	9.7	9.0–10.5
CAPS	3-(cyclohexylamino)-1-propanesulfonic acid	221.3	10.4	9.7–11.1

Data compiled from various sources, including *Biochemical and Reagents for Life Science Research* 1994 (Sigma-Aldrich) and references therein.

[a] pK_a = 9.0 for the second dissociation stage.

Phosphate Buffers (Gomori Buffers)

The most commonly used phosphate buffers are named after their inventor: Gomori (Gomori 1955). They consist of a mixture of monobasic dihydrogen phosphate and dibasic monohydrogen phosphate. By varying the amount of each salt, a range of buffers can be prepared that buffer well between pH 5.8 and pH 8.0 (please see Tables A1-3A and A1-3B). Phosphates have a very high buffering capacity and are highly soluble in water. However, they have a number of potential disadvantages:

- Phosphates inhibit many enzymatic reactions and procedures that are the foundation of molecular cloning, including cleavage of DNA by many restriction enzymes, ligation of DNA, and bacterial transformation.
- Because phosphates precipitate in ethanol, it is not possible to precipitate DNA and RNA from buffers that contain significant quantities of phosphate ions.
- Phosphates sequester divalent cations such as Ca^{2+} and Mg^{2+}.

TABLE A1-3A Preparation of 0.1 M Potassium Phosphate Buffer at 25°C

pH	Volume of 1 M K_2HPO_4 (ml)	Volume of 1 M KH_2PO_4 (ml)
5.8	8.5	91.5
6.0	13.2	86.8
6.2	19.2	80.8
6.4	27.8	72.2
6.6	38.1	61.9
6.8	49.7	50.3
7.0	61.5	38.5
7.2	71.7	28.3
7.4	80.2	19.8
7.6	86.6	13.4
7.8	90.8	9.2
8.0	94.0	6.0

Data from Green (1933).

Dilute the combined 1 M stock solutions to 1 liter with distilled H_2O. pH is calculated according to the Henderson-Hasselbalch equation:

$$pH = pK' + \log \left\{ \frac{\text{(proton acceptor)}}{\text{proton donor}} \right\}$$

where $pK' = 6.86$ at 25°C.

TABLE A1-3B Preparation of 0.1 M Sodium Phosphate Buffer at 25°C

pH	Volume of 1 M Na_2HPO_4 (ml)	Volume of 1 M NaH_2PO_4 (ml)
5.8	7.9	92.1
6.0	12.0	88.0
6.2	17.8	82.2
6.4	25.5	74.5
6.6	35.2	64.8
6.8	46.3	53.7
7.0	57.7	42.3
7.2	68.4	31.6
7.4	77.4	22.6
7.6	84.5	15.5
7.8	89.6	10.4
8.0	93.2	6.8

Data from ISCO (1982).

Dilute the combined 1 M stock solutions to 1 liter with distilled H_2O. pH is calculated according to the Henderson-Hasselbalch equation:

$$pH = pK' + \log \left\{ \frac{\text{(proton acceptor)}}{\text{proton donor}} \right\}$$

where $pK' = 6.86$ at 25°C.

ACIDS AND BASES

TABLE A1-4 Concentrations of Acids and Bases: Common Commercial Strengths

SUBSTANCE	FORMULA	M.W.	MOLES/ LITER[a]	GRAMS/ LITER	% BY WEIGHT	SPECIFIC GRAVITY	ml/LITER TO PREPARE 1 M SOLUTION
Acetic acid, glacial	CH_3COOH	60.05	17.4	1045	99.5	1.05	57.5
Acetic acid		60.05	6.27	376	36	1.045	159.5
Formic acid	HCOOH	46.02	23.4	1080	90	1.20	42.7
Hydrochloric acid	HCl	36.5	11.6	424	36	1.18	86.2
			2.9	105	10	1.05	344.8
Nitric acid	HNO_3	63.02	15.99	1008	71	1.42	62.5
			14.9	938	67	1.40	67.1
			13.3	837	61	1.37	75.2
Perchloric acid	$HClO_4$	100.5	11.65	1172	70	1.67	85.8
			9.2	923	60	1.54	108.7
Phosphoric acid	H_3PO_4	80.0	18.1	1445	85	1.70	55.2
Sulfuric acid	H_2SO_4	98.1	18.0	1766	96	1.84	55.6
Ammonium hydroxide	NH_4OH	35.0	14.8	251	28	0.898	67.6
Potassium hydroxide	KOH	56.1	13.5	757	50	1.52	74.1
			1.94	109	10	1.09	515.5
Sodium hydroxide	NaOH	40.0	19.1	763	50	1.53	52.4
			2.75	111	10	1.11	363.6

[a]With some acids and bases, stock solutions of different molarity/normality are in common use. These are often abbreviated "conc" for concentrated stocks and "dil" for dilute stocks.

TABLE A1-5 Approximate pH Values for Various Concentrations of Stock Solutions

SUBSTANCE	1 N	0.1 N	0.01 N	0.001 N
Acetic acid	2.4	2.9	3.4	3.9
Hydrochloric acid	0.10	1.07	2.02	3.01
Sulfuric acid	0.3	1.2	2.1	
Citric acid		2.1	2.6	
Ammonium hydroxide	11.8	11.3	10.8	10.3
Sodium hydroxide	14.05	13.07	12.12	11.13
Sodium bicarbonate		8.4		
Sodium carbonate		11.5	11.0	

PREPARATION OF BUFFERS AND STOCK SOLUTIONS FOR USE IN MOLECULAR BIOLOGY

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

pH Buffers

Phosphate-buffered Saline (PBS)

137 mM NaCl
2.7 mM KCl
10 mM Na_2HPO_4
2 mM KH_2PO_4

Dissolve 8 g of NaCl, 0.2 g of KCl, 1.44 g of Na_2HPO_4, and 0.24 g of KH_2PO_4 in 800 ml of distilled H_2O. Adjust the pH to 7.4 with HCl. Add H_2O to 1 liter. Dispense the solution into aliquots and sterilize them by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle or by filter sterilization. Store the buffer at room temperature.

PBS is a commonly used reagent that has been adapted for different applications. Note that the recipe presented here lacks divalent cations. If necessary, PBS may be supplemented with 1 mM $CaCl_2$ and 0.5 mM $MgCl_2$.

10x Tris EDTA (TE)

pH 7.4
100 mM Tris-Cl (pH 7.4)
10 mM EDTA (pH 8.0)

pH 7.6
100 mM Tris-Cl (pH 7.6)
10 mM EDTA (pH 8.0)

pH 8.0
100 mM Tris-Cl (pH 8.0)
10 mM EDTA (pH 8.0)

Sterilize solutions by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. Store the buffer at room temperature.

Tris-Cl (1 M)

Dissolve 121.1 g of Tris base in 800 ml of H_2O. Adjust the pH to the desired value by adding concentrated HCl <!>.

pH	*HCl*
7.4	70 ml
7.6	60 ml
8.0	42 ml

Allow the solution to cool to room temperature before making final adjustments to the pH. Adjust the volume of the solution to 1 liter with H_2O. Dispense into aliquots and sterilize by autoclaving.

If the 1 M solution has a yellow color, discard it and obtain Tris of better quality. The pH of Tris solutions is temperature-dependent and decreases ~0.03 pH units for each 1°C increase in temperature. For example, a 0.05 M solution has pH values of 9.5, 8.9, and 8.6 at 5°C, 25°C, and 37°C, respectively.

Tris Magnesium Buffer (TM)

50 mM Tris-Cl (pH 7.8)
10 mM $MgSO_4$

Tris-buffered Saline (TBS)

Dissolve 8 g of NaCl, 0.2 g of KCl, and 3 g of Tris base in 800 ml of distilled H_2O. Add 0.015 g of phenol red and adjust the pH to 7.4 with HCl. Add distilled H_2O to 1 liter. Dispense the solution into aliquots and sterilize them by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. Store the buffer at room temperature.

Enzyme Stocks and Buffers

Enzyme Stocks

Lysozyme (10 mg/ml)

Dissolve solid lysozyme at a concentration of 10 mg/ml in 10 mM Tris-Cl (pH 8.0) immediately before use. Make sure that the pH of the Tris solution is 8.0 before dissolving the protein. Lysozyme will not work efficiently if the pH of the solution is less than 8.0.

Lyticase (67 mg/ml)

Purchase from Sigma. Dissolve at 67 mg/ml (900 units/ml) in 0.01 M sodium phosphate containing 50% glycerol just before use.

Pancreatic DNase I (1 mg/ml)

Dissolve 2 mg of crude pancreatic DNase I (Sigma or equivalent) in 1 ml of

10 mM Tris-Cl (pH 7.5)
150 mM NaCl
1 mM $MgCl_2$

When the DNase I is dissolved, add 1 ml of glycerol to the solution and mix by gently inverting the closed tube several times. Take care to avoid creating bubbles and foam. Store the solution in aliquots of –20°C.

Pancreatic RNase (1 mg/ml)

Dissolve 2 mg of crude pancreatic RNase I (Sigma or equivalent) in 2 ml of TE (pH 7.6).

Proteinase K (20 mg/ml)

Purchase as a lyophilized powder and dissolve at a concentration of 20 mg/ml in sterile 50 mM Tris (pH 8.0), 1.5 mM calcium acetate. Divide the stock solution into small aliquots and store at –20°C. Each aliquot can be thawed and refrozen several times but should then be discarded. Unlike much cruder preparations of protease (e.g., pronase), proteinase K need not be self-digested before use. (Please see entry on Proteinase K in Appendix 4.)

Trypsin

Prepare bovine trypsin at a concentration of 250 µg/ml in 200 mM ammonium bicarbonate (pH 8.9) (Sequencer grade; Boehringer Mannheim). Store the solution in aliquots at –20°C.

Zymolyase 5000 (2 mg/ml)

Purchase from Kirin Breweries. Dissolve at 2 mg/ml in 0.01 M sodium phosphate containing 50% glycerol just before use.

Enzyme Dilution Buffers

DNase I Dilution Buffer

10 mM Tris-Cl (pH 7.5)
150 mM NaCl
1 mM $MgCl_2$

Polymerase Dilution Buffer

50 mM Tris-Cl (pH 8.1)
1 mM dithiothreitol
0.1 mM EDTA (pH 8.0)
0.5 mg/ml bovine serum albumin
5% (v/v) glycerol

Prepare solution fresh for each use.

Sequenase Dilution Buffer

10 mM Tris-Cl (pH 7.5)
5 mM dithiothreitol
0.5 mg/ml bovine serum albumin

Store the solution at –20°C.

***Taq* Dilution Buffer**

25 mM Tris (pH 8.8)
0.01 mM EDTA (pH 8.0)
0.15% (v/v) Tween-20
0.15% (v/v) Nonidet P-40

Enzyme Reaction Buffers

▲ **IMPORTANT** Wherever possible, use the 10x reaction buffer supplied by the manufacturer of the enzyme used. Otherwise, use the recipes given here.

10x Amplification Buffer

500 mM KCl
100 mM Tris-Cl (pH 8.3 at room temperature)
15 mM $MgCl_2$

Autoclave the 10x buffer for 10 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. Divide the sterile buffer into aliquots and store them at –20°C.

10x Bacteriophage T4 DNA Ligase Buffer

200 mM Tris-Cl (pH 7.6)
50 mM $MgCl_2$
50 mM dithiothreitol
0.5 mg/ml bovine serum albumin (Fraction V; Sigma) (*optional*)

Divide the buffer in small aliquots and store at –20°C. Add ATP when setting up the reaction. Add ATP to the reaction to an appropriate concentation (e.g., 1 mM).

10x Bacteriophage T4 DNA Polymerase Buffer

330 mM Tris-acetate (pH 8.0)
660 mM potassium acetate
100 mM magnesium acetate
5 mM dithiothreitol
1 mg/ml bovine serum albumin (Fraction V; Sigma)

Divide the 10x stock into small aliquots and store frozen at –20°C.

10x Bacteriophage T4 Polynucleotide Kinase Buffer

700 mM Tris-Cl (pH 7.6)
100 mM $MgCl_2$
50 mM dithiothreitol

Divide the 10x stock into small aliquots and store frozen at –20°C.

5x BAL 31 Buffer

3 M NaCl
60 mM $CaCl_2$
60 mM $MgCl_2$
100 mM Tris-Cl (pH 8.0)
1 mM EDTA (pH 8.0)

10x Dephosphorylation Buffer (for Use with CIP)

100 mM Tris-Cl (pH 8.3)
10 mM $MgCl_2$
10 mM $ZnCl_2$

10x Dephosphorylation Buffer (for Use with SAP)

200 mM Tris-Cl (pH 8.8)
100 mM $MgCl_2$
10 mM $ZnCl_2$

1x *Eco*RI Methylase Buffer

50 mM NaCl
50 mM Tris-Cl (pH 8.0)
10 mM EDTA
80 µM *S*-adenosylmethionine

Store the buffer in small aliquots at –20°C.

10x Exonuclease III Buffer

660 mM Tris-Cl (pH 8.0)
66 mM $MgCl_2$
100 mM β-mercaptoethanol <!>

Add β-mercaptoethanol just before use.

10x Klenow Buffer

0.4 M potassium phosphate (pH 7.5)
66 mM $MgCl_2$
10 mM β-mercaptoethanol <!>

10x Linker Kinase Buffer

600 mM Tris-Cl (pH 7.6)
100 mM $MgCl_2$
100 mM dithiothreitol
2 mg/ml bovine serum albumin

Prepare fresh just before use.

Nuclease S1 Digestion Buffer

0.28 M NaCl
0.05 M sodium acetate (pH 4.5)
4.5 mM $ZnSO_4 \cdot 7H_2O$

Store aliquots of nuclease S1 buffer at –20°C, and add nuclease S1 to a concentration of 500 units/ml just before use.

10x Proteinase K Buffer

100 mM Tris-Cl (pH 8.0)
50 mM EDTA (pH 8.0)
500 mM NaCl

10x Reverse Transcriptase Buffer

500 mM Tris-Cl (pH 8.3)
750 mM KCl
30 mM $MgCl_2$

RNase H Buffer

20 mM Tris-Cl (pH 7.6)
20 mM KCl
0.1 mM EDTA (pH 8.0)
0.1 mM dithiothreitol

Prepare fresh just before use.

5x Terminal Transferase Buffer

Most manufacturers supply a 5x reaction buffer, which typically contains:

500 mM potassium cacodylate (pH 7.2) <!>
10 mM $CoCl_2 \cdot 6H_2O$
1 mM dithiothreitol

5x terminal transferase (or tailing) buffer may be prepared according to the following method (Eschenfeldt et al. 1987):

1. Equilibrate 5 g of Chelex 100 (Bio-Rad) with 10 ml of 3 M potassium acetate at room temperature.
2. After five minutes, remove excess liquid by vacuum suction. Wash the Chelex three times with 10 ml of deionized H_2O.
3. Prepare a 1 M solution of potassium cacodylate. Equilibrate the cacodylate solution with the treated Chelex resin.
4. Recover the cacodylate solution by passing it through a Buchner funnel fitted with Whatman No. 1 filter paper.

5. To the recovered cacodylate add in order: H_2O, dithiothreitol, and cobalt chloride to make the final concentrations of 500 mM potassium cacodylate, 1 mM dithiothreitol, and 20 mM $CoCl_2$.

   Store the buffer in aliquots at –20°C.

10x Universal KGB (Restriction Endonuclease) Buffer

1 M potassium acetate
250 mM Tris-acetate (pH 7.6)
100 mM magnesium acetate tetrahydrate
5 mM β-mercaptoethanol <!>
0.1 mg/ml bovine serum albumin

Store the 10x buffer in aliquots at –20°C.

Hybridization Buffers

Alkaline Transfer Buffer (for Alkaline Transfer of DNA to Nylon Membranes)

0.4 N NaOH <!>
1 M NaCl

Church Buffer

1% (w/v) bovine serum albumin
1 mM EDTA
0.5 M phosphate buffer*
7% (w/v) SDS

*0.5 M phosphate buffer is 134 g of $Na_2HPO_4 \cdot 7H_2O$, 4 ml of 85% H_3PO_4 <!> (concentrated phosphoric acid), H_2O to 1 liter.

Denaturation Solution (for Neutral Transfer, Double-stranded DNA Targets Only)

1.5 M NaCl
0.5 M NaOH <!>

HCl (2.5 N)

Add 25 ml of concentrated HCl <!> (11.6 N) to 91 ml of sterile H_2O. Store the diluted solution at room temperature.

Hybridization Buffer with Formamide (for RNA)

40 mM PIPES (pH 6.8)
1 mM EDTA (pH 8.0)
0.4 M NaCl
80% (v/v) deionized formamide <!>

Use the disodium salt of PIPES to prepare the buffer, and adjust the pH to 6.4 with 1 N HCl.

Hybridization Buffer without Formamide (for RNA)

40 mM PIPES (pH 6.4)
0.1 mM EDTA (pH 8.0)
0.4 M NaCl

Use the disodium salt of PIPES to prepare the buffer, and adjust the pH to 6.4 with 1 N HCl.

Neutralization Buffer I (for Transfer of DNA to Uncharged Membranes)

1 M Tris-Cl (pH 7.4)
1.5 M NaCl

Neutralization Buffer II (for Alkaline Transfer of DNA to Nylon Membranes)

0.5 M Tris-Cl (pH 7.2)
1 M NaCl

Neutralizing Solution (for Neutral Transfer, Double-stranded DNA Targets Only)

0.5 M Tris-Cl (pH 7.4)
1.5 M NaCl

Prehybridization Solution (for Dot, Slot, and Northern Hybridization)

0.5 M sodium phosphate (pH 7.2)*
7% (w/v) SDS
1 mM EDTA (pH 7.0)

*0.5 M phosphate buffer is 134 g of $Na_2HPO_4 \cdot 7H_2O$, 4 ml of 85% H_3PO_4 <!>(concentrated phosphoric acid), H_2O to 1 liter.

Prehybridization and Hybridization Solutions

Prehybridization/Hybridization Solution (for Plaque/Colony Lifts)

50% (v/v) formamide (*optional*) <!>
6x SSC (or 6x SSPE)
0.05x BLOTTO

As an alternative to the above solution, use Church Buffer (please see recipe on p. A1.12). For advice on which hybridization solution to use, please see the panel on **PREHYBRIDIZATION AND HYBRIDIZATION SOLUTIONS** in Step 5 of Protocol 32 in Chapter 1. For advice on the use of formamide, please see the information panel on **FORMAMIDE AND ITS USES IN MOLECULAR CLONING** in Chapter 6.

Prehybridization/Hybridization Solution (for Hybridization in Aqueous Buffer)

6x SSC (or 6x SSPE)
5x Denhardt's reagent (see p. A1.15)
0.5% (w/v) SDS
1 µg/ml poly(A)
100 µg/ml salmon sperm DNA

Prehybridization/Hybridization Solution (for Hybridization in Formamide Buffers)

6x SSC (or 6x SSPE)
5x Denhardt's reagent (see A1.15)
0.5% (w/v) SDS
1 µg/ml poly(A)
100 µg/ml salmon sperm DNA
50% (v/v) formamide <!>

After a thorough mixing, filter the solution through a 0.45-µm disposable cellulose acetate membrane (Schleicher & Schuell Uniflow syringe membrane or equivalent). To decrease background when hybridizing under conditions of reduced stringency (e.g., 20–30% formamide), it is important to use formamide that is as pure as possible.

Prehybridization/Hybridization Solution (for Hybridization in Phosphate-SDS Buffer)

0.5 M phosphate buffer (pH 7.2)*
1 mM EDTA (pH 8.0)
7% (w/v) SDS
1% (w/v) bovine serum albumin

Use an electrophoresis grade of bovine serum albumin. No blocking agents or hybridization rate enhancers are required with this particular prehybridization/hybridization solution.
*0.5 M phosphate buffer is 134 g of $Na_2HPO_4 \cdot 7H_2O$, 4 ml of 85% H_3PO_4 <!> (concentrated phosphoric acid), H_2O to 1 liter.

20× SSC

Dissolve 175.3 g of NaCl and 88.2 g of sodium citrate in 800 ml of H_2O. Adjust the pH to 7.0 with a few drops of a 14 N solution of HCl <!>. Adjust the volume to 1 liter with H_2O. Dispense into aliquots. Sterilize by autoclaving. The final concentrations of the ingredients are 3.0 M NaCl and 0.3 M sodium acetate.

20× SSPE

Dissolve 175.3 g of NaCl, 27.6 g of $NaH_2PO_4 \cdot H_2O$, and 7.4 g of EDTA in 800 ml of H_2O. Adjust the pH to 7.4 with NaOH <!> (~6.5 ml of a 10 N solution). Adjust the volume to 1 liter with H_2O. Dispense into aliquots. Sterilize by autoclaving. The final concentrations of the ingredients are 3.0 M NaCl, 0.2 M NaH_2PO_4, and 0.02 M EDTA.

Blocking Agents

Blocking agents prevent ligands from sticking to surfaces. They are used in molecular cloning to stop nonspecific binding of probes in Southern, northern, and western blotting. If left to their own devices, these probes would bind tightly and nonspecifically to the supporting nitrocellulose or nylon membrane. Without blocking agents, it would be impossible to detect anything but the strongest target macromolecules.

No one knows for sure what causes nonspecific binding of probes. Hydrophobic patches, lignin impurities, excessively high concentrations of probe, overbaking or underbaking of nitrocellulose filters, and homopolymeric sequences in nucleic acid probes have all been blamed from time to time, together with a host of less likely culprits. Whatever the cause, the solution is generally simple: Treat the filters with a blocking solution containing a cocktail of substances that will compete with the probe for nonspecific binding sites on the solid support. Blocking agents work by brute force. They are used in high concentrations and generally consist of a cocktail of high-molecular-weight polymers (heparin, polyvinylpyrrolidine, nucleic acids), proteins (bovine serum albumin, nonfat dried milk), and detergents (SDS or Nonidet P-40). The following recommendations apply only to nylon and nitrocellulose filters. Charged nylon filters should be treated as described by the individual manufacturer.

Blocking Agents Used for Nucleic Acid Hybridization. Two blocking agents in common use in nucleic acid hybridization are Denhardt's reagent (Denhardt 1966) and BLOTTO (Bovine Lacto Transfer Technique Optimizer; Johnson et al. 1984). Usually, the filters carrying the immobilized target molecules are incubated with the blocking agents for an hour or two before the probe is added. In most cases, background hybridization is completely suppressed when filters are incubated with a blocking agent consisting of 6× SSC or SSPE containing 5× Denhardt's reagent, 1.0% SDS, and 100 mg/ml denatured, sheared salmon sperm DNA. This mixture should be used

whenever the ratio of signal to noise is expected to be low, for example, when carrying out northern analysis of low-abundance RNAs or Southern analysis of single-copy sequences of mammalian DNA. However, in most other circumstances (Grunstein-Hogness hybridization, Benton-Davis hybridization, Southern hybridization of abundant DNA sequences, etc.), a less expensive alternative is 6x SSC or SSPE containing 0.25–0.5% nonfat dried milk (BLOTTO; Johnson et al. 1984).

Blocking agents are usually included in both prehybridization and hybridization solutions when nitrocellulose filters are used. However, when the target nucleic acid is immobilized on nylon membranes, the blocking agents are often omitted from the hybridization solution. This is because high concentrations of protein are believed to interfere with the annealing of the probe to its target. Quenching of the hybridization signal by blocking agents is particularly noticeable when oligonucleotides are used as probes. This problem can often be solved by carrying out the hybridization step in a solution containing high concentrations of SDS (6–7%), sodium phosphate (0.4 M), bovine serum albumin (1%), and EDTA (0.02 M) (Church and Gilbert 1984).

Heparin is sometimes used instead of Denhardt's solution or BLOTTO when hybridization is carried out in the presence of the accelerator, dextran sulfate. It is used at a concentration of 500 µg/ml in hybridization solutions containing dextran sulfate. In hybridization solutions that do not contain dextran sulfate, heparin is used at a concentration of 50 µg/ml (Singh and Jones 1984). Heparin (Sigma porcine grade II or equivalent) is dissolved at a concentration of 50 mg/ml in 4x SSPE or SSC and stored at 4°C.

DENHARDT'S REAGENT

Denhardt's reagent is used for

- northern hybridization
- single-copy Southern hybridization
- hybridizations involving DNA immobilized on nylon membranes

Denhardt's reagent is usually made up as a 50x stock solution, which is filtered and stored at –20°C. The stock solution is diluted tenfold into prehybridization buffer (usually 6x SSC or 6x SSPE containing 1.0% SDS and 100 µg/ml denatured salmon sperm DNA). 50x Denhardt's reagent contains in H_2O (Denhardt 1966):

1% (w/v) Ficoll 400
1% (w/v) polyvinylpyrrolidone
1% (w/v) bovine serum albumin (Sigma, Fraction V)

BLOTTO <!>

BLOTTO is used for

- Grunstein-Hogness hybridization
- Benton-Davis hybridization
- all Southern hybridizations other than single-copy dot blots and slot blots

1x BLOTTO is 5% (w/v) nonfat dried milk dissolved in H_2O containing 0.02% sodium azide **<!>**. 1x BLOTTO is stored at 4°C and is diluted 10–25-fold into prehybridization buffer before use. BLOTTO should not be used in combination with high concentrations of SDS, which will cause the milk proteins to precipitate. If background hybridization is a problem, Nonidet P-40 may be added to a final concentration of 1% (v/v).

BLOTTO may contain high levels of RNase and should be treated with diethylpyrocarbonate (Siegel and Bresnick 1986) or heated overnight to 72°C (Monstein et al. 1992) when used in northern hybridizations and when RNA is used as a probe. BLOTTO is not as effective as Denhardt's solution when the target DNA is immobilized on nylon filters.

Blocking Agents Used for Western Blotting. The best and least expensive blocking reagent is nonfat dried milk (Johnson et al. 1984). It is easy to use and is compatible with all of the common immunological detection systems. The only time nonfat dried milk should not be used is when western blots are probed for proteins that may be present in milk.

One of the following recipes may be used to prepare blocking buffer. A blocking solution for western blots is phosphate-buffered saline containing 5% (w/v) nonfat dried milk, 0.01% Antifoam, and 0.02% sodium azide.

Blocking Buffer (TNT Buffer Containing a Blocking Agent)

10 mM Tris-Cl (pH 8.0)
150 mM NaCl
0.05% (v/v) Tween-20
blocking agent (1% [w/v] gelatin, 3% [w/v] bovine serum albumin, or 5% [w/v] nonfat dried milk)

Opinion about which of these blocking agents is best varies from laboratory to laboratory. We recommend carrying out preliminary experiments to determine which of them works best. Blocking buffer can be stored at 4°C and reused several times. Sodium azide <!> should be added to a final concentration of 0.05% (w/v) to inhibit the growth of microorganisms.

Extraction/Lysis Buffers and Solutions

Alkaline Lysis Solution I (Plasmid Preparation)

50 mM glucose
25 mM Tris-Cl (pH 8.0)
10 mM EDTA (pH 8.0)

Prepare Solution I from standard stocks in batches of ~100 ml, autoclave for 15 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle, and store at 4°C.

Alkaline Lysis Solution II (Plasmid Preparation)

0.2 N NaOH (freshly diluted from a 10 N stock) <!>
1% (w/v) SDS

Prepare Solution II fresh and use at room temperature.

Alkaline Lysis Solution III (Plasmid Preparation)

5 M potassium acetate	60.0 ml
glacial acetic acid <!>	11.5 ml
H_2O	28.5 ml

The resulting solution is 3 M with respect to potassium and 5 M with respect to acetate. Store the solution at 4°C and transfer it to an ice bucket just before use.

STET

10 mM Tris-Cl (pH 8.0)
0.1 M NaCl
1 mM EDTA (pH 8.0)
5% (v/v) Triton X-100

Make sure that the pH of STET is 8.0 after all ingredients are added. There is no need to sterilize STET before use.

Electrophoresis and Gel-loading Buffers

Commonly Used Electrophoresis Buffers

Buffer	Working Solution	Stock Solution/Liter
TAE	**1x**	**50x**
	40 mM Tris-acetate	242 g of Tris base
	1 mM EDTA	57.1 ml of glacial acetic acid <!>
		100 ml of 0.5 M EDTA (pH 8.0)
TBE[a]	**0.5x**	**5x**
	45 mM Tris-borate	54 g of Tris base
	1 mM EDTA	27.5 g of boric acid
		20 ml of 0.5 M EDTA (pH 8.0)
TPE	1x	**10x**
	90 mM Tris-phosphate	108 g of Tris base
	2 mM EDTA	15.5 ml of phosphoric acid <!>
		(85%, 1.679 g/ml)
		40 ml of 0.5 M EDTA (pH 8.0)
Tris-glycine[b]	**1x**	**5x**
	25 mM Tris-Cl	15.1 g of Tris base
	250 mM glycine	94 g of glycine (electrophoresis grade)
	0.1% SDS	50 ml of 10% SDS (electrophoresis grade)

[a]TBE is usually made and stored as a 5x or 10x stock solution. The pH of the concentrated stock buffer should be ~8.3. Dilute the concentrated stock buffer just before use and make the gel solution and the electrophoresis buffer from the same concentrated stock solution. Some investigators prefer to use more concentrated stock solutions of TBE (10x as opposed to 5x). However, 5x stock solution is more stable because the solutes do not precipitate during storage. Passing the 5x or 10x buffer stocks through a 0.22-μm filter can prevent or delay formation of precipitates.

[b]Use Tris-glycine buffers for SDS-polyacrylamide gels (see Appendix 8).

Specialized Electrophoresis Buffers

10x Alkaline Agarose Gel Electrophoresis Buffer

500 mM NaOH <!>
10 mM EDTA

Add 50 ml of 10 N NaOH and 20 ml of 0.5 M EDTA (pH 8.0) to 800 ml of H_2O and then adjust the final volume to 1 liter. Dilute the 10x alkaline agarose gel electrophoresis buffer with H_2O to generate a 1x working solution immediately before use. Use the same stock of 10x alkaline agarose gel electrophoresis buffer to prepare the alkaline agarose gel and the 1x working solution of alkaline electrophoresis buffer.

10x BPTE Electrophoresis Buffer

100 mM PIPES
300 mM Bis-Tris
10 mM EDTA

The final pH of the 10x buffer is ~6.5. The 10x buffer can be made by adding 3 g of PIPES (free acid), 6 g of Bis-Tris (free base), and 2 ml of 0.5 M EDTA to 90 ml of distilled H_2O and then treating the solution with diethylpyrocarbonate <!> (final concentration 0.1%; for more details, please see the information panel on **DIETHYLPYROCARBONATE** in Chapter 7).

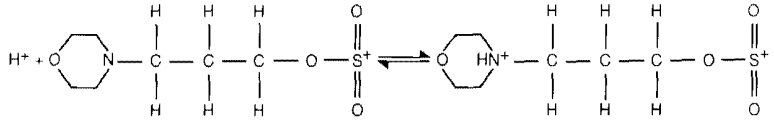
10x MOPS Electrophoresis Buffer

0.2 M MOPS (pH 7.0) <!>
20 mM sodium acetate
10 mM EDTA (pH 8.0)

Dissolve 41.8 g of MOPS in 700 ml of sterile DEPC-treated <!> H_2O. Adjust the pH to 7.0 with 2 N NaOH. Add 20 ml of DEPC-treated 1 M sodium acetate and 20 ml of DEPC-treated 0.5 M EDTA (pH 8.0). Adjust the volume of the solution to 1 liter with DEPC-treated H_2O. Sterilize the solution by passing it through a 0.45-μm Millipore filter, and store it at room temperature protected from light. The buffer yellows with age if it is exposed to light or is autoclaved. Straw-colored buffer works well, but darker buffer does not.

MOPS (3[*N*-MORPHOLINO]PROPANESULFONIC ACID

FW	pK_α (20°C)	ΔpK_α/°C	Molar strength of saturated solution at 0°C
209.3	7.15	−0.013	3.1



MOPS is one of the buffers developed by Robert Good's laboratories in the 1970s to facilitate isolation of chloroplasts and other plant organelles (for reviews, please see Good and Izawa 1972; Ferguson et al. 1980; please also see figure above). In molecular cloning, MOPS is a component of buffers used for the electrophoresis of RNA through agarose gels (Lehrach et al. 1977; Goldberg 1980).

TAFE Gel Electrophoresis Buffer

20 mM Tris-acetate (pH 8.2)
0.5 mM EDTA

Use acetic acid to adjust the pH of the Tris solution to 8.2, and use the free acid of EDTA, not the sodium salt. Concentrated solutions of TAFE buffer can also be purchased (e.g., from Beckman).

▲ **IMPORTANT** The TAFE gel electrophoresis buffer must be cooled to 14°C before use.

Gel-loading Buffers

6x Alkaline Gel-loading Buffer

300 mM NaOH <!>
6 mM EDTA
18% (w/v) Ficoll (Type 400, Pharmacia)
0.15% (w/v) bromocresol green
0.25% (w/v) xylene cyanol

Bromophenol Blue Solution (0.4%, w/v)

Dissolve 4 mg of solid bromophenol blue in 1 ml of sterile H_2O. Store the solution at room temperature.

TABLE A1-6 6× Gel-loading Buffers

BUFFER TYPE	6× BUFFER	STORAGE TEMPERATURE
I	0.25% (w/v) bromophenol blue 0.25% (w/v) xylene cyanol FF 40% (w/v) sucrose in H_2O	4°C
II	0.25% (w/v) bromophenol blue 0.25% (w/v) xylene cyanol FF 15% (w/v) Ficoll (Type 400; Pharmacia) in H_2O	room temperature
III	0.25% (w/v) bromophenol blue 0.25% (w/v) xylene cyanol FF 30% (v/v) glycerol in H_2O	4°C
IV	0.25% (w/v) bromophenol blue 40% (w/v) sucrose in H_2O	4°C

Bromophenol Blue Sucrose Solution

0.25% (w/v) bromophenol blue
40% (w/v) sucrose

Cresol Red Solution (10 mM)

Dissolve 4 mg of the sodium salt of cresol red (Aldrich) in 1 ml of sterile H_2O. Store the solution at room temperature.

10x Formaldehyde Gel-loading Buffer

50% (v/v) glycerol (diluted in DEPC-treated <!> H_2O)
10 mM EDTA (pH 8.0)
0.25% (w/v) bromophenol blue
0.25% (w/v) xylene cyanol FF

Formamide-loading Buffer

80% (w/v) deionized formamide <!>
10 mM EDTA (pH 8.0)
1 mg/ml xylene cyanol FF
1 mg/ml bromophenol blue

Purchase a distilled deionized preparation of formamide and store in small aliquots under nitrogen at –20°C. Alternatively, deionize reagent-grade formamide as described in Appendix 8.

RNA Gel-loading Buffer

95% (v/v) deionized formamide <!>
0.025% (w/v) bromophenol blue
0.025% (w/v) xylene cyanol FF
5 mM EDTA (pH 8.0)
0.025% (w/v) SDS

2x SDS Gel-loading Buffer

100 mM Tris-Cl (pH 6.8)
4% (w/v) SDS (electrophoresis grade)
0.2% (w/v) bromophenol blue
20% (v/v) glycerol
200 mM dithiothreitol or β-mercaptoethanol <!>

1x and 2x SDS gel-loading buffer lacking thiol reagents can be stored at room temperature. Add the thiol reagents from 1 M (dithiothreitol) or 14 M (β-mercaptoethanol) stocks just before the buffer is used.

2.5x SDS-EDTA Dye Mix

0.4% (v/v) SDS
30 mM EDTA
0.25% bromophenol blue
0.25% xylene cyanol FF
20% (w/v) sucrose

Special Buffers and Solutions

Elution Buffer (Qiagen)

50 mM Tris-Cl (pH 8.1–8.2)
1.4 M NaCl
15% (v/v) ethanol

KOH/Methanol Solution

This solution is for cleaning the glass plates used to cast sequencing gels. It is prepared by dissolving 5 g of KOH <!> pellets in 100 ml of methanol <!>. Store the solution at room temperature in a tightly capped glass bottle.

λ Annealing Buffer

100 mM Tris-Cl (pH 7.6)
10 mM $MgCl_2$

LB Freezing Buffer

36 mM K_2HPO_4 (anhydrous)
13.2 mM KH_2PO_4
1.7 mM sodium citrate
0.4 mM $MgSO_4 \cdot 7H_2O$
6.8 mM ammonium sulfate
4.4% (v/v) glycerol
in LB broth

LB freezing buffer (Zimmer and Verrinder Gibbins 1997) is best made by dissolving the salts in 100 ml of LB to the specified concentrations. Measure 95.6 ml of the resulting solution into a fresh container and then add 4.4 ml of glycerol. Mix the solution well and then sterilize by passing it through a 0.45-μm disposable Nalgene filter. Store the sterile freezing medium at a controlled room temperature (15–25°C).

$MgCl_2$–$CaCl_2$ Solution

80 mM $MgCl_2$
20 mM $CaCl_2$

P3 Buffer (Qiagen)

3 M potassium acetate (pH 5.5)

PEG–$MgCl_2$ Solution

40% (w/v) polyethylene glycol (PEG 8000)
30 mM $MgCl_2$

Dissolve 40 g of PEG 8000 in a final volume of 100 ml of 30 mM $MgCl_2$. Sterilize the solution by passing it through a 0.22-μm filter, and store it at room temperature.

QBT Buffer (Qiagen)

750 mM NaCl <!>
50 mM MOPS (pH 7.0) <!>
15% (v/v) isopropanol
0.15% (v/v) Triton X-100

Radioactive Ink <!>

Radioactive ink is made by mixing a small amount of ^{32}P with waterproof black drawing ink. We find it convenient to make the ink in three grades: very hot (>2000 cps on a hand-held minimonitor), hot (>500 cps on a hand-held minimonitor), and cool (>50 cps on a hand-held minimonitor). Use a fiber-tip pen to apply ink of the desired activity to the pieces of tape. Attach radioactive-warning tape to the pen, and store it in an appropriate place.

Sephacryl Equilibration Buffer

50 mM Tris-Cl (pH 8.0)
5 mM EDTA
0.5 M NaCl

SM and SM Plus Gelatin

Per liter:

NaCl	5.8 g
$MgSO_4 \cdot 7H_2O$	2 g
1 M Tris-Cl (pH 7.5)	50 ml
2% (w/v) gelatin solution	5 ml
H_2O	to 1 liter

Sterilize the buffer by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. After the solution has cooled, dispense 50-ml aliquots into sterile containers. SM may be stored indefinitely at room temperature. Discard each aliquot after use to minimize the chance of contamination.

Sorbitol Buffer

1 M sorbitol
0.1 M EDTA (pH 7.5)

STE

10 mM Tris-Cl (pH 8.0)
0.1 M NaCl
1 mM EDTA (pH 8.0)

Sterilize by autoclaving for 15 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. Store the sterile solution at 4°C.

10x TEN Buffer

0.1 M Tris-Cl (pH 8.0)
0.01 M EDTA (pH 8.0)
1 M NaCl

TES

10 mM Tris-Cl (pH 7.5)
1 mM EDTA (pH 7.5)
0.1% (w/v) SDS

Triton/SDS Solution

10 mM Tris-Cl (pH 8.0)
2% (v/v) Triton X-100
1% (w/v) SDS
100 mM NaCl
1 mM EDTA (pH 8.0)

Sterilize the solution by passing it through a 0.22-μm filter, and store it at room temperature.

Tris-Sucrose

50 mM Tris-Cl (pH 8.0)
10% (w/v) sucrose

Sterilize the solution by passing it through a 0.22-μm filter, and store it at room temperature. Solutions containing sucrose should not be autoclaved since the sugar tends to carbonize at high temperatures.

Wash Buffer (Qiagen)

50 mM MOPS-KOH <!> (pH 7.5–7.6)
0.75 M NaCl
15% (v/v) ethanol

When making this buffer, adjust the pH of a MOPS/NaCl solution before adding the ethanol.

Yeast Resuspension Buffer

50 mM Tris-Cl (pH 7.4)
20 mM EDTA (pH 7.5)

PREPARATION OF ORGANIC REAGENTS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Phenol

Most batches of commercial liquefied phenol <!> are clear and colorless and can be used in molecular cloning without redistillation. Occasionally, batches of liquefied phenol are pink or yellow, and these should be rejected and returned to the manufacturer. Crystalline phenol is not recommended because it must be redistilled at 160°C to remove oxidation products, such as quinones, that cause the breakdown of phosphodiester bonds or cause cross-linking of RNA and DNA.

Equilibration of Phenol

Before use, phenol must be equilibrated to a pH of >7.8 because the DNA partitions into the organic phase at acid pH. Wear gloves, full face protection, and a lab coat when carrying out this procedure.

1. Store liquefied phenol at –20°C. As needed, remove the phenol from the freezer, allow it to warm to room temperature, and then melt it at 68°C. Add hydroxyquinoline to a final concentration of 0.1%. This compound is an antioxidant, a partial inhibitor of RNase, and a weak chelator of metal ions (Kirby 1956). In addition, its yellow color provides a convenient way to identify the organic phase.

2. To the melted phenol, add an equal volume of buffer (usually 0.5 M Tris-Cl [pH 8.0] at room temperature). Stir the mixture on a magnetic stirrer for 15 minutes. Turn off the stirrer, and when the two phases have separated, aspirate as much as possible of the upper (aqueous) phase using a glass pipette attached to a vacuum line equipped with appropriate traps (please see Appendix 8, Figure A8-2).

3. Add an equal volume of 0.1 M Tris-Cl (pH 8.0) to the phenol. Stir the mixture on a magnetic stirrer for 15 minutes. Turn off the stirrer and remove the upper aqueous phase as described in Step 2. Repeat the extractions until the pH of the phenolic phase is >7.8 (as measured with pH paper).

4. After the phenol is equilibrated and the final aqueous phase has been removed, add 0.1 volume of 0.1 M Tris-Cl (pH 8.0) containing 0.2% β-mercaptoethanol <!>. The phenol solution may be stored in this form under 100 mM Tris-Cl (pH 8.0) in a light-tight bottle at 4°C for periods of up to 1 month.

Phenol:Chloroform:Isoamyl Alcohol (25:24:1)

A mixture consisting of equal parts of equilibrated phenol and chloroform:isoamyl alcohol <!> (24:1) is frequently used to remove proteins from preparations of nucleic acids. The chloroform denatures proteins and facilitates the separation of the aqueous and organic phases, and the isoamyl alcohol reduces foaming during extraction. Neither chloroform nor isoamyl alcohol requires treatment before use. The phenol:chloroform:isoamyl alcohol mixture may be stored under 100 mM Tris-Cl (pH 8.0) in a light-tight bottle at 4°C for periods of up to 1 month.

Deionization of Formamide

Many batches of reagent-grade formamide <!> are sufficiently pure to be used without further treatment. However, if any yellow color is present, deionize the formamide by stirring on a magnetic stirrer with Dowex XG8 mixed bed resin for 1 hour and filtering it twice through Whatman No. 1 paper. Store deionized formamide in small aliquots under nitrogen at –70°C.

Deionization of Glyoxal

Commercial stock solutions of glyoxal (40% or 6 M) contain various hydrated forms of glyoxal, as well as oxidation products such as glyoxylic acid, formic acid, and other compounds that can degrade RNA. These contaminants must be removed by treatment with a mixed-bed resin such as Bio-Rad AG-510-X8 until the indicator dye in the resin is exhausted. To deionize the glyoxal:

1. Immediately before use, mix the glyoxal with an equal volume mixed-bed ion-exchange resin (Bio-Rad AG-510-X8). Alternatively, pass the glyoxal through a small column of mixed bed resin, and then proceed to Step 3.
2. Separate the deionized material from the resin by filtration (e.g., through a Uniflow Plus filter; Schleicher & Schuell).
3. Monitor the pH of the glyoxal by mixing 200 µl of glyoxal with 2 µl of a 10 mg/ml solution of bromocresol green in H_2O, and observing the change in color. Bromocresol green is yellow at pH <4.8 and blue-green at pH >5.2.
4. Repeat the deionization process (Steps 1–2) until the pH of the glyoxal is >5.5.

Deionized glyoxal can be stored indefinitely at –20°C under nitrogen in tightly sealed microfuge tubes. Use each aliquot only once and then discard.

CHEMICAL STOCK SOLUTIONS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Acrylamide Solution (45% w/v)

acrylamide (DNA-sequencing grade) <!>	434 g
N,N'-methylenebisacrylamide <!>	16 g
H_2O	to 600 ml

Heat the solution to 37°C to dissolve the chemicals. Adjust the volume to 1 liter with distilled H_2O. Filter the solution through a nitrocellulose filter (e.g., Nalge, 0.45-micron pore size), and store the filtered solution in dark bottles at room temperature.

Actinomycin D (5 mg/ml)

Dissolve actinomycin D <!> in methanol <!> at a concentration of 5 mg/ml. Store the stock solution at –20°C in the dark. Please see the information panel on **ACTINOMYCIN D** in Chapter 7.

Adenosine Diphosphate (ADP) (1 mM)

Dissolve solid adenosine diphosphate in sterile 25 mM Tris-Cl (pH 8.0). Store small aliquots (~20 µl) of the solution at –20°C.

Ammonium Acetate (10 M)

To prepare a 1-liter solution, dissolve 770 g of ammonium acetate in 800 ml of H_2O. Adjust volume to 1 liter with H_2O. Sterilize by filtration. Alternatively, to prepare a 100-ml solution, dissolve 77 g of ammonium acetate in 70 ml of H_2O at room temperature. Adjust the volume to 100 ml with H_2O. Sterilize the solution by passing it through a 0.22-µm filter. Store the solution in tightly sealed bottles at 4°C or at room temperature. Ammonium acetate decomposes in hot H_2O and solutions containing it should not be autoclaved.

Ammonium Persulfate (10% w/v)

ammonium persulfate <!>	1 g
H_2O	to 10 ml

Dissolve 1 g ammonium persulfate in 10 ml of H_2O and store at 4°C. Ammonium persulfate decays slowly in solution, so replace the stock solution every 2–3 weeks. Ammonium persulfate is used as a catalyst for the copolymerization of acrylamide and bisacrylamide gels. The polymerization reaction is driven by free radicals generated by an oxido-reduction reaction in which a diamine (e.g., TEMED) is used as the adjunct catalyst (Chrambach and Rodbard 1972).

ATP (10 mM)

Dissolve an appropriate amount of solid ATP in 25 mM Tris-Cl (pH 8.0). Store the ATP solution in small aliquots at –20°C.

Calcium Chloride (2.5 M)

Dissolve 11 g of $CaCl_2 \cdot 6H_2O$ in a final volume of 20 ml of distilled H_2O. Sterilize the solution by passing it through a 0.22-µm filter. Store in 1-ml aliquots at 4°C.

Coomassie Staining Solution

Dissolve 0.25 g of Coomassie Brilliant Blue R-250 in 90 ml of methanol:H_2O <!> (1:1, v/v) and 10 ml of glacial acetic acid <!>. Filter the solution through a Whatman No. 1 filter to remove any particulate matter. Store at room temperature. Please see the entry on Coomassie Staining in Appendix 8.

Deoxyribonucleoside Triphosphates (dNTPs)

Dissolve each dNTP in H_2O at an approximate concentration of 100 mM. Use 0.05 M Tris base and a micropipette to adjust the pH of each of the solutions to 7.0 (use pH paper to check the pH). Dilute an aliquot of the neutralized dNTP appropriately, and read the optical density at the wavelengths given in the table below. Calculate the actual concentration of each dNTP. Dilute the solutions with H_2O to a final concentration of 50 mM dNTP. Store each separately at –70°C in small aliquots.

Base	Wavelength (nm)	Extinction Coefficient (E)($M^{-1}cm^{-1}$)
A	259	1.54×10^4
G	253	1.37×10^4
C	271	9.10×10^3
T	267	9.60×10^3

For a cuvette with a path length of 1 cm, absorbance = EM. 100 mM stock solutions of each dNTP are commercially available (Pharmacia).

For polymerase chain reactions (PCRs), adjust the dNTP solution to pH 8.0 with 2 N NaOH. Commercially available solutions of PCR-grade dNTPs require no adjustment.

Dimethylsulfoxide (DMSO)

Purchase a high grade of DMSO <!> (HPLC grade or better). Divide the contents of a fresh bottle into 1-ml aliquots in sterile tubes. Close the tubes tightly and store at –20°C. Use each aliquot only once and then discard.

Dithiothreitol (DTT, 1 M)

Dissolve 3.09 g of dithiothreitol in 20 ml of 0.01 M sodium acetate (pH 5.2) and sterilize by filtration. Dispense into 1-ml aliquots and store at –20°C. Under these conditions, dithiothreitol is stable to oxidation by air.

EDTA (0.5 M, pH 8.0)

Add 186.1 g of disodium EDTA·$2H_2O$ to 800 ml of H_2O. Stir vigorously on a magnetic stirrer. Adjust the pH to 8.0 with NaOH (~20 g of NaOH pellets<!>). Dispense into aliquots and sterilize by autoclaving. The disodium salt of EDTA will not go into solution until the pH of the solution is adjusted to ~8.0 by the addition of NaOH.

EGTA (0.5 M, pH 8.0)

EGTA is ethylene glycol bis(β-aminoethyl ether) *N,N,N′,N′*-tetraacteic acid. A solution of EGTA is made up essentially as described for EDTA above and sterilized by either autoclaving or filtering. Store the sterile solution at room temperature.

Ethidium Bromide (10 mg/ml)

Add 1 g of ethidium bromide <!> to 100 ml of H_2O. Stir on a magnetic stirrer for several hours to ensure that the dye has dissolved. Wrap the container in aluminum foil or transfer the solution to a dark bottle and store at room temperature.

Gelatin (2% w/v)

Add 2 g of gelatin to a total volume of 100 ml of H_2O and autoclave the solution for 15 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle.

Glycerol (10% v/v)

Dilute 1 volume of molecular-biology-grade glycerol in 9 volumes of sterile pure H_2O. Sterilize the solution by passing it through a prerinsed 0.22-µm filter. Store in 200-ml aliquots at 4°C.

IPTG (20% w/v, 0.8 M)

IPTG is isopropylthio-β-D-galactoside. Make a 20% solution of IPTG by dissolving 2 g of IPTG in 8 ml of distilled H_2O. Adjust the volume of the solution to 10 ml with H_2O and sterilize by passing it through a 0.22-µm disposable filter. Dispense the solution into 1-ml aliquots and store them at –20°C.

KCl (4 M)

Dissolve an appropriate amount of solid KCl in H_2O, autoclave for 20 minutes on liquid cycle and store at room temperature. Ideally, this solution should be divided into small (~100 µl) aliquots in sterile tubes and each aliquot thereafter used one time.

Lithium Chloride (LiCl, 5 M)

Dissolve 21.2 g of LiCl in a final volume of 90 ml of H_2O. Adjust the volume of the solution to 100 ml with H_2O. Sterilize the solution by passing it through a 0.22-µm filter, or by autoclaving for 15 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. Store the solution at 4°C.

$MgCl_2 \cdot 6H_2O$ (1 M)

Dissolve 203.3 g of $MgCl_2 \cdot 6H_2O$ in 800 ml of H_2O. Adjust the volume to 1 liter with H_2O. Dispense into aliquots and sterilize by autoclaving. $MgCl_2$ is extremely hygroscopic. Buy small bottles (e.g., 100 g) and do not store opened bottles for long periods of time.

$MgSO_4$ (1 M)

Dissolve 12 g of $MgSO_4$ in a final volume of 100 ml of H_2O. Sterilize by autoclaving or filter sterilization. Store at room temperature.

Maltose (20% w/v)

Dissolve 20 g of maltose in a final volume of 100 ml of H_2O and sterilize by passing it through a 0.22-µm filter. Store the sterile solution at room temperature.

NaOH (10 N)

The preparation of 10 N NaOH **<!>** involves a highly exothermic reaction, which can cause breakage of glass containers. Prepare this solution with extreme care in plastic beakers. To 800 ml of H_2O, slowly add 400 g of NaOH pellets **<!>**, stirring continuously. As an added precaution, place the beaker on ice. When the pellets have dissolved completely, adjust the volume to 1 liter with H_2O. Store the solution in a plastic container at room temperature. Sterilization is not necessary.

NaCl (Sodium Chloride, 5 M)

Dissolve 292 g of NaCl in 800 ml of H_2O. Adjust the volume to 1 liter with H_2O. Dispense into aliquots and sterilize by autoclaving. Store the NaCl solution at room temperature.

PEG 8000

Working concentrations of PEG <!> range from 13% to 40% (w/v). Prepare the appropriate concentration by dissolving PEG 8000 in sterile H_2O, warming if necessary. Sterilize the solution by passing it through a 0.22-μm filter. Store the solution at room temperature.

Polyethylene glycol (PEG) is a straight-chain polymer of a simple repeating unit $H(OCH_2\ CH_2)_nOH$. PEG is available in a range of molecular weights whose names reflect the number (*n*) of repeating units in each molecule. In PEG 400, for example, n = 8–9, whereas in PEG 4000, n ranges from 68 to 84. PEG induces macromolecular crowding of solutes in aqueous solution (Zimmerman and Minton 1993) and has a range of uses in molecular cloning, including:

- ***Precipitation of DNA molecules according to their size.*** The concentration of PEG required for precipitation is in inverse proportion to the size of the DNA frgaments (Lis and Schleif 1975a,b; Ogata and Gilbert 1977; Lis 1980); please see Chapter 1, Protocol 8, and Chapter 2, Protocol 6.
- ***Precipitation and purification of bacteriophage particles*** (Yamamoto et al. 1970).
- ***Increasing the efficiency of reassociation of complementary chains*** of nucleic acids during hybridization, blunt-end ligation of DNA molecules, and end-labeling of DNA with bacteriophage T4 polynucleotide kinase (Zimmerman and Minton 1993; please see the information panel on **CONDENSING AND CROWDING REAGENTS** in Chapter 1).
- ***Fusion of cultured cells with bacterial protoplasts*** (Schaffner 1980; Rassoulzadegan et al. 1982).

Potassium Acetate (5 M)

5 M potassium acetate	60 ml
glacial acetic acid <!>	11.5 ml
H_2O	28.5 ml

The resulting solution is 3 M with respect to potassium and 5 M with respect to acetate. Store the buffer at room temperature.

SDS (20% w/v)

Also called sodium lauryl sulfate. Dissolve 200 g of electrophoresis-grade SDS <!> in 900 ml of H_2O. Heat to 68°C and stir with a magnetic stirrer to assist dissolution. If necessary, adjust the pH to 7.2 by adding a few drops of concentrated HCl <!>. Adjust the volume to 1 liter with H_2O. Store at room temperature. Sterilization is not necessary. Do not autoclave.

Silver Stain. ***Please see staining section (Appendix 8).***

Sodium Acetate (3 M, pH 5.2 and pH 7.0)

Dissolve 408.3 g of sodium acetate·$3H_2O$ in 800 ml of H_2O. Adjust the pH to 5.2 with glacial acetic acid <!> or adjust the pH to 7.0 with dilute acetic acid. Adjust the volume to 1 liter with H_2O. Dispense into aliquots and sterilize by autoclaving.

Spermidine (1 M)

Dissolve 1.45 g of spermidine (free-base form) in 10 ml of deionized H_2O and sterilize by passing it through a 0.22-μm filter. Store the solution in small aliquots at –20°C. Make a fresh stock solution of this reagent every month.

SYBR Gold Staining Solution

SYBR Gold <!> (Molecular Probes) is supplied as a stock solution of unknown concentration in dimethylsulfoxide. Agarose gels are stained in a working solution of SYBR Gold, which is a 1:10,000 dilution of SYBR Gold nucleic acid stain in electrophoresis buffer. Prepare working stocks of SYBR Gold daily and store in the dark at regulated room temperature. For a discussion of staining agarose gels, please see Chapter 5, Protocol 2.

Trichloroacetic Acid (TCA; 100% solution)

To a previously unopened bottle containing 500 g of TCA <!>, add 227 ml of H_2O. The resulting solution will contain 100% (w/v) TCA.

X-gal Solution (2% w/v)

X-gal is 5-bromo-4-chloro-3-indolyl-β-D-galactoside. Make a stock solution by dissolving X-gal in dimethylformamide <!> at a concentration of 20 mg/ml solution. Use a glass or polypropylene tube. Wrap the tube containing the solution in aluminum foil to prevent damage by light and store at –20°C. It is not necessary to sterilize X-gal solutions by filtration. Please see the information panel on **X-GAL** in Chapter 1.

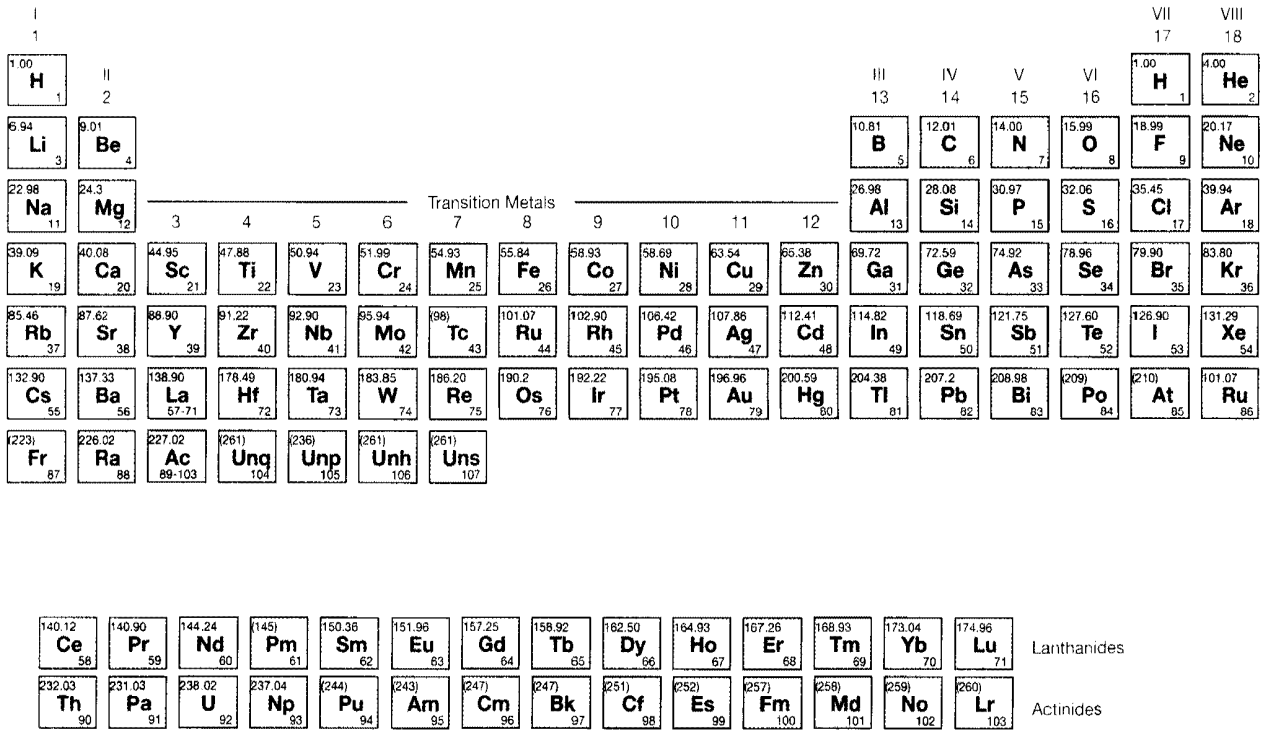


FIGURE A1-1 Periodic Table

Numbers in parentheses are the mass numbers of the most stable isotope of that element.

REAGENTS AND BUFFERS INDEX

Appendix 2

Media

LIQUID MEDIA FOR *E. COLI*

▲ **IMPORTANT** Use distilled deionized H_2O in all recipes. Unless otherwise stated, sterile media can be stored at room temperature.

GYT Medium (Tung and Chow 1995)

10% (v/v) glycerol
0.125% (w/v) yeast extract
0.25% (w/v) tryptone

Sterilize the medium by passing it through a prerinsed 0.22-µm filter. Store in 2.5-ml aliquots at 4°C.

LB Medium (Luria-Bertani Medium)

Per liter:

To 950 ml of deionized H_2O, add:	
tryptone	10 g
yeast extract	5 g
NaCl	10 g

Shake until the solutes have dissolved. Adjust the pH to 7.0 with 5 N NaOH (~0.2 ml). Adjust the volume of the solution to 1 liter with deionized H_2O. Sterilize by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle.

M9 Minimal Medium

Per liter:

To 750 ml of sterile H_2O (cooled to 50°C or less), add:	
5x M9 salts*	200 ml
1 M $MgSO_4$	2 ml
20% solution of the appropriate carbon source (e.g., 20% glucose)	20 ml
1 M $CaCl_2$	0.1 ml
sterile deionized H_2O	to 980 ml

If necessary, supplement the M9 medium with stock solutions of the appropriate amino acids and vitamins.

*5x M9 salts is made by dissolving the following salts in deionized H_2O to a final volume of 1 liter:

$Na_2HPO_4 \cdot 7H_2O$	64 g
KH_2PO_4	15 g
NaCl	2.5 g
NH_4Cl	5.0 g

Divide the salt solution into 200-ml aliquots and sterilize by autoclaving for 15 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle.

Prepare the $MgSO_4$ and $CaCl_2$ solutions separately, sterilize by autoclaving, and add the solutions after diluting the 5x M9 salts to 980 ml with sterile H_2O. Sterilize the glucose by passing it through a 0.22-µm filter before it is added to the diluted M9 salts.

When using *E. coli* strains that carry a deletion of the proline biosynthetic operon [Δ(*lac-proAB*)] in the bacterial chromosome and the complementing *proAB* genes on the F′ plasmid, supplement the M9 minimal medium with the following:

0.4% (w/v) glucose (dextrose)
5 mM $MgSO_4 \cdot 7H_2O$
0.01% thiamine

NZCYM Medium

Per liter:

To 950 ml of deionized H_2O, add:

NZ amine	10 g
NaCl	5 g
yeast extract	5 g
casamino acids	1 g
$MgSO_4 \cdot 7H_2O$	2 g

Shake until the solutes have dissolved. Adjust the pH to 7.0 with 5 N NaOH (~0.2 ml). Adjust the volume of the solution to 1 liter with deionized H_2O. Sterilize by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle.

NZ amine: Casein hydrolysate enzymatic (ICN Biochemicals). NZCYM, NZYM, and NZM are also available as dehydrated media from BD Biosciences.

NZYM Medium

NZYM medium is identical to NZCYM medium, except that casamino acids are omitted.

NZM Medium

NZM medium is identical to NZYM medium, except that yeast extract is omitted.

SOB Medium

Per liter:

To 950 ml of deionized H_2O, add:

tryptone	20 g
yeast extract	5 g
NaCl	0.5 g

Shake until the solutes have dissolved. Add 10 ml of a 250 mM solution of KCl. (This solution is made by dissolving 1.86 g of KCl in 100 ml of deionized H_2O.) Adjust the pH of the medium to 7.0 with 5 N NaOH **<!>** (~0.2 ml). Adjust the volume of the solution to 1 liter with deionized H_2O. Sterilize by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. Just before use, add 5 ml of a sterile solution of 2 M $MgCl_2$. (This solution is made by dissolving 19 g of $MgCl_2$ in 90 ml of deionized H_2O. Adjust the volume of the solution to 100 ml with deionized H_2O and sterilize by autoclaving for 20 minutes at 15 psi [1.05 kg/cm^2] on liquid cycle.)

SOC Medium

SOC medium is identical to SOB medium, except that it contains 20 mM glucose. After the SOB medium has been autoclaved, allow it to cool to 60°C or less. Add 20 ml of a sterile 1 M solution of glucose. (This solution is made by dissolving 18 g of glucose in 90 ml of deionized H_2O. After the sugar has dissolved, adjust the volume of the solution to 100 ml with deionized H_2O and sterilize by passing it through a 0.22-μm filter.)

Terrific Broth (also known as TB; Tartof and Hobbs 1987)

Per liter:

To 900 ml of deionized H_2O, add:

tryptone	12 g
yeast extract	24 g
glycerol	4 ml

Shake until the solutes have dissolved and then sterilize by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. Allow the solution to cool to 60°C or less, and then add 100 ml of a sterile solution of 0.17 M KH_2PO_4, 0.72 M K_2HPO_4. (This solution is made by dissolving 2.31 g of KH_2PO_4 and 12.54 g of K_2HPO_4 in 90 ml of deionized H_2O. After the salts have dissolved, adjust the volume of the solution to 100 ml with deionized H_2O and sterilize by autoclaving for 20 minutes at 15 psi [1.05 kg/cm^2] on liquid cycle.)

2× YT Medium

Per liter:

To 900 ml of deionized H_2O, add:

tryptone	16 g
yeast extract	10 g
NaCl	5 g

Shake until the solutes have dissolved. Adjust the pH to 7.0 with 5 N NaOH <!>. Adjust the volume of the solution to 1 liter with deionized H_2O. Sterilize by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle.

MEDIA CONTAINING AGAR OR AGAROSE

▲ **IMPORTANT** Use distilled deionized H_2O in all recipes.

Prepare liquid media according to the recipes given above. Just before autoclaving, add one of the following:

Bacto Agar (for plates)	15 g/liter
Bacto Agar (for top agar)	7 g/liter
agarose (for plates)	15 g/liter
agarose (for top agarose)	7 g/liter

Sterilize by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. When the medium is removed from the autoclave, swirl it gently to distribute the melted agar or agarose evenly throughout the solution. *Be careful!* The fluid may be superheated and may boil over when swirled. Allow the medium to cool to 50–60°C before adding thermolabile substances (e.g., antibiotics). To avoid producing air bubbles, mix the medium by swirling. Plates can then be poured directly from the flask; allow ~30–35 ml of medium per 90-mm plate. To remove bubbles from medium in the plate, flame the surface of the medium with a Bunsen burner before the agar or agarose hardens. Set up a color code (e.g., two red stripes for LB-ampicillin plates; one black stripe for LB plates, etc.) and mark the edges of the plates with the appropriate colored markers.

When the medium has hardened completely, invert the plates and store them at 4°C until needed. The plates should be removed from storage 1–2 hours before they are used. If the plates are fresh, they will "sweat" when incubated at 37°C. When this condensation drops on the agar/agarose surface, it allows bacterial colonies or bacteriophage plaques to spread and increases the chances of cross-contamination. This problem can be avoided by wiping off the condensation from the lids of the plates and then incubating the plates for several hours at 37°C in an inverted position before they are used. Alternatively, remove the liquid by shaking the lid with a single, quick motion. To minimize the possibility of contamination, hold the open plate in an inverted position while removing the liquid from the lid.

STORAGE MEDIA

▲ **IMPORTANT** Use distilled deionized H_2O in all recipes.

Liquid Cultures

Bacteria growing on plates, or in liquid culture, can be stored in aliquots of LB medium containing 30% (v/v) sterile glycerol. Aliquots of 1 ml of LB with glycerol should be prepared and vortexed to ensure that the glycerol is completely dispersed. Alternatively, bacterial strains may be stored in LB freezing buffer:

LB freezing buffer:
36 mM K_2HPO_4 (anhydrous)
13.2 mM KH_2PO_4
1.7 mM sodium citrate
0.4 mM $MgSO_4 \cdot 7H_2O$
6.8 mM ammonium sulfate
4.4% (v/v) glycerol
in LB

LB freezing buffer (Zimmer and Verrinder Gibbins 1997) is best made by dissolving the salts in 100 ml of LB to the specified concentrations. Measure 95.6 ml of the resulting solution into a fresh container, and then add 4.4 ml of glycerol. Mix the solution well and then sterilize by passing it through a 0.45-µm disposable Nalgene filter. For more information on storage of bacterial cultures, please see Appendix 8.

Stab Cultures

Prepare stab cultures in glass vials (2–3 ml) with screw-on caps fitted with rubber gaskets. Add molten LB agar until the vials are two-thirds full. Autoclave the partially filled vials (with their caps loosely screwed on) for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. Remove the vials from the autoclave, let them cool to room temperature, and then tighten the caps. Store the vials at room temperature until needed.

ANTIBIOTICS

TABLE A2-1 Commonly Used Antibiotic Solutions

	Stock Solution[a]		Working Concentration	
	Concentration	Storage	Stringent Plasmids	Relaxed Plasmids
Ampicillin	50 mg/ml in H_2O	–20°C	20 µg/ml	50 µg/ml
Carbenicillin	50 mg/ml in H_2O	–20°C	20 µg/ml	60 µg/ml
Chloramphenicol	34 mg/ml in ethanol	–20°C	25 µg/ml	170 µg/ml
Kanamycin	10 mg/ml in H_2O	–20°C	10 µg/ml	50 µg/ml
Streptomycin	10 mg/ml in H_2O	–20°C	10 µg/ml	50 µg/ml
Tetracycline[b]	5 mg/ml in ethanol	–20°C	10 µg/ml	50 µg/ml

Magnesium ions are antagonists of tetracycline. Use media without magnesium salts (e.g., LB medium) for selection of bacteria resistant to tetracycline.

[a]Sterilize stock solutions of antibiotics dissolved in H_2O by filtration through a 0.22-µm filter.

[b]Antibiotics dissolved in ethanol need not be sterilized. Store solutions in light-tight containers.

TABLE A2-2 Antibiotic Modes of Action

Antibiotic	Molecular Weight	Mode of Action	Further Information
Actinomycin C_1 (actinomycin D)	1255.4	Inhibits synthesis of RNA by binding to double-stranded DNA.	
Amphotericin	924.1	Broad-spectrum antifungal agent from *Streptomyces*.	
Ampicillin	349.4	Inhibits cell-wall synthesis by interfering with peptidoglycan cross-linking.	Please see the information panel on **AMPICILLIN AND CARBENICILLIN** at the end of Chapter 1.
Bleomycin	n.a.[a]	Inhibits DNA synthesis; cleaves single-stranded DNA.	
Carbenicillin (disodium salt)	422.4	Inhibits bacterial wall synthesis.	
Chloramphenicol	323.1	Inhibits translation by blocking peptidyl transferase on the 50S ribosomal subunit; at higher concentrations can inhibit eukaryotic DNA synthesis.	Please see the information panel on **CHLORAMPHENICOL** at the end of Chapter 1.
Geneticin (G418 geneticin disulfate)	692.7	Aminoglycoside is toxic to a broad range of cell types (bacterial, higher plant, yeast, mammaliam, protozoans, helminths); used in selection of eukaryotic cells transformed with neomycin resistance genes.	
Gentamycin	692.7	Inhibits protein synthesis by binding to L6 protein of the 50S ribosomal subunit.	
Hygromycin B	527.5	Inhibits protein synthesis.	
Kanamycin monosulfate	582.6	Broad-spectrum antibiotic; binds to 70S ribosomal subunit and inhibits growth of gram-positive and gram-negative bacteria and mycoplasmas.	Please see the information panel on **KANAMYCINS** at the end of Chapter 1.
Methotrexate	454.45	A folic acid analog; a powerful inhibitor of the enzyme dihydrofolate reductase.	
Mitomycin C	334.33	Inhibits DNA synthesis; antibacterial to gram-positive, gram-negative, and acid-fast bacilli.	
Neomycin B sulfate	908.9	Binds to 30S ribosomal subunit and inhibits bacterial protein synthesis.	
Novobiocin sodium salt	634.62	Bacteriostatic antibiotic; inhibits growth of gram-positive bacteria.	
Penicillin G sodium salt	356.4	Inhibits peptidoglycan synthesis in bacterial cell walls.	
Puromycin dihydrochloride	544.4	Inhibits protein synthesis by acting as an analog of aminoacyl tRNA (causes premature chain termination).	
Rifampicin	823.0	Strongly inhibits prokaryotic RNA polymerase and mammalian RNA polymerase to a lesser degree.	
Streptomycin sulfate	1457.4	Inhibits protein synthesis; binds to 30S ribosomal subunit.	
Tetracycline hydrochloride	480.9	Inhibits bacterial protein synthesis; blocks ribosomal binding of aminoacyl-tRNA.	Please see the information panel on **TETRACYCLINE** at the end of Chapter 1.

[a]n.a. indicates not available.

SOLUTIONS FOR WORKING WITH BACTERIOPHAGE λ

▲ **IMPORTANT** Use distilled deionized H_2O in all recipes.

Maltose

Maltose, an inducer of the gene (*lamB*) that encodes the bacteriophage λ receptor, is often added to the medium during growth of bacteria that are to be used for plating bacteriophage λ. Add 1 ml of a sterile 20% maltose solution for every 100 ml of medium. For a further discussion of the use of maltose, please see the Materials list in Chapter 2, Protocol 1. Make up a sterile 20% stock solution of maltose as follows:

maltose	20 g
H_2O	to 100 ml

Sterilize the solution by passing it through a 0.22-μm filter. Store the sterile solution at room temperature.

SM

This buffer is used for storage and dilution of bacteriophage λ stocks.

Per liter:

NaCl	5.8 g
$MgSO_4 \cdot 7H_2O$	2 g
1 M Tris-Cl (pH 7.5)	50 ml
2% gelatin solution	5 ml
H_2O	to 1 liter

Sterilize the buffer by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. After the solution has cooled, dispense 50-ml aliquots into sterile containers. SM may be stored indefinitely at room temperature.

A 2% gelatin solution is made by adding 2 g of gelatin to a total volume of 100 ml of H_2O and autoclaving the solution for 15 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle.

TM

Per liter:

1 M Tris-Cl (pH 7.5)	50 ml
$MgSO_4 \cdot 7H_2O$	2 g
H_2O	to 1 liter

Sterilize the buffer by autoclaving for 20 minutes at 15 psi (1.05 kg/cm^2) on liquid cycle. After the solution has cooled, dispense 50-ml aliquots into sterile containers. TM may be stored indefinitely at room temperature.

MEDIA FOR THE PROPAGATION AND SELECTION OF YEAST*

▲ **CAUTION Please see Appendix 12 for appropriate handling of materials marked with <!>,**

▲ **IMPORTANT** Use distilled deionized H_2O in all recipes. Unless otherwise stated, media and solutions are sterilized by autoclaving at 15 psi (1.05 kg/cm^2) for 15–20 minutes.

Complete Minimal (CM) or Synthetic Complete (SC) and Drop-out Media

To test the growth requirements of strains, it is useful to have media in which each of the commonly encountered auxotrophies is supplemented except the one of interest (drop-out media). Dry growth supplements are stored premixed. CM (or SC) is a medium in which the drop-out mix contains all possible supplements (i.e., nothing is "dropped out").

yeast nitrogen base without amino acids*	6.7 g
glucose	20 g
Bacto Agar	20 g
drop-out mix	2 g
H_2O	to 1000 ml

*Yeast nitrogen base without amino acids (YNB) is sold either with or without ammonium sulfate. This recipe is for YNB with ammonium sulfate. If the bottle of YNB is lacking ammonium sulfate, add 5 g of ammonium sulfate and only 1.7 g of YNB.

Drop-out Mix

Combine the appropriate ingredients, minus the relevant supplements, and mix in a sealed container. Turn the container end-over-end for at least 15 minutes; add a few clean marbles to help mix the solids.

Adenine	0.5 g
Alanine	2.0 g
Arginine	2.0 g
Asparagine	2.0 g
Aspartic acid	2.0 g
Cysteine	2.0 g
Glutamine	2.0 g
Glutamic acid	2.0 g
Glycine	2.0 g
Histidine	2.0 g
Inositol	2.0 g
Isoleucine	2.0 g
Leucine	10.0 g
Lysine	2.0 g
Methionine	2.0 g
para-Aminobenzoic acid	0.2 g
Phenylalanine	2.0 g
Proline	2.0 g
Serine	2.0 g
Threonine	2.0 g
Tryptophan	2.0 g
Tyrosine	2.0 g
Uracil	2.0 g
Valine	2.0 g

*Reprinted from Adams et al. (1998).

TABLE A2-3 Components of Supplemented Minimal Media

Constituent	Stock Concentration (g/100 ml)	Volume for 1 Liter of Stock of Medium (ml)	Final Concentration in Medium (mg/liter)	Volume of Stock to Spread on Plate (ml)
Adenine sulfate	0.2[a]	10	20	0.2
Uracil	0.2[a]	10	20	0.2
L-Tryptophan	1	2	20	0.1
L-Histidine HCl	1	2	20	0.1
L-Arginine HCl	1	2	20	0.1
L-Methionine	1	2	20	0.1
L-Tyrosine	0.2	15	30	0.2
L-Leucine	1	10	100	0.1
L-Isoleucine	1	3	30	0.1
L-Lysine HCl	1	3	30	0.1
L-Phenylalanine	1[a]	5	50	0.1
L-Glutamic acid	1[a]	10	100	0.2
L-Aspartic acid	1[a,b]	10	100	0.2
L-Valine	3	5	150	0.1
L-Threonine	4[a,b]	5	200	0.1
L-Serine	8	5	400	0.1

[a]Store at room temperature.
[b]Add after autoclaving the medium.

Supplemented Minimal Medium (SMM)

SMM is SD (please see below) to which various growth supplements have been added. These solutions can then be stored for extended periods. Some should be stored at room temperature, in order to prevent precipitation, whereas the other solutions may be refrigerated. Wherever applicable, HCl salts of amino acids are preferred.

Prepare the medium by adding the appropriate volumes of the stock solutions to the ingredients of SD medium and then adjusting the total volume to 1 liter with distilled H_2O. Add threonine and aspartic acid solutions separately to the medium after it is autoclaved.

Alternatively, it is often more convenient to prepare the medium by spreading a small quantity of the supplement(s) on the surface of an SD plate. Allow the solution(s) to then dry thoroughly onto the plate before inoculating it with yeast strains.

Table A2-3 provides the concentrations of the stock solutions, the volume of stock solution necessary for mixing 1 liter of medium, the volume of stock solution to spread on SD plates, and the final concentration of each constituent in SMM.

Synthetic Dextrose Minimal Medium (SD)

SD is a synthetic minimal medium containing salts, trace elements, vitamins, a nitrogen source (yeast nitrogen base without amino acids), and glucose.

yeast nitrogen base without amino acids*	6.7 g
glucose	20 g
Bacto Agar	20 g
H_2O	1000 ml

*Please see note to recipe for CM on p. A2.9.

X-Gal Indicator Plates for Yeast

Because 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal <!>) does not work for yeast at the normal acidic pH of SD medium, a medium at neutral pH medium is used. This

choice is clearly a trade-off as many yeast strains will not grow well at neutral pH. For each liter of X-gal indicator plates, prepare the following solutions:

Solution I

10x phosphate-buffered stock solution	100 ml
1000x mineral stock solution (see recipe below)	1 ml
drop-out mix (see Drop-out Mix above)	2 g

Adjust the volume to 450 ml with distilled H_2O if the medium is to contain glucose or to 400 ml if it is to contain galactose.

10x Phosphate-buffered Stock Solution

KH_2PO_4 (1 M)	136.1 g
$(NH_4)_2SO_4$ (0.15 M)	19.8 g
KOH (0.75 N) <!>	42.1 g
H_2O	1000 ml

Adjust the pH to 7.0 and autoclave.

1000x Mineral Stock Solution

$FeCl_3$ (2 mM)	32 mg
$MgSO_4 \cdot 7H_2O$ (0.8 M)	19.72 g
H_2O	100 ml

Autoclave and store. This solution will develop a fine yellow precipitate, which should be resuspended before use.

Solution II

Mix in a 2-liter flask:

Bacto Agar	20 g
H_2O	500 ml

- Autoclave Solutions I and II separately.
- After cooling to below 65°C, add the following components to Solution I:

glucose or other sugar to a final concentration of 2%	
X-gal (20 mg/ml dissolved in dimethylformamide <!>)	2 ml
100x vitamin stock solution	10 ml

- Include any other heat-sensitive supplements at this point.
- Mix Solutions I and II together and pour ~30 ml/plate.

100x Vitamin Stock Solution

thiamine (0.04 mg/ml)	4 mg
biotin (2 µg/ml)	0.2 mg
pyridoxine (0.04 mg/ml)	4 mg
inositol (0.2 mg/ml)	20 mg
pantothenic acid (0.04 mg/ml)	4 mg
H_2O	100 ml

Sterilize by passing the solution through a 0.22-µm filter.

X-Gal Plates for Lysed Yeast Cells on Filters

These plates are used for checking β-galactosidase activity in cells that have been lysed and are immobilized on Whatman 3MM filters.

Bacto Agar	20 g
1 M Na_2HPO_4	57.7 ml
1 M NaH_2PO_4	42.3 ml
$MgSO_4$	0.25 g
H_2O	900 ml

After autoclaving, add 6 ml of X-gal solution (20 mg/ml in dimethylformamide).

YPD (YEPD) Medium

YPD is a complex medium for routine growth of yeast.

yeast extract	10 g
peptone	20 g
glucose	20 g
H_2O	to 1000 ml

To prepare plates, add 20 g of Bacto Agar (2%) before autoclaving.

Appendix 3

Vectors and Bacterial Strains

TABLE A3-1 Vectors

PLASMIDS/PHAGEMIDS	SUPPLIER	APPLICATION	SELECTION	PROMOTERS
pACYC	NEB	Cloning (low copy number)	Kan, Amp	
pBC KS +/–	Stratagene	DNA sequencing; in vitro transcription	Cam	T3, T7
pBluescript II	Stratagene	DNA sequencing; in vitro transcription	Amp	T3, T7
pBluescript SK⁻ (Chapter 11, Figure 11-7)	Stratagene	Cloning; expression; in vitro transcription	Amp	T3, T7
pBR313	ATCC	Cloning	Amp	
pBR322 (Chapter 8, Figure 8-11)	Various (NEB)	Molecular-weight markers; subcloning (low copy number)	Amp/Tet	
pBR327	ATCC	Cloning	Amp	(derived pBR322)
pBS	Stratagene	DNA rescue; DNA sequencing; expression of fusion proteins	Amp	T7, T3
pCR1000	Invitrogen	Cloning	Amp	(derived pUC19)
pET series (Chapter 15)	Various (Novagen)	Protein expression; protein purification	Amp	T7
pGEM series	Promega	DNA sequencing; in vitro mutagenesis		T7/SP6
pGEM T	Promega	Cloning PCR products	Amp	T7/SP6
PGEM-3Z (Chapter 17, Figure 17-5)	Promega	General cloning; in vitro transcription	Amp	
pGEMZF	Promega	General cloning; in vitro transcription; production of single-stranded DNA	Amp	
pGEX series	ATCC	Cloning; GST fusion vector	Amp	P*tac*
pMAL series (Chapter 15)	NEB	Protein expression	Amp	P*tac*
pMB1	–	Cloning	Tet	
pMB1/colE1	–	Cloning	Tet	
pMOB45	ATCC	Cloning	Cam	
pPCR-Script Direct	Stratagene			
pSC101	ATCC	Cloning	Tet	
pSE280	Invitrogen			
PSP18/19	–	Cloning	Amp	T7/SP6
pSPORT1	Life Technologies	cDNA cloning; in vitro transcription; and subtraction library procedures	Amp	T7/SP6, *lacI*
pTrx (Chapter 15)	–	*E. coli* expression; thioredoxin fusion vector	β-lactamase	λp_L
pTrxFus (Chapter 15)	–	*E. coli* expression; thioredoxin fusion vector	β-lactamase	λp_L
pTZ18	Pharmacia, USB	Cloning; mutagenesis; transcription	B/w screen	
pUC vectors	Various	General cloning	Amp	
pUC17	Stratagene	General cloning	Amp	

pUC19 (Chapter 1, Figure 1-11)	Various	General cloning	Amp
pXf3	ATCC	Cloning	Amp
pZLI (Chapter 11, Figure 11-9)	Life Technologies	Cloning; cDNA expression; transcription	Amp
R6K	ATCC	Cloning	Amp

λ VECTORS	SUPPLIER	APPLICATION	CAPACITY	FEATURES
λ2001	ATCC	Cloning (replacement vector)	10–23 kb	Generates Spi⁻ CI⁻ Gam⁻ Int⁻ Red⁻ recombinants
λDASH	Stratagene	Genomic libraries; restriction mapping; chromosome walking	9–23 kb	T3/T7 promoters
λ EMBL series	Various	Cloning; library construction		Spi/P2 selection against non-recombinants
λFIX	Stratagene	Genomic libraries; restriction mapping; chromosome walking	9–23 kb	T3/T7 promoters
λgt10	NEB, Stratagene	cDNA libraries	7.6 kb	Carries *imm434* cI (disrupted by insertion)
λgt11-23	Various (NEB, Stratagene)	cDNA libraries	7.6 kb	Fusion protein under control of *lac* promoter
λZAP Express	Stratagene	cDNA libraries	12 kb	
λZipLox	Life Technologies	cDNA libraries	7 kb	*lac* promoter; *loxP* sequences
φX174	NEB	Best molecular-weight standard for electron microscopy	n.a.[a]	
Charon 32-35 and 40	–	Cloning recombinogenic sequences		*red gam*⁻ recombinants
SurfZAP	Stratagene	Peptide display; expression; in vitro transcription		T3/T7, *lac* promoters

MAMMALIAN VECTORS	SUPPLIER	APPLICATION	FEATURES
pBK-CMV (Chapter 11, Figure 11-8)	Stratagene	Construction of cDNA libraries	CMV promoter, T3/T7 promoters
pβ-gal series (Chapter 17, Figure 17-5)	CLONTECH	Expression: β-galactosidase reporter; analysis of cloned regulatory elements	SV40 early promoter and enhancer
pCAT3 series (Chapter 17, Figure 17-3)	Stratagene	Expression: chloramphenicol acetyltransferase reporter; analysis of cloned regulatory elements	SV40 promoter and enhancer
pcDNA3.1 (+/–) (Chapter 11, Figure 11-13)	Invitrogen	High-level expression	CMV promoter, Neo selection
pCMV1,2,3,4,5,6	ATCC	Construction of cDNA expression libraries; high-level expression	CMV promoter, SV40 early promoter and enhancer
pCMVScript (Chapter 11, Figure 11-12)	Stratagene	Expression	CMV immediate early promoter

(*Continued on the following pages.*)

TABLE A3-1 (*Continued*)

Mammalian Vectors	Supplier	Application	Features
pCMV-SPORT-β-gal (Chapter 16, Figure 16-2)	Life Technologies	Reporter vector used to monitor transfection efficiency	CMV promoter preceding the *E. coli* β-galactosidase gene
pd2EGFP series (Chapter 17, Figure 17-13)	CLONTECH	Expression: GFP reporter; analysis of cloned regulatory elements	SV40 early promoter and enhancer
pGL3 series (Chapter 17, Figure 17-4)	Promega	Expression: luciferase reporter; analysis of cloned regulatory elements	SV40 early promoter and enhancer
pIND(SP1)/V5-His A (Chapter 17, Figure 17-10)	Invitrogen	Inducible expression: ecdysone regulation; expression of target gene controlled by modified ecdysone (glucocorticoid) receptor	Heat shock minimal and SV40 promoter; ecdysone/glucocorticoid response elements; SP1 enhancer
pSPL3 (Chapter 11, Figure 11-17)	Life Technologies	Exon trapping	Splice donor/splice acceptor sequences
pTet-Splice (Chapter 17, Figure 17-9)	Life Technologies	Inducible expression: tetracycline regulation	Tet promoter
pTet-tTak (Chapter 17, Figure 17-9)	Life Technologies	Inducible expression: produces tetracycline *trans*-activator	Basal CMV promoter and Tet operator
pVgRXR	Invitrogen	Inducible expression: ecdysone regulation; produces the modified ecdysone (glucocorticoid) receptor	RSV, CMV, and SV40 promoters

Yeast Vectors	Supplier	Application	Selection	Regulatory Elements
pB42AD (pJG-4)	CLONTECH	Two-hybrid selection: library construction of genes fused to the B42 activation domain	*TRP1*	GAL1 promoter
PEG202	OriGene, CLONTECH	Two-hybrid selection: fusion of LexA DNA-binding domain to bait	*HIS3*	ADH1 promoter
pGILDA	CLONTECH	Two-hybrid selection: fusion of LexA DNA-binding domain to bait	*HIS3*	GAL1 promoter
pRFHM1	OriGene	Two-hybrid (control): nonactivating fusion of bicoid homeodomain to LexA	*HIS3*	ADH1 promoter
pSH17-4	OriGene	Two-hybrid (control): activating fusion of *GAL4* domain to LexA	*HIS3*	ADH1 promoter
YAC (e.g., pYAC4) (Chapter 4, Figure 4-12)	ATCC, Sigma	Artificial chromosome cloning vector	Amp, *ARS1*	*Tetrahymena* telomere sequence

Shuttle Vectors (replicate in both *E. coli* and *S. cerevisiae*)	Supplier	Application	Selection	Regulatory Elements
YIp (yeast integrating plasmid; e.g., pRS303, 304, 305, 306)	ATCC	Plasmid must integrate into yeast chromosome to be maintained	Yeast: *HIS3*, *TRP1*, *LEU2*, *URA3*, respectively; *E. coli*: Amp	*E. coli*: pBluescript backbone (f1 +/– origin; colE1 origin)

YCp (yeast centromere plasmid; e.g., pRS313, 314, 315, 316)	ATCC	Maintained at low copy number; very stable propagation	Yeast: *HIS3*, *TRP1*, *LEU2*, *URA3*, respectively; *E. coli*: Amp	Yeast: low copy number due to centromere *E. coli*: pBluescript backbone (f1 +/– origin; colE1 origin)
YEp (yeast episomal plasmid; e.g., pRS323, 324, 325, 326)	ATCC	High-copy-number propagation of cloned genes	Yeast: *HIS3*, *TRP1*, *LEU2*, *URA3*, respectively; *E. coli*: Amp	*E. coli*: pBluescript backbone (f1 +/– origin; colE1 origin)
YRp (yeast replicating plasmid)	ATCC	Yeast genetic analysis		
pJK101	Origene	Two hybrid selection: test in repression assay	Yeast: *URA3* *E. coli*: Amp	
pSH18-34 (or pMW111, 112)	OriGene, Invitrogen, CLONTECH	Two hybrid selection: test for transcriptional activation assay	Yeast: *URA3* *E. coli*: Amp (pMW-Kan)	

OTHER VECTORS	SUPPLIERS	APPLICATION	SELECTION	FEATURES
BAC (e.g., pBeloBAC11) (Chapter 4, Figure 4-2)	NEB	Genomic cloning: 120–300 kb	Cam	*loxP*, *cosN* sites
Cosmid (e.g., SuperCos1) (Chapter 4, Figures 4-4 and 4-8)	Stratagene	Genomic cloning: 30–45 kb	Amp, Neo	T3, T7 promoters, SV40 *ori*
Cosmid (e.g., pJB8) (Chapter 4, Figure 4-7)	ATCC	Genomic cloning: 30–45 kb	Amp	
M13 (Phagescript SK) (Chapter 3, Figure 3-1)	Stratagene	Cloning; sequencing; mutagenesis		T3, T7 promoters
P1 (e.g., pAD10SacBII) (Chapter 4, Figure 4-3)	–	Genomic cloning: 70–100 kb	Kan	*loxP* sites
PAC: (e.g., pCYPAC1) (Chapter 4, Figure 4-4)	–	Genomic cloning: 130–150 kb	Kan, *sacB*	*loxP* sites

The details of Table A3-1 have been assembled from various sources, including the American Type Culture Collection (ATTC) Web Site (www.atcc.org), the Vector Database (www.vectordb.atcg.com), and various company (supplier) Web Sites. Please note that vectors sold by different companies, even with the same name, may vary slightly. Figures for vectors that are included in this manual are cited; for other vector figures, please see appropriate supplier.

"n.a. indicates not available.

TABLE A3-2 Bacterial Strains

STRAIN	RELEVANT GENOTYPE	REMARKS
71/18	*supE thi* Δ(*lac-proAB*) F′ [*ProAB*$^+$ *lacI*^{q} *lacZ* ΔM15]	A strain used for growth of phagemids. The F′ in this strain carries *lacZ* ΔM15, which permits α-complementation with the amino terminus of β-galactosidase encoded in λZAP. The strain makes high levels of Lac repressor and can be used for inducible expression of genes that are under the control of the *lac* promoter and for detection of recombinants expressing β-galactosidase fusion proteins (Messing et al. 1977; Dente et al. 1983; Rüther and Müller-Hill 1983).
BB4	*supF*58 *supE*44 *hsdR*514 *galK*2 *galT*22 *trpR*55 *metB*1 *tonA* Δ*lac*U169 F′ [*ProAB*$^+$ *lacI*^{q} *lacZ* ΔM15 Tn*10*(*tet*^{r})]	A *recA*$^-$ strain used for growth of λZAP and other λ bacteriophages. The F′ in this strain carries *lacZ* ΔM15, which permits α-complementation with the amino terminus of β-galactosidase encoded in λZAP. The F′ allows superinfection with an M13 helper bacteriophage, a step required for converting a recombinant λZAP to a pBluescript plasmid (Bullock et al. 1987).
BHB2688	(N205 *recA* [λ*imm*434 λc*Its b2 red* Eam Sam/λ])	A bacteriophage λ lysogen used to prepare packaging extracts (Hohn and Murray 1977; Hohn 1979).
BHB2690	(N205 *recA* [λ*imm*434 c*Its b2 red* Dam Sam/λ])	A bacteriophage λ lysogen used to prepare packaging extracts (Hohn and Murray 1977; Hohn 1979).
BL21 (DE3 a.k.a. "Origami")	*hsdS gal* (λc*Its*857 *ind*1 Sam7 *nin*5 *lac*UV5-T7 gene *1*)	A strain used for high-level expression of genes cloned into expression vectors containing the bacteriophage T7 promoter (e.g., the pET series). The gene encoding bacteriophage T7 RNA polymerase is carried on the bacteriophage λDE3, which is integrated into the chromosome of BL21 (Studier and Moffatt 1986).
BNN102 (C600*hflA*)	*supE*44 *hsdR thi*-1 *thr*-1 *leuB*6 *lacY*1 *tonA*21 *hflA*150 [*chr*::Tn*10* (*tet*^{r})]	An *hflA* strain used to select λgt10 recombinants. The *hflA* mutation suppresses plaque formation by c*I*$^+$ bacteriophages but allows plaque formation by recombinant c*I*$^-$ bacteriophages (Young and Davis 1983a).
C-1a	Wild-type strain	A wild-type clone of *E. coli* strain C that has been maintained on minimal medium for several years. *E. coli* C is F$^-$ and lacks host restriction and modification activity. It is a nonsuppressing host strain used in complementation tests with amber mutants of bacteriophage λ (Bertani and Weigle 1953; Borck et al. 1976).
C600 (BNN93)	*supE*44 *hsdR thi*-1 *thr*-1 *leuB*6 *lacY*1 *tonA*21	An amber suppressing strain often used for making lysates (Appleyard 1954) and for propagation of λgt10 (Young and Davis 1983a).
CES200	*sbcB*15 *recB*21 *recC*22 *hsdR*	A strain used for growth of Spi$^-$ bacteriophages (Nader et al. 1985).
CES201	*recA sbcB*15 *recB*21 *recC*22 *hsdR*	A recombination-deficient strain used for growth of Spi$^-$ bacteriophages (Wyman and Wertman 1987).
CJ236	*dut*1 *ung*1 *thi*-1 *relA*1/pCJ105(*cam*^{r} F′)	A *dut ung* strain used to prepare uracil-containing DNA for use as templates in in vitro mutagenesis (Kunkel et al. 1987). pCJ105 carries an F′ and *cam*^{r}; growth of CJ236 in the presence of chloramphenicol selects for retention of the F′.
CSH18	*supE thi* Δ(*lac-pro*) F′ [*ProAB*$^+$ *lacZ*$^-$]	An amber suppressing strain used to screen recombinants made in bacteriophage λ vectors carrying a *lacZ* gene in the stuffer fragments. These vectors give rise to blue plaques in the presence of the chromogenic substrate X-gal; recombinants in which the stuffer fragment has been replaced by foreign DNA give rise to white plaques (Miller 1972; Williams and Blattner 1979).
DE3 (Origami)	Please see BL21	
DH1	*supE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1	A recombination-deficient amber suppressing strain used for plating and growth of bacteria transformed by plasmids and cosmids (Low 1968; Meselson and Yuan 1968; Hanahan 1983).
DH5	*supE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1	A recombination-deficient amber suppressing strain used for plating and growth of plasmids and cosmids (Low 1968; Meselson and Yuan 1968; Hanahan 1983). This strain has a higher transformation efficiency than DH1.

DH5α	*supE44 ΔlacU169 (φ80 lacZ* ΔM15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	A recombination-deficient amber suppressing strain used for plating and growth of plasmids and *hsdR17 recA1 endA1 gyrA96 thi-1 relA1* cosmids. The φ80 *lacZ* ΔM15 mutation permits α-complementation with the amino terminus of β-galactosidase encoded in pUC vectors (Hanahan 1983).
DH5αMCR	*F⁻ mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80 *lacZ* ΔM15 (*lacZYA-argF*)*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	A host for pUC and other α-complementation vectors; used for generating genomic libraries containing methylated cytosine or adenine residues.
DH10B	Δ(*mrr-hsd RMS-mcrBC*) *mcrA recA1*	A recombination-defective strain used for the propagation of BACs.
DH11S	Δ(*lac-proAB*) Δ(*recA1398*) Δ(*mrr-hsd RMS-mcrBC mcrA deoR rpsL srl⁻*) F′[*proAB lacI^q lacZ* ΔM15]	A recombination-deficient strain used for growth of phagemids; the strain facilitates cloning of methylated genomic DNA and enhances transformation by large plasmids (Lin et al. 1992).
DP50*supF*	*supE44 supF58 hsdS3*($r_B m_B$) *dapD8 lacY1 glnV44* Δ(*gal-uvrB*)*47 tyrT58 gyrA29 tonA* Δ(*thyA57*)	A strain used for isolation and propagation of bacteriophage λ recombinants (Leder et al. 1977; B. Bachmann, pers. comm.).
ED8654	*supE supF hsdR metB lacY gal trpR*	An amber suppressing strain commonly used to propagate bacteriophage λ vectors and their recombinants (Borck et al. 1976; Murray et al. 1977).
ED8767	*supE44 supF58 hsdS3*($r_B^- m_B^-$) *recA56 galK2 galT22 metB1*	A recombination-deficient suppressing strain used for propagation of bacteriophage λ vectors (Murray et al. 1977).
HB101	*supE44 hsdS20*($r_B^- m_B$) *recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1*	An amber suppressing strain commonly used for large-scale production of plasmids. The hybrid strain is isolated from an *E. coli* K12 x *E. coli* B cross, and can be transformed efficiciently by plasmids (Boyer and Roulland-Dussoix 1969; Bolivar and Backman 1979).
HMS174	*recA1 hsdR rif*	A recombination-deficient nonsuppressing strain used for high-level expression of genes cloned into expression vectors containing bacteriophage T7 promoter. Bacteriophage T7 RNA polymerase is provided by infection with a bacteriophage λ that carries bacteriophage T7 gene *1* (Campbell et al. 1978; Studier and Moffatt 1986).
JM101[a]	*supE thi* Δ(*lac-proAB*) F′ [*traD36 proAB⁺ lacI^q lacZ* ΔM15]	An amber suppressing F′ strain that will support growth of bacteriophage M13 vectors (Messing 1979).
JM105	*supE endA sbcB15 hsdR4 rpsL thi* Δ(*lac-proAB*) F′ [*traD36 proAB⁺ lacI^q lacZ* ΔM15]	An amber suppressing F′ strain that will support growth of bacteriophage M13 vectors and will modify but not restrict transfected DNA (Yanisch-Perron et al. 1985). The F′ in this strain carried *lacZ* ΔM15, which permits α-complementation with the amino terminus of β-galactosidase encoded in λZAP.
JM107[b]	*supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ(*lac-proAB*) F′ [*traD36 proAB⁺ lacI^q lacZ* ΔM15]	An amber suppressing F′ strain that will support growth of bacteriophage M13 vectors and will modify but not restrict transfected DNA (Yanisch-Perron et al. 1985). The F′ in this strain carried *lacZ* ΔM15, which permits α-complementation with the amino terminus of β-galactosidase encoded in λZAP.
JM109[b,c]	*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ(*lac-proAB*) F′ [*traD36 proAB⁺ lacI^q lacZ* ΔM15]	An amber suppressing F′ recombination-deficient strain that will support growth of bacteriophage M13 vectors and will modify but not restrict transfected DNA (Yanisch-Perron et al. 1985). The F′ in this strain carried *lacZ* ΔM15, which permits α-complementation with the amino terminus of β-galactosidase encoded in λZAP.
JM110	*dam dcm supE44 hsdR17 thi leu rpsL lacY galK galT ara tonA thr tsx* Δ(*lac-proAB*) F′ [*traD36 proAB⁺ lacI^q lacZ* ΔM15]	An amber suppressing F′strain that will not methylate adenine in GATC sequences and will support growth of bacteriophage M13 vectors (Yanisch-Perron et al. 1985). The F′ in this strain carried *lacZ* ΔM15, which permits α-complementation with the amino terminus of β-galactosidase encoded in λZAP.

(Continued on following pages.)

TABLE A3-2 (*Continued*)

STRAIN	RELEVANT GENOTYPE	REMARKS
K802	*supE hsdR gal metB*	An amber suppressing strain used to propagate bacteriophage λ vectors and their recombinants (Wood 1966).
KC8	*hsdR leuB600 trpC9830 pyrF::*Tn*5 hisB463 lacΔ X74 strA galU galK*	A strain for selective rescue of yeast plasmids carrying either the activation domain or DNA-binding domain derived from a GAL4 or LexA two-hybrid screening. KC8 carries the *trpC*, *leuB*, and *hisB* mutations that may be used for complementation with yeast *TRP1*, *LEU2*, and *HIS3* wild-type markers. Note that the Tn5 transposon confers kanamycin resistance.
KK2186	*supE sbcB*15 *hsdR*4 *rpsL thi* Δ(*lac-proAB*) F′ [*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ* ΔM15]	A strain that will support growth of vectors carrying amber mutations and will modify but not restrict transfected DNA (Zagursky and Berman 1984).
KW251	*supE hsdR mcrB mcrA recD*	A recombination-deficient strain that is permissive for vectors carrying amber mutations. Used for propagation of high-titer lysates of bacteriophage λ. Will modify but not restrict DNA at *Eco*RI sites; permits propagation of cytosine-methylated DNA.
LE392	*supE*44 *supF*58 *hsdR*514 *galK*2 *galT*22 *metB*1 *trpR*55 *lacY*1	An amber suppressing strain commonly used to propagate bacteriophage λ vectors and their recombinants. LE392 is a derivative of ED8654 (Borck et al. 1976; Murray et al. 1977).
LG90	Δ(*lac-proAB*)	A strain from which *lacZ* is deleted that is used for detection of recombinants expressing β-galactosidase fusion proteins (Guarente and Ptashne 1981).
M5219	*lacZ trpA rpsL* (λ*bio*252 *cIts*857 ΔH1)	A strain used for regulated expression of genes cloned downstream from the bacteriophage λ p_L promoter. It contains a defective λ prophage that encodes the bacteriophage λ *cIts*857 repressor and N protein, which is an antagonist of transcription termination (Remaut et al. 1981; Simatake and Rosenberg 1981).
MBM7014.5	*hsdR*2 *mcrB*1 *zjj*202::Tn*10* (*tet*^{R}) *araD*139 *araCU*25am Δ*lac*U169	An *mcrB* strain used for construction of libraries in λORF8. Libraries are made with DNA treated with methylases to protect *Hin*dIII and *Bam*HI sites. M·*Alu*I methylase is used to protect *Hin*dIII sites since M·*Hin*dIII methylase is not available commercially. This strain is defective in the restriction system that recognizes *Alu*I-methylated DNA sites (Raleigh and Wilson 1986).
MC1061	*hsdR mcrB araD*139 Δ(*araABC-leu*)7697 Δ*lacX74 galU galK rpsL thi*	An *mcrB* strain used for λORF8 primary libraries as described for the strain MBM7014.5 (Meissner et al. 1987).
MM294	*supE*44 *hsdR endA*1 *pro thi*	An amber suppressing strain used for large-scale production of plasmids. It is transformed efficiently by plasmids (Meselson and Yuan 1968).
MV1184[d]	*ara* Δ(*lac-proAB*) *rpsL thi* (φ80 *lacZ* ΔM15) Δ(*srl-recA*)306::Tn*10* (*tet*^{R}) F′ [*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ* ΔM15]	A recombination-deficient strain used to propagate phagemids pUC118/pUC119 and to obtain single-stranded copies of phagemids (Vieira and Messing 1987).
MV1193	Δ(*lac-proAB*) *rpsL thi endA spcB*15 *hsdR*4 Δ(*srl-recA*)306::Tn*10* (*tet*^{R}) F′ [*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ* ΔM15]	A recombination-deficient strain used to propagate phagemids pUC118/pUC119 and to obtain single-stranded copies of phagemids (Zoller and Smith 1984, 1987).
MZ-1	*galK*Δ8*attL*Δ*BamN*$_7$*N*$_{53}$ *cIts*857Δ H1 *his ilv bio N*^{1}	A temperature-sensitive λ-lysogenic strain used as a host for plasmids containing the bacteriophage λ p_L promoter (Nagai and Thøgersen 1984).
NM519	*hsdR recBC sbcA*	A recombination-deficient strain used for growth of Spi$^-$ λ bacteriophages; will modify but not restrict DNA at *Eco*RI sites (Arber et al. 1983).

NM522	*supE* Δ(*lac proAB*) Δ*mcrB hsdSM*)5 F′[*proAB lacI*^{q} *lacZ* ΔM15]	A strain deficient in modification and restriction used for the growth of phagemids; allows enhanced cloning of methylated genomic DNA (Gough and Murray 1983).
NM531	*supE supF hsdR trpR lacY recA*13 *metB gal*	A recombination-deficient suppressing strain used for propagation of bacteriophage λ vectors (Arber et al. 1983).
NM538	*supF hsdR trpR lacY*	A strain used for assay and propagation of bacteriophage λ (Frischauf et al. 1983).
NM539	*supF hsdR lacY* (P2cox)	A strain used for selection of Spi$^-$ λ bacteriophages. NM539 is a derivative of NM538 (Frischauf et al. 1983).
NM554	*AraD*139 *galK galU hsd R*2 (*rK*$^-$*Km*$^+$) *recA*13 *rpsL thi*-1 Δ(*ara-leu*)7696 Δ*lacX*74 *F*$^-$	A strain used for plating an unamplified cosmid library for screening.
NS3516	*Cre*$^+$ parent of NS3529	A Cre$^+$ host strain used to prepare large quantities of recombinant DNA cloned in bacteriophage P1 (Sternberg et al. 1994).
NS3529	*recA mcrA* Δ(*mrr-hsdRMS-mcrBC*) (λ*imm*434 *nin*5X1-*Cre*) (λ*imm*λ*LP*1)	A host strain for preparing bacteriophage P1 libraries. The *imm434* prophage produces Cre constitutively; the *imm*λ prophage produces the *lacI*^{q} repressor constitutively (Sternberg et al. 1994).
Q358	*supE hsdR* φ80^r	An amber suppressing host used for growth of bacteriophage λ vectors (Karn et al. 1980).
Q359	*supE hsdR* φ80^r (P2)	An amber suppressing host used to select Spi$^-$ λ recombinants (Karn et al. 1980).
R594	*galK*2 *galT*22 *rpsL*179 *lac*	A nonsuppressing strain used as a nonpermissive host for bacteriophage λ vectors containing amber or ochre mutations (Campbell 1965).
RB791	W3110 *lacI*^{q}L8	A strain that makes high levels of Lac repressor and is used for inducible expression of genes under the control of the *lac* and *tac* promoters (Brent and Ptashne 1981).
RR1	*supE*44 *hsdS*20(r_b^- m_b^-) *ara*-14 *proA*2 *lacY*1*galK*2 *rpsL*20 *xyl*-5 *mtl*-1	A *recA*$^+$ derivative of HB101 that can be transformed with high efficiency (Bolivar et al. 1977; Peacock et al. 1981; B. Bachmann, pers. comm.).
SMR10	*E. coli* C (λ*cos*2 ΔB *xis*1 *red*3 *gam* am210 c*I*ts857 *nin*5 Sam7/λ)	A bacteriophage λ lysogen used to prepare packaging extracts (Rosenberg 1985).
TAP90	*supE*44 *supF*58 *hsdR pro leuB thi*-1 *rpsL lacY*1 *tonA*1 *recD*1903::mini-*tet*	A strain used for production of high-titer bacteriophage λ lysates. This restriction-deficient *supE supF* strain has a mini-*tet* insertion in *recD*, which improves growth of Spi$^-$ λ bacteriophages (Patterson and Dean 1987).
TG1	*supE hsd*Δ5 *thi* Δ(*lac-proAB*) F′ [*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ* ΔM15]	An *Eco*K$^-$ derivative of JM101 that neither modifies nor restricts transfected DNA (Gibson 1984).
TG2	*supE hsd*Δ5 *thi* Δ(*lac-proAB*) Δ(*srl-recA*)306::Tn*10* (*tet*^{r}) F′ [*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ* ΔM15]	A recombination-deficient derivative of TG1 (M. Biggin, pers. comm.).
XL1-Blue	*supE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*46 *thi relA*1 *lac*$^-$ F′ [*proAB*$^+$ *lacI*^{q} *lacZ* ΔM15 Tn*10* (*tet*^{r})]	A recombination-deficient strain that will support the growth of vectors carrying some amber mutations, but not those with the Sam100 mutation (e.g., λZAP). Transfected DNA is modified but not restricted. XL1-Blue is used to propagate λZAPII recombinants, which are unstable in BB4. The F′ in this strain allows blue/white screening on X-gal and permits superinfection with bacteriophage M13 (Bullock et al. 1987).

(*Continued on following page.*)

TABLE A3-2 (*Continued*)

STRAIN	RELEVANT GENOTYPE	REMARKS
XL1-Blue MRF′	*supE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*46 *thi relA*1 *lac*⁻ Δ(*mcrA*)183 Δ(*mcrCB-hsdSMR-mrr*)173 F′ [*proAB*⁺ *lacI*^{q} *lacZ* ΔM15 Tn*10* (*tet*^{r})]	A recombination-deficient strain that is permissive for growth of vectors carrying amber mutations. Will modify but not restrict DNA at *Eco*RI sites; permits α-complementation of β-galactosidase in a *recA*⁻ strain (Jerpseth et al. 1992). It is resistant to tetracycline.
XS101	*recA*1 *hsdR rpoB*331 F′ [*kan*]	A recombination-deficient strain that modifies but does not restrict transfected DNA. It carries an episome conferring resistance to kanamycin and is used for growth of phagemids (Levinson et al. 1984).
XS127	*gyrA thi rpoB*331 Δ(*lac-proAB*) *argE* F′ [*traD*36 *proAB*⁺ *lacI*^{q} *lacZ* ΔM15]	A strain used for growth of phagemids (Levinson et al. 1984).
Y1089	*araD*139 Δ*lac*U169 *proA*⁺ Δ*lon* *rpsL hflA*150 [*chr*::Tn*10* (*tet*^{r})] PMC9	A strain used for protein production from λgt11 and λgt18-23 recombinants. Expression of the foreign protein is controlled by high levels of Lac repressor made by pMC9, which carries *lacI*^{q}. Y1089 is deficient in the *lon* protease, which may enhance stability of the foreign proteins. Lysogens are formed at a high frequency in this strain (Young and Davis 1983b).
Y1090hsdR	*supF hsdR araD* 139 Δ*lon* Δ*lac*U169 *rpsL trpC*22::Tn*10* (*tet*^{r}) pMC9	A strain used for immunological screening of expression libraries and propagation of λgt11 and λgt18-23 (Young and Davis 1983b; Jendrisak et al. 1987). Expression of the foreign protein is controlled by the high levels of Lac repressor made by pMC9, which carries *lacI*^{q}. Detection of proteins toxic to *E. coli* can be achieved by adding IPTG several hours after initiation of plaque formation. Some proteins are unstable in *E. coli*. Y1090 *hsdR* is deficient in the *lon* protease, which may enhance stability of antigens and facilitate antibody screening. The *supF* marker suppresses Sam100 to allow cell lysis (Young and Davis 1983b).
YK537	*supE*44 *hsdR hsdM recA*1 *phoA*8 *leuB*6 *thi lacY rpsL*20 *galK*2 *ara*-14 *xyl*-5 *mtl*-1	A recombination-deficient suppressing strain used for regulated expression of genes cloned downstream from the *phoA* promoter (Oka et al. 1985).

[a]Strain JM103 (Messing et al. 1981) is a restrictionless derivative of JM101 that has been used to propagate bacteriophage M13 recombinants. However, some cultivars of JM103 have lost the *hsdR4* mutation (Felton 1983) and are lysogenic for bacteriophage P1 (which codes for its own restriction/modification system). JM103 is therefore no longer recommended as a host for bacteriophage M13 vectors. Strain KK2186 (Zagursky and Berman 1984) is genetically identical to JM103 except that it is nonlysogenic for bacteriophage P1.

[b]Strains JM106 and JM108 are identical to JM107 and JM109, respectively, except that they do not carry an F′ episome. These strains will not support the growth of bacteriophage M13 but may be used to propagate plasmids. However, JM106 and JM108 do not carry the *lacI*^{q} marker (normally present on the F′ episome) and are therefore unable effectively to suppress the synthesis of potentially toxic products encoded by foreign DNA sequences cloned into plasmids carrying the *lacZ* promoter.

[c]Strains JM108 and JM109 are defective for synthesis of bacterial cell walls and form mucoid colonies on minimal media. This trait does not affect their ability to support the growth of bacteriophage M13.

[d]The original strain of MV1184, constructed by M. Volkert (pers. comm.), did not carry an F′ episome. However, the strain of MV1184 distributed by the Messing laboratory clearly carries an F′ episome. It is therefore advisable to check strains of MV1184 on their arrival in the laboratory for their ability to support the growth of male-specific bacteriophages.

Appendix 4

Enzymes Used in Molecular Cloning

MODIFICATION/RESTRICTION SYSTEMS IN *E. COLI*

Modification of DNA is used by *Escherichia coli* to distinguish between its own genome and foreign DNAs introduced by bacteriophage infection, plasmid transfer, or transfection. DNAs with familiar patterns of methylation on adenosine or cytosine residues are immune to attack, whereas unmethylated DNAs or DNAs with unfamiliar patterns of methylation are earmarked for degradation by endogenous restriction enzymes.

Methylating Enzymes in *E. coli*

Wild-type strains of *E. coli* express DNA methyltransferases that transfer a methyl group from *S*-adenosylmethionine to an adenine or cytosine residue within a defined target sequence. Described below are several types of methylating enzymes.

dam *Methyltransferase*

In *dam*$^+$ strains of *E. coli*, adenine residues embedded in the sequence ^{5}...GATC...^{3} carry a methyl group attached to the N^6 atom (Hattman et al. 1978). More than 99% of these modified adenine bases, which are found on both strands of the palindromic recognition sequence, are formed by action of DNA adenine methylase, the product of the *dam* gene. DNA adenine methylase is a single-subunit nucleotide-independent (type II) DNA methyltransferase that transfers a methyl group from *S*-adenosylmethionine to adenine residues in the recognition sequence ^{5}.....GATC.....^{3} (for review, please see Marinus 1987; Palmer and Marinus 1994). In *E. coli*, *dam* methylation is required for efficient DNA mismatch repair, for accurate initiation of DNA replication at *oriC*, for segregation and partition of chromosomes carrying *oriC*, and for modulation of gene expression (for review, please see Palmer and Marinus 1994). The transfer of a methyl group to the N^6 atom of adenine places a bulky alkyl residue in the major groove of B-form DNA and prevents cleavage in vitro by some restriction enzymes whose recognition sites contain the sequence ^{5}.....GATC.....^{3}. By contrast, other restriction enzymes require methylation at -GATC- residues to cleave DNA.

The recognition sites of several restriction enzymes (*Pvu*I, *Bam*HI, *Bcl*I, *Bgl*II, *Xho*II, *Mbo*I, and *Sau*3AI) contain this sequence, as do a proportion of the sites recognized by *Cla*I (1 site in 4), *Xba*I (1 site in 16), *Taq*I (1 site in 16), *Mbo*II (1 site in 16), and *Hph*I (1 site in 16). The inhibition of *Mbo*I digestion of prokaryotic DNA presents no practical problem because the restriction enzyme *Sau*3AI recognizes exactly the same sequence as *Mbo*I but is unaffected by *dam* methylation. (*Note:* Mammalian DNA is not methylated at the N^6 position of adenine, and thus, either *Mbo*I or *Sau*3AI can be used effectively.) However, when it is necessary to cleave prokaryotic DNA at every possible site with *Cla*I, *Xba*I, *Taq*I, *Mbo*II, or *Hph*I or to cleave it at all with *Bcl*I, the DNA must be prepared from strains of *E. coli* that are *dam*$^-$ (Backman 1980; Roberts et al. 1980; McClelland 1981).

Lists of restriction enzymes whose pattern of cleavage is affected by *dam* methylation have been assembled by Kessler and Manta (1990) and McClelland and Nelson (1992); additional information is available in the brochures of most commercial suppliers of enzymes and in a database of restriction and modification enzymes (REBASE) that is accessible at rebase.neb.com/rebase.

dcm *Methyltransferase*

dcm introduces methyl groups at the C^5 position of the internal cytosine in the sequence ^{5}...CCAGG...^{3} or ^{5}...CCTGG...^{3} and therefore prevents or suppresses cleavage by restriction

enzymes such as *Eco*RI (Marinus and Morris 1973; May and Hattman 1975). For most purposes, this problem can be avoided by using *Bst*NI, which recognizes exactly the same sequence as *Eco*RII (although it cuts the DNA at a different location within the sequence). If *Bst*NI cannot be substituted for *Eco*RII, the DNA must be prepared from strains of *E. coli* that are *dcm*$^-$ (Marinus 1973; Backman 1980; Roberts et al. 1980). Certain other enzymes may cleave at sequences that partially overlap the modified *dcm* recognition sequence. Detailed information on the methylation sensitivity of individual restriction enzymes is provided by the REBASE Web Site (rebase.neb.com/rebase). *dcm*$^-$ mutants of *E. coli* show no phenotype, and the biological significance of *dcm* methylation is obscure (for review, please see Palmer and Marinus 1994).

The Modification Component of Modification/Restriction Systems

- ***Type I modification/restriction systems.*** The classical type I modification/restriction system in many wild-type strains of *E. coli* is encoded by the three *hsd* (*Eco*K) genes. Two of the polypeptides encoded by these genes (*hsdM* and *hsdS*) are needed to transfer methyl groups to the N^6 position of two adenines in the recognition sequence $^{5'}$AAC(N_6)GTGC$^{3'}$ (Kan et al. 1979). This type of modification, which accounts for ~1% of N^6 adenine methylation in *E. coli*, protects DNA against cleavage by the heterotrimeric *Eco*K restriction endonuclease (encoded by the *hsdRMS* genes), a type I restriction endonuclease which, at the expense of ATP hydrolysis, produces double-stranded breaks at a variable distance from the recognition site (Bickle 1987). Many strains of *E. coli* used for molecular cloning carry mutations in the *hsd* genes.
- ***Type II modification/restriction systems.*** Classical (type II) modification/restriction systems have two components, a restriction endonuclease and a DNA methyltransferase; both components recognize the same target sequence. In vivo, type II restriction endonucleases cleave unmethylated sequences, whereas methyltransferases prefer hemimethylated substrates generated during DNA replication. The modifying enzymes of many prokaryotic type II modification/restriction systems are monomers that transfer a methyl group to the 5-carbon of the pyrimidine ring of cytosine, creating 5-methylcytosine. The catalytic mechanism of this reaction involves nucleophilic attack on C^6 of the target cytosine by a cysteine residue to generate a covalent intermediate. The addition of the nucleophile allows the transfer of a methyl group from *S*-adenosylmethionine to the activated 5-carbon. Abstraction of the proton at the 5-position yields an enzyme that undergoes conjugative elimination to yield the product 5-methylcytosine (Wu and Santi 1987; for reviews, please see Kumar et al. 1994; Verdine 1994; Winkler 1994).

Contacts between type II restriction enzymes and their recognition sequence involve amino acids that are widely separated in the primary protein sequence, resulting in bending and kinking of the target DNA (Kim et al. 1990; Winkler et al. 1993). Type II methylases produce a more radical change in DNA structure by forcing the target cytosine to flip out of the plane of the DNA helix and into the active site of the enzyme. This flipping mechanism, which was first established for M·*Hha*I (Klimasauskas et al. 1994), is believed to be universal among 5-methylcytosine transferases, all of which share a well-conserved set of sequence motifs and a common architecture (for review, please see Winkler 1994).

Methylation-dependent Restriction Systems in *E. coli*

E. coli K contains at least three different methylation-dependent restriction systems that recognize different target sequences only when methylated: *mrr* (6-methyladenine [^{m6}A]), *mcr*A (^{m5}CG [^{m5}C = 5-methylcytosine]), and *mcr*B (PU^{m5}C) (Raleigh and Wilson 1986; Heitman and Model

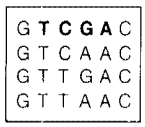
1987; Raleigh et al. 1988). DNAs that are methylated at such sites are inefficiently cloned into wild-type strains of *E. coli* (Whittaker et al. 1988). For example, human DNA, which is extensively methylated in vivo at ^{m5}CG, is restricted by *mcr*A. Nonrestricting strains of *E. coli* (Raleigh and Wilson 1986; Raleigh et al. 1988) are therefore preferred for transformation and cloning of methylated DNA.

M·*Eco*RI methylase catalyzes the transfer of methyl groups from *S*-adenosyl L-methionine (SAM) to the adenines marked with a star (*) in the *Eco*RI recognition sequence in the figure at the right. The modification of adenine to 6-methylaminopurine protects the DNA from cleavage by *Eco*RI (Greene et al. 1975).

Modification of Restriction Sites by DNA Methylation

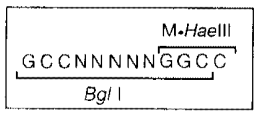
For many of the type II restriction enzymes, a corresponding methylase has been isolated that modifies the cognate recognition sequence and renders it resistant to cleavage. These methylases, a number of which are available from commercial suppliers, are useful in a number of tasks in molecular cloning. For example, in the construction of genomic DNA libraries (Maniatis et al. 1978), random fragments of genomic DNA generated by partial cleavage with the restriction enzymes *Alu*I and *Hae*III can be treated with M·*Eco*RI methylase prior to the addition of synthetic *Eco*RI linkers. When the linkers attached to genomic DNA are subsequently digested with *Eco*RI, the natural restriction sites within the genomic DNA are protected from cleavage. The same strategy may be used to spare natural restriction sites when preparing double-stranded cDNA for cloning.

Methylases can also be used to alter the apparent cleavage specificity of certain restriction enzymes (Nelson et al. 1984; Nelson and Schildkraut 1987). These alterations are accomplished in vitro by methylation of a subset of the sequences recognized by certain restriction enzymes. Only the methylated subsets will be resistant to cleavage. For example, the restriction enzyme *Hinc*II recognizes the degenerate sequence GTPyPuAC and will therefore cleave the four sequences shown at the right.



The M·*Taq*I methylase recognizes only the sequence TCGA and methylates the adenine residue (McClelland 1981). The subset of *Hinc*II recognition sequences that contains the internal sequence TCGA will therefore be resistant to cleavage after methylation by M·*Taq*I, whereas the other three recognition sequences will remain sensitive to *Hinc*II.

A second class of overlapping methylation and restriction sites occurs at the boundaries of the recognition sequences of a restriction enzyme and a methylase. For example, a *Bam*HI site (GGATCC) that happens to be preceded by CC or followed by GG partially overlaps a site (CCGG) that can be methylated by the enzyme M·*Msp*I. Because M·*Msp*I methylates the 5′ cytosine of its recognition sequence (^{m5}CCGG), the *Bam*HI site becomes methylated at an internal cytosine residue (GGAT^{m5}CCGG) and therefore is resistant to cleavage by *Bam*HI. Another example is provided by *Bgl*I and M·*Hae*III illustrated at the left. As in the cases of *Hinc*II/M·*Taq*I and *Bam*HI/M·*Msp*I, the methylation blocks the cleavage of a previously existing restriction site. Finally, certain adenine methylases can be used in conjunction with the methylation-dependent restriction enzyme *Dpn*I to produce highly specific cleavages at sequences 8–12 bp in length (McClelland et al. 1984; McClelland and Nelson 1987; Weil and McClelland 1989).



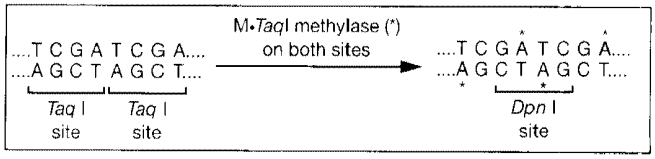


TABLE A4-1 Isoschizomer Pairs That Differ in Their Sensitivity to Sequence-specific Methylation

	Isoschizomer Pairs[a]	
Methylated Sequence[b]	Cut by	Not Cut by
T^{m5}CCGGA	*Acc*III	*Bsp*MII
TC^{m5}CGGA	*Acc*III	*Bsp*MII
GGWC^{m5}C	*Afl*I	*Ava*II (*Eco*47I)
TCCGG^{m6}A	*Bsp*MII	*Acc*III
C^{m5}CWGG	*Bst*NI (*Mva*I)	*Eco*RII
GGTAC^{m5}C	*Kpn*I	*Asp*718I
C^{m5}CGG	*Msp*I	*Hpa*II (*Hap*II)
C^{m4}CGG	*Msp*I	*Hpa*II
G^{m6}ATC	*Sau*3AI (*Fnu*EI)	*Mbo*I (*Nde*II)
TCGCG^{m6}A	*Sbo*13I	*Nru*I
RG^{m6}ATCY	*Xho*II	*Mfl*I
CC^{m5}CGGG	*Xma*I (*Cfr*9I)	*Sma*I

In each row, the first column lists a methylated sequence, the second column lists an isoschizomer that cuts this sequence, and the third column lists an isoschizomer that does not cut this sequence. For references, please see McClelland and Nelson (1988). (Reprinted, with permission, from McClelland and Nelson 1988.)

[a]An enzyme is classified as insensitive to methylation if it cuts the methylated sequence at a rate that is at least one tenth the rate at which it cuts the unmethylated sequence. An enzyme is classified as sensitive to methylation if it is inhibited at least 20-fold by methylation relative to the unmethylated sequence.

[b]Sequences are in 5′→3′ order. R = G or A; Y = C or T; W = A or T; ^{m4}C = 4-methylcytosine, ^{m5}C = 5-methylcytosine, and ^{m6}A = 6-methyladenine.

This example differs from those above in that the strategy *creates* a site in DNA that otherwise would not be cleaved by *Dpn*I. This general strategy of using specific methylases in conjunction with restriction enzymes has produced >60 new cleavage specificities and many more are possible (for references, please see McClelland and Nelson 1988). A compilation of the sensitivities of individual restriction enzymes to site-specific modifications may be found at the REBASE Web Site: rebase.neb.com/rebase.

Several pairs of isoschizomers differ in their sensitivity to site-specific methylation (see Table A4-1). Such endonuclease pairs are useful for studying the level and distribution of site-specific methylation in cellular DNA, for example, ^{m5}CG in mammals, ^{m5}CG and ^{m5}CNG in plants, and G^{m6}ATC in enterobacteria (Waalwijk and Flavell 1978; McClelland 1983; Bird et al. 1985).

Sensitivity to site-specific DNA methylation is clearly not limited to restriction enzymes but is a property of DNA-binding proteins in general (see Sternberg 1985; Wang et al. 1986). ^{m4}C, ^{m5}C, ^{hm5}C, and ^{m6}A site-specific modification at "noncanonical" sites will block several type II methylases. The data are summarized in Table A4-2.

Influence of Methylation on DNA Mapping

Mammalian DNA contains ^{m5}C residues in addition to the four normal bases. These residues are found primarily at the 5′ side of G residues. Although only a portion of CpG doublets are methylated, the pattern of methylation is highly cell-type-specific (Bird and Southern 1978); any given CpG doublet is methylated in the majority of cells of a given population or in only a few of them.

Nearly all restriction enzymes used for long-range mapping of mammalian chromosomal DNA recognize sequences that contain CpG. Because this dinucleotide occurs approximately five-

TABLE A4-2 Methylation Sensitivity of Type II DNA Methylases

Methylase (Specificity)[a]	Not Blocked by Prior Modification at[b]	Blocked by Prior Modification at[b]
M·*Alu*I (AG^{m5}CT)		AG^{m4}CT
M·*Bam*HI (GGAT^{m4}CC)	GG^{m6}ATCC	GGATC^{m5}C
M·*Cfr*6I (CAG^{m5}CTG)		CAG^{m5}CTG
M·*Cla*I (ATCG^{m6}AT)	^{m6}ATCGAT	
	AT^{m5}CGAT	
M·*Cvi*BIII (TCG^{m6}A)	T^{m5}CGA	
M·*Eco*RI (GA^{m6}ATTC)	GAATT^{m5}C	G^{m6}AATTC
M·*Eco*RII (C^{m5}CWGG)		C^{m4}CWGG
E. coli dam (G^{m6}ATC)	GAT^{m5}C[c]	
	GAT^{hm5}C	
M·*Fok*IA (GG^{m6}ATG)	CATC^{m5}C	CAT^{m5}CC
M·*Hha*I (G^{m5}CGC)	GCG^{m5}C	
M·*Hha*II (G^{m6}ANTC)	GANT^{m5}C	
M·*Hpa*II (C^{m5}CGG)		^{m5}CCGG
M·*Hph*I (T^{m5}CACC)	GGTG^{m6}A	
M·*Mbo*I (G^{m6}ATC)	GAT^{m5}C	
M·*Mbo*II (GAAG^{m5}A)	T^{m5}CTT^{m5}C	
M·*Msp*I (^{m5}CCGG)		C^{m5}CGG
M·*Mva*I (C^{m4}CWGG)	C^{m5}CWGG	
M·*Pvu*II (CAG^{m4}CTG)		CAG^{m5}CTG
M·T2 *dam* (G^{m6}ATC)	GAT^{hm5}C	
M·T4 *dam* (G^{m6}ATC)	GAT^{hm5}C	
M·*Taq*I (TCG^{m6}A)	T^{m5}CGA	

Reprinted, with permission, from McClelland and Nelson (1988). For references, please see McClelland and Nelson (1988).

[a]Sequences are in 5′→3′ order. W = A or T; N = A or C or G or T; ^{m4}C = 4-methylcytosine, ^{m5}C = 5-methylcytosine, ^{hm5}C = 5-hydroxymethylcytosine, and ^{m6}A = 6-methyladenine.

[b]An enzyme is classified as insensitive to methylation if it methylates the methylated sequence at a rate that is at least one tenth the rate at which it methylates the unmethylated sequence. An enzyme is classified as sensitive to methylation if it is inhibited at least 20-fold by methylation relative to the unmethylated sequence.

[c]*E. coli dam* modifies GAT^{m5}C at a reduced rate.

fold less frequently in mammalian DNA than expected (Normore 1976; Setlow 1976; Shapiro 1976), restriction enzyme recognition sites that contain the CpG dinucleotide are extremely rare (Lindsay and Bird 1987; McClelland and Nelson 1987). Furthermore, most of these dinucleotides are methylated, and almost all enzymes with CpG in the recognition sequence fail to cleave ^{m5}CpG-methylated DNA (Nelson and McClelland 1987); for example, *Bsp*MII, *Cla*I, *Csp*I, *Eag*I, *Mlu*I, *Nae*I, *Nar*I, *Not*I, *Nru*I, *Pvu*I, *Rsr*II, *Sal*I, *Xho*I, and *Xor*II are all sensitive to ^{m5}CpG methylation. Finally, methylation of CpG dinucleotides in preparations of mammalian DNA is rarely complete. This variability in methylation of sites that are recognized by rarely cutting restriction enzymes can be a serious problem in mapping of mammalian DNAs by pulsed-field gel electrophoresis. Among the known restriction enzymes suitable for generating very large fragments of mammalian DNA, only a handful are capable of cleaving DNA modified at ^{m5}CpG doublets. These include *Acc*III, *Asu*II, *Cfr*9I, *Sfi*I, and *Xma*I. Propagation of mammalian DNAs in *E. coli* will free CpG dinucleotides from methylation. The pattern of cleavage of the same segment of mammalian genomic DNA will therefore differ in cloned and uncloned preparations.

The CpG dinucleotide is not as rare in many other species as it is in mammalian DNA, and it is not methylated in *Drosophila* and *Caenorhabditis* DNAs. Thus, the fragments produced by digestion of these DNAs with rarely cutting restriction enzymes are less than half the size of those produced from mammalian DNA (see Table A4-3). Enzymes that are not sensitive to certain site-specific methylations are particularly useful for achieving complete digestion of modified DNA. For procedures such as the physical mapping of heavily methylated plant DNA, it is desirable to

TABLE A4-3 Average Sizes of DNA Fragments Generated by Cleavage with Restriction Enzymes

ENZYME	SEQUENCE	CEL	DRO	ECO	HUM	MUS	YSC	XEL
*Apa*I	GGGCCC	40,000	6,000	15,000	2,000	3,000	20,000	5,000
*Asc*I	GGCGCGCC	400,000	60,000	20,000	80,000	100,000	500,000	200,000
*Avr*II	CCTAGG	20,000	20,000	150,000	8,000	7,000	20,000	15,000
*Bam*HI	GGATTC	9,000	4,000	5,000	5,000	4,000	9,000	5,000
*Bgl*I	$GCCN_5GGC$	25,000	4,000	3,000	3,000	4,000	15,000	6,000
*Bgl*II	GCGCGC	4,000	4,000	6,000	3,000	3,000	4,000	3,000
*Bss*HII	GCGCGC	30,000	6,000	2,000	10,000	15,000	30,000	20,000
*Dra*I	TTTAAA	1,000	1,000	2,000	2,000	3,000	1,000	2,000
*Eag*I	CGGCCG	20,000	3,000	4,000	10,000	15,000	20,000	15,000
*Eco*RI	GAATTC	2,000	4,000	5,000	5,000	5,000	3,000	4,000
*Hin*dIII	AAGCTT	3,000	4,000	5,000	4,000	3,000	3,000	3,000
*Nae*I	GCCGGC	15,000	3,000	2,000	4,000	6,000	15,000	6,000
*Nar*I	GGCGCC	15,000	3,000	2,000	4,000	6,000	15,000	7,000
*Nhe*I	GCTAGC	30,000	10,000	25,000	10,000	10,000	10,000	10,000
*Not*I	GCGGCCGC	600,000	30,000	200,000	100,000	200,000	450,000	200,000
*Pac*I	TTAATTAA	20,000	25,000	50,000	60,000	100,000	15,000	50,000
*Pme*I	GTTTAAAC	40,000	40,000	40,000	70,000	80,000	50,000	50,000
*Rsr*II	CGGWCCG	50,000	15,000	10,000	60,000	60,000	60,000	70,000
*Sac*I	GAGCTC	4,000	4,000	10,000	3,000	3,000	9,000	4,000
*Sac*II	CCGCGG	20,000	5,000	3,000	6,000	8,000	20,000	15,000
*Sal*I	GTCGAC	8,000	5,000	5,000	20,000	20,000	10,000	15,000
*Sfi*I	$GGCCN_5GGCC$	1,000,000	60,000	150,000	30,000	40,000	350,000	100,000
*Sgr*AI	CXCCGGXG	100,000	20,000	8,000	70,000	80,000	90,000	90,000
*Sma*I	CCCGGG	30,000	10,000	6,000	4,000	5,000	50,000	5,000
*Spe*I	ACTAGT	8,000	9,000	60,000	10,000	15,000	6,000	8,000
*Sph*I	GCATGC	15,000	5,000	4,000	6,000	6,000	10,000	6,000
*Srf*I	GCCCGGGC	1,000,000	90,000	50,000	50,000	90,000	600,000	100,000
*Sse*I	CCTGCAGG	200,000	50,000	40,000	15,000	15,000	150,000	30,000
*Ssp*I	AATATT	1,000	1,000	2,000	2,000	3,000	1,000	2,000
*Swa*I	ATTTAAAT	9,000	15,000	40,000	30,000	60,000	15,000	30,000
*Xba*I	TCTAGA	4,000	9,000	70,000	5,000	8,000	4,000	6,000
*Xho*I	CTCGAG	5,000	4,000	15,000	7,000	7,000	15,000	10,000

Average size fragments predicted for *Caenorhabditis elegans* (CEL), *Drosophila melanogaster* (DRO), *Escherichia coli* (ECO), human (HUM), mouse (MUS), *Saccharomyces cerevisiae* (YSC), and *Xenopus laevis* (XEL).

Listed are those restriction enzymes that are known or predicted to cleave infrequently in seven commonly studied genomes. Factors affecting the ability of restriction enzymes to cleave a particular genome include (1) percentage G+C content, (2) specific dinucleotide, trinucleotide, and/or tetranucleotide frequencies, and (3) methylation. Using available information on percentage G+C content, dinucleotide frequencies, and a few kilobases of DNA sequence, predictions can be made about potential cleavage with restriction enzymes.

Modified, with permission, from New England Biolabs.

choose restriction enzymes that are insensitive to ^{m5}CG and ^{m5}CNG. Examples of such enzymes are *Bcl*I, *Bst*EII, *Bst*NI, *Cvi*QI, *Eco*RV, *Hin*cII, *Hpa*I, *Kpn*I, *Mbo*II, *Nde*I, *Nde*II, *Rsa*I, *Spe*I, *Sph*I, *Taq*I, *Tth*HBI, and *Xmn*I. The *Ase*I, *Dra*I, *Mse*I, and *Ssp*I enzymes have recognition sequences that do not contain cytosine, so they can be used to cleave heavily cytosine-methylated DNA.

Restriction Endonucleases

Restriction enzymes bind specifically to and cleave double-stranded DNA at specific sites within, or adjacent to, a particular sequence known as the recognition sequence. These enzymes have been classified into three groups or types. Type I and type III enzymes carry modification (methylation) and ATP-dependent restriction activities in the same protein. Type III enzymes cut the DNA at the recognition site and then dissociate from the substrate. However, type I enzymes bind to the recognition sequence but cleave at random sites when the DNA loops back to the bound enzyme. Neither type I nor type III restriction enzymes are widely used in molecular cloning.

Type II modification/restriction systems are binary systems consisting of a restriction endonuclease that cleaves a specific sequence of nucleotides and a separate methylase that modifies the same recognition sequence. An increasingly large number of type II restriction endonucleases have been isolated and characterized, a great many of which are useful for cloning and other molecular manipulations.

The vast majority of type II restriction endonucleases recognize specific sequences that are four, five, or six nucleotides in length and display twofold symmetry. A few enzymes, however, recognize longer sequences or sequences that are degenerate. The location of cleavage sites within the axis of dyad symmetry differs from enzyme to enzyme: Some cleave both strands exactly at the axis of symmetry, generating fragments of DNA that carry blunt ends, whereas others cleave each strand at similar locations on opposite sides of the axis of symmetry, creating fragments of DNA that carry protruding single-stranded termini.

The restriction enzyme database, REBASE, contains a complete listing of all known restriction endonucleases, including the recognition sequences, methylation sensitivity, commercial availability, and references. The database, updated daily, is available at rebase.neb.com/rebase.

DNA POLYMERASES

Many steps in molecular cloning involve the synthesis of DNA in in vitro reactions catalyzed by DNA polymerases. Most of these enzymes require a template and synthesize a product whose sequence is complementary to that of the template. Most polymerases strongly prefer DNA templates, but they will also copy RNA, albeit at lower efficiencies. The most frequently used DNA-dependent DNA polymerases are *E. coli* DNA polymerase I (holoenzyme), the large fragment of *E. coli* DNA polymerase I (Klenow fragment), the DNA polymerases encoded by bacteriophages T4 and T7, modified bacteriophage T7 DNA polymerases (Sequenase and Sequenase version 2.0), and thermostable DNA polymerases. One polymerase, reverse transcriptase (RNA-dependent DNA polymerase), prefers to copy RNA; it will also accept DNA templates and can therefore be used to synthesize double-stranded DNA copies of RNA templates. Finally, one DNA polymerase does not copy a template at all but adds nucleotides only to the termini of existing DNA molecules. This DNA polymerase is called terminal transferase (terminal deoxynucleotidyl transferase).

The properties of the template-dependent polymerases are summarized in Table A4-4 and are described in greater detail in the following pages. The data presented in this table have been collected from several publications. The values serve as accurate guidelines for using the enzymes, but optimal conditions will always vary slightly with the enzyme preparation (degree of purity) and the DNA preparation or when carrying out a sequence of enzymatic reactions in one mixture. For a discussion of the relative advantages of various DNA polymerases used in sequencing reactions, please see Chapter 12.

TABLE A4-4 Comparison of Template-dependent DNA Polymerases

	E. coli DNA Polymerase I (Holoenzyme)	Klenow Fragment of *E. coli* DNA Polymerase I	Bacteriophage T4 DNA Polymerase	Sequenase[a] (Bacteriophage T7 DNA Polymerase)	*Taq* DNA Polymerase (AmpliTaq)	Reverse Transcriptase
Reactions						
DNA polymerase 5′→3′	+	+	+	+	+	+
Exonuclease						
double-stranded DNA 5′→3′	+	–	–	–	+[b]	–
single-stranded DNA 3′→5′	+	+	+	–	–	–
double-stranded DNA 3′→5′	+	+	+	–	–	–
RNase H activity	+	–	–	–	n.i.[c]	+
Displacement of strand from double-stranded fragment	+	+	–	–	n.i.	n.i.
Nick translation	+	–	–	–	n.i.	–
Exchange	+	+	+	–	–	–
Template						
Intact double-stranded DNA	–	–	–	–	–	–
Primed single strands of DNA	+	+	+	+	+	+
Duplex with gaps or single-stranded protruding 5′ termini	+	+	+	+	+	+
Nicked double-stranded DNA	+	+	–	–	n.i.	–
Unprimed single-stranded DNA (hairpin)	+	+	n.i.	n.i.	n.i.	+
Single-stranded RNA	+[d]	+[d]	–	–	n.i.	+
Requirements						
Divalent cation	Mg^{2+}	Mg^{2+}	Mg^{2+}	Mg^{2+}[e]	Mg^{2+}[e]	Mg^{2+}
pH optimum	7.4	7.4 (phosphate) 8.4 (Tris)	8.0–9.0 (50% as active at pH 7.5)	7.6–7.8	8.3 at room temp. (Tris) (50% as active in phosphate [pH 7])	7.6 (Mo-MLV) 8.3 (AMV)
Sulfhydryl reagents	+	+	+	+	–	+
Structure						
Molecular mass (kD)	109	76	114	92	94	84 (Mo-MLV) 170 (AMV)
Number of subunits	1	1	1	2	1	1 (Mo-MLV) 2 (AMV)

Reprinted, with permission, from Kornberg and Kornberg (1974) and Lehman (1981).

[a]Sequenase is a derivative of bacteriophage T7 DNA polymerase that has been modified by chemical treatment (Sequenase) or by genetic engineering (Sequenase version 2.0) to suppress 3′→5′ exonuclease activity that is a potent component of wild-type bacteriophage T7 DNA polymerase. Consequently, neither version of Sequenase can be used in exchange reactions.

[b]AmpliTaq and *Taq* DNA polymerase have a polymerization-dependent 5′→3′ exonuclease activity.

[c]n.i. indicates no information available.

[d]Greater activity in the presence of Mn^{2+} instead of Mg^{2+}.

[e]Mn^{2+} may substitute for Mg^{2+}, although rates are lower and the specificity of the polymerases may be changed.

DNA Polymerase I (Holoenzyme)

(*E. coli*)

DNA polymerase I consists of a single polypeptide chain (M_r = 109,000) that can function as a 5′→3′ DNA polymerase, a 5′→3′ exonuclease, and a 3′→5′ exonuclease (Kelley and Stump 1979) and that has an inherent RNase H activity. The RNase H activity is essential for cell viability in *E. coli* but has not been used in molecular cloning. For further details, please see the information panel on ***E. COLI* DNA POLYMERASE I AND THE KLENOW FRAGMENT** in Chapter 9.

Uses

1. Labeling of DNA by nick translation (please see Figure A4-1). Of all the polymerases, only *E. coli* DNA polymerase I can carry out this reaction, since it alone has a 5′→3′ exonuclease activity that can remove nucleotides from the DNA strand ahead of the advancing enzyme.

2. The holoenzyme was originally used for synthesis of the second strand of cDNA in cDNA cloning (Efstratiadis et al. 1976), but it has since been superseded by reverse transcriptase and the Klenow fragment of *E. coli* DNA polymerase I, which do not have 5′→3′ exonuclease activities. The 5′→3′ exonuclease of *E. coli* DNA polymerase I degrades oligonucleotides that may serve as primers for the synthesis of the second strand of cDNA.

3. End-labeling of DNA molecules with protruding 3′ tails. This reaction works in two stages. First, the 3′→5′ exonuclease activity removes protruding 3′ tails from the DNA and creates a recessed 3′ terminus. Then, in the presence of high concentrations of one radiolabeled precursor, exonucleolytic degradation is balanced by incorporation of dNTPs at the 3′ terminus. This reaction, which consists of cycles of removal and replacement of the 3′-terminal nucleotides from recessed or blunt-ended DNA, is sometimes called an exchange or replacement reaction. If this type of reaction is used for the end-labeling, bacteriophage T4 DNA polymerase is the enzyme of choice. Although both *E. coli* DNA polymerase I and bacteriophage T4 DNA polymerase can carry out this type of reaction, the bacteriophage enzyme carries a more potent 3′→5′ exonuclease activity.

   In many cases, a single buffer can be used both for cleavage of DNA with a restriction enzyme and for the subsequent end-labeling. However, not all restriction enzymes work in buffers used for DNA polymerase reactions, and it is advisable to carry out pilot reactions with the particular batch of enzyme on hand. If the restriction enzyme does not work in the DNA polymerase buffer, it will be necessary to carry out the restriction enzyme digestion and end-labeling in two separate steps. In this case, cleave the DNA in the appropriate restriction enzyme buffer, remove the restriction enzyme by extraction with phenol:chloroform, precipitate the DNA with ethanol, dissolve it in TE, and add the appropriate volume of a 10x DNA polymerase buffer.

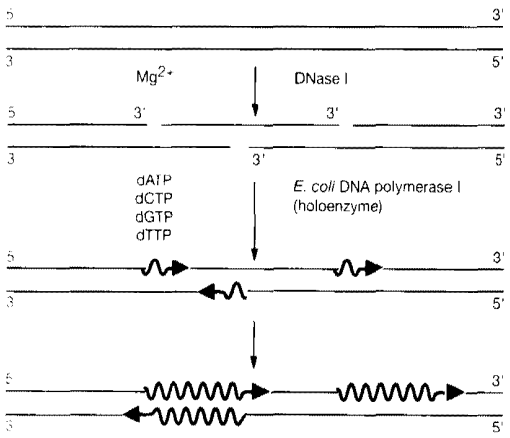


FIGURE A4-1 Nick Translation Using *E. coli* DNA Polymerase

Single-stranded nicks are introduced into the DNA by treatment with DNase I. *E. coli* DNA polymerase (Pol I) binds to the nick or short gap in duplex DNA, and the 5′→3′ exonuclease activity of DNA polymerase I then removes nucleotides from the one strand of the DNA, creating a template for simultaneous synthesis of the growing strand of DNA. The original nick is therefore translated along the DNA molecule by the combined action of the 5′→3′ exonuclease and the 5′→3′ polymerase. In the reaction presented here, the nick in the upper strand of the duplex DNA is translated from left to right by *E. coli* DNA polymerase in the presence of dNTPs. In the lower strand of duplex DNA, nick translation occurs from right to left. The stretches of newly synthesized DNA are represented by the colored arrows.

E. COLI DNA POLYMERASE I (HOLOENZYME)

Activity: 5′→3′ DNA polymerase

Substrate: Single-stranded DNA template with a DNA primer bearing a 3′-hydroxyl group.

Reaction:

$$DNA_{OH} \xrightarrow[\substack{Mg^{2+}\\ dATP,\ dTTP,\ dGTP,\ dCTP}]{\substack{E.\ coli\\ DNA\ polymerase\ I}} DNA - ({}_pdN)_n + nPP_i$$

For example:

$$^{5'}\ldots {}_pC\,{}_pC\,{}_pG_{OH}\ ^{3'}$$

$$^{3'}\ldots G_p\,G_p\,C_p\,T_p\,A_p\,T_p\,C_p\,G_p\,A_p\ldots {}^{5'}$$

$$\underset{\substack{dATP,\ dTTP\\ dGTP,\ dCTP}}{Mg^{2+}} \Big\downarrow \ E.\ coli\ DNA\ polymerase\ I$$

$$^{5'}\ldots {}_pC\,{}_pC\,{}_pG\,{}_pA\,{}_pT\,{}_pA\,{}_pG\,{}_pC\,{}_pT\ldots {}^{3'}$$

$$^{3'}\ldots G_p\,G_p\,C_p\,T_p\,A_p\,T_p\,C_p\,G_p\,A_p\ldots {}^{5'}$$

Activity: 5′→3′ Exonuclease

Substrate: Double-stranded DNA or RNA-DNA hybrids. Degrades double-stranded DNA from the 5′ termini; also degrades the RNA component of an RNA-DNA hybrid (i.e., this nuclease possesses inherent RNase H activity).

Reaction:

$$\text{double-stranded DNA} \xrightarrow[Mg^{2+}]{\substack{E.\ coli\\ DNA\ polymerase\ I}} {}^{5'}{}_pN_{OH} + {}^{5'}{}_pN({}_pN)_{np}N_{OH} + \text{single-stranded DNA}$$

For example:

$$^{5'}\ {}_pC\,{}_pG\,{}_pC\,{}_pA\,{}_pT\,{}_pC\,{}_pT\ldots {}^{3'}$$

$$^{3'}\ G_p\,C_p\,G_p\,C_p\,G_p\,T_p\,A_p\,G_p\,A_p\ldots {}^{5'}$$

$$\underset{}{Mg^{2+}} \Big\downarrow \ E.\ coli\ DNA\ polymerase\ I$$

$$^{5'}\ {}_pC\,{}_pA\,{}_pT\,{}_pC\,{}_pT\ldots {}^{3'}$$

$$^{3'}\ G_p\,C_p\,G_p\,C_p\,G_p\,T_p\,A_p\,G_p\,A_p\ldots {}^{5'}$$

$$+\ {}^{5'}{}_pC + {}^{5'}{}_pG \text{ or } {}^{5'}{}_pC\ {}_pG$$

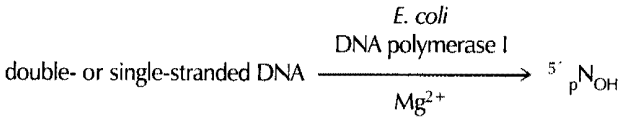
(Continued on following page.)

E. COLI DNA POLYMERASE I (HOLOENZYME)

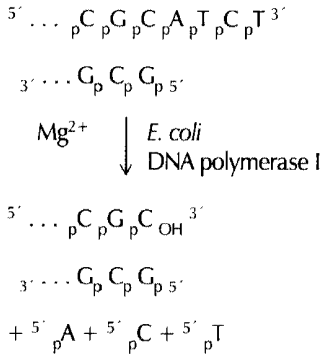
Activity: 3′→5′ Exonuclease

Substrate: Double-stranded or single-stranded DNA containing 3′-hydroxyl termini. Degrades DNA from 3′-hydroxyl termini. Exonuclease activity on double-stranded DNA is blocked by 5′→3′ polymerase activity and is inhibited by dNMPs with 5′ phosphates.

Reaction:



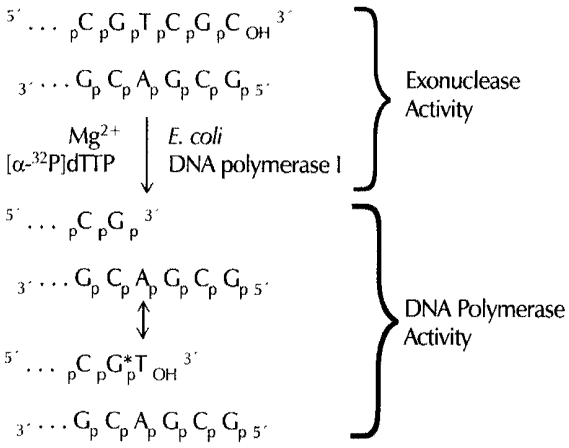
For example:



Activity: Exchange (replacement) reaction

Substrate: If only one dNTP is present, the 3′→5′ exonuclease activity will degrade double-stranded DNA from the 3′-hydroxyl terminus until a base is exposed that is complementary to the dNTP. A continuous series of synthesis and exchange reactions will then occur at that position.

Reaction:



Large Fragment of DNA Polymerase I (Klenow Fragment)

(*E. coli*)

The 5′→3′ exonuclease activity of *E. coli* DNA polymerase I is often troublesome because it degrades the 5′ terminus of primers that are bound to DNA templates and removes 5′ phosphates from the termini of DNA fragments that are to be used as substrates for ligation. The 5′→3′ exonuclease activity can be removed proteolytically from the holoenzyme without affecting either the polymerase activity or the 3′→5′ exonuclease activity (Klenow and Henningsen 1970). The Klenow fragment of *E. coli* DNA polymerase I that is available today from commercial sources consists of a single polypeptide chain (M_r = 76,000) produced by cleavage of intact DNA polymerase I with subtilisin or by cloning (Jacobsen et al. 1974; Joyce and Grindley 1983). For further details, please see the information panel on ***E. COLI* DNA POLYMERASE I AND THE KLENOW FRAGMENT** in Chapter 9.

Uses

1. Filling the recessed 3′ termini created by digestion of DNA with restriction enzymes. In many cases, a single buffer can be used both for cleavage of DNA with a restriction enzyme and for the subsequent filling of recessed 3′ termini (or end-labeling of DNA molecules with protruding 3′ tails [please see Use 3]). The Klenow fragment works well in virtually all buffers used for digestion of DNA with restriction enzymes. However, not all restriction enzymes work in buffers used for DNA polymerase reactions, and it is advisable to carry out pilot reactions with the particular batch of enzyme on hand. If the restriction enzyme does not work in the DNA polymerase buffer, it will be necessary to carry out the restriction enzyme digestion and filling of recessed 3′ termini in two separate steps. In this case, cleave the DNA in the appropriate restriction enzyme buffer, remove the restriction enzyme by extraction with phenol:chloroform, precipitate the DNA with ethanol, dissolve it in TE, and add the appropriate volume of a 10× DNA polymerase buffer.

2. Labeling the termini of DNA fragments by using [^{32}P]dNTPs to fill recessed 3′ termini (end-labeling).

3. End-labeling of DNA molecules with protruding 3′ tails. This reaction works in two stages. First, the 3′→5′ exonuclease activity removes protruding 3′ tails from the DNA and creates a recessed 3′ terminus. Then, in the presence of high concentrations of one radiolabeled precursor, exonucleolytic degradation is balanced by incorporation of dNTPs at the 3′ terminus. This reaction, which consists of cycles of removal and replacement of the 3′-terminal nucleotides from recessed or blunt-ended DNA, is sometimes called an exchange or replacement reaction. If this type of reaction is used for end-labeling, bacteriophage T4 DNA polymerase is the enzyme of choice. Although both the Klenow fragment of *E. coli* DNA polymerase I and bacteriophage T4 DNA polymerase can carry out an exchange reaction, the bacteriophage enzyme carries a more potent 3′→5′ exonuclease activity.

4. Synthesis of the second strand of cDNA in cDNA cloning.

5. Synthesis of double-stranded DNA from single-stranded templates during in vitro mutagenesis. The Klenow fragment can displace hybridized oligonucleotide primers from the template, leading to low frequencies of mutagenesis. This problem can be avoided by using bacteriophage T4 DNA polymerase, which does not cause strand displacement (Nossal 1974).

6. Sequencing of DNA using the Sanger dideoxy-mediated chain-termination method (Sanger et al. 1977).

7. At one time, the 3′→5′ exonuclease activity of the Klenow fragment was used to digest protruding 3′ termini created by some restriction enzymes. Lately, bacteriophage T4 DNA polymerase has become the enzyme of choice for this purpose because of its greater 3′→5′ exonuclease activity.

8. The Klenow fragment has also been used in the polymerase chain reaction to amplify genomic DNA sequences in vitro that are to be used as probes or for direct cloning of mutant alleles of known genes. However, the *Taq* DNA polymerase has now become the enzyme of choice for this purpose because it is stable in heat and therefore need not be freshly added after each round of synthesis and denaturation.

NOTES

- End-labeling with the Klenow fragment provides an alternative to the use of bacteriophage T4 polynucleotide kinase for generating labeled DNA fragments that can be used as size markers during gel electrophoresis. Because DNA fragments are labeled in proportion to their molar concentrations and not their sizes, both small and large fragments in a restriction digest become labeled to an equal extent. It is therefore possible to use autoradiography to locate bands of DNA that are too small to be visualized by staining with ethidium bromide or SYBR dyes.
- The end-filling and end-labeling reactions work well on relatively crude DNA preparations (e.g., minipreparations of plasmids).

E. COLI DNA POLYMERASE I KLENOW FRAGMENT

Activity: 5′→3′ DNA polymerase

Substrate: Single-stranded DNA template with a primer containing a free 3′-hydroxyl group.

Reaction:

$$DNA_{OH} \xrightarrow[\text{dATP, dTTP, dGTP, dCTP}]{\substack{\text{Klenow fragment of } \textit{E. coli} \text{ DNA polymerase I} \\ Mg^{2+}}} DNA - ({}_pdN)_n + nPP_i$$

For example:

$$^{5'}\ldots {}_pC\ {}_pC\ {}_pG\ {}_{OH}{}^{3'}$$

$$^{3'}\ldots G_p\ G_p\ C_p\ T_p\ A_p\ T_p\ C_p\ G_p\ A_p\ldots {}^{5'}$$

Mg^{2+}, dATP, dTTP, dGTP, dCTP; ↓ Klenow fragment of *E. coli* DNA polymerase I

$$^{5'}\ldots {}_pC\ {}_pC\ {}_pG\ {}_pA\ {}_pT\ {}_pA\ {}_pG\ {}_pC\ {}_pT\ldots {}^{3'}$$

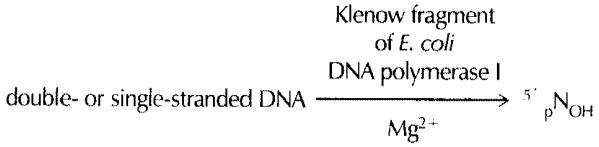
$$^{3'}\ldots G_p\ G_p\ C_p\ T_p\ A_p\ T_p\ C_p\ G_p\ A_p\ldots {}^{5'}$$

E. COLI DNA POLYMERASE I KLENOW FRAGMENT

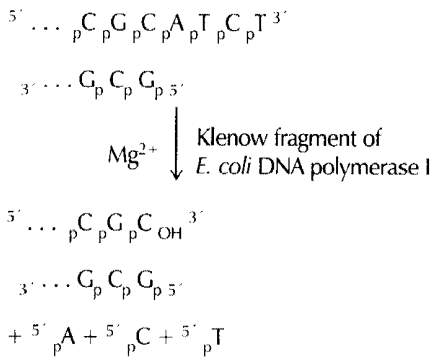
Activity: 3′→5′ Exonuclease

Substrate: Double-stranded or single-stranded DNA degrades from free 3′-hydroxyl termini; exonuclease activity on double-stranded DNAs is blocked by 5′→3′ polymerase activity.

Reaction:



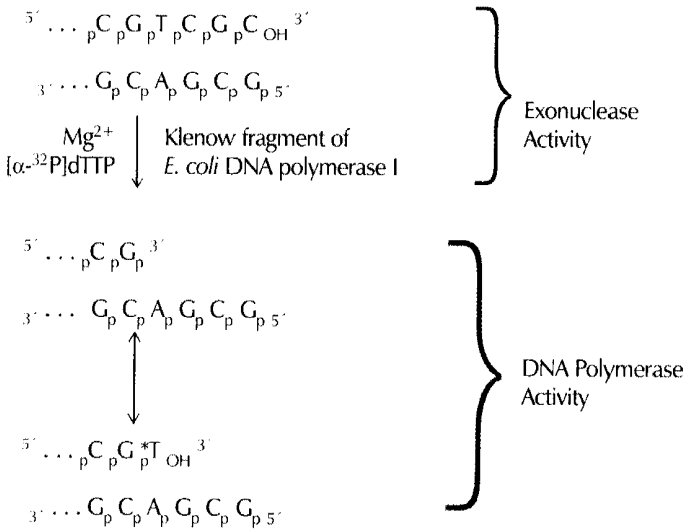
For example:



Activity: Exchange (replacement) reaction

Substrate: If only one dNTP is present, 3′→5′ exonuclease activity will degrade double-stranded DNA from the 3′-hydroxyl terminus until a base is exposed that is complementary to the dNTP. A continuous series of synthesis and exchange reactions will then occur at that position.

Reaction:



Bacteriophage T4 DNA Polymerase

(Bacteriophage T4-infected *E. coli*)

Bacteriophage T4 DNA polymerase (M_r = 114,000) and the Klenow fragment of *E. coli* polymerase I are similar in that each possesses a 5′→3′ polymerase activity and a 3′→5′ exonuclease activity that is more active on single-stranded DNA than on double-stranded DNA. However, the exonuclease activity of bacteriophage T4 DNA polymerase is more than 200 times that of the Klenow fragment. Because it does not displace oligonucleotide primers from single-stranded DNA templates (Nossal 1974), bacteriophage T4 DNA polymerase works more efficiently than the Klenow fragment in mutagenesis reactions in vitro.

Uses

1. Filling or labeling the recessed 3′ termini created by digestion of DNA with restriction enzymes. Labeling reactions must be carried out in the presence of high concentrations of dNTPs in order for the polymerization (filling) reaction to overwhelm the powerful 3′→5′ exonuclease activity.

2. End-labeling of DNA molecules with protruding 3′ tails. This reaction works in two stages. First, the potent 3′→5′ exonuclease activity removes protruding 3′ tails from the DNA and creates a recessed 3′ terminus. Then, in the presence of high concentrations of one radiolabeled precursor, exonucleolytic degradation is balanced by incorporation of dNTPs at the 3′ terminus. This reaction, which consists of cycles of removal and replacement of the 3′-terminal nucleotides from recessed or blunt-ended DNA, is sometimes called an exchange or replacement reaction.

   In many cases, a single buffer can be used both for cleavage of DNA with a restriction enzyme and for the subsequent end-labeling. Bacteriophage T4 DNA polymerase will function at ~50% of maximal activity in many buffers that are commonly used for digestion of DNA with restriction enzymes. However, not all restriction enzymes work in buffers used for DNA polymerase reactions, and it is advisable to carry out pilot reactions with the particular batch of enzyme on hand. If the restriction enzyme does not work in the DNA polymerase buffer, it will be necessary to carry out the restriction enzyme digestion and end-labeling in two separate steps. In this case, cleave the DNA in the appropriate restriction enzyme buffer, remove the restriction enzyme by extraction with phenol:chloroform, precipitate the DNA with ethanol, dissolve it in TE, and add the appropriate volume of a 10x DNA polymerase buffer. Filling and end-labeling reactions with bacteriophage T4 DNA polymerase can be carried out at 12°C to maximize the ratio of polymerase activity to exonuclease activity. However, these reactions are often carried out at room temperature or at 37°C without adverse effects.

3. Labeling DNA fragments for use as hybridization probes. The recessed 3′ termini created by partial digestion of double-stranded DNA with the 3′→5′ exonuclease activity are filled with [^{32}P]dNTPs (replacement synthesis) (O'Farrell et al. 1980). Hybridization probes prepared by this technique have two advantages over probes prepared by nick translation. First, they lack the artifactual hairpin structures that can be produced during nick translation. Second, they can easily be converted into strand-specific probes by cleavage with suitable restriction enzymes (please see Figure A4-2).

   By contrast to nick translation, however, this method does not produce a uniform distribution of label along the length of the DNA. Furthermore, 3′ exonuclease activity degrades single-stranded DNA much faster than it degrades double-stranded DNA, so that after a molecule has been digested to its midpoint, it will dissociate into two half-length single strands that will

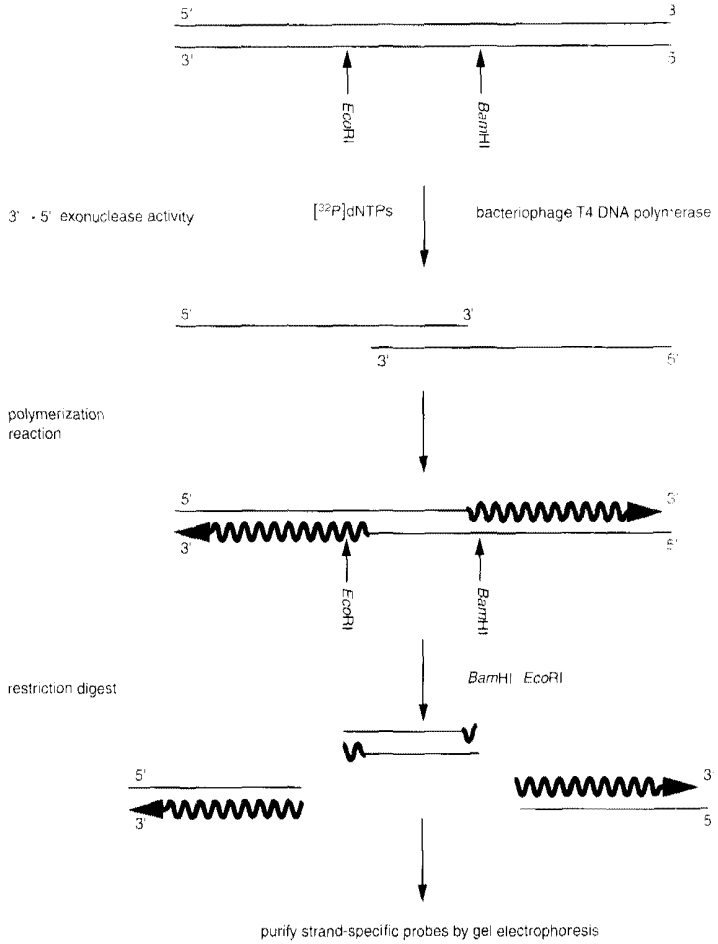


FIGURE A4-2 Production of Strand-specific Hybridization Probes

be rapidly degraded. It is therefore important to stop the exonuclease reaction before the enzyme reaches the center of the molecule. Consequently, the replacement synthesis method yields a population of molecules that are fully labeled at their termini but contain progressively decreasing quantities of label toward their centers.

4. Conversion of termini of double-stranded DNA to blunt-ended molecules. Protruding 3′ termini will be removed from double-stranded DNA by the potent 3′→5′ exonuclease activity of bacteriophage T4 DNA polymerase. In the presence of high concentrations of dNTPs, further degradation of the double-stranded region of the template will be balanced by synthesis. The ability to convert protruding 3′ termini to blunt ends is an extremely valuable reaction that is frequently used when preparing DNAs for addition of synthetic linkers. As described above, molecules with recessed 3′ termini can be repaired by bacteriophage T4 DNA polymerase in a filling reaction similar to that catalyzed by the Klenow fragment. Thus, DNAs with a mixture of protruding 5′ and 3′ termini (e.g., double-stranded cDNAs synthesized from RNA templates) can be converted to blunt-ended molecules (polished) by bacteriophage T4 DNA polymerase in the presence of high concentrations of dNTPs.

5. Extension of mutagenic oligonucleotide primers that are bound to single-stranded DNA templates. Bacteriophage T4 DNA polymerase is preferred in this reaction to the Klenow fragment because it cannot displace the short oligonucleotide from the template. The efficiency of mutagenesis is therefore increased approximately twofold.

BACTERIOPHAGE T4 DNA POLYMERASE

Activity: 5′→3′ DNA polymerase

Substrate: 3′-hydroxyl, single-stranded DNA template-primer complex. The enzyme cannot displace the strand ahead of the nick. However, addition of bacteriophage T4 gene *32* protein allows the enzyme to begin synthesis at a nick in buffers of low ionic strength.

Reaction:

$$DNA_{OH} \xrightarrow[Mg^{2+}\ dATP,\ dTTP,\ dGTP,\ dCTP]{\text{bacteriophage T4 DNA polymerase}} DNA - ({}_pdN)_n + nPP_i$$

For example:

$$^{5'}\ldots {}_pC\ {}_pC\ {}_pG\ {}_{OH}{}^{3'}$$

$$^{3'}\ldots G_p\ G_p\ C_p\ T_p\ A_p\ T_p\ C_p\ G_p\ A_p\ldots {}^{5'}$$

Mg^{2+}, dATP, dTTP, dGTP, dCTP ↓ bacteriophage T4 DNA polymerase

$$^{5'}\ldots {}_pC\ {}_pC\ {}_pG\ {}_pG\ {}_pA\ {}_pT\ {}_pA\ {}_pG\ {}_pC\ {}_pT\ldots {}^{3'}$$

$$^{3'}\ldots G_p\ G_p\ C_p\ T_p\ A_p\ T_p\ C_p\ G_p\ A_p\ldots {}^{5'}$$

Activity: 3′→5′ Exonuclease

Substrate: Considerably more active on single-stranded DNA than on double-stranded DNA; exonuclease activity on double-stranded DNAs is blocked by 5′→3′ DNA polymerase activity.

Reaction:

$$\text{single-stranded DNA} \xrightarrow[Mg^{2+}]{\text{bacteriophage T4 DNA polymerase}} {}^{5'}{}_pN_{OH}$$

For example:

$$^{5'}\ldots {}_pC\ {}_pG\ {}_pC\ {}_pA\ {}_pT\ {}_pC\ {}_pT\ {}^{3'}$$

$$^{3'}\ldots G_p\ C_p\ G_{p\ 5'}$$

Mg^{2+} ↓ bacteriophage T4 DNA polymerase

$$^{5'}\ldots {}_pC\ {}_pG\ {}_pC\ {}_{OH}{}^{3'}$$

$$^{3'}\ldots G_p\ C_p\ G_{p\ 5'}$$

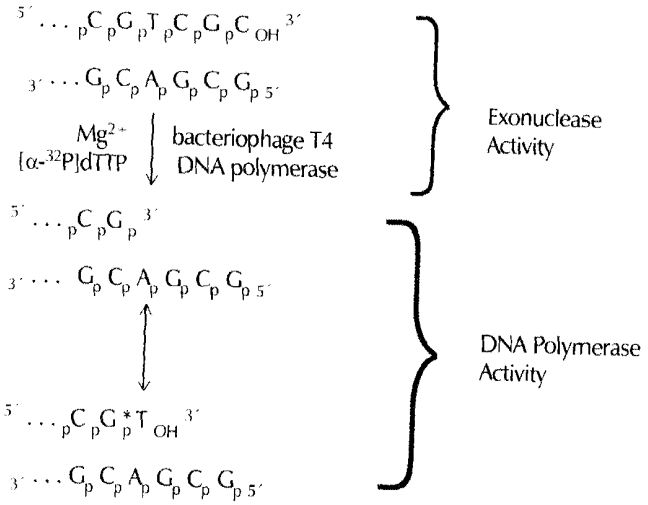
$$+\ {}^{5'}{}_pA + {}^{5'}{}_pC + {}^{5'}{}_pT$$

BACTERIOPHAGE T4 DNA POLYMERASE

Activity: Exchange (replacement) reaction

Substrate: If only one dNTP is present, 3′→5′ exonuclease activity will degrade double-stranded DNA from the 3′-hydroxyl terminus until a base is exposed that is complementary to the dNTP. A continuous series of synthesis and exchange reactions will then occur at that position.

Reaction:



Bacteriophage T7 DNA Polymerase

(Bacteriophage T7-infected *E. coli*)

The DNA polymerase synthesized after infection of *E. coli* by bacteriophage T7 is a complex of two tightly bound proteins, the bacteriophage T7 gene 5 protein and the host protein thioredoxin. This complex is the most processive of all known DNA polymerases. In other words, the average length of DNA synthesized by a single molecule of bacteriophage T7 DNA polymerase is greater than that of DNAs synthesized by other thermolabile DNA polymerases. This property has considerable advantages, for example, when sequencing DNA by the Sanger dideoxy chain-termination method (Sanger et al. 1977) (please see Chapter 12). Bacteriophage T7 DNA polymerase, like the Klenow fragment and the holoenzyme of *E. coli* DNA polymerase I and bacteriophage T4 DNA polymerase, has a 3′→5′ exonuclease activity (encoded by bacteriophage T7 gene 5). As in bacteriophage T4 DNA polymerase, the 3′→5′ exonuclease activity is potent. The degree of activity of the 3′→5′ exonuclease of bacteriophage T7 DNA polymerase is ~1000 times that of the Klenow fragment. Bacteriophage T7 DNA polymerase does not have a 5′→3′ exonuclease activity (Tabor et al. 1987). For illustrations of the 5′→3′ DNA polymerase and 3′→5′ exonuclease activities and exchange reaction encoded by bacteriophage T7 DNA polymerase, please see the panel on **BACTERIOPHAGE T4 DNA POLYMERASE** (p. A4.20). For more information on wild-type and modified T7 DNA polymerase (Sequenase and Sequenase version 2.0), please see the information panel on **SEQUENASE** in Chapter 12.

USES

1. Primer-extension reactions that require the copying of long stretches of template.
2. Rapid end-labeling by either filling or exchange (replacement) reactions such as those described for bacteriophage T4 DNA polymerase.

Thermostable DNA-dependent DNA Polymerases

Thermostable DNA-dependent DNA polymerases have been purified and characterized from a number of organisms, primarily from the thermophilic and hyperthermophilic eubacteria Archaebacteria, whose most abundant DNA polymerases are reminiscent of DNA polymerase I of mesophilic bacteria, and thermophilic Archaea, whose chief DNA polymerases belong to the polymerase α family. Although there is considerable structural variation among thermostable DNA polymerases, all are monomeric with molecular mass values ranging from 60 kD to 100 kD. The significant properties and activities of thermostable DNA polymerases are given in Table A4-5.

The first well-characterized thermostable DNA polymerases were isolated from the extreme thermophile *Thermus aquaticus* (Chien et al. 1976). One of these early isolates, *Taq* DNA polymerase, has, along with a number of genetically engineered variants, become an indispensable component of reactions to amplify specific sequences of DNA in vitro by polymerase chain reaction and in DNA sequencing (please see the introductions to Chapters 8 and 12, respectively). *Taq* and its derivatives have a 5′→3′ polymerization-dependent exonuclease activity. For nucleotide incorporation, the enzyme works best at 75–80°C, depending on the target sequence; its polymerase activity is reduced by a factor of 2 at 60°C and by a factor of 10 at 37°C. In many cases, however, it is necessary to initiate polymerization reactions at suboptimal temperatures in order to prevent dissociation of the primer from the template. For further information on *Taq* and other thermostable polymerases, please see the introduction to Chapter 8 and the information panel on ***TAQ* DNA POLYMERASE** in Chapter 8.

TABLE A4-5 Properties of Thermostable DNA Polymerases

Enzyme	Manufacturer[a]	Organism	Optimum Temperature (°C)	Exonuclease Activity	Error Rate $\times 10^{-6}$	Stability (Minutes at Specified Temperature)	K_m dNTP (μM)	K_m DNA (nM)
Taq	BM, LT, Pro, Strat, P-E, T	*T. aquaticus*	75–80	5′–3′	20–100	9 min at 97.5°C	10–16	2
Taq Stoffel fragment	P-E	*T. aquaticus*	75–80	none	50	21 min at 97.5°C	–	2
rTth	BM, ET, P-E	*T. thermophilus*	75–80	5′–3′	~20	20 min at 95°C	115	–
Tfl	Pro	*T. flavus*	70	none	100	120 min at 70°C	63	–
Hot Tub	Amr	*T. ubiquitus*	–	none	–	–	–	–
Tbr	Amr, Finnz	*T. brockianus*	75–80	5′–3′	–	150 min at 96°C	–	–
UlTma	P-E, Roche	*Thermotoga maritima*	75–80	3′–5′	–	50 min at 95°C	–	–
rBst	ET	*Bacillus sterothermophilus*	60–65	5′–3′ (3′–5′)[b]	–	–	–	–
Isotherm *Bst* large fragment	ET, Bio-Rad	*Bacillus sterothermophilus*	60–65	none	–	–	7–85	–
Pwo	BM	*Pyrococcus woesei*	60–65	3′–5′	3.2	>2 hr at 100°C	–	–
Tli	Pro	*Thermococcus litoralis*	70–80	3′–5′	20–45	100 min at 100°C	66	0.1
DeepVent	NEB	*Pyrococcus* (strain GB-D)	70–80	3′–5′	20–45	480 min at 100°C	50	0.01
Pfu	Strat	*Pyrococcus furiosus*	72–78	3′–5′	1.6	240 min at 95°C	–	–

Data for this table were taken from reviews by Perler et al. (1996) and Bej and Mahbubani (1994), from Internet sources, and from literature distributed by commercial manufacturers. For details of the reaction conditions that are optimal for each enzyme, please consult the instructions supplied with the enzyme by the manufacturer.

[a](BM) Boehringer Mannheim; (ET) Epicentre Technologies; (LT) Life Technologies; (Pro) Promega; (NEB) New England Biolabs; (P-E) Perkin-Elmer; (T) TaKaRa; (Strat) Stratagene; (Amr) Amresco; (Finnz) Finnzymes OY.

[b]Suspected activity.

TAQ DNA POLYMERASE

Activity: 5′→3′ DNA polymerase

Substrate: Single-stranded template, primer with 3′-hydroxyl.

Reaction:

$$DNA_{OH} \xrightarrow[Mg^{2+}\ dATP,\ dTTP,\ dGTP,\ dCTP]{Taq\ DNA\ polymerase} DNA - ({}_pdN)_n + nPP_i$$

Reverse Transcriptase (RNA-dependent DNA Polymerase)

(Murine and avian retroviruses)

Two forms of reverse transcriptases are commercially available: a preparation made from purified avian myeloblastosis virus (AMV) and an enzyme isolated from a strain of *E. coli* that expresses a cloned copy of the reverse transcriptase gene of the Moloney murine leukemia virus (Mo-MLV). Both enzymes lack a 3′→5′ exonuclease active on DNA and catalyze the reactions shown below in the panel on **REVERSE TRANSCRIPTASE**. The murine and avian reverse transcriptases differ from each other in a number of respects:

- The avian enzyme consists of two polypeptide chains that carry both a polymerase activity and a powerful RNase H activity (Verma 1981). The murine enzyme, a single polypeptide chain of M_r = 84,000, has a polymerase activity and a comparatively weak RNase H activity (Gerard 1983). This weak RNase activity is a considerable advantage when attempting to synthesize

REVERSE TRANSCRIPTASE

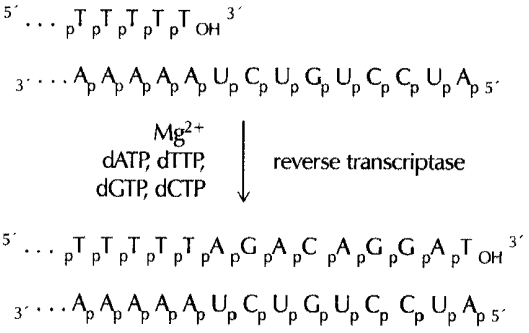
Activity: 5′→3′ DNA polymerase

Substrate: RNA or DNA template with an RNA or DNA primer bearing a 3′-hydroxyl group.

Reaction:

$$DNA_{OH} \text{ or } RNA_{OH} \xrightarrow[\substack{Mg^{2+}\\ dATP,\ dTTP,\ dGTP,\ dCTP}]{\text{reverse transcriptase}} \begin{matrix} DNA-({}_pdN)_n + nPP_i \\ RNA-({}_pdN)_n + nPP_i \end{matrix}$$

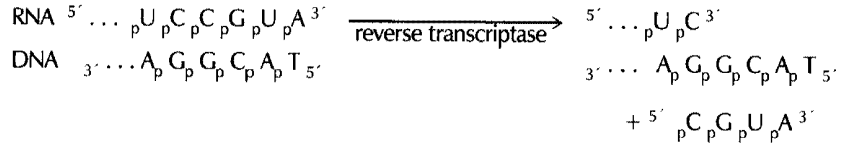
For example:



Activity: RNase H (5′→3′ and 3′→5′ exoribonuclease)

Substrate: Reverse transcriptase specifically degrades RNA in an RNA-DNA hybrid by a processive mechanism.

Reaction: Degradation of substrates with free ends yielding ribonucleotide products that are 4–20 nucleotides in length and contain 5′-phosphate and 3′-hydroxyl termini.



cDNAs complementary to long mRNAs. At the beginning of the reaction, the hybrids formed between the primer and the template mRNA are substrates for RNase H. Thus, at the beginning of cDNA synthesis, there is a competition between degradation of the template mRNA and initiation of DNA synthesis (Berger et al. 1983). In addition, RNase H can cleave the template near the 3′ terminus of the growing DNA strand if reverse transcriptase pauses during synthesis (Kotewicz et al. 1988). In consequence, the high level of RNase H activity in preparations of the avian enzyme tends to suppress the yield of cDNA and to restrict its length.

- The avian enzyme works efficiently at 42°C (the normal body temperature of chickens), whereas the wild-type murine enzyme is rapidly inactivated at this temperature. RNAs rich in secondary structure are therefore copied more efficiently by the avian enzyme than by the murine enzyme. However, preparations of the avian enzyme can be contaminated by an endonuclease that cleaves DNA. This is now less of a problem than it was in the early 1980s, when cDNA libraries generated with comparatively impure preparations of the avian enzyme seldom exceeded 1 kb in length.
- The avian enzyme works more efficiently at pH 8.3 than at pH 7.6, the pH preferred by the murine enzyme. The length of the cDNA synthesized by either enzyme is greatly reduced when reactions are carried out at a pH that differs from the optimum by as little as 0.2 unit. Since the pH of Tris changes with temperature, it is essential to check that the pH of the reaction mixture is correct at the temperature chosen for incubation.

For more information, please see the information panel on **Mo-MLV REVERSE TRANSCRIPTASE** in Chapter 11.

USES

1. Reverse transcriptase is used chiefly to transcribe mRNA into double-stranded cDNA that can be inserted into prokaryotic vectors. However, reverse transcriptase can also be used with either single-stranded DNA or RNA templates to make probes for use in hybridization experiments. Three types of primers are used in these reactions:
    - ***Oligo(dT)***$_{12\text{-}18}$, which binds to the poly(A) tract at the 3′ terminus of mammalian mRNAs and primes the synthesis of the first strand of cDNA. Depending on the quality of the reverse transcriptase and the reaction conditions, the sequences at the 3′ terminus of the template may be overrepresented in the cDNA.
    - ***Oligonucleotides of random sequence*** (Taylor et al. 1976). The aim is to use a population of oligonucleotides whose sequence diversity is so large that at least some individual oligonucleotides will anneal to the template and serve as primers for reverse transcriptase. Because different oligonucleotides bind to different sequences, a large proportion of the sequences of the template will be copied by the enzyme, and if all of the primers are present at equal concentrations, all sequences of the template should be copied at equal frequencies. Oligonucleotides of random sequence can be synthesized on an automated DNA synthesizer or can be generated by hydrolysis of high-molecular-weight DNA.
    - ***Oligonucleotides of defined sequence.*** Oligonucleotides of defined sequence can be used to prime the synthesis of cDNA corresponding to a particular mRNA. Because the newly synthesized DNA is complementary to the sequences of the mRNA that lie upstream of the primer, this method (primer extension) provides an accurate measurement of the distance between a fixed point on an mRNA and its 5′ terminus.
2. Labeling the termini of DNA fragments with protruding 5′ termini (filling reaction).

3. The enzyme can also be used to sequence DNAs by the dideoxy chain termination method (Sanger et al. 1977) when other enzymes (e.g., the Klenow fragment or Sequenase) yield unsatisfactory results.

NOTES

- Reverse transcriptase lacks 3′→5′ exonuclease activity, which acts as an editing function in *E. coli* DNA polymerase I, and is therefore prone to error. In the presence of high concentrations of dNTPs and Mn^{2+}, ~1 base in every 500 is misincorporated.
- Because the K_m of reverse transcriptase for its dNTP substrates is very high — in the millimolar range — it is essential to include high concentrations of dNTPs in this reaction to prevent premature termination of newly synthesized DNA chains.
- Reverse transcriptase can be used to synthesize single-stranded copies of DNA templates using oligonucleotide primers (please see Figure A4-3, part A). However, both double-stranded and single-stranded cDNAs are generated from RNA templates (please see Figure A4-3, part B). Self-primed synthesis is much less efficient than synthesis from the added oligonucleotide primers. Therefore, self-complementary hairpin molecules usually constitute only a small fraction of the synthesized double-stranded cDNA (see Chapter 11). If necessary, both self-primed and exogenously primed second-strand synthesis can be inhibited by including actinomycin D in the reaction mixture at a final concentration of 50 μg/ml.

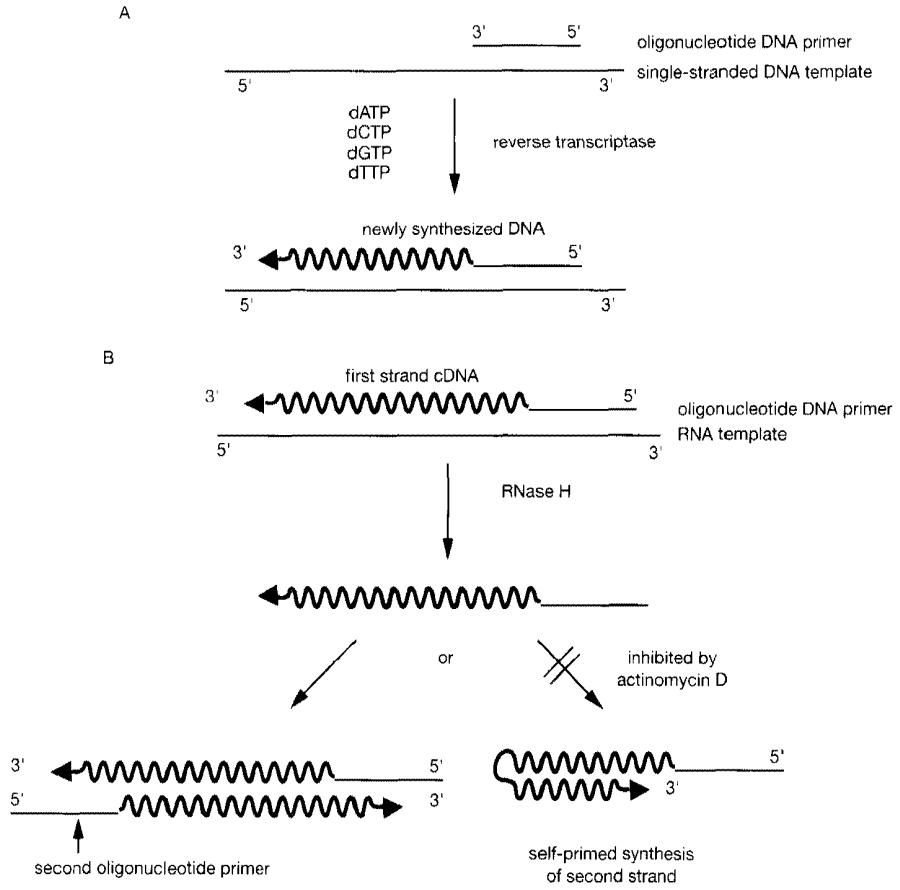


FIGURE A4-3 Use of Reverse Transcriptase to Generate Probes from DNA (*A*) and RNA (*B*) Templates

Terminal Transferase (Terminal Deoxynucleotidyl Transferase)

(Calf thymus)

Terminal transferase (M_r = 60,000) is an unusual DNA polymerase found only in prelymphocytes in early stages of lymphoid differentiation (Chang and Bollum 1986). In the presence of a divalent cation, the purified enzyme catalyzes the addition of dNTPs to the 3′-hydroxyl termini of DNA molecules (Bollum 1974). When the nucleotide to be added is a purine, Mg^{2+} is the preferred cation; when the nucleotide is a pyrimidine, Co^{2+} is used instead. The minimum chain length of the acceptor DNA is three dNTPs, and as many as several thousand dNTPs can be incorporated if the ratio of acceptor to nucleotide is adjusted correctly. Single nucleotides can be added to the 3′ termini of DNA if modified nucleotides (e.g., ddNTPs or cordycepin triphosphate) are used as substrates. Homopolymers of rNTPs can also be synthesized at the 3′ termini of DNA molecules in the presence of Co^{2+} (for references, please see Chang and Bollum 1986). The enzyme strongly prefers to use DNAs with protruding 3′ termini as acceptors. However, blunt or recessed 3′ termini are used, albeit less efficiently, in buffers of low ionic strength that contain Co^{2+} or Mn^{2+} (Roychoudhury et al. 1976; Nelson and Brutlag 1979; Roychoudhury and Wu 1980; Michelson and Orkin 1982; Deng and Wu 1983). For further information, please see the information panel on **TERMINAL TRANSFERASE** in Chapter 8.

Uses

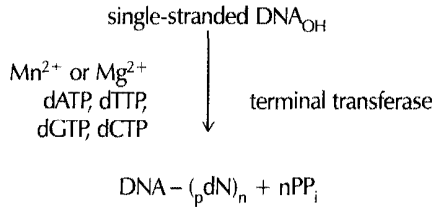
1. Addition of complementary homopolymeric tails to vector and cDNA.
2. Labeling the 3′ termini of DNA fragments with a ^{32}P-labeled dNTP (Tu and Cohen 1980), a ddNTP (Cozzarelli et al. 1969), or an rNTP (Wu et al. 1976). For labeling with rNTPs, [α-^{32}P]rNTP is used in the presence of Co^{2+}, followed by treatment with alkali (please see the panel on **TERMINAL TRANSFERASE** below).

TERMINAL TRANSFERASE

Activity: Terminal transferase

Substrate: Single-stranded DNA with a 3′-hydroxyl terminus or double-stranded DNA with a protruding 3′-hydroxyl terminus is preferred. Blunt-ended, double-stranded DNA or DNA with a recessed 3′-hydroxyl terminus serves as a template if Co^{2+} is supplied as a cofactor (Roychoudhury et al. 1976; Nelson and Brutlag 1979; Roychoudhury and Wu 1980; Michelson and Orkin 1982; Deng and Wu 1983).

Reaction:



DNA-DEPENDENT RNA POLYMERASES

Bacteriophage SP6 and Bacteriophages T7 and T3 RNA Polymerases
(Bacteriophage SP6-infected *Salmonella typhimurium* LT2 and bacteriophage T7- or T3-infected *E. coli*)

Bacteriophage SP6 synthesizes a DNA-dependent RNA polymerase that recognizes and initiates synthesis of RNA on double-stranded DNA templates carrying the appropriate bacteriophage-specific promoter. The polymerase is used in vitro to generate large quantities of RNA complementary to one strand of foreign DNA that has been placed immediately downstream from the promoter. Vectors carrying the promoter are available to synthesize RNA complementary to either strand of the template by changing the orientation of the promoter with respect to the cloned foreign DNA sequences (Butler and Chamberlin 1982; Melton et al. 1984). Alternatively, the promoter can be added to the target DNA during PCR primed by oligonucleotides equipped with a concensus promoter (please see Chapter 9, Protocol 6).

Bacteriophages T7 and T3 also synthesize DNA-dependent RNA polymerases that recognize and initiate synthesis of RNA on double-stranded DNA templates that carry the appropriate bacteriophage-specific promoter (for further details, please see the information panel on **PROMOTER SEQUENCES RECOGNIZED BY BACTERIOPHAGE-ENCODED RNA POLYMERASES** in Chapter 7). These polymerases are used in vitro just like the bacteriophage SP6 RNA polymerase. Bacteriophages T7 and T3 RNA polymerases have been cloned and expressed in *E. coli* (Davanloo et al. 1984; Tabor and Richardson 1985; Morris et al. 1986), and bacteriophage T7 RNA polymerase has been cloned and expressed in yeast (Chen et al. 1987). Vectors carrying the bacteriophage T7 promoter may therefore be used to express cloned genes in vivo (please see point 2 under Uses below).

USES

1. Synthesis of single-stranded RNA for use as hybridization probes, functional mRNAs for in vitro translation systems, or substrates for in vitro splicing reactions. Each of the three RNA polymerases has a high degree of specificity for its cognate promoter.

2. The bacteriophage T7 transcription system has been used to express cloned genes in bacteria (Tabor and Richardson 1985; Studier and Moffatt 1986) and in yeast (Chen et al. 1987). Two types of bacteriophage T7 expression systems have been developed for *E. coli*. In the first system, stable lysogens are established with bacteriophage λ carrying the bacteriophage T7 RNA polymerase gene under the control of the *E. coli lac*UV5 promoter. Plasmids containing the gene of interest under the control of the bacteriophage T7 promoter are then introduced into the lysogens containing the bacteriophage T7 RNA polymerase gene. Activation of the bacteriophage T7 promoter is then achieved by isopropylthio-β-D-galactoside induction of the *lac*UV5 promoter driving the bacteriophage T7 RNA polymerase gene. In the second system, the bacteriophage T7 promoter/plasmid carrying the gene of interest is introduced into bacteria, and the bacteriophage T7 promoter is activated by infecting the bacteria with bacteriophage λ containing the bacteriophage T7 RNA polymerase gene.

   In yeast, the bacteriophage T7 RNA polymerase gene is placed under the control of a yeast promoter and stably introduced into yeast cells on an autonomously replicating vector. Expression is achieved by introducing into the yeast cells a second plasmid that contains the gene of interest under the control of the bacteriophage T7 promoter (Chen et al. 1987).

BACTERIOPHAGES SP6, T3, AND T7 DNA-DEPENDENT RNA POLYMERASES

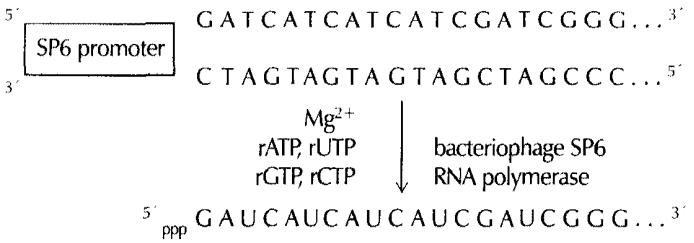
Activity: 5′→3′ RNA polymerase

Substrate: Double-stranded DNA molecules containing bacteriophage SP6, T3, or T7 promoters.

Reaction:

$$\text{double-stranded DNA} \xrightarrow[\substack{Mg^{2+}\\ \text{rATP, rUTP, rGTP, rCTP}}]{\text{bacteriophage SP6, T3, or T7 RNA polymerase}} \text{RNA} + PP_i + \text{template}$$

For example:



LIGASES, KINASES, AND PHOSPHATASES

DNA ligases catalyze end-to-end joining of pieces of DNA. The ligases used most often in cloning are encoded by bacteriophage T4, although there is a less versatile enzyme available from uninfected *E. coli*. Both types of ligases are used primarily on DNA substrates with 5′-terminal phosphate groups.

RNA ligase is a bacteriophage T4 enzyme that is capable of covalently joining single-stranded RNA (or DNA) molecules containing 5′-phosphate and 3′-hydroxyl termini. However, the primary use of this enzyme has been in 3′ end-labeling of RNA. This is accomplished using ^{32}P-labeled mononucleoside 3′,5′-bisphosphate (pNp), which is added to the 3′-hydroxyl terminus of RNA.

The DNA ligases used in molecular cloning differ in their abilities to ligate noncanonical substrates, such as blunt-ended duplexes, DNA-RNA hybrids, or single-stranded DNAs. These and other properties are summarized in Table 1-12 in Chapter 1 and discussed below. For additional information on DNA ligases, please see the information panel on **DNA LIGASES** in Chapter 1.

DNAs that lack the required phosphate residues can be prepared for ligation by phosphorylation with bacteriophage T4 polynucleotide kinase. Conversely, DNAs can be rendered resistant to ligation by enzymatic removal of phosphate residues from their 5′ termini with phosphatases. The properties of bacteriophage T4 polynucleotide kinase are summarized in Table A4-6 and described in greater detail below. The data in this table have been collected from numerous papers published over the years. The values serve as accurate guidelines for using the enzymes, but optimal conditions will always vary slightly with the enzyme preparation (degree of purity) and the DNA preparation, or when carrying out a sequence of enzymatic reactions (e.g., digestion, end-filling, and ligation) in one mixture.

TABLE A4-6 Properties of Bacteriophage Polynucleotide Kinase

	FORWARD REACTION	EXCHANGE REACTION
Nucleic acid substrate	double-stranded DNA single-stranded RNA or DNA nick or gap oligonucleotide 3′ dNMP	double-stranded DNA single-stranded RNA or DNA nick or gap oligonucleotide
K_m	double-stranded DNA, 7.6 μM 5′ dNMP, 22–143 μM ATP, 14–140 μM[b]	ADP, 300 μM[a] ATP, 10 μM[a]
pH optimum	7.4–8.0 (Tris-Cl)	6.4 (imidazole)
Sulfhydryl requirement	+	+
Mg^{2+} requirement	+	+
Effect of ionic strength	stimulated by NaCl and KCl excess KCl inhibits on all substrates except single-stranded DNA	no information
Inhibitors (50% inhibition)	$(NH_4)_2SO_4$, 7 mM P_i, 20 mM PP_i, 5 mM	P_i, 50 mM
Activators	polyamine, 2 mM, 300%	no information

[a]These are concentrations that give optimal activity at pH 6.4, not K_ms.
[b]K_m for ATP varies with substrate.

Bacteriophage T4 DNA Ligase

(Bacteriophage T4-infected *E. coli*)

Bacteriophage T4 DNA ligase, a polypeptide of M_r = 68,000, catalyzes the formation of phosphodiester bonds between adjacent 3′-hydroxyl and 5′-phosphate termini in DNA (Weiss et al. 1968). The properties of the enzymes are summarized in Table 1-12 in Chapter 1.

USES

1. Joining DNA molecules with compatible cohesive termini. Intermolecular ligation is stimulated by low concentrations of agents, such as polyethylene glycol, that promote the efficient interaction of macromolecules in aqueous solutions (please see the information panel on **CONDENSING AND CROWDING REAGENTS** in Chapter 1).

2. Joining blunt-ended double-stranded DNA molecules to one another or to synthetic linkers. This reaction is much slower than ligation of cohesive termini. However, the rate of blunt-end ligation is improved greatly by the addition of monovalent cations (150–200 mM NaCl) and low concentrations of polyethylene glycol (Pheiffer and Zimmerman 1983; Hayashi et al. 1986).

NOTES

- At least three different assays are used to measure the activity of bacteriophage T4 DNA ligase. Most manufacturers (apart from New England Biolabs) now calibrate the enzyme in Weiss units (Weiss et al. 1968). One Weiss unit is the amount of enzyme that catalyzes the exchange of 1 nmole of ^{32}P from pyrophosphate into [γ, β-^{32}P]ATP in 20 minutes at 37°C. One Weiss unit corresponds to 0.2 unit determined in the exonuclease resistance assay (Modrich and Lehman 1970) and to 60 cohesive-end units (as defined by New England Biolabs). 0.015 Weiss unit of bacteriophage T4 DNA ligase therefore will ligate 50% of the *Hin*dIII fragments of bacteriophage λ (5 µg) in 30 minutes at 16°C. Throughout this manual, bacteriophage T4 DNA ligase is given in Weiss units.

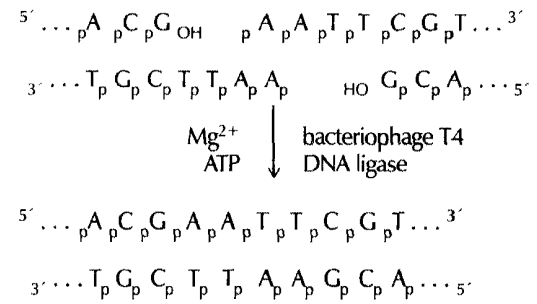
- Bacteriophage T4 DNA ligase is not inhibited by the presence of dNTPs and works adequately in virtually all buffers used for digestion of DNA with restriction enzymes.

BACTERIOPHAGE T4 DNA LIGASE

Activity: Ligation of cohesive DNA termini or nicks

Substrate: Active on double-stranded DNA with complementary cohesive termini that base pair to bring together 3′-hydroxyl and 5′-phosphate termini. In addition, the enzyme is active on nicked DNA and active, albeit far less efficiently, on RNA substrates. (For a more complete description of substrates, see Engler and Richardson 1982.)

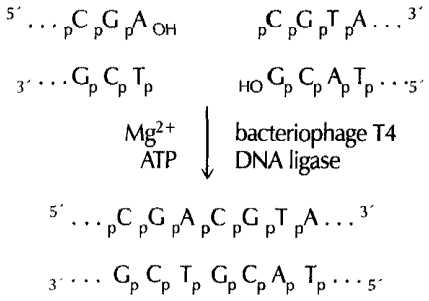
Reaction:



Activity: Ligation of blunt ends

Substrate: High concentrations of blunt-ended, double-stranded DNA containing 5′-phosphate and 3′-hydroxyl termini.

Reaction:



E. coli DNA Ligase

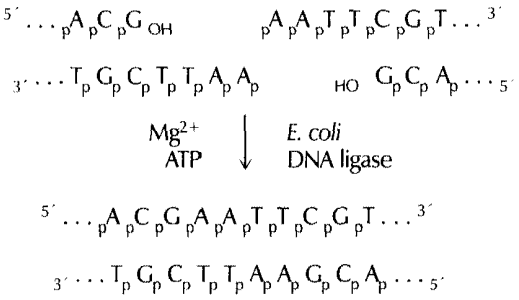
E. coli DNA ligase catalyzes the formation of phosphodiester bonds in double-stranded DNA containing complementary protruding 5′ or 3′ termini (Panasenko et al. 1977, 1978). The reaction requires NAD^+ as a cofactor. Initial studies indicated that this enzyme would not ligate blunt-ended double-stranded DNA, but subsequent studies revealed that blunt-end ligation can be achieved in the presence of polyethylene glycol or Ficoll, compounds that act as volume excluders (please see the information panel on **CONDENSING AND CROWDING REAGENTS** in Chapter 1). This effectively increases the concentration of DNA termini and enzyme (Zimmerman and Pheiffer 1983). *E. coli* DNA ligase has been used in cDNA cloning methods based on replacement synthesis, as described by Okayama and Berg (1982), because of its inability to join adjacent RNA and DNA segments that arise during the synthesis of the second strand of cDNA. It is not, however, widely used in other molecular cloning procedures, since bacteriophage T4 DNA ligase is capable of efficiently joining blunt ends in the absence of volume excluders. *E. coli* DNA ligase does not ligate RNA.

E. COLI DNA LIGASE

Activity: Ligation of cohesive DNA termini or nicks

Substrate: Active on double-stranded DNA with complementary cohesive termini that base pair to bring together 3′-hydroxyl and 5′-phosphate termini. In addition, the enzyme is active on nicked DNA. Blunt termini can be ligated in the presence of crowding reagents.

Reaction:



Bacteriophage T4 RNA Ligase

(Bacteriophage T4-infected *E. coli*)

Bacteriophage T4 RNA ligase catalyzes the covalent joining of 5′-phosphate termini in single-stranded DNA or RNA to 3′-hydroxyl termini in single-stranded DNA or RNA (Uhlenbeck and Gumport 1982).

USES

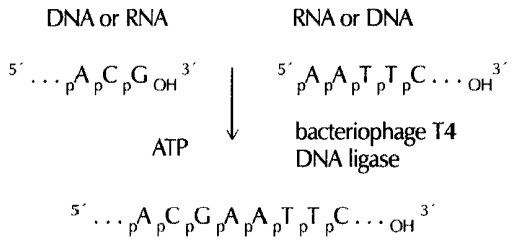
1. Because small molecules (e.g., pNp) are effective substrates, bacteriophage T4 RNA ligase can be used to radiolabel the 3′ termini of RNA molecules in vitro (Uhlenbeck and Gumport 1982).
2. Ligation of oligodeoxyribonucleotides.
3. Bacteriophage T4 RNA ligase has been reported to stimulate the activity of bacteriophage T4 DNA ligase (Sugino et al. 1977). However, agents such as polyethylene glycol that increase macromolecular crowding are equally effective and much less expensive.

BACTERIOPHAGE T4 RNA LIGASE

Activity: RNA ligase

Substrate: 5′-phosphate acceptors include single-stranded DNA and RNA. Phosphate donors include single-stranded DNA and RNA and nucleotides such as pNp.

Reaction:



Thermostable DNA Ligases

The genes encoding thermostable ligases from several thermophilic bacteria have been cloned, sequenced, and expressed to high levels in *E. coli* (e.g., please see Takahashi et al. 1984; Barany and Gelfand 1991; Lauer et al. 1991; Jónsson et al. 1994). Several of these enzymes are available from commercial sources. Like the *E. coli* enzyme, almost all thermostable ligases use NAD^+ as a cofactor and work preferentially at nicks in double-stranded DNA. In addition, thermostable ligases, like their mesophilic homolog, can catalyze blunt-end ligation in the presence of crowding agents, even at elevated temperatures (Takahashi and Uchida 1986). Because thermostable ligases retain activity after multiple rounds of thermal cycling, they are used extensively in the ligase amplification reaction to detect mutations in mammalian DNAs.

Bacteriophage T4 DNA Polynucleotide Kinase

(Bacteriophage T4-infected *E. coli*)

Bacteriophage T4 polynucleotide kinase catalyzes the transfer of the γ-phosphate of ATP to a 5′ terminus of DNA or RNA (Richardson 1971). Two types of reactions are commonly used. In the *forward* reaction, the γ-phosphate is transferred to the 5′ terminus of dephosphorylated DNA (Richardson 1971). In the *exchange* reaction, an excess of ADP causes bacteriophage T4 polynucleotide kinase to transfer the terminal 5′-phosphate from phosphorylated DNA to ADP; the DNA is then rephosphorylated by transfer of a radiolabeled γ-phosphate from [γ-^{32}P]ATP (Berkner and Folk 1977). In addition to its phosphorylation activity, bacteriophage T4 polynucleotide kinase carries a 3′ phosphatase activity (Richardson 1981). The properties of this enzyme are summarized in Table A4-6 and in Richardson (1981).

USES

1. Radiolabeling 5′ termini in DNA for sequencing by the Maxam-Gilbert technique (Maxam and Gilbert 1977), for nuclease S1 analysis, and for other uses requiring terminally labeled DNA.
2. Phosphorylating synthetic linkers and other fragments of DNA that lack terminal 5′ phosphates in preparation for ligation.

NOTES

- Bacteriophage T4 polynucleotide kinase is difficult to purify from infected cells, and impure preparations are not uncommon. Wherever possible, use bacteriophage T4 polynucleotide kinase that has been purified from cells expressing high levels of a cloned copy of the bacteriophage T4 gene.
- When setting up reactions involving the termini of nucleic acid molecules, the concentration of the reacting species can be calculated using Table A4-7 as a guide.
- Spermidine stimulates incorporation of [γ-^{32}P]ATP and inhibits a nuclease present in some preparations of bacteriophage T4 polynucleotide kinase.
- ATP should be present at a concentration of at least 1 μM in the forward reaction and at least 2 μM in the exchange reaction. Maximum enzyme activity requires still higher concentrations (please see Table A4-6).
- The DNA to be phosphorylated should be rigorously purified by gel electrophoresis, density gradient centrifugation, or chromatography on columns of Sepharose CL-4B in order to remove low-molecular-weight nucleic acids. Although such contaminants may make up only a small fraction of the weight of the nucleic acids in the preparation, they provide a much larger proportion of the 5′ termini. Unless steps are taken to remove them, contaminating low-molecular-weight DNAs and RNAs can be the predominant species of nucleic acids that are labeled in bacteriophage T4 polynucleotide kinase reactions.
- Ammonium ions are strong inhibitors of bacteriophage T4 polynucleotide kinase. Therefore, DNA should not be dissolved in, or precipitated from, buffers containing ammonium salts prior to treatment with the kinase.
- Low concentrations of phosphate also inhibit bacteriophage T4 polynucleotide kinase. Imidazole buffer (pH 6.4) is therefore the buffer of choice for the exchange reaction, and Tris buffer is the buffer of choice for the forward reaction.

TABLE A4-7 Concentration of Ends in Kinasing Reactions

Size of Double-stranded DNA (in Base Pairs)	Amount of DNA Required to Contribute 1 pmole of 5′ Termini (in μg)
50	1.7×10^{-2}
100	3.3×10^{-2}
250	8.4×10^{-2}
500	1.7×10^{-1}
1000	3.3×10^{-1}
2500	8.4×10^{-1}
5000	1.7

BACTERIOPHAGE T4 POLYNUCLEOTIDE KINASE

Activity: Kinase (forward reaction)

Substrate: Single- or double-stranded DNA with 5′-hydroxyl terminus; RNA with a 5′-hydroxyl terminus. The enzyme phosphorylates protruding 5′ single-stranded termini more rapidly than blunt ends or recessed 5′ termini; however, with sufficient enzyme and ATP, such termini can be completely phosphorylated. The reaction at nicks or gaps in double-stranded DNA is less efficient than for single-stranded termini; however, with sufficient concentrations of ATP and enzyme, gaps can be completely phosphorylated and nicks can be phosphorylated to 70%.

Reaction:

$$DNA_{OH}{}^{5'} \text{ or } RNA_{OH}{}^{5'} \xrightarrow[\substack{[\gamma\text{-}^{32}P]ATP \\ \text{dithiothreitol} \\ Mg^{2+}}]{\text{bacteriophage T4 polynucleotide kinase}} {}^{5'}[^{32}P]DNA \text{ or } {}^{5'}[^{32}P]RNA + ADP$$

For example:

$$^{5'}HO\,C_pG_pC\ldots{}^{3'}$$

$$[\gamma\text{-}^{32}P]dATP,\ \text{dithiothreitol},\ Mg^{2+} \downarrow \text{bacteriophage T4 polynucleotide kinase}$$

$$^{5'}{}^{*}_{p}C_pG_pC\ldots{}^{3'} + ADP$$

Activity: Kinase (exchange reaction)

Substrate: Single-stranded DNA with a 5′-phosphate terminus is most efficiently labeled (96%). Recessed 5′-phosphate termini are labeled to 70% with sufficient enzyme. 5′-phosphate groups at nicks are labeled 30-fold less efficiently than single-stranded 5′-phosphate termini.

Reaction:

$$^{5'}{}_{p}C_pG_pC\ldots{}^{3'} + \text{excess ADP}$$

$$[\gamma\text{-}^{32}P]dATP,\ \text{dithiothreitol},\ Mg^{2+} \downarrow \text{bacteriophage T4 polynucleotide kinase}$$

$$^{5'}{}^{*}_{p}C_pG_pC\ldots{}^{3'} + ADP + ATP$$

Alkaline Phosphatases

(*E. coli*, calf intestine, and shrimp)

The three alkaline phosphatases — bacterial alkaline phosphatase (BAP), calf intestinal alkaline phosphatase (CIP), and shrimp alkaline phosphatase (SAP) — catalyze the removal of 5′-phosphate residues from DNA, RNA, rNTPs, and dNTPs. For further details, please see the information panel on **ALKALINE PHOSPHATASE** in Chapter 9.

USES

1. Removing 5′ phosphates from DNA or RNA prior to labeling 5′ termini with ^{32}P.
2. Removing 5′ phosphates from fragments of DNA to prevent self-ligation.

NOTES

- BAP is the most active of the three enzymes, but it is also far more resistant to heat and detergents. It is therefore difficult to inhibit BAP completely at the end of dephosphorylation reactions.
- Proteinase K is used to digest CIP, which must be completely removed if subsequent ligations are to work efficiently. An alternative method is to inactivate the CIP by heating to 65°C for 1 hour (or 75°C for 10 minutes) in the presence of 5 mM EDTA (pH 8.0) and then to purify the dephosphorylated DNA by extraction with phenol:chloroform.
- SAP is extremely heat-labile and can be denatured completely and irreversibly by heating to 65°C for 15 minutes.

ALKALINE PHOSPHATASES

Activity: Phosphatase

Substrate: Single- or double-stranded DNA and RNA; rNTPs and dNTPs.

Reaction:

$$^{5'}{}_{p}\text{DNA or } ^{5'}{}_{p}\text{RNA} \xrightarrow{\text{alkaline phosphatase}} {}^{5'}{}_{OH}\text{DNA or } {}^{5'}{}_{OH}\text{RNA}$$

NUCLEASES

Ribonuclease H

Ribonuclease H (RNase H) catalyzes the endonucleolytic degradation of the RNA moiety of DNA-RNA hybrids, generating oligoribonucleotides of varying chain lengths with 3′-hydroxyl and 5′-phosphate termini. RNase H was first recognized and isolated from calf thymus (Stein and Hausen 1969; Hausen and Stein 1970), but the enzyme is now known to be present in a wide variety of mammalian tissues, yeasts, prokaryotes, and virus particles. Many types of cells contain more than one RNase H.

In many retroviruses, RNase H is associated with the multifunctional enzyme reverse transcriptase and carries out important functions at several stages during the transcription of the viral genome into DNA. In eubacteria, RNase H is believed to be required for the removal of RNA primers from Okazaki fragments, for processing of transcripts into primers that are used by DNA polymerase I to initiate DNA synthesis, and to remove R-loops that provide sites for opportunistic initiation of unregulated DNA synthesis at the chromosomal origin of replication in *E. coli*. RNase H is presumed to carry out similar functions in eukaryotic cells.

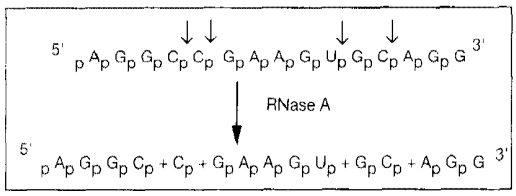
RNase H has been reported to increase markedly the inhibition of gene expression by antisense oligodeoxynucleotides. Hybrids between these oligonucleotides and specific sequences in mRNAs are sensitive to degradation by the enzyme. RNase H is required for initiation of replication at the origin (ori) of colicin E1 (colE1)-type plasmids in vitro. The enzyme also seems to suppress initiation of DNA synthesis at sites other than ori.

X-ray crystallographic analysis shows that *E. coli* RNase H consists of two domains, one of which contains an Mg^{2+}-binding site enmeshed in β strands, a fold previously recognized in DNase I. For further information and references, please see Crouch (1990), Wintersberger (1990), Hostomsky et al. (1993), Jung and Lee (1995), Kanaya and Ikehara (1995), Rice et al. (1996), Crooke (1998), and the information panel on **RIBONUCLEASE H** in Chapter 8.

Ribonuclease A (Pancreatic)

(Bovine pancreas)

Ribonuclease A (RNase A) is an endoribonuclease that specifically attacks single-stranded RNA 3′ to pyrimidine residues and cleaves the phosphate linkage to the adjacent nucleotide. The end products are pyrimidine 3′ phosphates and oligonucleotides with terminal pyrimidine 3′ phosphates (Davidson 1972). RNase A, which works in the absence of cofactors and divalent cations, can be inhibited by placental RNase inhibitor (Blackburn et al. 1977) or by vanadyl-ribonucleoside complexes (Puskas et al. 1982).



USES

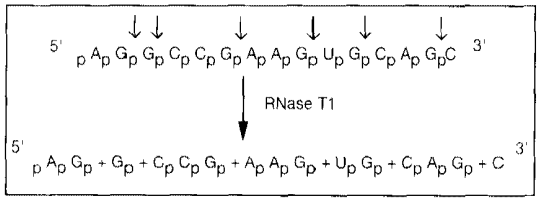
1. Removing unhybridized regions of RNA from DNA-RNA or RNA-RNA hybrids.
2. Mapping single-base mutations in DNA or RNA (Myers et al. 1985; Winter et al. 1985). In this method, single-base mismatches in RNA-DNA or RNA-RNA hybrids are recognized and cleaved by RNase A. A ^{32}P-labeled RNA probe complementary to wild-type DNA or RNA is synthesized in vitro using a plasmid containing a bacteriophage SP6 or T7 promoter. The RNA probe is then annealed to test DNA or RNA containing a single-base substitution. The resulting single-base mismatch is cleaved by RNase A, and the location of the mismatch is then determined by analyzing the sizes of the cleavage products by gel electrophoresis; ~50% of all possible single-base mismatches can be detected by this method.

PREPARATION OF RNASE THAT IS FREE OF DNASE

Dissolve pancreatic RNase (RNase A) at a concentration of 10 mg/ml in 0.01 M sodium acetate (pH 5.2). Heat to 100°C for 15 minutes. Allow it to cool slowly to room temperature. Adjust the pH by adding 0.1 volume of 1 M Tris-Cl (pH 7.4). Dispense into aliquots and store at –20°C. RNase precipitates when concentrated solutions are heated to 100°C at neutral pH.

Ribonuclease T1

Ribonuclease T1 (RNase T1) is an endoribonuclease that specifically attacks the 3′-phosphate groups of guanine nucleotides and cleaves the 5′-phosphate linkage to the adjacent nucleotide. The end products are guanosine 3′ phosphates and oligonucleotides with terminal guanosine-3′-phosphate groups (Davidson 1972).



USE

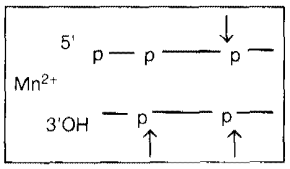
Removing unhybridized regions of RNA from DNA-RNA or RNA-RNA hybrids.

Deoxyribonuclease I (Pancreatic)

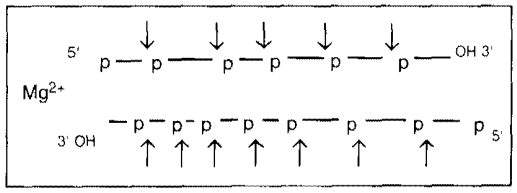
(Bovine pancreas)

The endonuclease DNase I from bovine pancreas is a glycoprotein that degrades double-stranded DNA by an in-line SN-2 mechanism involving nucleophilic attack on the scissile phosphodiester bond. The phosphate group remains attached to the new 5′ terminus. DNase I requires divalent metal ions for DNA hydrolysis (Kunitz 1950) and displays maximal activity in the presence of Ca^{2+} and either Mg^{2+} or Mn^{2+} (Price 1972). These cations work synergistically; in the presence of both Ca^{2+} (0.1 mM) and Mg^{2+} (10 mM), the rate of hydrolysis is greater than the sum of the rates for either cation alone. However, high concentrations of Mg^{2+} (>50 mM), but not of Ca^{2+}, are inhibitory.

The mode of action and the specificity of the enzyme are affected by the type of divalent cations used. In the presence of Mn^{2+}, or when very high concentrations of the enzyme are used in the absence of monovalent cations, DNase I breaks both strands of superhelical double-stranded DNA simultaneously at approximately the same site (Melgar and Goldthwait 1968; Campbell and Jackson 1980).



In the presence of Mg^{2+}, DNase I works in a totally different fashion and introduces nicks into each strand of double-stranded DNA independently. As more and more nicks accumulate, the number of base pairs between adjacent nicks on opposite strands gradually decreases and is eventually insufficient to hold the molecule together. The terminal products of this reaction are a complex mixture of acid-soluble 5′-phosphorylated oligonucleotides.



Whether this nicking activity of DNase I displays sequence specificity remains a surprisingly murky topic, with conflicting data from several groups. On the one hand, biochemical analysis of the digestion products of bulk *E. coli* DNA shows only weak sequence specificity (Ehrlich et al. 1973; Bernardi et al. 1975). On the other hand, footprints generated by digestion of double-stranded DNA with DNase I often show evidence of preferential cleavage by the enzyme. In addition, Scheffler et al. (1968) showed that DNase I has a very marked preference for cleaving poly(d[A-T])·poly(d[A-T]) to the 5′ side of T residues, a result that was confirmed with homopolymeric poly(d[A-T]) by Lomonossoff et al. (1981). However, this apparent specificity may be due more to a special alternating B conformation thought to be associated with the homopolymer, rather than to preference of the enzyme for particular sequences (Klug et al. 1979). Support for this idea comes from analysis of the crystal structure of complexes between DNase I and short double-stranded oligonucleotides (Suck et al. 1984; Suck and Oefner 1986). An exposed loop of the enzyme binds in the minor groove of B-DNA with both strands of the nucleic acid bending to make contact with the enzyme. Suck et al. (1988) have suggested that, in contrast to the results of Scheffler et al. (1968) and Lomonossoff et al. (1981), A-T tracts in double-stranded DNA might be relatively resistant to cleavage with DNase I.

In summary, whether DNase I nicks double-stranded DNA in a sequence-dependent fashion appears to be influenced by the structure of the template. The enzyme introduces nicks into complex DNAs with limited regard for sequence. However, on synthetic oligonucleotides of defined sequence, the enzyme shows a higher degree of preference, perhaps because the conformation of these substrates limits access of the enzyme to certain classes of phosphodiester bonds.

When very short double-stranded DNAs are used as substrates, DNase I also exhibits a topographical specificity that is manifested as "end-effects." The probability that a particular phosphodiester bond will be cleaved increases as a function of its distance from the 5′ end of a DNA strand, at least as far as the eighth phosphodiester bond (Galas and Schmitz 1978; Lomonossoff et al. 1981). This preference exists because the enzyme efficiently cleaves only when it can interact with three or four nucleotides that are 5′ to the cleavage site. The predominant products of complete digestion of double-stranded DNA are therefore 5′-phosphorylated tetranucleotides (Bernardi et al. 1975). Two methods have been used in molecular cloning to limit the action of DNase I to a single endonucleolytic cleavage per molecule of template DNA:

- As discussed above, in the presence of a transition metal ion such as Mn^{2+}, DNase I cleaves both strands of superhelical DNA at approximately the same site (Melgar and Goldthwait 1968) to yield fragments that are blunt-ended or that have protruding termini only one or two nucleotides in length. The resulting linear molecules are relatively resistant to further cleavage by the nuclease. This reaction has been used to generate random deletions in a segment of target DNA cloned in a plasmid or bacteriophage M13 vector (Frischauf et al. 1980; Anderson 1981; Hong 1982). Because the sites of double-stranded cleavage are distributed in a statistically random fashion, the population of closed circular DNAs is converted into a permuted set of linear molecules. These are then digested with a restriction enzyme whose unique site of cleavage lies at one end of the target DNA. Recircularization of the resulting population generates clones that lack sequences lying between the site of DNase I cleavage and the restriction site.
- In the presence of subsaturating quantities of an intercalating dye such as ethidium bromide, DNase I randomly introduces a single nick into one strand of closed circular DNAs. The resulting relaxed circular molecules are then relatively resistant to further cleavage by the enzyme. DNase I can therefore be used to introduce a single nick into closed circular DNAs in preparation for resecting prior to bisulfite-mediated mutagenesis (Greenfield et al. 1975).

USES

1. To remove DNA templates from in vitro transcription reactions and from preparations of mRNA. These RNAs are contaminated with large amounts of DNA that must be removed before analysis by northern hybridization, construction of cDNA libraries, reverse transcriptase (RT)-PCR, etc. Removal of contaminating DNA is particularly important when purifying RNAs from transfected cells or cells infected with DNA viruses. Unfortunately, many commercial preparations of pancreatic DNase I are contaminated with significant amounts of RNase. DNase I that is free of RNase can be obtained commercially but at great cost. If RNase-free DNase is used on a regular basis, please see the panel on **PREPARATION OF DNASE THAT IS FREE OF RNASE** on the following page.
2. To digest DNA that is left unprotected by interaction with proteins (DNA footprinting) (Galas and Schmitz 1978; Schmitz and Galas 1979). DNase I was the reagent originally used to develop DNA footprinting and despite the subsequent discovery of elegant chemical methods to cleave DNA in a sequence-independent fashion, DNA footprinting with DNase I remains by far the most popular way to localize specific interactions between proteins and DNA.
3. To introduce random single-stranded nicks into double-stranded DNA to generate templates for nick-translation reactions (Maniatis et al. 1975; Rigby et al. 1977). In this case, very small amounts of the enzyme are used to prevent wholesale destruction of the template DNA.

PREPARATION OF DNASE THAT IS FREE OF RNASE

mRNA prepared from uninfected mammalian cells contains only small amounts of DNA that generally do not compromise northern hybridization or other types of RNA analysis. However, mRNAs prepared from transfected mammalian cells or from cells infected with DNA viruses are contaminated with large amounts of DNA that must be removed by digestion with DNase I. Unfortunately, many commercial preparations of pancreatic DNase I, even those that claim to be RNase-free, are contaminated with significant amounts of RNase. In addition, the use of commercially prepared RNase-free DNase I can become expensive when many samples are prepared. DNase I can be treated by heating in the presence of iodoacetate to reduce RNase activity by ~98%. The resulting preparations of DNase are acceptable for all but the most stringent applications (e.g., this method should not be used to remove DNA from RNA to be used in construction of a cDNA library). DNase purified in this manner should always be used in the presence of a protein inhibitor of RNase (please see the information panel on **INHIBITORS OF RNASES** in Chapter 7).

1. Dissolve 10 mg of pancreatic DNase (Sigma) in 10 ml of 0.1 M iodoacetic acid, 0.15 M sodium acetate (pH 5.2).
2. Heat the solution to 55°C for 45 minutes. Cool the solution to 0°C, and add 1 M $CaCl_2$ to a final concentration of 5 mM.
3. Dispense the DNase I into small aliquots and store at –20°C.

BAL 31 Nuclease

(*Alteromonas espejiana* BAL 31)

BAL 31 is predominantly a 3′ exonuclease that removes mononucleotides from both 3′ termini of the two strands of linear DNA. BAL 31 is also an endonuclease; thus, the single-stranded DNA generated by the 3′ exonuclease activity is degraded by the endonuclease. The mechanisms of these reactions are complex and are summarized in the information panel on **BAL 31** in Chapter 13.

Degradation is absolutely dependent on the presence of calcium, and the reaction can therefore be stopped at different stages by the addition of the chelating agent EGTA. Because degradation occurs relatively uniformly from the termini of DNA, digestion with BAL 31 can be used to map restriction sites in small fragments of DNA (Legerski et al. 1978). DNA is digested with BAL 31, and samples are withdrawn at different times and placed in a solution containing EGTA. After digestion of these samples with the restriction enzyme of interest, restriction fragments can be seen to disappear in a defined order. By using a DNA consisting of vector sequences at one terminus (for which the restriction map is known) and unmapped sequences at the other, it is possible to distinguish fragments from the two termini and to deduce the order of the fragments in the unmapped DNA.

BAL 31 can also be used to remove unwanted sequences from the termini of DNAs before cloning. After treatment with the exonuclease/endonuclease, the termini of the DNA are repaired with bacteriophage T4 DNA polymerase or the Klenow fragment of *E. coli* DNA polymerase I. Synthetic linkers are added to the DNA, which is then inserted into a suitable vector. In this way, it is possible to generate a set of deletions from a defined endpoint in DNA. Although the enzyme is predominantly a 3′ exonuclease, it also has a DNA endonuclease activity and cleaves internally in single-stranded regions of DNA or in double-stranded DNA that contains helical distortions (Lau and Gray 1979; Gray et al. 1981; Wei et al. 1983). BAL 31 will also digest RNA, albeit inefficiently.

USES

1. Removing nucleotides from the termini of double-stranded DNA in a controlled manner. The shortened molecules can be used for a variety of purposes such as to produce deletions, to position a desired sequence next to a promoter or other controlling element, or to attach synthetic linkers at desired sites in the DNA.
2. Mapping restriction sites in DNA (Legerski et al. 1978).
3. Mapping secondary structure in DNA, for example, junctions between B-DNA and Z-DNA or sites of covalent or noncovalent modifications in double-stranded DNA (Gray et al. 1981; Wei et al. 1983).
4. Removing nucleotides from double-stranded RNA in preparing recombinant RNAs (Miele et al. 1983).

NOTES

- When the products of BAL 31 digestion are to be ligated, it is important to consider that the 3′ exonuclease activity of the enzyme works ~20-fold more efficiently than the DNA endonuclease. Thus, the average length of single-stranded tails created by digestion of linear double-stranded DNA is dependent on the enzyme concentration. At high enzyme concentrations (2–5 units/ml), an average of five nucleotides of single-stranded DNA remain per terminus and 10–20% of the molecules can be ligated to blunt-ended DNA without further treatment. At low enzyme concentrations (0.1–0.2 unit/ml), the single-stranded termini may be very long and the efficiency of blunt-end ligation is very low. Repair with bacteriophage T4 DNA poly-

merase (or, in some cases, the Klenow fragment) is almost obligatory before cloning DNAs treated with either high or low concentrations of BAL 31.

- Most commercial preparations of BAL 31 contain two kinetically distinct forms of the enzyme, a fast and a slow form. The slow form is a proteolytic degradation product of the fast form. Pure preparations of the fast form are available, but they are expensive (Wei et al. 1983). Results using mixed preparations will vary, depending on the relative amounts of the two forms in the initial preparation and the rate of conversion of the fast to the slow form during the assay.

  Preparations rich in the fast form are preferred for such tasks as removal of long (>1000 bp) segments from the termini of double-stranded DNA; degradation of double-stranded RNA; and mapping of restriction sites, B-Z DNA junctions, and lesions in double-stranded DNA. The slow form of the enzyme is used to remove short segments (10–100 bp) from the termini of double-stranded DNA. Mixed preparations of the enzyme can be used for any of these tasks, although the results will vary as mentioned above.

- BAL 31 works asynchronously, generating a population of DNA molecules whose termini have been resected to various extents and whose single-stranded tails vary in length. Following digestion with BAL 31 and repair with bacteriophage T4 DNA polymerase (or the Klenow fragment), it is often more efficient to isolate DNAs of the required size by gel electrophoresis rather than to screen very large numbers of randomly generated clones.

- BAL 31 degrades AT-rich sequences significantly more rapidly than it degrades GC-rich regions. Thus, molecules that terminate in AT-rich regions are underrepresented in populations of DNAs that have been digested with the enzyme.

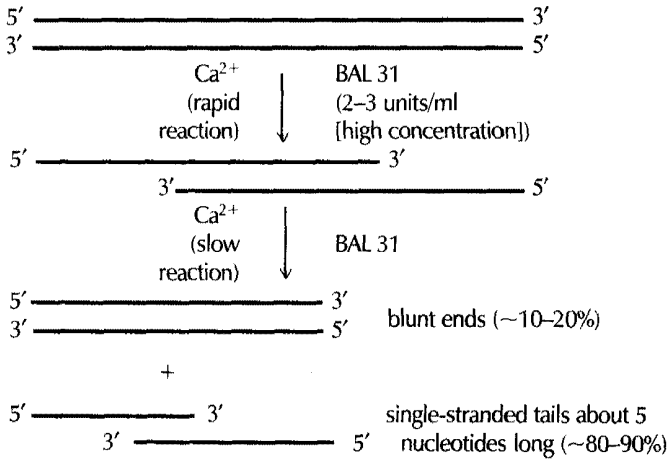
- BAL 31 should not be frozen. Store the enzyme at 4°C.

BAL 31 NUCLEASE

Activity: Exonuclease/endonuclease

Substrate: BAL 31 degrades double-stranded DNA sequentially from both termini. The mechanism is thought to involve a rapid exonucleolytic degradation followed by a slow endonucleolytic reaction on the complementary strand. Double-stranded DNA with blunt or protruding 3′-hydroxyl termini are degraded to shorter double-stranded molecules. The enzyme is also active at nicks, on single-stranded DNA with 3′-hydroxyl termini, and on double-stranded RNA molecules.

Reaction:

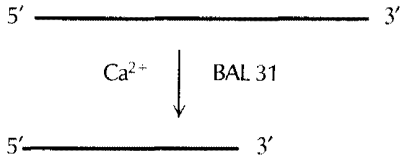


BAL 31 NUCLEASE

Activity: Exonuclease (shortens single-stranded DNA)

Substrate: Single-stranded DNA with 3′-hydroxyl termini.

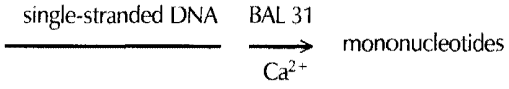
Reaction:



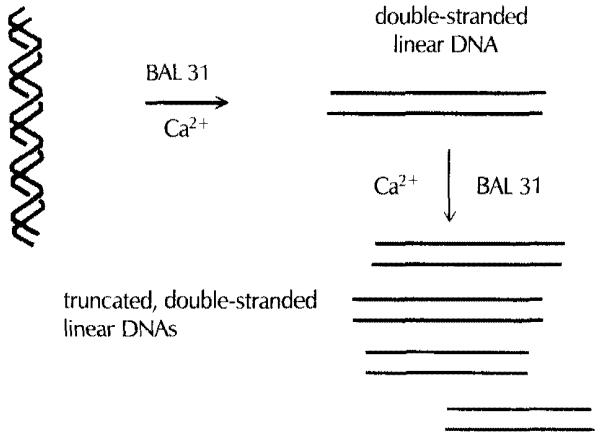
Activity: Endonuclease

Substrate: Single-stranded DNA; supercoiled DNA; DNA with B-DNA, Z-DNA junctions and other non-B-DNA comformations.

Reaction:



superhelical double-stranded DNA



Nuclease S1

(Aspergillus oryzae)

Nuclease S1 degrades single-stranded DNA or RNA (Vogt 1973) to yield 5′-phosphate mono- or oligonucleotides. Double-stranded DNA, double-stranded RNA, and DNA-RNA hybrids are relatively resistant to the enzyme. However, double-stranded nucleic acids are digested completely by nuclease S1 if they are exposed to very large amounts of the enzyme. Moderate amounts of the enzyme will cleave double-stranded nucleic acids at nicks or small gaps (Kroeker and Kowalski 1978). For further information, please see both the introduction to Protocol 10 and the information panel on **NUCLEASE S1** in Chapter 7.

USES

1. Analyzing the structure of DNA-RNA hybrids (Berk and Sharp 1977; Favaloro et al. 1980).
2. Removing single-stranded tails from DNA fragments to produce blunt ends.
3. Opening the hairpin loop generated during synthesis of double-stranded cDNA.

NOTE

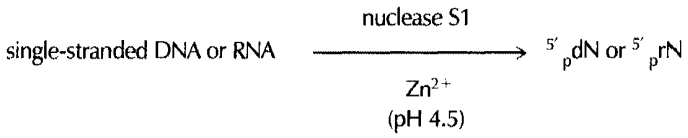
- Since the enzyme works at low pH, depurination often occurs, which limits the usefulness of nuclease S1 for some applications.

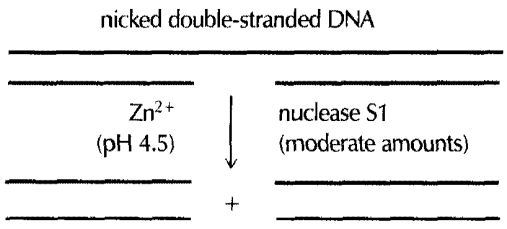
NUCLEASE S1

Activity: Single-strand-specific nuclease

Substrate: Single-stranded DNA or RNA; more active on DNA than on RNA.

Reaction:





Mung Bean Nuclease

(Mung bean sprouts)

Mung bean nuclease degrades single-stranded DNA to mono- or oligonucleotides with phosphate groups at their 5′ termini (Laskowski 1980). Double-stranded DNA, double-stranded RNA, and DNA-RNA hybrids are relatively resistant to the enzyme. However, double-stranded nucleic acids are digested completely by mung bean nuclease if they are exposed to very large amounts of the enzyme (Kroeker and Kowalski 1978).

Although mung bean nuclease and nuclease S1 are similar to each other in their physical and catalytic properties, mung bean nuclease may be less severe in its action than nuclease S1. For example, nuclease S1 has been shown to cleave the DNA strand opposite a nick in a duplex, whereas mung bean nuclease will only attack the nick after it has been enlarged to a gap several nucleotides in length (Kroeker and Kowalski 1978). For further details, please see the information panel on **MUNG BEAN NUCLEASE** in Chapter 7.

USES

1. Converting protruding termini of DNA to blunt ends.
2. Analyzing the structure of DNA-RNA hybrids.

Exonuclease III

(*E. coli*)

Exonuclease III catalyzes the stepwise removal of 5′ mononucleotides from the 3′-hydroxyl termini of double-stranded DNA (Weiss 1976). Linear double-stranded DNA and circular DNAs containing nicks or gaps are substrates. The activity of the enzyme results in the formation of long single-stranded regions in double-stranded DNA. The enzyme also carries three other activities: an endonuclease specific for apurinic DNA, an RNase H activity (Rogers and Weiss 1980), and a 3′ phosphatase activity, which removes 3′-phosphate termini but does not cleave internal phosphodiester bonds. The exonuclease will not degrade single-stranded DNA or double-stranded DNA with a protruding 3′ terminus (Rogers and Weiss 1980).

Exonuclease III is nonprocessive and typically generates populations of molecules that have been resected to similar extents. This property simplifies the task of isolating DNA molecules whose lengths have been reduced by the desired amount. For further details, please see the information panel on **EXONUCLEASE III** in Chapter 13.

USES

1. Generating partially resected DNAs that can be used as substrates for the Klenow fragment (e.g., in the preparation of strand-specific probes) (for a similar application for bacteriophage T4 DNA polymerase, please see Figure A4-2 [p. A4.19]).
2. Generating nested sets of deletions of the terminal sequences of double-stranded linear DNAs. This reaction is usually carried out in conjunction with mung bean nuclease or nuclease S1 and is an alternative to using BAL 31. Because exonuclease III will degrade DNA with recessed 3′ termini but not termini with protruding 3′ single strands, it can be used to create unidirectional sets of deletions (Henikoff 1984). Thus, if the substrate molecule carries a protruding 3′ terminus at one end (e.g., created by digestion with *Pst*I) and a recessed 3′ terminus at the other, exonuclease III will digest only in one direction (from the recessed 3′ terminus). After

removal of the resulting single-stranded segments (with nuclease S1 or mung bean nuclease) and repair with bacteriophage T4 DNA polymerase, the deleted molecules can be inserted into an appropriate vector by blunt-end ligation.

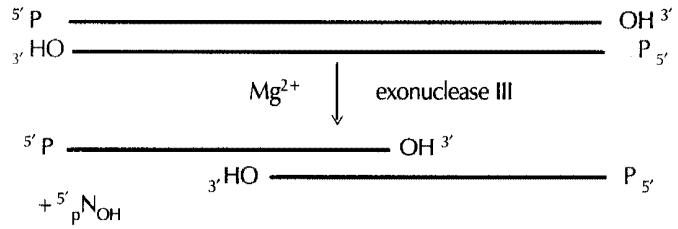
3. Some methods of site-specific mutagenesis use thiophosphate derivatives of the dNTPs for second-strand synthesis primed by the mutagenic primer. The parental template strand can be preferentially degraded with exonuclease III, increasing the frequency of mutants obtained upon transformation of *E. coli*, since exonuclease III will not cleave thioester bonds (please see the information panel on **SELECTING AGAINST WILD-TYPE DNA IN SITE-DIRECTED MUTAGENESIS** in Chapter 13).

EXONUCLEASE III

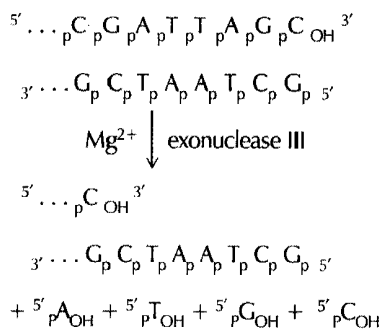
Activity: 3′ Exonuclease

Substrate: This enzyme is active on 3′-hydroxyl termini of double-stranded DNA with blunt ends or with ends containing unpaired 5′ termini and recessed 3′ termini. 3′-hydroxyl termini at nicks in double-stranded DNA are also substrates. The DNA must contain phosphodiester bonds; thioesters are not cleaved.

Reaction:



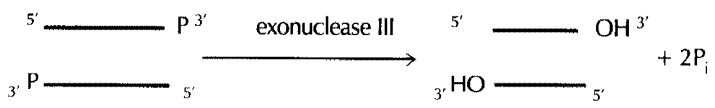
For example:



Activity: 3′ Phosphatase

Substrate: Double- or single-stranded DNA with a 3′-phosphate terminus; internal phosphodiester bonds are not cleaved.

Reaction:



Bacteriophage λ Exonuclease

(Bacteriophage λ-infected *E. coli*)

Bacteriophage λ exonuclease catalyzes the processive, stepwise release of 5′ mononucleotides from double-stranded DNA. Although the preferred substrate is double-stranded DNA with a terminal 5′ phosphate (Little et al. 1967), the enzyme will also work, albeit 100-fold less efficiently on single-stranded DNA. Double-stranded DNAs with nicks or gaps will not serve as substrates.

USE

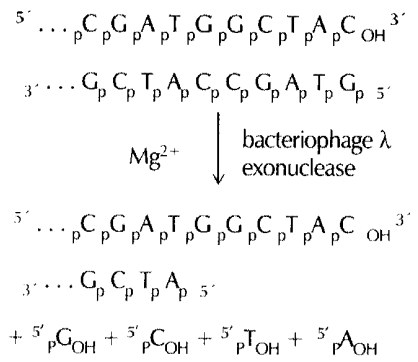
Bacteriophage λ exonuclease was used for a wide variety of purposes in the early days of molecular cloning (for review, please see Little 1981). Today, it is used chiefly to modify the 5′-phosphate termini of DNAs that are to be used as substrates for other enzymes (e.g., terminal transferase).

BACTERIOPHAGE λ EXONUCLEASE

Activity: 5′ Exonuclease

Substrate: Double-stranded DNA with 5′-phosphate termini or with protruding 5′ termini.

Reaction:



PROTEOLYTIC ENZYMES

Proteinase K

Proteinase K is a highly active serine protease of the subtilisin type (for reviews, please see Siezen et al. 1991; Siezen and Leunissen 1997) that is secreted by stationary cultures of the mold *Tritirachium album* var. Limber (Ebeling et al. 1974). The K in the enzyme's name indicates that the protease can supply the mold's total requirement for carbon and nitrogen by hydrolysis of keratin. Proteinase K catalyzes hydrolysis of a wide variety of peptide bonds but exhibits a preference for peptide bonds carboxy-terminal to aromatic and uncharged amino acids.

The mature enzyme consists of 279 amino acid residues (M_r = 28,930) (Jany et al. 1986; Gunkel and Gassen 1989) and has two binding sites for Ca^{2+}, which lie some distance from the catalytic site. The Ca^{2+} ions are not directly involved in catalysis, but they contribute to structural stability of the enzyme (Betzel et al. 1988; Müller et al. 1994). When Ca^{2+} is removed from the enzyme, some of the catalytic activity is lost because of long-range structural changes (Bajorath et al. 1988, 1989). Because the residual activity is sufficient to degrade most proteins, digestion with proteinase K is usually carried out in the presence of EDTA. In addition, proteinase K remains active in the presence of urea (1–4 M) and detergents that are routinely used to lyse mammalian cells (e.g., 0.5% SDS or 1% Triton X-100). Because proteinase K efficiently digests native proteins, it can rapidly inactivate DNases and RNases in cell lysates, which facilitates the isolation of high-molecular-weight DNA and intact RNA (Wiegers and Hilz 1971, 1972; Hilz et al. 1975).

Proteinase K is purchased as a lyophilized powder and should be dissolved at a concentration of 20 mg/ml in sterile 50 mM Tris-Cl (pH 8.0), 1.5 mM calcium acetate. The stock solution should be divided into small aliquots and stored at –20°C. An aliquot can be thawed and refrozen several times, but it should then be discarded. Unlike much cruder preparations of protease (e.g., pronase), proteinase K need not be self-digested before use. The activity of proteinase K is severalfold higher at 50°C than at 37°C. Please see Table A4-8.

TABLE A4-8 Proteolytic Enzymes

	Stock Solution	Storage Temperature	Concentration in Reaction	Reaction Buffer	Temperature	Pretreatment
Pronase[a]	20 mg/ml in H_2O	–20°C	1 mg/ml	0.01 M Tris-Cl (pH 7.8) 0.01 M EDTA 0.5% SDS	37°C	self-digestion[b]
Proteinase K	20 mg/ml in H_2O	–20°C	50 µg/ml	0.01 M Tris-Cl (pH 7.8) 0.005 M EDTA 0.5% SDS	37–56°C	none required

[a]Pronase is a mixture of serine and acid proteases isolated from *Streptomyces griseus*.

[b]Self-digestion eliminates contamination with DNase and RNase. Self-digested pronase is prepared by dissolving powdered pronase in 10 mM Tris-Cl (pH 7.5), 10 mM NaCl to a final concentration of 20 mg/ml and incubating for 1 hour at 37°C. Store the self-digested pronase in small aliquots at –20°C in tightly capped tubes.

OTHER ENZYMES

Lysozymes

Lysozymes are a family of enzymes that catalyze the acid-base hydrolysis of β-(1,4) linkages between *N*-acetylglucosamine and *N*-acetylmuramic acid residues in the proteoglycan of bacterial cell walls (Blake et al. 1967; Fursht 1985). They were discovered by Alexander Fleming (1922), who identified an enzyme activity that rapidly lysed suspensions of bacteria. The products of digestion by vertebrate lysozymes of the cell walls of bacteria were identified in the late 1950s, which allowed the structure and composition of the substrate to be deduced (for review, please see Jollés 1960).

Lysozymes are widely distributed in nature and are expressed wherever there is a need to lyse bacterial cells, for example, during release of bacteriophages from infected cells, on the surface of vertebrate mucosa, and in a great number of secretions of different animals, both vertebrate and invertebrate. There is no structural similarity between vertebrate and bacteriophage-encoded lysozymes such as bacteriophage λ endolysin and bacteriophage T4 endoacetylmuramidase.

In molecular cloning, vertebrate lysozymes (e.g., egg-white lysozyme) are used at pH 8.0 in combination with EDTA and detergents to liberate cosmid and plasmid DNAs from their bacterial hosts (Godson and Vapnek 1973). For further details, please see the information panel on **LYSOZYMES** in Chapter 1.

Agarase

Agarase enzymatically digests molten agarose to produce soluble monomers and oligosaccharides. It is used in the isolation of DNA fragments from agarose gels. The enzyme is available from a number of manufacturers and is active in both Tris-acetate and Tris-borate buffers. Optimum activity is achieved at pH 6.0. Isolation of DNA using agarase is a very gentle method, which makes it suitable for the purification of large fragments. For a protocol on the purification of DNA from pulsed-field gels after agarase digestion, please see Chapter 5, Protocol 19.

Uracil DNA Glycosylase

Uracil DNA glycosylase (UDG) cleaves the *N*-glycosidic bond between the uracil base and the phosphodiester backbone of DNA. UDG removes the uracil residues from the sugar moiety of single- and double-stranded DNAs without destroying the phosphodiester backbone, thus preventing its uses as a hybridization target or template for DNA polymerase. The resulting apyrimidinic DNA becomes susceptible to hydrolysis at high temperatures. UDG will not remove uracil from RNA, and it specifically attacks uracil-containing DNA.

Treatment of uracil-containing DNA with UDG disrupts base pairing in the DNA duplex. The use of dUTP in amplification reactions introduces uracil at the sites opposite to adenine. Subsequent treatment of the duplex with UDG results in single-stranded termini that may be used in cloning the amplified fragments of DNA.

5′- CUACUCAUCAA_____ATGATGAGATGAT-3′
3′- GATGAGTAGTT_____UACUACUCUACUA-5′

+UDG

5′-C-AC-CA-CAA_____ATGATGAGATGAT-3′
3′-GATGAGTAGTT_____-AC-AC-C-AC-A-5′

Topoisomerase I

Topoisomerase I (Topo I) is a crucial component of DNA replication. The enzyme works by introducing transient nicks into one strand of the backbone of both positively and negatively supercoiled DNAs, thus allowing the structure to untwist, and then resealing the breaks (for review, please see Wang 1996). During the course of the reaction catalyzed by topoisomerase I, a transient covalent intermediate is formed between a specific tyrosine residue of the enzyme and one end of the break in its DNA substrate. This bond between the enzyme and its substrate is broken by the activity of tyrosine-DNA phosphodiesterase, the gene for which has been cloned and shown to be highly conserved among higher eukaryotes (Pouliot et al. 1999).

In the absence of Mg^{2+}, topoisomerase I works to generate relaxed, covalently closed circles (for review, see Kornberg and Baker 1992). The enzyme is sometimes used in molecular cloning to enhance the electrophoretic separation of plamid DNAs. Thus, fractionation by gel electrophoresis of closed circular DNAs that have been relaxed by treatment with topoisomerase I will resolve molecules that differ in length by a single nucleotide pair (Wang 1979; Luckow et al. 1987). Topoisomerase I is commercially available from Life Technologies, or it can be purified readily from calf thymus as described by Prell and Vosberg (1980).

Appendix 5

Inhibitors of Enzymes

TABLE A5-1 Protease Inhibitors

PROTEASE INHIBITORS	M.W.	INHIBITS	DOES NOT INHIBIT
Antipain dihydrochloride	677.6	Inhibits papain, trypsin, cathepsins A and B, and to a small extent plasmin.	
Antithrombin III	58,000	Inhibits all serine proteases of the blood coagulation system, as well as trypsin and chymotrypsin.	Does not inhibit cysteine proteases, aspartic proteases, or metalloproteases.
APMSF	252.7	Specific and irreversible inhibitor of serine proteases.	Does not inhibit chymotrypsin or acetylcholinesterase.
Aprotinin	~6500	Serine protease inhibitor; inhibits plasmin, chymotrypsin, kallikrein, and trypsin.	Does not inhibit thrombin or Factor X.
Bestatin	344.8	Inhibits metalloproteases, primarily aminopeptidases.	Does not inhibit carboxypeptidases.
Calpain inhibitor I	383.5	Strong competitive inhibitor of calpain I and to a lesser extent calpain II; also inhibits papain, cathepsins B and L, and to a small extent cathepsin H and α-chymotrypsin.	Does not inhibit trypsin.
Calpain inhibitor II	401.6	Strong competitive inhibitor of calpain I and to a lesser extent calpain II; also inhibits papain, cathepsins B and L, and to a small extent cathepsin H; weakly inhibits α-chymotrypsin.	Does not inhibit trypsin.
Chymostatin	607.7	Inhibits serine and cysteine proteases; specific inhibitor of α-, β-, γ-, δ-chymotrypsin, papain, and cathepsins A, B, and C.	
3,4-Dichloroisocoumarin	215.0	Inhibitor of serine proteases.	
Elastatinal	512.6	Irreversible inhibitor of elastase.	
Hirudin	6963.5	Inhibits thrombin.	
Iodoacetic acid	186.0	Inhibits cysteine proteases.	
Leupeptin	475.6	Inhibits serine and cysteine proteases	
α_2-Macroglobulin	~725,000	Universal protease inhibitor blocking all classes of endoproteinases.	Does not inhibit endoproteinases that are highly specific for one or a limited number of sequences, e.g., tissue kallikrein, urokinase, coagulation Factor XIIA, and endoproteinase Lys-C.
Pefabloc SC, AEBSF, 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride	239.5	Inhibits serine proteases and prevents nonspecific covalent modification of proteins.	
Pepstatin A	685.9	Inhibits acid proteases, e.g., pepsin, renin, cathepsin D, chymosin, and many microbial acid proteases.	
PMSF (phenylmethylsulfonyl fluoride)	174.2	Inhibits serine proteases, e.g., chymotrypsin, thrombin, and papin.	
TIMP-2		Inhibits matrix metalloproteinase activity in enzymatic assays and in vitro malignant invasion assays.	
TLCK; 1-chloro-3-tosylamido-7-amino-1-2-heptanone; Nα-*p*-tosyl-l-lysine chloromethyl ketone hydrochloride	369.3	Irreversibly inhibits trypsin as well as many other serine and cysteine proteases, e.g., bromelain, ficin, or papain.	Does not inhibit chymotrypsin.
TPCK; 1-chloro-3-(4-tosylamido)-4-phenyl-2-butanone; N-tosyl-l-phenylalanine chloromethyl ketone	351.9	Irreversibly inhibits chymotrypsin, as well as many other serine and cysteine proteases, e.g., bromelain, ficin, or papain.	Does not inhibit trypsin.

Appendix 6

Nucleic Acids

VITAL STATISTICS OF DNA

TABLE A6-1 Genome Comparisons

Organism	Size of DNA (bp)	Weight of DNA (Daltons)	Reference
Mammals	~3.0×10^9	~1.9×10^{12}	
Drosophila melanogaster	~1.8×10^8	~7.8×10^{10}	Adams et al. (2000)
Caenorhabditis elegans	~9.70×10^7	6.59×10^6	*C. elegans* Sequencing Consortium (1998)
Saccharomyces cerevisiae	1.30×10^7	8.44×10^9	Goffeau et al. (1996)
Haemophilus influenzae Rd	1.83×10^6	1.19×10^9	Fleischmann et al. (1995)
Mycoplasma genitalium	0.58×10^6	3.76×10^8	Fraser et al. (1995)
Methanococcus jannaschii	1.66×10^6	1.08×10^9	Bult et al. (1996)
Synechocystis sp.	3.57×10^6	2.32×10^9	Kaneko et al. (1996)
Mycoplasma pneumoniae	8.10×10^5	0.53×10^8	Himmelreich et al. (1996)
Helicobacter pylori	1.66×10^6	1.08×10^9	Tomb et al. (1997)
Escherichia coli	4.60×10^6	3.00×10^9	Blattner et al. (1997)
Methanobacterium thermoautotrophicum	1.75×10^6	1.14×10^9	Smith et al. (1997)
Bacillus subtilis	4.20×10^6	2.73×10^9	Kunst et al. (1997)
Archaeoglobus fulgidus	2.18×10^6	1.40×10^9	Klenk et al. (1997)
Borrelia burgdorferi	1.44×10^6	9.35×10^8	Fraser et al. (1997)
Aquifex aeolicus	~1.50×10^6	9.74×10^9	Deckert et al. (1998)
Pyrococcus horikoshii	1.80×10^6	1.17×10^9	Kawarabayasi et al. (1998)
Mycobacterium tuberculosis	4.40×10^6	2.90×10^9	Cole et al. (1998)
Treponema pallidum	1.14×10^6	7.40×10^8	Fraser et al. (1998)
Chlamydia trachomatis	1.05×10^6	6.80×10^8	Stephens et al. (1998)
Rickettsia prowazekii	1.10×10^6	7.10×10^8	Andersson et al. (1998)
Helicobacter pylori	1.64×10^6	1.06×10^9	Alm et al. (1999)
Chlamydia pneumoniae	1.23×10^6	7.98×10^8	Kalman et al. (1999)
Deinococcus radiodurans	3.28×10^6	2.13×10^9	White et al. (1999)
Thermotoga maritima	1.80×10^6	1.17×10^9	Nelson et al. (1999)
Bacteriophage T2	~2.0×10^5	~1.3×10^8	
Bacteriophage λ	48,514	3.1×10^7	Daniels et al. (1983)
pBR322	4,363	2.8×10^6	Sutcliffe (1978, 1979)
pUC18/pUC19	2,686	1.7×10^6	Yanisch-Perron et al. (1985)

Source: www.tigr.org For updates, see TIGR Web Site (www.tigr.org)

TABLE A6-2 Frequency of Restriction Endonuclease Sites in the Human Genome

Enzyme	Sequence	Average Fragment Size (kb)	Estimated Number of Sites
*Apa*I	GGGCC	2	1.5×10^6
*Asc*I	GGCGCGCC	80	3.75×10^4
*Avr*II	CCTAGG	8	3.75×10^5
*Bam*HI	GGATCC	5	6×10^5
*Bgl*I	$GCCN_5GGC$	3	1×10^6
*Bgl*II	AGATCT	3	1×10^6
*Bss*HII	GCGCGC	10	3×10^5
*Dra*I	TTTAAA	2	1.5×10^6
*Eag*I	CGGCCG	10	3×10^5
*Eco*RI	GAATTC	5	6×10^5
*Hin*dIII	AAGCTT	4	7.5×10^5
*Nae*I	GCCGGC	4	7.5×10^5
*Nar*I	GGCGCC	4	7.5×10^5
*Nhe*I	GCTAGC	10	3×10^5
*Not*I	GCGGCCGC	100	3×10^4
*Pac*I	TTAATTAA	60	5×10^4
*Pme*I	GTTTAAAC	70	4.3×10^4
*Rsr*I	CGGWCCG	60	5×10^4
*Sac*I	GAGCTC	3	1×10^6
*Sac*II	CCGCGG	6	5×10^5
*Sal*I	GTCGAC	20	1.5×10^5
*Sbf*I	CCTGCAGG	15	5.33×10^5
*Sfi*I	$GGCCN_5GGCC$	30	1×10^5
*Sgr*AI	CRCCGGYC	70	4.3×10^4
*Sma*I	CCCGGG	4	7.5×10^5
*Spe*I	ACTAGT	10	3×10^5
*Sph*I	GCATGC	6	5×10^5
*Srf*I	GCCCGGGC	50	6×10^4
*Ssp*I	AATATT	2	1.5×10^6
*Swa*I	ATTTAAAT	30	1×10^5
*Xba*I	TCTAGA	5	6×10^5
*Xho*I	CTCGAG	7	4.3×10^5

(Adapted, with permission, from 1998/99 New England Biolabs Catalog [©NEB].)

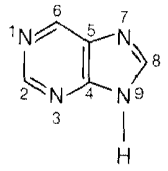
TABLE A6-3 Restriction Endonuclease Cleavage at the End of a DNA Fragment

Enzyme	Base Pairs from End	% Cleavage Efficiency	Enzyme	Base Pairs from End	% Cleavage Efficiency
*Aat*II	3	88	*Kpn*I	2	100
	2	100		2	100
	1	95		1	99
*Acc*65	2	99	*Mlu*I	2	99
	1	75			
*Afl*II	1	13	*Mun*I	2	100
*Age*I	1	100	*Nco*I	2	100
	1	100			
*Apa*I	2	100	*Ngo*M IV	2	100
*Asc*I	1	97	*Nhe*I	1	100
				2	82
*Avr*II	1	100	*Not*I	7	100
				4	100
				1	98
*Bam*HI	1	97	*Nsi*I	3	100
				3	77
				2	95
*Bgl*II	3	100	*Pac*I	1	76
*Bsi*WI	2	100	*Pme*I	1	94
*Bsp*EI	2	100	*Pst*I	3	98
	1	8		2	50
				1	37
*Bsr*GI	2	99	*Sac*I	1	99
	1	88			
*Bss*HII	2	100	*Sal*I	3	89
				2	23
				1	61
*Eag*I	2	100	*Spe*I	2	100
				2	100
*Eco*RI	1	100	*Sph*I	2	99
	1	88		2	97
	1	100		1	92
*Eco*RV	1	100	*Xba*I	1	99
				1	94
*Hin*dIII	3	90	*Xho*I	1	97
	2	91			
	1	0			
*Kas*I	2	97	*Xma*I	2	98
	1	93		2	92

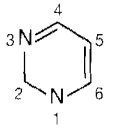
The results represent the ability of various restriction endonucleases to cleave close to the end of a DNA fragment. The cleavage efficiency is given for a particular recognition site placed at x bases from the end of a linearized vector. As a general rule, enzymes not listed above require 6 base pairs on either side of their recognition site to cleave efficiently. (Reproduced, with permission, 1998/99 New England Biolabs Catalog [©NEB].)

TABLE A6-4 Concentration of Double-stranded DNA in Solution

Double-stranded DNA (50 μg/ml)	bp/Molecule	Molecular Mass of DNA (Daltons)	50 μg/ml Solution				
			Molecules DNA/ml	Moles/ml	Molar Concentration of DNA	Phosphate	Termini
Bacteriophage λ	48,514	3.20×10^7	9.41×10^{11}	1.56×10^{-12}	1.56 nM	157 μM	3.12 nM
pAd10SacBII	30,300	2.00×10^7	1.51×10^{12}	2.50×10^{-12}	2.50 nM	157 μM	5.00 nM
pCYPAC1	19,600	1.29×10^7	2.33×10^{12}	3.88×10^{-12}	3.88 nM	157 μM	7.76 nM
pYAC4	11,400	7.52×10^6	4.00×10^{12}	6.65×10^{-12}	6.65 nM	157 μM	13.1 nM
pBeloBACII	7,400	4.88×10^6	6.17×10^{12}	1.03×10^{-11}	10.3 nM	157 μM	20.6 nM
pBR322	4,363	2.88×10^6	1.05×10^{13}	1.74×10^{-11}	17.4 nM	157 μM	34.8 nM
pUC18/pUC19	2,686	1.77×10^6	1.70×10^{13}	2.83×10^{-11}	28.3 nM	157 μM	56.6 nM
Segment of DNA (1 kb)	1,000	6.60×10^5	4.56×10^{13}	7.58×10^{-11}	75.8 nM	157 μM	152 nM
Octameric double-stranded linker	8	5.28×10^3	5.70×10^{15}	9.47×10^{-9}	9.47 μM	157 μM	18.9 nM



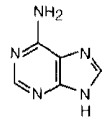
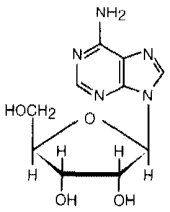
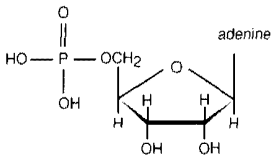
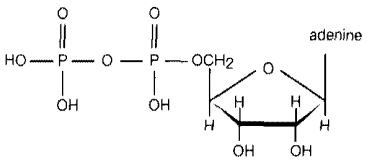
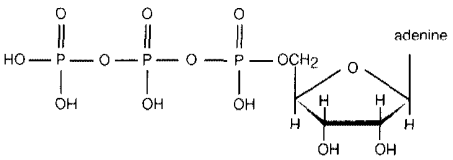
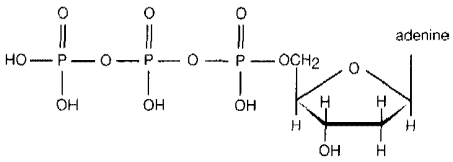
the purine ring system



the pyrimidine ring system

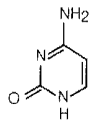
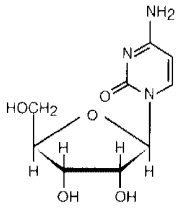
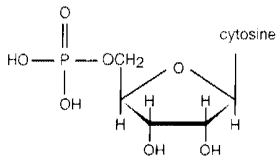
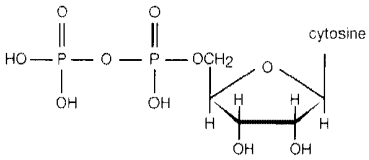
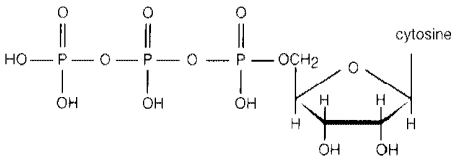
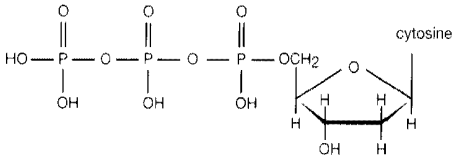
FIGURE A6-1 Numbering of Atoms on Purine and Pyrimidine Molecules

TABLE A6-5 Adenine and Related Compounds

	Structure[a]	M.W.	λ_{max} (nm)	E_{max} ($\times 10^{-3}$)	OD_{280}/OD_{260}
Adenine		135.1	260.5	13.4	0.13
Adenosine		267.2	260	14.9	0.14
Adenosine 5′-phosphate (5′-AMP)		347.2	259	15.4	0.16
Adenosine 5′-diphosphate (5′-ADP)		427.2	259	15.4	0.16
Adenosine 5′-triphosphate (5′-ATP)		507.2	259	15.4	0.15
2′-Deoxyadenosine 5′-triphosphate (5′-dATP)		491.2	259	15.4	0.15

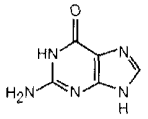
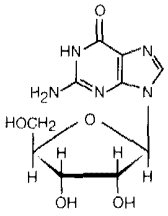
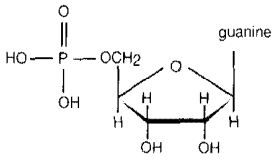
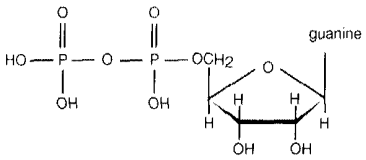
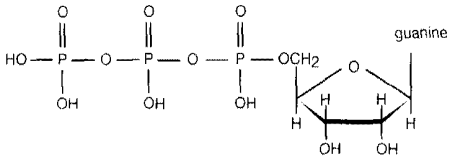
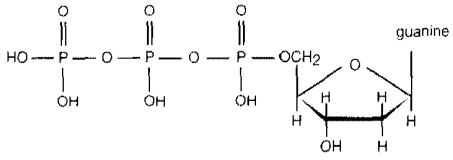
[a] At physiological pH, all of the hydroxyls bound to phosphate are ionized.

TABLE A6-6 Cytosine and Related Compounds

	Structure[a]	M.W.	λ_{max} (nm)	E_{max} ($\times 10^{-3}$)	OD_{280}/OD_{260}
Cytosine		111.1	267	6.1	0.58
Cytidine		243.2	271	8.3	0.93
Cytidine 5′-phosphate (5′-CMP)		323.2	271	9.1	0.98
Cytidine 5′-diphosphate (5′-CDP)		403.2	271	9.1	0.98
Cytidine 5′-triphosphate (5′-CTP)		483.2	271	9.0	0.97
2′-Deoxycytidine 5′-triphosphate (5′-dCTP)		467.2	272	9.1	0.98

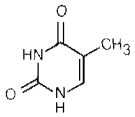
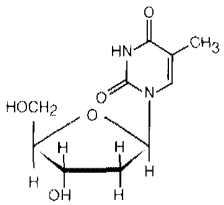
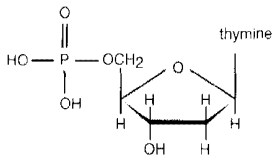
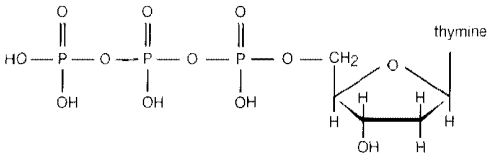
[a]At physiological pH, all of the hydroxyls bound to phosphate are ionized.

TABLE A6-7 Guanine and Related Compounds

	Structure[a]	M.W.	λ_{max} (nm)	E_{max} ($\times 10^{-3}$)	OD_{280}/OD_{260}
Guanine		151.1	276	8.15	1.04
Guanosine		283.2	253	13.6	0.67
Guanosine 5′-phosphate (5′-GMP)		363.2	252	13.7	0.66
Guanosine 5′-diphosphate (5′-GDP)		443.2	253	13.7	0.66
Guanosine 5′-triphosphate (5′-GTP)		523.2	253	13.7	0.66
2′-Deoxyguanosine 5′-triphosphate (5′-dGTP)		507.2	253	13.7	0.66

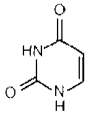
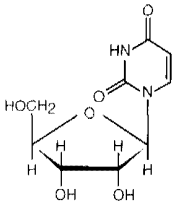
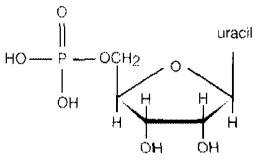
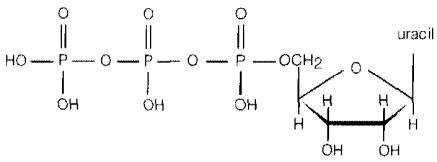
[a] At physiological pH, all of the hydroxyls bound to phosphate are ionized.

TABLE A6-8 Thymine and Related Compounds

	Structure[a]	M.W.	λ_{max} (nm)	E_{max} ($\times 10^{-3}$)	OD_{280}/OD_{260}
Thymine		126.1	264.5	7.9	0.53
2′-deoxythymidine		242.2	267	9.7	0.70
2′-deoxythymidine 5′-phosphate (5′-TMP)		322.2	267	9.6	0.73
2′-deoxythymidine 5′-triphosphate (5′-TTP)		482.2	267	9.6	0.71

[a]At physiological pH, all of the hydroxyls bound to phosphate are ionized.

TABLE A6-9 Uracil and Related Compounds

	Structure[a]	M.W.	λ_{max} (nm)	E_{max} ($\times 10^{-3}$)	OD_{280}/OD_{260}
Uracil		112.1	259	8.2	0.17
Uridine		244.2	262	10.1	0.35
Uridine 5′-phosphate (5′-UMP)		324.2	260	10.0	0.38
Uridine 5′-triphosphate (5′-UTP)		484.2	260	10.0	0.38

[a]At physiological pH, all of the hydroxyls bound to phosphate are ionized.

TABLE A6-10 Unusual Bases

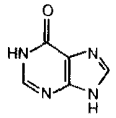
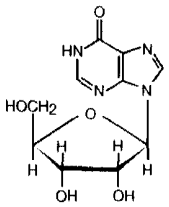
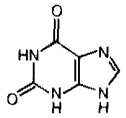
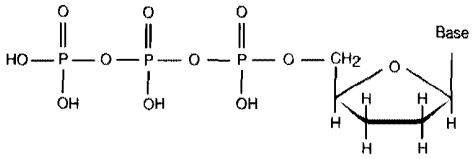
	STRUCTURE	M.W.	λ_{max} (nm)	E_{max} ($\times 10^{-3}$)	OD_{280}/OD_{260}
Hypoxanthine		136.1	249.5	10.7	0.09
Inosine		268.2	248.5	12.3	0.25
Xanthine		152.1	267	10.3	0.61

TABLE A6-11 Nucleoside Analogs Used as Chain Terminators in DNA Sequencing

2′,3′-DIDEOXYRIBONUCLEOSIDE 5′TRIPHOSPHATES	STRUCTURE[a]	M.W.
		
2′,3′-Dideoxyadenosine (ddATP)	Base = adenine	475.2
2′,3′-Dideoxycytidine (ddCTP)	Base = cytosine	451.2
2′,3′-Dideoxyguanosine (ddGTP)	Base = guanine	491.2
2′,3′-Dideoxythymidine (ddTTP)	Base = thymine	466.2 $Na_4 \cdot H_2O$, 608.2

[a]At physiological pH, all of the hydroxyls bound to phosphate are ionized.

OLIGONUCLEOTIDES

TABLE A6-12 Commonly Used Oligonucleotide Primers

PRIMER	SEQUENCE
λgt10 forward primer	5′-AGCAAGTTCAGCCTGGTTAAG-3′
λgt10 reverse primer	5′-CTTATGAGTATTTCTTCCAGGGTA-3′
λgt11 forward primer	5′-GGTGGCGACGACTCCTGGAGCCCG-3′
λgt11 reverse primer	5′-TTGACACCAGACCAACTGGTAATG-3′
pUC/M13-40 forward primer	5′-GTTTTCCCAGTCACGACG-3′
pUC/M13-48 reverse primer	5′-AGCGGATAACAATTTCACACAGG-3′
pUC/M13-20 forward primer	5′-GTAAAACGACGGCCAGT-3′
pUC/M13-20 reverse primer	5′-GGAAACAGCTATGACCATG-3′
SP6 universal primer	5′-ATTTAGGTGACACTATAG-3′
T7 universal primer	5′-TAATACGACTCACTATAGGG-3′
T3 promoter primer	5′-ATTAACCCTCACTAAAGGGA-3′

For advice on custom primer design, please see the introduction to Chapter 10.

TABLE A6-13 Molecular Conversions for Oligonucleotides

SIZE OF OLIGONUCLEOTIDE (NUCLEOTIDES)	MOLECULAR MASS (DALTONS)	MOLECULES OF DNA IN 1 μg	MOLES OF DNA IN 1 μg
8	2.64×10^3	2.28×10^{14}	379 pmoles
10	3.30×10^3	1.82×10^{14}	303 pmoles
12	3.96×10^3	1.52×10^{14}	253 pmoles
14	4.62×10^3	1.30×10^{14}	216 pmoles
16	5.28×10^3	1.14×10^{14}	190 pmoles
18	5.94×10^3	1.01×10^{14}	168 pmoles
20	6.60×10^3	9.12×10^{13}	152 pmoles

NOMOGRAMS

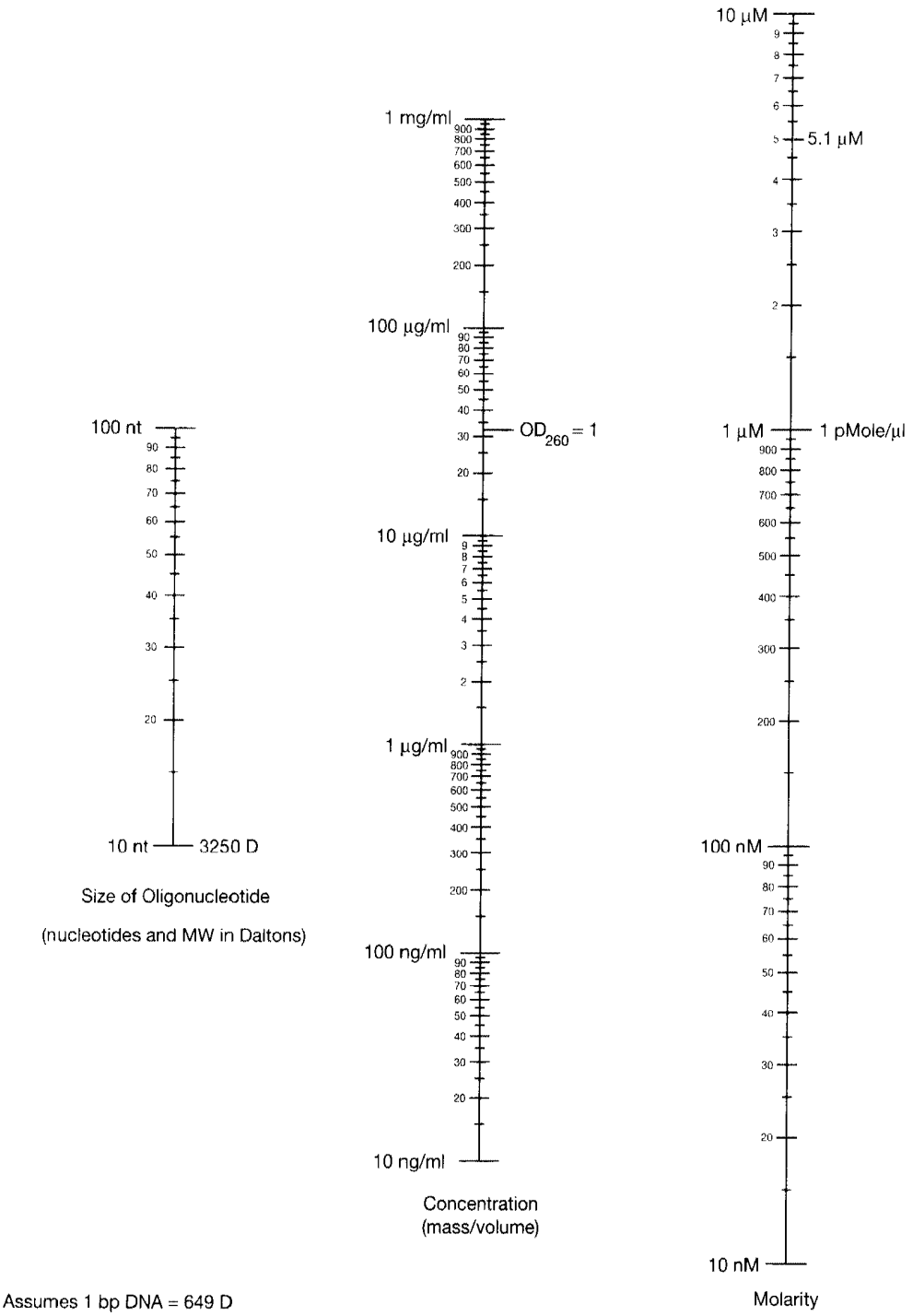


FIGURE A6-2 Nomogram for Single-stranded DNA

This nomogram can be used for the conversion of concentration between different conventions and to obtain approximate values of DNA concentration from OD_{260} readings. For example, a solution of an oligonucleotide 20 nucleotides in length that produces an $OD_{260} = 1$ has a concentration of 33 µg/ml. To calculate the molarity of the solution, draw a line from the size of the molecule (in nucleotides or molecular weight) on the left-hand scale, through the point of known concentration (33 µg/ml) on the middle scale. Extrapolate the line through the third scale and read off the molarity (5.1 µM is equivalent to 5.1 pmoles/µl). (Figure kindly provided by Siân Curtis.)

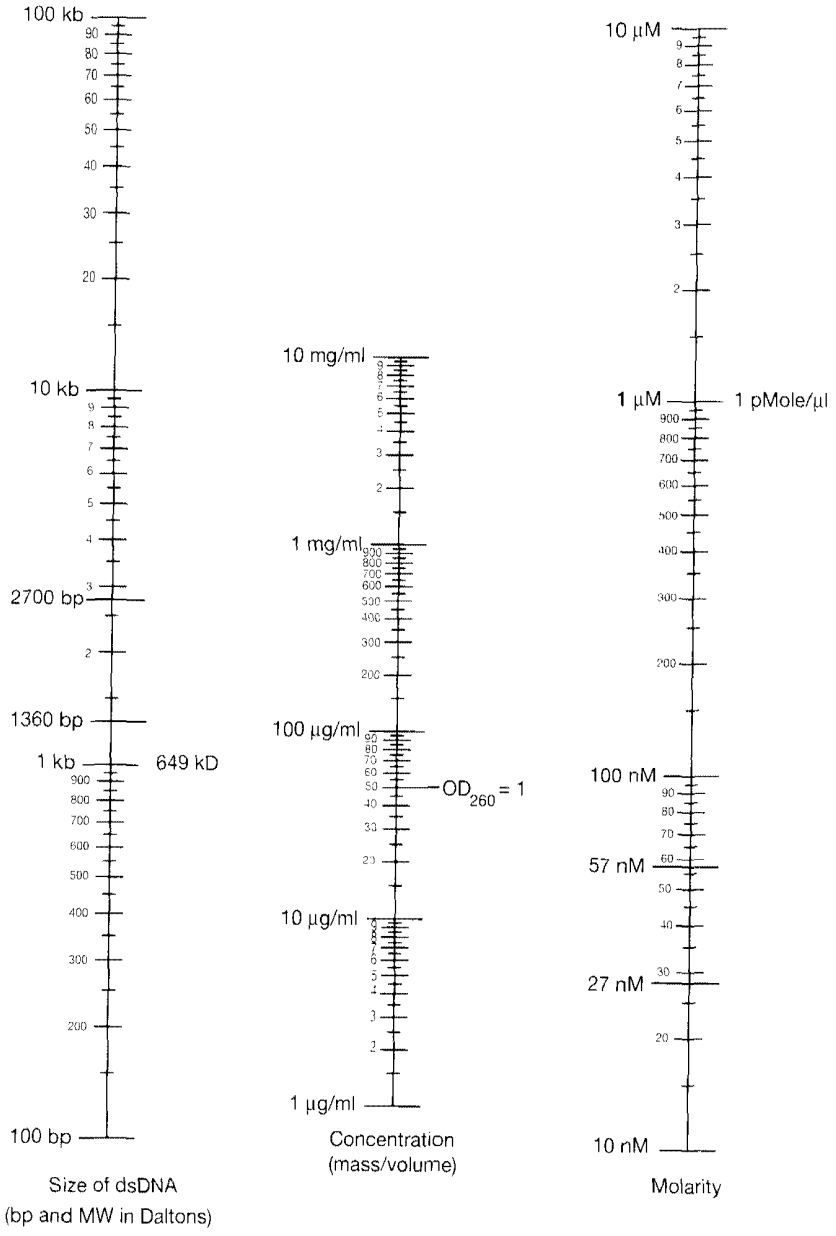


FIGURE A6-3 Nomogram for Double-stranded DNA

This nomogram can be used for the conversion of concentration between different conventions, and also to obtain approximate values of DNA concentration from OD_{260} readings. A solution containing 50 µg/ml of *double-stranded* DNA has an absorbance of 1 at 260 nm, i.e., $A_{260} = 1 = 50$ µg/ml of double-stranded DNA. For example, a solution of pUC 18/19 (2686 bp) that produces an $OD_{260} = 1$ has a concentration of 50 µg/ml. To calculate the molarity of the solution, draw a line from the size of the molecule (base pairs or molecular weight) on the left-hand scale through the point of known concentration (50 µg/ml) on the middle scale. Extrapolate the line through the third scale and read off the molarity (57 nM is equivalent to 57 fmoles/µl). (Figure kindly provided by Siân Curtis.)

SIZE MARKERS

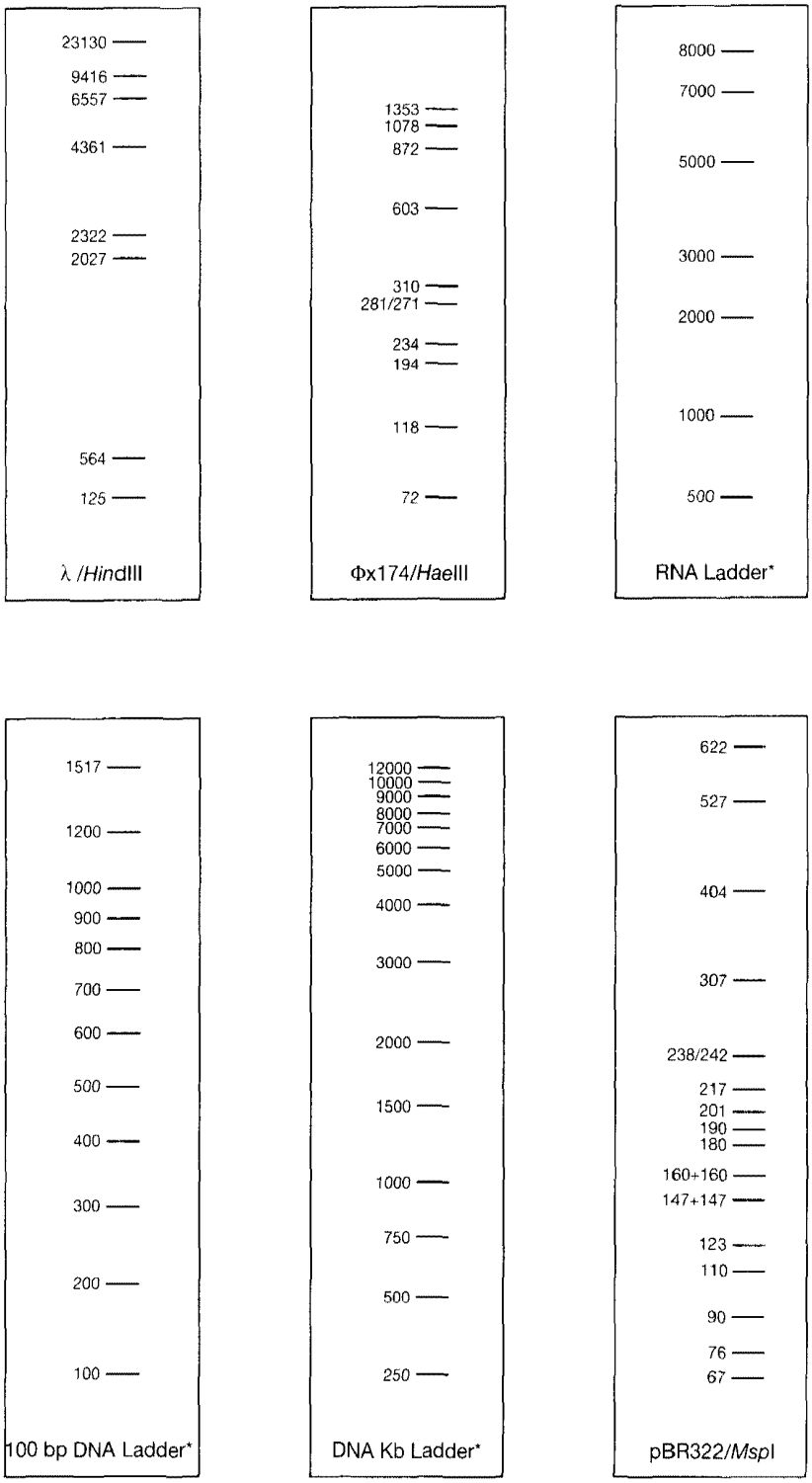


FIGURE A6-4 DNA Size Markers

(100-bp DNA ladder, RNA ladder, and pBR322/*MspI* are reproduced, with permission, from 1998/99 New England Biolabs Catalog [©NEN]). (DNA Size Markers are reproduced with the express permission of Stratagene. Copyright 1999, Stratagene. All rights reserved.)

Appendix 7

Codons and Amino Acids

CODONS AND CODON USAGE

The genetic code is redundant, using 61 codons to specify 20 amino acids. Only two amino acids (Met and Trp) are specified by a single codon, whereas the remaining 18 amino acids are each specified by multiple codons (please see Figure A7-1).

The synonymous codons that specify a single amino acid are not used with equal frequency (Grantham et al. 1980, 1981). Instead, biased usage of synonymous codons is the rule in all species and is the norm in most genes. In addition, there is considerable variation in codon usage between genes in a single species (Bennetzen and Hall 1982; Gouy and Gautier 1982). This results from variation between genes in the G+C content of the third position of synonymous codons (Ikemura 1985; Mirouchoud and Gautier 1988). The pattern of codon usage is the result of selective forces, mutational bias, and genetic drift. For example:

- ***Selection during translation so that the codons used most frequently match the most abundant tRNAs.*** Highly expressed and weakly expressed genes display different patterns of codon usage in both prokaryotes and eukaryotes (Shields and Sharp 1987; Sharp and Devine 1989; Sharp and Cowe 1991). Direct assays of translatability of mRNAs lend further support to this view (e.g., please see Ikemura 1982; Sørensen et al. 1989).
- ***"Knock-on" or context effects, in which mutation of one base influences other bases in the neighborhood*** (Bulmer 1990; Eyre-Walker 1991).
- ***Any selection that results in a change in GC content in codons.*** Bernardi and Bernardi (1986) have argued that regions of the genome with increased (G+C) content are thermodynamically more stable and that mRNAs rich in (G+C) have more secondary structures and are therefore more stable. If correct, this relationship would have effects both on amino acid composition of proteins and on codon usage. There would be selection for "strong" codons that contain G and C at the first two positions (Gly, Ala, Pro) and a selection against weak codons that have A and T at both positions (Phe, Ile, Lys, Asn, Tyr) (Karlin and Bucher 1992). In addition, there would be a slow drift to replace A and T in the third position with G or C. Over the course of time, this gentle evolutionary wind might blow codons, like fallen leaves, into local clumps within individual genes or organisms. However, there are also stabilizing forces that protect codons rich in (A+T). Chief among these is the strong selection against the sequence CG in mammalian genomes. The cytosine of this sequence tends to be methylated and is then prone to mutation. One result is a bias against C in the third position of codons that precede codons beginning with G (e.g., please see Zhang et al. 1991).

Codon usage is a concern in molecular cloning if a sequence of amino acids is used to design an oligonucleotide for screening cDNA or genomic libraries. No rules guarantee selection of the correct codon at a position of ambiguity. However, a probe of 30 or more oligonucleotides would be expected to have at least 76% homology with its target sequence even if all codon choices were made on a random basis (Lathe 1985). If substitutions are chosen on the statistical basis of known codon utilization in the species of interest, the expected homology increases to 82%; it rises still further (to 86%) if regions lacking Leu, Arg, and Ser are chosen (each of these amino acids is specified by six codons). In yeast, still higher accuracy can be achieved by taking into account whether the gene is highly or weakly expressed. In mammalian cells, this latter refinement is not generally used because of the great variation in the level of expression of genes between tissues.

A database of codon usage in different organisms is available at http://www.kazusa.or.jp/codon/. The list is based on data derived from an analysis of complete coding sequences in GenBank (Nakamura et al. 1999). For a more detailed analysis of the fluctuation of codon usage in different yeast genes, please see Sharp and Cowe (1991). Codon usage in humans is shown in Table A7-1.

TABLE A7-1 Codon Usage in Humans

Amino Acid	Frequency in Human Proteins (%)[a]	Codons and Their Usage in Human Proteins (%)[b]	
Alanine	6.99	GCU (28.0)	GCC (41.6)
		GCA (20.0)	GCG (10.3)
Arginine	5.28	CGU (8.9)	CGC (21.4)
		CGA (5.4)	CGG (10.4)
		AGA (9.9)	AGG (11.1)
Asparagine	3.92	AAU (42.3)	AAC (57.7)
Aspartic Acid	5.07	GAU (42.8)	GAC (57.2)
Cysteine	2.44	UGU (40.6)	UGC (59.4)
Glutamic Acid	6.82	GAA (39.2)	GAG (60.7)
Glutamine	4.47	CAA (24.8)	CAG (75.2)
Glycine	7.10	GGU (15.8)	GGC (35.8)
		GGA (24.1)	GGG (24.3)
Histidine	2.35	CAU (39.6)	CAC (60.4)
Isoleucine	4.50	AUU (33.1)	AUC (54.0)
		AUA (12.9)	
Leucine	9.56	UUA (5.5)	UUG (11.5)
		CUU (11.1)	CUC (20.8)
		CUA (6.5)	CUG (44.5)
Lysine	5.71	AAA (38.9)	AAG (61.1)
Methionine	2.23	AUG (100)	
Phenylalanine	3.84	UUU (41.1)	UUC (58.2)
Proline	5.67	CCU (27.3)	CCC (35.2)
		CCA (25.7)	CCG (11.6)
Serine	7.25	UCU (18.3)	UCC (23.7)
		UCA (12.9)	UCG (5.9)
		AGU (13.2)	AGC (25.9)
Threonine	5.68	ACU (22.4)	ACC (40.5)
		ACA (25.4)	ACG (11.8)
Tryptophan	1.38	UGG (100)	
Tyrosine	3.13	UAU (40.0)	UAC (60.0)
Valine	6.35	GUU (16.4)	GUC (25.7)
		GUA (9.3)	GUG (48.7)

[a] Calculated from an analysis of 1490 human genes (601,683 codons) by Wada et al. (1992). The codon usage for other mammals does not differ greatly from those presented here for human genes.

[b] The figures in parentheses show the frequency with which a particular codon is used to specify a certain amino acid. For example, of the alanine residues present in 1490 human proteins, 28% are specified by GCU, 41.6% by GCC, etc.

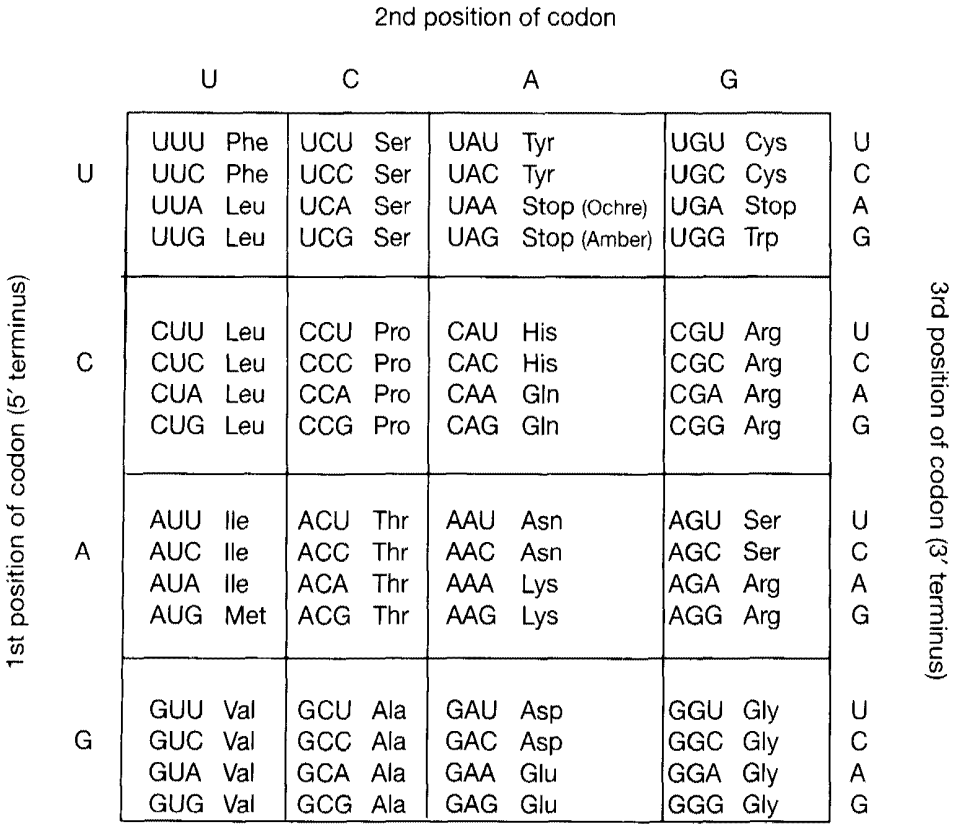


FIGURE A7-1 The Genetic Code

AMBER MUTANTS AND AMBER SUPPRESSORS

In amber mutants, a codon that specifies an amino acid in a protein is replaced by the chain-terminating codon, UAG. Because there is no tRNA that recognizes UAG, translation of mRNA ceases at the position of the chain-terminating codon, thereby generating an incomplete fragment of the protein.

Some strains of *E. coli* can suppress the chain-terminating phenotype of amber codons, irrespective of the gene in which the amber mutation is located. Such strains carry a mutation that changes the sequence of the anticodon loop in a particular species of tRNA. The suppressor tRNA recognizes the amber codon and inserts its cognate amino acid at the chain-terminating codon, allowing protein synthesis to continue. The efficiency of this process is not absolute. In a strain carrying a strong amber suppressor, suppression of polypeptide chain termination might occur 50% of the time. In weakly suppressing strains, the efficiency may be 10% or less. All of the suppressor strains used in molecular cloning are strong suppressors. Different suppressor tRNAs insert different amino acids at the chain-terminating UAG codon. A few amber suppressors can also suppress ochre (UAA) mutations because of wobble in the third position of the codon.

At one time, it was mandatory for bacteriophage λ vectors to carry amber mutations in genes encoding coat proteins. It was believed that such mutations might reduce the risk of recombinant bacteriophages spreading from the laboratory into field strains of *E. coli*. Hosts for these vectors carry one or two strong amber suppressors — *supE* and *supF* — that insert glutamine and tyrosine, respectively, at UAG codons. These suppressors are not interchangeable. Some amber mutations in bacteriophage λ vectors are suppressed only by *supE* (e.g., *P*am3) and others are suppressed only by *supF* (e.g., *S*am7 and *S*am100). The presence of suppressors does not generally affect the growth of the bacteriophage λ vectors that do not carry amber mutations. Most bacteriophage λ vectors can therefore be assayed and propagated on a strain of *E. coli* such as LE392 that carries *supE* and *supF*.

HISTORICAL FOOTNOTE

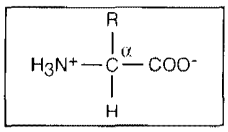
Conditional mutations are a special class of mutations that can occur in a great variety of genes in a single organism. Although temperature-sensitive mutants of *Drosophila*, *Neurospora*, and *E. coli* had been isolated and characterized years before, the full power of temperature-sensitive and amber mutants only became apparent in the early 1960s when Edgar, Epstein, and their colleagues isolated and analyzed a large collection of conditional lethal mutations of bacteriophage T4 (Edgar and Lielausis 1964; Epstein et al. 1964). The first conditional lethal mutants of bacteriophage T4 were isolated quite serendipitously by a CalTech graduate student, Harris Bernstein, during a fruitless search for an entirely different class of phage mutants (Edgar 1966). Bernstein, then a student of *Neurospora* genetics, had wandered into the Epstein-Steinberg laboratory one evening hoping to persuade someone to go with him to the movies. Instead he found himself picking bacteriophage T4 plaques as part of a hunt for "anti-rII" mutants that could grow on *E. coli* K(λ) but not on *E. coli* B. During the course of the evening, there was much debate about whether mutants of this class could exist, with Bernstein arguing strongly that they should. Epstein and Steinberg were more skeptical but, as encouragement to Bernstein, promised that any mutants he isolated would be named after his mother. The next day, when the results of the hunt were analyzed, ~20 mutants with the expected phenotype were found. The promise was kept by translating the German word "Bernstein" into its English equivalent "amber."

Subsequent work showed that the "anti-rII" mutants were not what they first appeared to be, since they grew in many strains of *E. coli* K not lysogenic for λ. Genetic mapping showed that amber mutants were widely distributed over the bacteriophage T4 genome, and physiological tests showed that their replication was blocked at many different stages in nonpermissive hosts. The hunt for parochial "anti-rII" mutants had therefore uncovered a general class of conditionally lethal, suppressor-sensitive mutants located in bacteriophage genes that were previously completely inaccessible. Strangely enough, none of the first group of 20 mutants were picked by Bernstein. It turned out that he had been flaming his bacterial wire too enthusiastically and had killed all of the bacteriophages!

TABLE A7-2 Amber Suppressors Used in Molecular Cloning

SUPPRESSOR	CODON RECOGNIZED	AMINO ACID INSERTED	TRNA GENE FROM WHICH SUPPRESSOR IS DERIVED
supD (*sul*)	amber (UAG)	serine	*serU*
supE (*sulI*)	amber (UAG)	glutamine	*glnU*
supF (*sulII*)	amber (UAG)	tyrosine	*tyrT*
supB (*suB*)	amber (UAG) and ochre (UAA)	glutamine	*glnU*
supC (*suC*)	amber (UAG) and ochre (UAA)	tyrosine	*tyrT*

AMINO ACIDS



Of the 20 standard α-amino acids (Table A7-3) that are incorporated into proteins, 19 have the general structure shown in the figure at the right. R represents the side chain of the amino acid. The twentieth acid, proline, is really an imino acid in which the side chain is bonded to the nitrogen atom of the peptide group. Except in glycine, where the side chain is a hydrogen atom, the α-carbon is asymmetric and is always the L-isomer. The α-amino acids can be arranged into several groups according to the chemical properties of their side chains (Table A7-4). The amino acids can also be grouped according to various other criteria, including size and hydrophilicity (Chothia 1976; Kyte and Doolittle 1982; Taylor 1986). The relationships among these groupings of amino acids can be represented in a Venn diagram (Figure A7-2), which is based on the mutational matrix of Dayhoff (1972). The Venn diagram shows the relationships among 20 common amino acids. Cysteine is shown in two locations: The reduced form, cystine (CH), contains a polarizable S-H bond and is therefore similar in some ways to serine (which carries an O-H bond). The oxidized form, cysteine (CS-S), contains no polarizable bond and is therefore more hydrophobic in nature.

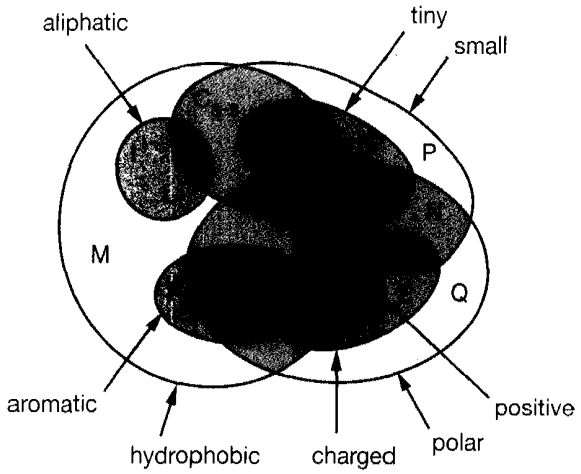


FIGURE A7-2 Venn Diagram

Venn diagram showing the relationships among 20 common amino acids. (Redrawn, with permission, from Taylor 1986.)

TABLE A7-3 Nomenclature of the 20 Standard α-Amino Acids

Amino Acid	Three-letter Symbol	One-letter Symbol	Mnemonic for One-letter Symbol
Alanine	Ala	A	Alanine
Arginine	Arg	R	ARginine
Asparagine	Asn	N	AsparagiNe
Aspartic acid	Asp	D	AsparDic
Cysteine	Cys	C	Cysteine
Glutamic acid	Glu	E	GluEtamic
Glutamine	Gln	Q	Q-tamine
Glycine	Gly	G	Glycine
Histidine	His	H	Histidine
Isoleucine	Ile	I	Isoleucine
Leucine	Leu	L	Leucine
Lysine	Lys	K	before L
Methionine	Met	M	Methionine
Phenylalanine	Phe	F	Fenylalanine
Proline	Pro	P	Proline
Serine	Ser	S	Serine
Threonine	Thr	T	Threonine
Tryptophan	Trp	W	TWyptophan
Tyrosine	Tyr	Y	TYrosine
Valine	Val	V	Valine

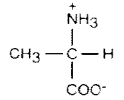
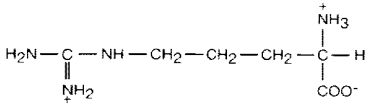
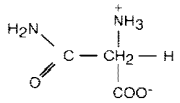
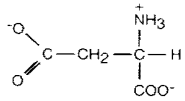
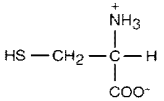
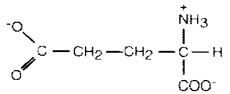
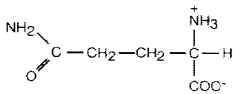
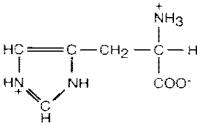
TABLE A7-4 Properties of L α-Amino Acids

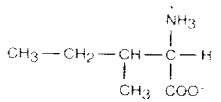
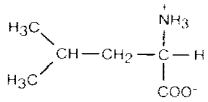
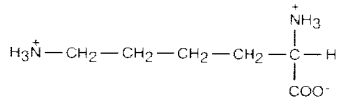
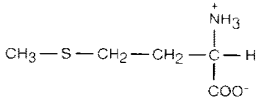
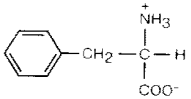
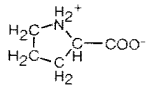
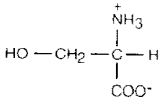
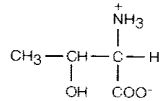
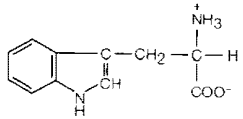
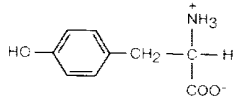
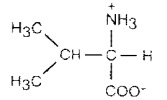
Major Properties of Side Chains	Amino Acids
No side chain	Gly
Aliphatic	Ala, Val, Leu, Ile, Pro
Hydroxyl group	Ser, Thr
Acidic group	Asp, Glu
Amide group	Asn, Gln
Basic group	Lys, Arg
Imidazole group	His
Aromatic group	Phe, Tyr, Trp
Sulfur-containing	Met, Cys

TABLE A7-5 Molar Conversions for Proteins

Molecular Weight of Unmodified Protein	Approximate Number of Residues	Molecules of Protein in 1 ml of a Solution Containing 1 mg/ml	Moles of Protein in 1 ml of a Solution Containing 1 mg/ml	Molar Concentration of Protein Solution Containing 1 mg/ml
100,000	917	6×10^{15}	10^{-8}	10^{-5} M
80,000	734	7.5×10^{15}	1.25×10^{-8}	1.25×10^{-5} M
60,000	550	10^{16}	1.66×10^{-8}	1.66×10^{-5} M
40,000	367	1.5×10^{16}	2.50×10^{-8}	2.5×10^{-5} M
20,000	183	3×10^{16}	50×10^{-8}	5.0×10^{-5} M
10,000	92	6×10^{16}	10^{-7}	10^{-4} M

TABLE A7-6 Properties of Amino Acids

Amino Acid	M.W. of Residue in Protein at pH 7.0	FW	pK_a -COOH	pK_a $-NH_2$	pK_a R-Group	Hydropathy Index	Structure
Alanine	71	89.10	2.35	9.87		1.8	
Arginine	157	174.20	1.82	8.99	12.48	−4.5	
Asparagine	114	132.12	2.14	8.72		−3.5	
Aspartic Acid	115	133.11	1.99	9.90	3.90	−3.5	
Cysteine	103	121.16	1.92	10.70	8.37	2.5	
Glutamic Acid	129	147.13	2.10	9.47	4.07	−3.5	
Glutamine	128	146.15	2.17	9.13		−3.5	
Glycine	57	75.07	2.35	9.78		−0.4	
Histidine	137	155.16	1.80	9.33	6.04	−3.2	

Isoleucine	113	131.18	2.32	9.76		4.5	
Leucine	113	131.18	2.33	9.74		3.8	
Lysine	128	146.19	2.16	9.06	10.54	−3.9	
Methionine	131	149.21	2.13	9.28		1.9	
Phenylalanine	147	165.19	2.20	9.31		2.8	
Proline	97	115.13	1.95	10.64		−1.6	
Serine	87	105.09	2.19	9.21		−0.8	
Threonine	101	119.12	2.09	9.10		−0.7	
Tryptophan	186	204.23	2.46	9.41		−0.9	
Tyrosine	163	181.19	2.20	9.21	10.46	−1.3	
Valine	99	117.15	2.29	9.74		4.2	

Data used from Kyte and Doolittle (1982).

Appendix 8

Commonly Used Techniques in Molecular Cloning

PREPARATION OF GLASSWARE AND PLASTICWARE

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

All glassware should be sterilized by autoclaving or baking. Some, but not all, plasticware can be autoclaved, depending on the type of plastic. Many items of sterilized plasticware are commercially available. All of the procedures commonly used in molecular cloning should be carried out in sterile glassware or plasticware; there is no significant loss of material by adsorption onto the surfaces of the containers. However, for certain procedures (e.g., handling very small quantities of single-stranded DNA or sequencing by the Maxam-Gilbert technique), it is best to use glassware or plasticware that has been coated with a thin film of silicone. A simple procedure for siliconizing small items such as pipettes, tubes, and beakers is given below. To siliconize large items such as glass plates, please refer to the note at the end of the protocol.

Siliconizing Glassware, Plasticware, and Glass Wool

The following method was supplied by Brian Seed (Massachusetts General Hospital).

1. Place the items to be siliconized inside a large, glass desiccator.
2. Add 1 ml of dichlorodimethylsilane <!> to a small beaker inside the desiccator.
3. Attach the desiccator, through a trap, to a vacuum pump. Turn on the vacuum and continue to apply suction until the dichlorodimethylsilane begins to boil. Immediately clamp the connection between the vacuum pump and the desiccator. Switch off the vacuum pump. The desiccator should maintain a vacuum.

   It is essential to turn off the vacuum pump as soon as the dichlorodimethylsilane begins to boil. Otherwise, the volatile agent will be sucked into the pump and cause irreparable damage to the vacuum seals.

4. When the dichlorodimethylsilane has evaporated (1–2 hours), open the desiccator in a chemical fume hood. After the fumes of dichlorodimethylsilane have dispersed, remove the glassware or plasticware. Bake glassware and glass wool for 2 hours at 180°C before use. Rinse plasticware extensively with H_2O before use; do not autoclave.

NOTES

- Large items of glassware can be siliconized by soaking or rinsing them in a 5% solution of dichlorodimethylsilane in chloroform or heptane. Commercial preparations for siliconizing are also available (e.g., Sigmacoat).
- As the organic solvent evaporates, the dichlorodimethylsilane is deposited on the glassware, which must be rinsed numerous times with H_2O or baked for 2 hours at 180°C before use.

Preparation of RNase-free Glassware

Guidelines for the treatment of glassware for use with RNA are given in the information panel on **HOW TO WIN THE BATTLE WITH RNASE** in Chapter 7.

PREPARATION OF DIALYSIS TUBING

The separation of molecules across a semipermeable membrane is driven by the concentration differential between the solutions on either side of the membrane and is constrained by the size (molecular weight) of the molecules relative to the size of the pores within the membrane. The pore size determines the molecular-weight cut-off, defined as the molecular weight at which 90% of the solute will be retained by the membrane. The exact permeability of a solute is dependent not only on the size of the molecule, but also on the shape of the molecule, its degree of hydration, and its charge. Each of these parameters may be influenced by the nature of the solvent, its pH, and its ionic strength. As a consequence, the molecular-weight cut-off should be used as a guide and not an absolute predictor of performance with every type of solute and solvent. Dialysis membranes are available in an enormous range of pore sizes (from 100 daltons to 2000 kD). For dialysis of most plasmid DNAs and many proteins, a molecular-weight cut-off of 12,000 to 14,000 is suitable.

1. Cut the tubing into pieces of convenient length (10–20 cm).
2. Boil the tubing for 10 minutes in a large volume of 2% (w/v) sodium bicarbonate and 1 mM EDTA (pH 8.0).
3. Rinse the tubing thoroughly in distilled H_2O.
4. Boil the tubing for 10 minutes in 1 mM EDTA (pH 8.0).
5. Allow the tubing to cool, and then store it at 4°C. Be sure that the tubing is always submerged.

   ▲ **IMPORTANT** From this point onward, always handle the tubing with gloves.
6. Before use, wash the tubing inside and out with distilled H_2O.

NOTE

- Instead of boiling for 10 minutes in 1 mM EDTA (pH 8.0) (Step 4), the tubing may be autoclaved at 20 psi (1.40 kg/cm^2) for 10 minutes on liquid cycle in a loosely capped jar filled with H_2O.

STORAGE OF BACTERIAL CULTURES

Stab Cultures

To store a bacterial culture in solid medium, pick a single, well-isolated colony with a sterile inoculating needle and stab the needle several times through the agar to the bottom of a stab vial (for the preparation of stab vials, please see Appendix 2). Replace and tighten the cap, and label both the vial and the cap. Store the vial in the dark at room temperature.

Cultures Containing Glycerol

Storage of Bacterial Cultures Growing in Liquid Media

1. To 1.5 ml of bacterial culture, add 0.5 ml of sterile 60% glycerol (sterilized by autoclaving for 20 minutes at 15 psi [1.05 kg/cm^2] on liquid cycle).
2. Vortex the culture to ensure that the glycerol is evenly dispersed.
3. Transfer the culture to a labeled storage tube equipped with a screw cap and an air-tight gasket.
4. Freeze the culture in ethanol–dry ice or in liquid nitrogen, and then transfer the tube to –70°C for long-term storage.
5. To recover the bacteria, scrape the frozen surface of the culture with a sterile inoculating loop, and then immediately streak the bacteria that adhere to the needle onto the surface of an LB agar plate containing the appropriate antibiotic. Return the frozen culture to storage at –70°C. Incubate the plate overnight at 37°C.

Storage of Bacterial Cultures Growing on Agar Plates

1. Scrape the bacteria growing on the surface of an agar plate into 2 ml of LB medium in a sterile tube. Add an equal volume of LB medium containing 30% sterile glycerol.
2. Vortex the mixture to ensure that the glycerol is completely dispersed.
3. Dispense aliquots of the glycerinated culture into sterile tubes equipped with screw caps and air-tight gaskets. Freeze the cultures as described above.

   This method is useful for storing copies of cDNA libraries established in plasmid vectors (for discussion, please see Hanahan 1985).

ESTIMATION OF CELL NUMBER*

The number of mammalian cells in a defined volume of medium can be measured using a hemocytometer. Automated methods using cell-counting devices such as those produced by Coulter are desirable when large numbers of individual samples are to be counted. A method to estimate the number of live cells in a population by staining with a vital dye also is provided here.

Hemocytometry Counting

A hemocytometer contains two chambers, each of which when filled and coverslipped contains a total volume of 9 µl. Each chamber is ruled into nine major squares, and each square is 1 x 1 mm with a depth of 0.1 mm. Thus, when coverslipped, the volume of each square is 0.1 mm^3 or 0.1 µl. Additional subdivisions of the major nine squares are not necessary for counting and can be ignored. A representation of the marking on a hemocytometer is shown in Figure A8-1.

1. Trypsinize the cells (please see Chapter 17, Protocol 8: Stage 1, Step 18) and resuspend them in growth medium.

2. Use Pasteur pipettes to remove two independent samples from the cell suspension to be counted. Deliver each sample of cell suspension into one side of the coverslipped hemocytometer by capillary action.

   Fluid should just fill the chamber and not overflow into the troughs outside the counting face. Load the first sample into one chamber and the second sample into the second chamber.

3. Count the total number of cells in five of the nine large squares in each of two sides of the hemocytometer for a total of ten squares.

   The microscope field using a 10x objective and a 10x ocular should encompass the majority of one of the nine squares of the chamber and is a convenient magnification to use for counting. Cells that overlap the border on two sides of the square should be included in the cell count and not counted on the other two sides. If the initial dilution results in more than 50–100 cells/square, make a further dilution to improve counting accuracy and speed the process of determining cell numbers.

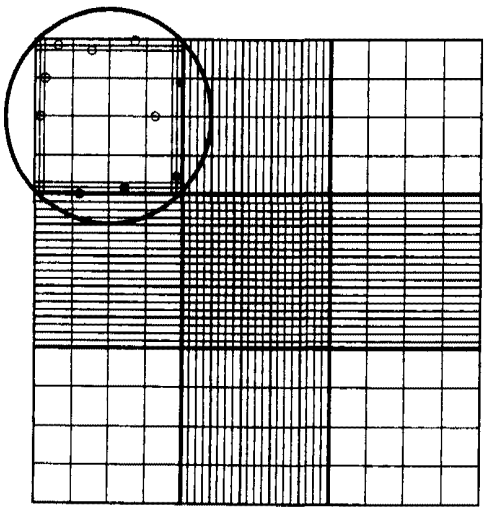


FIGURE A8-1 Standard Hemocytometer Chamber

The circle indicates the approximate area covered at 100x microscope magnification (10x ocular and 10x objective). Count the cells on top and left touching the middle line (*open circles*). Do not count the cells touching the middle line at bottom and right (*closed circles*). Count the 4 corner squares and the middle square in both chambers (only one chamber is represented here).

*Adapted from Spector et al. (1998 *Cells: A Laboratory Manual*).

4. Add the number of cells in a total of ten chambers (five from one side and five from the other) to give the number of cells in 1×10^{-3} ml (1×10^{-4} ml/square x 10 squares = a volume of 10^{-3} ml). Multiply the total number of cells by 1000 to give the number of cells/ml in the sample counted.

   If dilutions from the original cell suspension have been made, this factor must also be incorporated.

   EXAMPLE:

   1 ml of a 10-ml suspension of cells is diluted with 4 ml of medium. The diluted suspension is then sampled with a Pasteur pipette twice. The first sample is delivered to one chamber of the hemocytometer. The second sample is delivered to the second side. Five squares are counted from each side of the hemocytometer.

   Number of cells/square: 45, 37, 52, 40, 60, 48, 54, 70, 58, 60
   Total count: 524
   Dilution factor: (1 + 4)/1 = 5

   Cells/ml (in original): $524 \times 10^3 \times 5 = 2.62 \times 10^6$/ml cells

5. Immediately after use, clean the hemocytometer and coverslip by rinsing in distilled H_2O followed by 70% ethanol. Dry with lens paper.

   ▲ **IMPORTANT** Do not allow the cell suspension to dry on the hemocytometer.

NOTES

Errors that may result from using hemocytometer counts are due to:

- ***Variable sampling from the original cell suspension.*** The cell suspension must be agitated; do not allow the cells to settle to the bottom of the container.
- ***Inadequate or excessive filling of the hemocytometer chamber.*** The volume in the chambers counted is based on the coverslip resting on the sides of the hemocytometer. Overflow increases the volume counted.
- ***Cell clumping.*** Large clumps of cells may be too large to enter the chamber through capillary action and will be excluded from the cell count. Small clumps that are able to enter the chamber are difficult to count with accuracy. It is important to have a monodisperse suspension of cells for accurate counting. The cells must be thoroughly mixed to achieve uniformity.

Viability Staining

Various manipulations of cells, including passaging, freezing, and dissociation from primary tissue, can result in cell death. Exclusion of the dye, Trypan Blue, can be used to determine the number of surviving cells in a population (Phillips 1973). Normal healthy cells are able to exclude the dye, but Trypan Blue diffuses into cells in which membrane integrity has been lost. The dye exclusion method is an approximate estimate of cell viability and often does not distinguish within a 10–20% difference. Additionally, cells that exclude dye are not necessarily capable of attachment and prolonged survival or proliferation.

1. Trypsinize the cells (please see Chapter 17, Protocol 8: Stage 1, Step 18) and aseptically dilute 0.5 ml of cells into phosphate-buffered saline to a concentration from 2×10^5 to 4×10^5 cells/ml.
2. Aseptically transfer 0.5 ml of the diluted cell suspension in phosphate-buffered saline to a fresh tube and add 0.5 ml of a solution of Trypan Blue (0.4% w/v).

3. Allow the cells to remain in the dye solution for no less than 3 minutes and no longer than 10 minutes. Use a Pasteur pipette to sample the cells in dye and deliver them to a hemocytometer by capillary action.
4. Count a total of at least 500 cells, keeping a separate count of blue cells. Determine the frequency of those that are blue, i.e., have not excluded the dye.
5. Determine the percent viability from the number of cells that have not excluded the dye.

EXAMPLE:

A monolayer culture is trypsinized and resuspended in 5 ml of medium; 0.5 ml of cells is mixed with 4.5 ml of PBS, and 0.5 ml of the suspended cells in PBS is transferred to a small tube and mixed with 0.5 ml of Trypan Blue solution. In the sample transferred to the hemocytometer, 540 cells are counted; 62 of the cells fail to exclude the dye and are blue. The percent viability equals 88.5%.

$540 - 62$ = the number of cells that excluded the dye

540 = the total number of cells counted

$$\frac{540 - 62 \times 100}{540} = 88.5\% \text{ viability}$$

PURIFICATION OF NUCLEIC ACIDS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Perhaps the most basic of all procedures in molecular cloning is the purification of nucleic acids. The key step, the removal of proteins, can often be carried out simply by extracting aqueous solutions of nucleic acids with phenol:chloroform <!> and chloroform <!>. Such extractions are used whenever it is necessary to inactivate and remove enzymes that are used in one step of a cloning operation before proceeding to the next. However, additional measures are required when nucleic acids are purified from complex mixtures of molecules such as cell lysates. In these cases, it is usual to remove most of the protein by digestion with proteolytic enzymes such as pronase or proteinase K (please see Appendix 4, Table A4-8), which are active against a broad spectrum of native proteins, before extracting with organic solvents.

Extraction with Phenol:Chloroform

The standard way to remove proteins from nucleic acid solutions is to extract first with phenol:chloroform (optionally containing hydroxyquiniline at 0.1%) and then with chloroform. This procedure takes advantage of the fact that deproteinization is more efficient when two different organic solvents are used instead of one. Furthermore, although phenol denatures proteins efficiently, it does not completely inhibit RNase activity, and it is a solvent for RNA molecules that contain long tracts of poly(A) (Brawerman et al. 1972). Both of these problems can be circumvented by using a mixture of phenol:chloroform:isoamyl alcohol (25:24:1). The subsequent extraction with chloroform removes any lingering traces of phenol from the nucleic acid preparation. Extraction with ether, which was widely used for this purpose for many years, is no longer required or recommended for routine purification of DNA.

1. Transfer the sample to a polypropylene tube and add an equal volume of phenol:chloroform.

   The nucleic acid will tend to partition into the organic phase if the phenol has not been adequately equilibrated to a pH of 7.8–8.0.

2. Mix the contents of the tube until an emulsion forms.

3. Centrifuge the mixture at 80% of the maximum speed that the tubes can bear for 1 minute at room temperature. If the organic and aqueous phases are not well separated, centrifuge again for a longer time.

   Normally, the aqueous phase forms the upper phase. However, if the aqueous phase is dense because of salt (>0.5 M) or sucrose (>10%), it will form the lower phase. The organic phase is easily identifiable because of the yellow color contributed by the 8-hydroxyquinoline that is added to phenol during equilibration (please see Appendix 1).

4. Use a pipette to transfer the aqueous phase to a fresh tube. For small volumes (<200 µl), use an automatic pipettor fitted with a disposable tip. Discard the interface and organic phase.

   To achieve the best recovery, the organic phase and interface may be "back-extracted" as follows: After the first aqueous phase has been transferred as described above, add an equal volume of TE (pH 7.8) to the organic phase and interface. Mix well. Separate the phases by centrifugation as in Step 3. Combine this second aqueous phase with the first, and proceed to Step 5.

5. Repeat Steps 1–4 until no protein is visible at the interface of the organic and aqueous phases.

6. Add an equal volume of chloroform and repeat Steps 2–4.

7. Recover the nucleic acid by standard precipitation with ethanol.

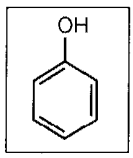
   Occasionally, ether <!> is used to remove traces of chloroform from preparations of high-molecular-weight DNA (please see the Notes below).

NOTES

The organic and aqueous phases may be mixed by vortexing when isolating small DNA molecules (<10 kb) or by gentle shaking when isolating DNA molecules of moderate size (10–30 kb). When isolating large DNA molecules (>30 kb), the following precautions must be taken to avoid shearing (please see also Chapter 6).

- Mix the organic and aqueous phases by rotating the tube slowly (20 rpm) on a wheel.
- Use large-bore pipettes to transfer the DNA from one tube to another.

HISTORICAL FOOTNOTE ON PHENOL (C_6 H_6O, F.W. = 94.11)



Until the mid 1950s, the standard method of purifying DNA involved stripping protein from the nucleic acid with detergent and strong salt solutions (e.g., perchlorate). Final deproteinization was achieved by several extractions with chloroform laced with isoamyl alcohol (Sevag 1934; Sevag et al. 1938). The first reported use of phenol to purify nucleic acids was published by Kirby (1956), who was aware of the power of phenol to extract proteins from aqueous solution (Grassmann and Deffner 1953). In his initial paper, Kirby showed that extraction of homogenates of mammalian tissue with a two-phase phenol-H_2O mixture at room temperature led to partitioning of RNA into the aqueous layer. DNA remained associated with protein at the interface. Kirby quickly realized that replacement of H_2O by solutions of anionic salts released both RNA and DNA into the aqueous phase (Kirby 1957; for review, please see Kirby 1964). Although the use of anionic salts to release proteins from DNA was quickly abandoned in favor of strong anionic detergents such as SDS, Kirby's original description of phenol extraction forms the basis of many purification methods in common use today. The function of the phenol is probably the same as that of a protein solvent: It extracts protein that has been dissociated from nucleic acids by anionic salts or detergents. So efficient is this process that pure preparations of nucleic acid are obtained after just two or three extractions with phenol.

Purified phenol has a specific gravity of 1.07 and therefore forms the lower phase when mixed with H_2O. However, the organic and aqueous phases may be difficult to separate or may invert when phenol is used to extract protein from aqueous solutions containing high concentrations of solutes. This problem is largely alleviated when a 50:50 mixture of phenol:chloroform is used, because the higher density of chloroform (1.47) ensures separation of the two phases. Denatured proteins collect at the interface between the two phases while lipids partition efficiently into the organic layer. Isoamyl alcohol is often added to the phenol:chloroform mixture to reduce foaming.

Pure phenol is supplied as a white crystalline mass (mp 43°C). However, on exposure to air and light, phenol is prone to redden, a process that is accelerated by alkalinity. Crystalline phenol is not recommended because it must be redistilled at 182°C to remove oxidation products such as quinones that cause the breakdown of phosphodiester bonds or promote cross-linking of nucleic acids.

The liquefied form of phenol provided by many manufacturers contains ~8% H_2O and can be stored frozen at –20°C. Liquefied phenol, if colorless, can be used in molecular cloning without redistillation. Today, only occasional batches of liquefied phenol are pink or yellow, and these should be rejected and returned to the manufacturer.

Before use, phenol must be saturated with H_2O and equilibrated with Tris to a pH of >7.8 to suppress partitioning of DNA into the organic phase, which occurs at acidic pH.

Drop Dialysis

Low-molecular-weight contaminants, which may inhibit restriction digestion or DNA sequencing, can be removed from DNA in solution by drop dialysis.

1. Spot a drop (~50 µl) of DNA in the center of a Millipore Series V membrane (0.025 µm), floating shiny side up on 10 ml of sterile H_2O in a 10-cm diameter Petri dish.

2. Dialyze the DNA for 10 minutes.

3. Remove the drop to a clean microfuge tube, and use aliquots of the dialyzed DNA for restriction enzyme digestion and/or DNA sequencing.

CONCENTRATING NUCLEIC ACIDS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Ethanol Precipitation

Precipitation with ethanol is the standard method to recover nucleic acids from aqueous solutions. It is rapid, virtually foolproof, and efficient: Subnanogram amounts of DNA and RNA can be quantitatively precipitated with ethanol, collected by centrifugation, and redissolved within minutes.

Ethanol depletes the hydration shell from nucleic acids and exposes negatively charged phosphate groups. Counterions such as Na^+ bind to the charged groups and reduce the repulsive forces between the polynucleotide chains to the point where a precipitate can form. Ethanol precipitation can therefore only occur if cations are available in sufficient quantity to neutralize the charge on the exposed phosphate residues. The most commonly used cations are shown in Table A8-1 and are described below.

- ***Ammonium acetate*** is frequently used to reduce the coprecipitation of unwanted contaminants (e.g., dNTPs or oligosaccharides) with nucleic acids. For example, two sequential precipitations of DNA in the presence of 2 M ammonium acetate result in the removal of >99% of the dNTPs from preparations of DNA (Okayama and Berg 1982). Ammonium acetate is also the best choice when nucleic acids are precipitated after digestion of agarose gels with agarase. The use of this cation reduces the possibility of coprecipitation of oligosaccharide digestion products. However, ammonium acetate should not be used when the precipitated nucleic acid is to be phosphorylated, since bacteriophage T4 polynucleotide kinase is inhibited by ammonium ions.
- ***Lithium chloride*** is frequently used when high concentrations of ethanol are required for precipitation (e.g., when precipitating RNA). LiCl is very soluble in ethanolic solutions and is not coprecipitated with the nucleic acid. Small RNAs (tRNAs and 5S RNAs) are soluble in solutions of high ionic strength (without ethanol), whereas large RNAs are not. Because of this difference in solubility, precipitation in high concentrations of LiCl (0.8 M) can be used to purify large RNAs.
- ***Sodium chloride*** (0.2 M) should be used if the DNA sample contains SDS. The detergent remains soluble in 70% ethanol.
- ***Sodium acetate*** (0.3 M, pH 5.2) is used for most routine precipitations of DNA and RNA.

Until a few years ago, ethanol precipitation was routinely carried out at low temperature (e.g., in a dry-ice/methanol bath). This is now known to be unnecessary. At 0°C in the absence of carrier, DNA concentrations as low as 20 ng/ml will form a precipitate that can be quantitatively recovered by centrifugation in a microfuge. However, when lower concentrations of DNA or very small fragments (<100 nucleotides in length) are processed, more extensive centrifugation may be necessary to cause the pellet of nucleic acid to adhere tightly to the centrifuge tube. Centrifugation at 100,000*g* for 20–30 minutes allows the recovery of picogram quantities of nucleic acid in the absence of carrier.

TABLE A8-1 Salt Solutions

SALT	STOCK SOLUTION (M)	FINAL CONCENTRATION (M)
Ammonium acetate	10.0	2.0–2.5
Lithium chloride	8.0	0.8
Sodium chloride	5.0	0.2
Sodium acetate	3.0 (pH 5.2)	0.3

When dealing with small amounts of DNA, it is prudent to save the ethanolic supernatant from each step until all of the DNA has been recovered. This retention is especially important after precipitates of DNA have been washed with 70% ethanol, a treatment that often loosens the precipitates from the wall of the tube.

Dissolving DNA Precipitates

Until a few years ago, DNA precipitates recovered after ethanol precipitation were dried under vacuum before being redissolved. This practice has now been abandoned (1) because desiccated pellets of DNA dissolve slowly and inefficiently and (2) because small fragments of double-stranded DNA (<400 bp) become denatured upon drying, probably as a result of loss of the stabilizing shell of bound water molecules (Svaren et al. 1987).

These days, the best practice is to remove ethanol from the nucleic acid pellet and from the sides of the tube by gentle aspiration and then to store the open tube on the bench for ~15 minutes to allow most of the residual ethanol to evaporate. The still-damp pellet of nucleic acid can then be dissolved rapidly and completely in the appropriate buffer. If necessary, the open tube containing the redissolved DNA can be incubated for 2–3 minutes at 45°C in a heating block to allow any traces of ethanol to evaporate.

The precipitated DNA is not all found at the bottom of the tube after centrifugation in an angle-head rotor. In the case of microfuge tubes, for example, at least 40% of the precipitated DNA is plastered on the wall of the tube. To maximize recovery of DNA, use a pipette tip to roll a bead of solvent several times over the appropriate segment of the wall. If the sample of DNA is radioactive, check that no detectable radioactivity remains in the tube after the dissolved DNA has been removed.

Carriers

Carriers (or coprecipitants) are inert substances that are used to improve the recovery of small quantities of nucleic acids during ethanol precipitation. Insoluble in ethanolic solutions, carriers form a precipitate that traps the target nucleic acids. During centrifugation, carriers generate a visible pellet that facilitates handling of the target nucleic acids. This may be their major virtue: As discussed above, ethanol precipitation — even of small amounts of nucleic acids in dilute solution — is remarkably efficient. Carriers do little, other than provide visual clues to the location of the target nucleic acid. Three substances are commonly used as carriers: yeast tRNA, glycogen, and linear polyacrylamide. Their advantages and disadvantages are listed in Table A8-2.

TABLE A8-2 Carriers

Carrier	Working Concentration	Advantages/Disadvantages
Yeast tRNA	10–20 µg/ml	Yeast tRNA is inexpensive, but it has the disadvantage that it cannot be used for precipitating nucleic acids that will be used as substrates in reactions catalyzed by polynucleotide kinase or terminal transferase. The termini of yeast RNA are excellent substrates for these enzymes and would compete with the termini contributed by the target nucleic acid.
Glycogen	50 µg/ml	Glycogen is usually used as a carrier when nucleic acids are precipitated with 0.5 M ammonium acetate and isopropanol. Glycogen is not a nucleic acid and therefore does not compete with the target nucleic acids in subsequent enzymatic reactions. However, it can interfere with interactions between DNA and proteins (Gaillard and Strauss 1990).
Linear polyacrylamide	10–20 µg/ml	Linear polyacrylamide is an efficient neutral carrier for precipitating picogram amounts of nucleic acids with ethanol and proteins with acetone (Strauss and Varshavsky 1984; Gaillard and Strauss 1990).

HISTORICAL FOOTNOTE

Ethanol precipitation predates molecular cloning by ~50 years. It was first used as a method to concentrate biologically active nucleic acid by J. Lionel Alloway, who worked at the Rockefeller Institute in the early 1930s. His project was to prepare active cell-free extracts of S-type *Streptococcus pneumoniae* that would permit bacterial transformation of R-type organisms in vitro. At that time, transformation had been achieved only with intact, heat-killed donor cells. After many frustrating failures, Alloway reported in 1932 that he could get the substance responsible for transformation into solution by heating a freeze/thaw extract of the S organisms to 60°C, removing particulate matter by centrifugation, and passing the solution through a filter made of porous porcelain (Alloway 1932). This last step was included to silence skeptics who believed that transformation was an artifact caused by an occasional S-type organism that survived the extraction procedure.

Alloway's success at eliminating the need for heat-killed donor cells was a major step on the road that eventually led to the discovery of DNA as the transforming material (Avery et al. 1944). However, not all of Alloway's cell-free preparations worked, and, even when transformation was obtained, the efficiency was very low. Alloway must have realized that these problems were caused by the dilute nature of the extract, for he began to search for different ways to lyse the pneumococci and for different methods to concentrate the transforming activity (Alloway 1933). Maclyn McCarty (1985) described Alloway's discovery of ethanol precipitation as follows:

> Alloway then introduced another new procedure that became an indispensable part of all work on the transforming substance from that time forward. He added pure alcohol in a volume five times that of the extract which resulted in precipitation of most of the material that had been released from the pneumococci... . The precipitated material could be redissolved in salt solution and shown to contain the active substance in transformation tests. Alcohol precipitation and resolution could be repeated at will without loss of activity.

Alloway was certainly not the first person to precipitate nucleic acids with ethanol. This technique had already been used as a purification step by several generations of organic chemists who were puzzling over the structure of the bases in DNA. However, Alloway was the first to use ethanol precipitation to prepare material that could change the phenotype of recipient cells. Final proof that the transforming factor was DNA still lay a dozen or more years into the future. But Alloway could fairly claim to be the inventor of a technique that is now second nature to us all.

Standard Ethanol Precipitation of DNA in Microfuge Tubes

1. Estimate the volume of the DNA solution.

2. Adjust the concentration of monovalent cations either by dilution with TE (pH 8.0) if the DNA solution contains a high concentration of salts or by addition of one of the salt solutions shown in Table A8-1.

   If the volume of the final solution is 400 µl or less, carry out precipitation in a single microfuge tube. Larger volumes can be divided among several microfuge tubes, or the DNA can be precipitated and centrifuged in tubes that will fit in a medium-speed centrifuge or ultracentrifuge.

3. Mix the solution well. Add exactly 2 volumes of ice-cold ethanol and again mix the solution well. Store the ethanolic solution on ice to allow the precipitate of DNA to form.

   Usually 15–30 minutes is sufficient, but when the size of the DNA is small (<100 nucleotides) or when it is present in small amounts (<0.1 µg/ml), extend the period of storage to at least 1 hour and add $MgCl_2$ to a final concentration of 0.01 M.

   DNA can be stored indefinitely in ethanolic solutions at 0°C or at –20°C.

4. Recover the DNA by centrifugation at 0°C.

   For most purposes, centrifugation at maximum speed for 10 minutes in a microfuge is sufficient. However, as discussed above, when low concentrations of DNA (<20 ng/ml) or very small fragments are being processed, more extensive centrifugation may be required.

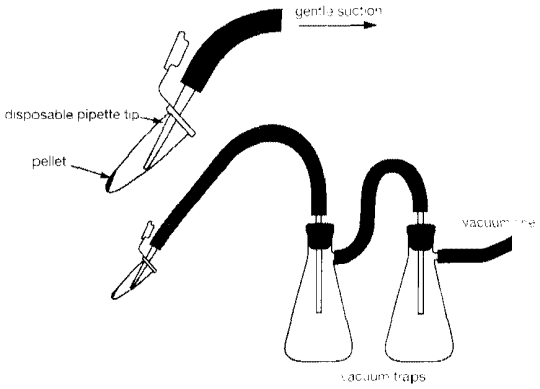


FIGURE A8-2 Aspiration of Supernatants

Hold the open microfuge tube at an angle, with the pellet on the upper side. Use a disposable pipette tip attached to a vacuum line to withdraw fluid from the tube. Insert the tip just beneath the miniscus on the lower side of the tube. Move the tip toward the base of the tube as the fluid is withdrawn. Use a gentle suction to avoid drawing the pellet into the pipette tip. Keep the end of the tip away from the pellet. Finally, vacuum the walls of the tube to remove any adherent drops of fluid.

5. Carefully remove the supernatant with an automatic micropipettor or with a disposable pipette tip attached to a vacuum line (please see Figure A8-2). Take care not to disturb the pellet of nucleic acid (which may be invisible). Use the pipette tip to remove any drops of fluid that adhere to the walls of the tube.

   It is best to save the supernatant from valuable DNA samples until recovery of the precipitated DNA has been verified.

6. Fill the tube half way with 70% ethanol and recentrifuge at maximum speed for 2 minutes at 4°C in a microfuge.

7. Repeat step 5.

8. Store the open tube on the bench at room temperature until the last traces of fluid have evaporated.

   It was once common practice to dry pellets of nucleic acid in a lyophilizer. This step is not only unnecessary, but also undesirable, since it causes denaturation of small (<400-nucleotide) fragments of DNA (Svaren et al. 1987) and greatly reduces the recovery of larger fragments of DNA.

9. Dissolve the DNA pellet (which is often invisible) in the desired volume of buffer (usually TE [pH between 7.6 and 8.0]). Rinse the walls of the tube well with the buffer.

NOTES

- After centrifugation in a microfuge, not all of the DNA is deposited on the bottom of the microfuge tube. Up to 50% of the DNA is smeared on the wall of the tube. To recover all of the DNA, it is necessary to work a bead of fluid backward and forward over the appropriate quadrant of wall. This step can easily be done by pushing the bead of fluid over the surface with a disposable pipette tip attached to an automatic micropipettor.

- One volume of isopropanol <!> may be used in place of 2 volumes of ethanol to precipitate DNA. Precipitation with isopropanol has the advantage that the volume of liquid to be centrifuged is smaller. However, isopropanol is less volatile than ethanol and is therefore more difficult to remove; moreover, solutes such as sucrose or sodium chloride are more easily coprecipitated with DNA when isopropanol is used. In general, precipitation with ethanol is preferable, unless it is necessary to keep the volume of fluid to a minimum.

- In general, DNA precipitated from solution by ethanol can be redissolved easily in buffers of low ionic strength, such as TE (pH 8.0). Occasionally, difficulties arise when buffers containing $MgCl_2$ or >0.1 M NaCl are added directly to the DNA pellet. It is therefore preferable to dissolve the DNA in a small volume of low-ionic-strength buffer and to adjust the composi-

tion of the buffer later. If the sample does not dissolve easily in a small volume, add a larger volume of buffer and repeat the precipitation with ethanol. The second precipitation may help eliminate additional salts or other components that may be preventing dissolution of the DNA.

Precipitation of RNA with Ethanol

RNA is efficiently precipitated with 2.5–3.0 volumes of ethanol from solutions containing 0.8 M LiCl, 5 M ammonium acetate, or 0.3 M sodium acetate. The choice among these salts is determined by the way in which the RNA will be used later. Since the potassium salt of dodecyl sulfate is extremely insoluble, avoid potassium acetate if the precipitated RNA is to be dissolved in buffers that contain SDS, for example, buffers that are used for chromatography on oligo(dT)-cellulose. For the same reason, avoid potassium acetate if the RNA is already dissolved in a buffer containing SDS. Avoid LiCl when the RNA is to be used for cell-free translation or reverse transcription. LiCl ions inhibit initiation of protein synthesis in most cell-free systems and suppress the activity of RNA-dependent DNA polymerase.

NOTE

- Solutions used for precipitation of RNA must be free of RNase (please see Chapter 7).

Precipitation of Large RNAs with Lithium Chloride

Whereas small RNAs (tRNAs and 5S RNAs) are soluble in solutions of high ionic strength, large RNAs (e.g., rRNAs and mRNAs) are insoluble and can be removed by centrifugation.

1. Measure the volume of the sample and add 0.2 volume of RNase-free 8 M LiCl. Mix the solution well and store it on ice for at least 2 hours.
2. Centrifuge the solution at 15,000*g* for 20 minutes at 0°C. Discard the supernatant, and dissolve the precipitated high-molecular-weight RNA in 0.2 volume of H_2O.
3. Repeat Steps 1 and 2.
4. Recover the high-molecular-weight RNA from the resuspended pellet by precipitation with 2 volumes of ethanol.

Concentrating and Desalting Nucleic Acids with Microconcentrators

Ultrafiltration is an alternative to ethanol precipitation for the concentration and desalting of nucleic acid solutions. It requires no phase change and is particularly useful for dealing with very low concentrations of nucleic acids. The Microcon cartridge, supplied by Millipore, is a centrifugal ultrafiltration device that can desalt and concentrate nucleic acid samples efficiently. The protocol presented below and the accompanying notes have been adapted from those provided on the Millipore Web Site (www.millipore.com). Complete directions may be found on this Web Site.

1. Select a Microcon unit with a nucleotide cut-off equal to or smaller than the molecular size of the nucleic acid of interest (please see Table A8-3).
2. Insert the Microcon cartridge into one of the two vials supplied, as shown in Figure A8-3.
3. To concentrate (without affecting salt concentration), pipette up to 500 µl of sample (DNA or RNA) into the reservoir. Centrifuge for the recommended time, not exceeding the *g* force shown in Table A8-3.

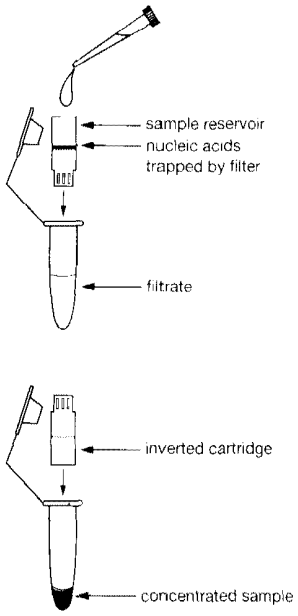


FIGURE A8-3 Concentration/Desalting of Nucleic Acid Solutions Using Micron Ultracentrifugation

4. To exchange salt, add the proper amount of appropriate diluent to bring the concentrated sample to 500 µl. Centrifuge for the recommended time, not exceeding the *g* force shown in Table A8-3. To achieve a lower salt concentration, repeat the entire step as necessary. Please see the footnote below to Table A8-3.

   ▲ **IMPORTANT** Do not overfill the filtrate vial.

5. Remove the reservoir from the vial and invert the reservoir into a new vial (save the filtrate until the sample has been analyzed).

6. Centrifuge at 500–1000*g* for 2 minutes in a microfuge to recover nucleic acid in the vial.

7. Remove reservoir. Cap the vial to store the sample.

TABLE A8-3 Nucleotide Cut-offs for Microcon Concentrators

Microcon Model	Color Code	Nucleotide Cut-off[a]		Maximum Recommended *g* Force	Spin Time in Minutes	
		ss	ds		4°C	25°C
3	yellow	10	10	14,000	185	95
10	green	30	20	14,000	50	35
30	clear	60	50	14,000	15	8
50	rose	125	100	14,000	10	6
100	blue	300	125	500	25	15

Note that ultrafiltration alone does not change buffer composition. The salt concentration in a sample concentrated by spinning in a Microcon will be the same as that in the original sample. For desalting, the concentrated sample is diluted with H_2O or buffer to its original volume and spun again (called discontinuous diafiltration). This removes the salt by the concentration factor of the ultrafiltration. For example, if a 500-µl sample containing 100 mM salt is concentrated to 25 µl (20x concentration factor), 95% of the total salt in the sample will be removed. The salt concentration in the sample will remain at 100 mM. Rediluting the sample to 500 µl in H_2O will bring the salt concentration to 5 mM. Concentrating to 25 µl once more will remove 99% of the original total salt. The concentrated sample will now be in 0.25 mM salt. For more complete salt removal, an additional redilution and spinning cycle will remove 99.9% of the initial salt content.

[a]ss indicates single-stranded and ds indicates double-stranded.

Concentrating Nucleic Acids by Extraction with Butanol

During extraction of aqueous solutions with solvents such as secondary butyl alcohol (isobutanol) or *n*-butyl alcohol (*n*-butanol<!>), some of the water molecules are partitioned into the organic phase. By carrying out several cycles of extraction, the volume of a nucleic acid solution can be reduced significantly. This method of concentration is used to reduce the volume of dilute solutions to the point where the nucleic acid can be recovered easily by precipitation with ethanol.

1. Measure the volume of the nucleic acid solution and add an equal volume of isobutanol. Mix the solution well by vortexing.

   Addition of too much isobutanol can result in removal of all the H_2O and precipitation of the nucleic acid. If this happens, add H_2O to the organic phase until an aqueous phase (which should contain the nucleic acid) reappears.

2. Centrifuge the solution at maximum speed for 20 seconds at room temperature in a microfuge or at 1600*g* for 1 minute in a benchtop centrifuge. Use an automatic micropipettor to remove and discard the upper (isobutanol) phase.

3. Repeat Steps 1 and 2 until the desired volume of aqueous phase is achieved.

   Because isobutanol extraction does not remove salt, the salt concentration increases in proportion to the reduction in the volume of the solution. The nucleic acid can be transferred to the desired buffer by spun-column chromatography or by precipitation with ethanol.

QUANTITATION OF NUCLEIC ACIDS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Two types of methods are widely used to measure the amount of nucleic acid in a preparation. If the sample is pure (i.e., without significant amounts of contaminants such as proteins, phenol, agarose, or other nucleic acids), spectrophotometric measurement of the amount of UV irradiation absorbed by the bases is simple and accurate. If the amount of DNA or RNA is very small or if the sample contains significant quantities of impurities, the amount of nucleic acid can be estimated from the intensity of fluorescence emitted by ethidium bromide or Hoechst 33258. A summary of the methods commonly used to measure the concentrations of DNA in solution are listed in Table A8-4. More detailed discussion of the methods follows after the table.

TABLE A8-4 Measuring Nucleic Acid Concentrations

METHOD	INSTRUMENT	COMMENTS
Absorbance at 260 nm	spectrophotometer	Useful only for highly purified preparations of nucleic acid, since it detects any compound that absorbs significantly at 260 nm, which includes, for example, DNA, RNA, EDTA, and phenol. The ratio of absorbance at 260 nm and 280 nm is often used as a test for contamination of a preparation of DNA and RNA with protein. Despite its popularity, this test is of questionable worth. Nucleic acids absorb so strongly at 260 nm that only a significant level of protein contamination will cause a significant change in the ratio of absorbance at the two wavelengths (Warburg and Christian 1942; Glasel 1995; Manchester 1995, 1996; Wilfinger et al. 1997) (please see the panel on **ABSORPTION SPECTROSCOPY OF NUCLEIC ACIDS** on the following page). The specific absorption coefficients of both DNA and RNA are affected by the ionic strength and the pH of the solution (Beaven et al. 1955; Wilfinger et al. 1997). Accurate measurements of concentration can be made only when the pH is carefully controlled and the ionic strength of the solution is low. It is difficult to measure the absorbance of small volumes of solution and the method is reliable only over a fairly narrow range of concentrations (5 µg/ml to 90 µg/ml).
Emission at 458 nm in the presence of Hoechst 33258	fluorometer	Hoechst 33258 is one of a class of *bis*-benzimidazole fluorescent dyes that bind nonintercalatively and with high specificity to double-stranded DNA. After binding, the fluorescent yield increases from 0.01 to 0.6 (Latt and Wohlleb 1975); Hoechst 33258 is therefore a useful fluorochrome for fluorometric detection and quantitation of double-stranded DNA. Hoechst 33258 interacts preferentially with A/T-rich regions of the DNA helix, with the $\log_{10}$ of the intensity of fluorescence increasing in proportion to the A+T content of the DNA (Daxhelet et al. 1989). The fluorescent yield of Hoechst 33258 is approximately threefold lower with single-stranded DNA (Hilwig and Gropp 1975). Fluorometry assays with Hoechst 33258 do not work at extremes of pH and are affected by both detergents and salts (Van Lancker and Gheyssens 1986). Assays are therefore usually carried out in 0.2 M NaCl, 10 mM EDTA at pH 7.4. The concentration of DNA in the unknown sample is estimated from a standard curve constructed using a set of reference DNAs (10–250 ng/ml) whose base composition is the same as the unknown sample. The intensity of emission is nearly linear over a 1000-fold range of DNA concentrations. The DNAs must be of high molecular weight since Hoechst 33258 does not bind efficiently to small fragments of DNA. All DNAs and solutions must be free of ethidium bromide, which quenches the fluorescence of Hoechst 33258. However, because Hoechst 33258 has little affinity for proteins or rRNA, measurements can be carried out using cell lysates or purified preparations of DNA (Cesarone et al. 1979, 1980; Labarca and Paigen 1980).
Dipstick (a kit from Invitrogen)		This method is good only for solutions containing low concentrations of DNA and RNA (<10 µg/ml), and is both expensive and relatively slow (30–40 minutes).
Ethidium bromide spot test	UV transilluminator	A fast and sensitive method that utilizes the UV-induced fluorescence emitted by intercalated ethidium bromide molecules. The DNA preparations under test are spotted onto an agarose plate containing 0.5 µg/ml ethidium bromide. A series of DNAs of known concentration are used as standards. Because the amount of fluorescence is proportional to the total mass of DNA, the quantity of DNA can be estimated by comparing the light emitted at 590 nm by the test preparations and the standards. The results of the assay can be recorded on film. In a similar, older test, developed in the early 1970s, DNA samples and standards are spotted onto a sheet of Saran Wrap, mixed with a dilute solution of ethidium bromide, and photographed. The chief problem with the method is that it is sensitive to interference by RNA.

Spectrophotometry of DNA or RNA

For quantitating the amount of DNA or RNA, readings are taken at wavelengths of 260 nm and 280 nm. The reading at 260 nm allows calculation of the concentration of nucleic acid in the sample. An OD of 1 corresponds to ~50 µg/ml for double-stranded DNA, 40 µg/ml for single-stranded DNA and RNA, and ~33 µg/ml for single-stranded oligonucleotides. The ratio between the readings at 260 nm and 280 nm (OD_{260}:OD_{280}) provides an estimate of the purity of the nucleic acid. Pure preparations of DNA and RNA have OD_{260}:OD_{280} values of 1.8 and 2.0, respectively. If there is significant contamination with protein or phenol, the OD_{260}:OD_{280} will be less than the values given above, and accurate quantitation of the amount of nucleic acid will not be possible.

Because it is rapid, simple, and nondestructive, absorption spectroscopy has long been the method of choice to measure the amount of DNA and RNA in concentrated pure solutions. However, absorption spectroscopy is comparatively insensitive and, with most laboratory spectrophotometers, nucleic acid concentrations of at least 1 µg/ml are required to obtain reliable estimates of A_{260}. In addition, absorption spectroscopy cannot readily distinguish between DNA and RNA, and it cannot be used with crude preparations of nucleic acids. Because of these limitations, a number of alternative methods have been devised to measure the concentration of DNA and RNA (please see Table A8-4).

ABSORPTION SPECTROSCOPY OF NUCLEIC ACIDS

Purines and pyrimidines in nucleic acids absorb UV light. As described by the Beer-Lambert Law, the amount of energy absorbed at a particular wavelength is a function of the concentration of the absorbing material:

$$I = I_0 10^{-\varepsilon dc}$$

where I = intensity of transmitted light
I_0 = intensity of incident light
ε = molar extinction coefficient (also known as the molar absorbtion coefficient)
d = optical path length (in cm)
c = concentration of absorbing material (mole/liter)

ε is numerically equal to the absorbance of a 1 M solution in a 1-cm light path and is therefore expressed in M^{-1} cm^{-1}. Absorbance data are collected using a UV spectrometer and are generally reported as the absorbance A ($\log I/I_0$). When D = 1 cm, A is called the optical density or OD at a particular wavelength, λ.

$$OD_\lambda = \varepsilon\ c$$

Because the absorption spectra of DNA and RNA are maximal at 260 nm, absorbance data for nucleic acids are almost always expressed in A_{260} or OD_{260} units. For double-stranded DNA, one A_{260} or OD_{260} unit corresponds to a concentration of 50 µg/ml. The Beer-Lambert law is valid at least to an OD = 2 and the concentration of a solution of nucleic acid is therefore easily calculated by simple interpolation. For example, a solution whose OD_{260} = 0.66 contains 33 µg/ml of double-stranded DNA. For nucleic acids, ε decreases as the ring systems of adjacent purines and pyrimidines become stacked in a polynucleotide chain. The value of ε therefore decreases in the following series:

free base
↓
small oligonucleotides
↓
single-stranded nucleic acids
↓
double-stranded nucleic acids

This means that single-stranded nucleic acids have a higher absorbance at 260 nm than double-stranded nucleic acids. Thus, the molar extinction coefficient of double-stranded DNA at 260 nm is 6.6, whereas the molar extinction coefficient of single-stranded DNA and RNA is ~7.4. Note, however, that the extinction coefficients of both DNA and RNA are affected by the ionic strength and the pH of the solution (Beaven et al. 1955; Wilfinger et al. 1997). Accurate measurements of concentration can be made only when the pH is carefully controlled and the ionic strength of the solution is low.

(Continued on facing page.)

The extinction coefficients of nucleic acids are the sum of the extinction coefficients of each of their constituent nucleotides. For large molecules, where it is both impractical and unnecessary to sum the coefficients of all the nucleotides, an average extinction coefficient is used. For double-stranded DNA, the average extinction coefficient is 50 $(\mu g/ml)^{-1}$ cm^{-1}; for single-stranded DNA or RNA, the average coefficient is 38 $(\mu g/ml)^{-1}$ cm^{-1}. These values mean that

1 OD_{260} unit equals

50 µg/ml double-stranded DNA

or

38 µg/ml single-stranded DNA or RNA

For small molecules such as oligonucleotides, it is best to calculate an accurate extinction coefficient from the base composition. Because the concentrations of oligonucleotides are commonly reported as mmole/liter, a millimolar extinction coefficient (E) is conventionally used in the Beer-Lambert equation.

$$E = A\,(15.3) + G\,(11.9) + C\,(7.4) + T\,(9.3)$$

where A, G, C, and T are the number of times each nucleotide is represented in the sequence of the oligonucleotide. The numbers in parentheses are the molar extinction coefficients for each deoxynucleotide at pH 7.0.

OD_{260}:OD_{280} Ratios

Although it is possible to estimate the concentration of solutions of nucleic acids and oligonucleotides by measuring their absorption at a single wavelength (260 nm), this is not good practice. The absorbance of the sample should be measured at several wavelengths since the ratio of absorbance at 260 nm to the absorbance at other wavelengths is a good indicator of the purity of the preparation. Significant absorption at 230 nm indicates contamination by phenolate ion, thiocyanates, and other organic compounds (Stulnig and Amberger 1994), whereas absorption at higher wavelengths (330 nm and higher) is usually caused by light scattering and indicates the presence of particulate matter. Absorption at 280 nm indicates the presence of protein, because aromatic amino acids absorb strongly at 280 nm.

For many years, the ratio of the absorbance at 260 nm and 280 nm (OD_{260}:OD_{280}) has been used as a measure of purity of isolated nucleic acids. This method dates from Warburg and Christian (1942) who showed that the ratio is a good indicator of contamination of protein preparations by nucleic acids. The reverse is not true! Because the extinction coefficients of nucleic acids at 260 nm and 280 nm are so much greater than that of proteins, significant contamination with protein will not greatly change the OD_{260}:OD_{280} ratio of a nucleic acid solution (please see Table A8-5). Nucleic acids absorb so strongly at 260 nm that only a significant level of protein contamination will cause a significant change in the ratio of absorbance at the two wavelengths (Warburg and Christian 1942; Glasel 1995; Manchester 1995, 1996; Wilfinger et al. 1997).

TABLE A8-5 Absorbance of Nucleic Acids and Proteins

% Protein	% Nucleic Acid	OD_{260}:OD_{280}	% Protein	% Nucleic Acid	OD_{260}:OD_{280}
100	0	0.57	45	55	1.89
95	5	1.06	40	60	1.91
90	10	1.32	35	65	1.93
85	15	1.48	30	70	1.94
80	20	1.59	25	75	1.95
75	25	1.67	20	80	1.97
70	30	1.73	15	85	1.98
65	35	1.78	10	90	1.98
60	40	1.81	5	95	1.99
55	45	1.84	0	100	2.00
50	50	1.87			

Using the predicted values in this table, Glasel (1995) derived an empirical equation to describe $\%N$ for a range of OD_{260}:OD_{280} ratios: $\%N = F([11.16R - 6.32\sqrt{[2.16 - R]}])$, where $R = OD_{260}$:OD_{280}. Note that estimates of purity of nucleic acids based on OD_{260}:OD_{280} ratios are accurate only when the preparations are free of phenol. Water saturated with phenol absorbs with a characteristic peak at 270 nm and an OD_{260}:OD_{280} ratio of 2 (Stulnig and Amberger 1994). Nucleic acid preparations free of phenol should have OD_{260}:OD_{270} ratios of ~1.2.

Fluorometric Quantitation of DNA Using Hoechst 33258

Measuring the concentration of DNA using fluorometry is simple and more sensitive than spectrophotometry, and allows the detection of nanogram quantities of DNA. The assay can only be used to measure the concentration of DNAs whose sizes exceed ~1 kb, as Hoechst 33258 binds poorly to smaller DNA fragments. In this assay, DNA preparations of known and unknown concentrations are incubated with Hoechst 33258 fluorochrome. Absorption values for the unknown sample are compared with those observed for the known series, and the concentration of the unknown sample is estimated by interpolation.

HOECHST 33258

Hoechst 33258 is one of a class of *bis*-benzimidazole fluorescent dyes that bind nonintercalatively and with high specificity into the minor groove of double-stranded DNA. After binding, the fluorescent yield increases from 0.01 to 0.6 (Latt and Wohlleb 1975), and Hoechst 33258 can therefore be used for fluorometric detection and quantification of double-stranded DNA in solution. Hoechst 33258 is preferred to ethidium bromide for this purpose because of its greater ability to differentiate double-stranded DNA from RNA and single-stranded DNA (Loontiens et al. 1990).

Like many other nonintercalative dyes (Müller and Gautier 1975), Hoechst 33258 binds preferentially to A/T-rich regions of the DNA helix (Weisblum and Haenssler 1974), with the $\log_{10}$ of the intensity of fluorescence increasing in proportion to the A+T content of the DNA (Daxhelet et al. 1989). The fluorescent yield of Hoechst 33258 is approximately threefold lower with single-stranded DNA (Hilwig and Gropp 1975).

Facts and Figures

- Hoechst 33258 in free solution has an excitation maximum at ~356 nm and an emission maximum at 492 nm. However, when bound to DNA, Hoechst 33258 absorbs maximally at 365 nm and emits maximally at 458 nm (Cesarone et al. 1979, 1980).
- Fluorometry assays with Hoechst 33258 do not work at extremes of pH and are affected by both detergents and salts (Van Lancker and Gheyssens 1986). Assays are therefore usually carried out under standard conditions (0.2 M NaCl, 10 mM EDTA at pH 7.4). However, two different salt concentrations are required to distinguish double-stranded from single-stranded DNA and RNA (Labarca and Paigen 1980). The concentration of DNA in the unknown sample is estimated from a standard curve constructed using a set of reference DNAs (10–250 ng/ml) whose base composition is the same as the unknown sample. The DNAs must be of high molecular weight since Hoechst 33258 does not bind efficiently to small fragments of DNA. Measurements should be carried out rapidly to minimize photobleaching and shifts in fluorescence emission due to changes in temperature. Either a fixed wavelength fluorometer (e.g., Hoefer model TKO 100) or a scanning fluorescence spectrometer (e.g., Hitachi Perkin-Elmer model MPF-2A) can be used.
- The concentration of Hoechst 33258 (M_r = 533.9) in the reaction should be kept low (5×10^{-7} M to 2.5×10^{-6} M), since quenching of fluorescence occurs when the ratio of dye to DNA is high (Stokke and Steen 1985). However, two concentrations of dye are sometimes used to extend the dynamic range of the assay.
- All DNAs and solutions should be free of ethidium bromide, which quenches the fluorescence of Hoechst 33258. However, because Hoechst 33258 has little affinity for proteins or rRNA, measurements can be carried out using cell lysates or purified preparations of DNA (Cesarone et al. 1979; Labarca and Paigen 1980).
- Unlike ethidium bromide, Hoechst dyes are cell-permeant.

1. Turn the fluorometer on 1 hour before the assay is carried out to allow the machine to warm up and stabilize.

   When bound to high-molecular-weight double-stranded DNA, Hoechst 33258 dye absorbs maximally at 365 nm and emits maximally at 458 nm.

2. Prepare an appropriate amount of diluted Hoechst 33258 dye solution by combining 50 µl of concentrated Hoechst 33258 dye solution per 100 ml of fluorometry buffer (please see Appendix 1). Each tube in the DNA assay requires 3 ml of diluted Hoechst 33258 dye solution. Transfer 3 ml of diluted dye solution to an appropriate number of clean glass tubes. Include six extra tubes for a blank and the standard curve.

The concentrated Hoechst 33258 dye solution is prepared in H_2O at 0.2 mg/ml and can be stored at room temperature in a foil-wrapped test tube.

3. Prepare a standard curve by adding 100, 200, 300, 400, and 500 ng of DNA from the reference stock solution to individual tubes. Mix and read the absorbance on the prewarmed fluorometer of each tube immediately after addition of the DNA.

   The reference stock solution of DNA is prepared in TE to a concentration of 100 μg/ml. Because the binding of Hoechst 33258 dye to DNA is influenced by the base composition, the DNA used to construct the standard curve should be from the same species as the test sample.

4. Add 0.1 μl (i.e., 1 μl of a 1:10 dilution), 1.0 μl, and 10 μl of the preparation of DNA, whose concentration is being determined, to individual tubes containing diluted dye solution. Immediately read the fluorescence.

5. Construct a standard curve plotting fluorescence on the ordinate and weight of reference DNA (in ng) on the abscissa. Estimate the concentration of DNA in the unknown sample by interpolation.

   If the reading for the unknown DNA solution falls outside that of the standard curve, read the fluorescence of a larger sample or make an appropriate dilution of the sample and repeat the assay.

NOTES

- Binding of Hoechst 33258 is adversely influenced by pH extremes, the presence of detergents near or above their critical micelle concentrations, and salt concentrations above 3 M. If these conditions or reagents are used to prepare the DNA and improbable results are obtained in the fluorometry assay, precipitate an aliquot of the DNA with ethanol, rinse the pellet of nucleic acid in 70% ethanol, dissolve the dried pellet in TE, and repeat the assay.
- If the preparation of test DNA is highly viscous, sampling with standard yellow tips may be so inaccurate that the dilutions of unknown DNA will not track with the standard curve. In this case, the best solution is to withdraw two samples (10–20 μl) with an automatic pipettor equipped with a cut-off yellow tip. Each sample is then diluted with ~0.5 ml of TE (pH 8.0) and vortexed vigorously for 1–2 minutes. Different amounts of the diluted samples can then be transferred to the individual tubes containing diluted dye solution. The results obtained from the two sets of samples should be consistent.
- Use scissors or a dog nail clipper (e.g., Fisher) to generate cut-off yellow tips. Alternatively, the tips can be cut with a sharp razor blade. Sterilize the cut-off tips before use, either by autoclaving or by immersion in 70% alcohol for 2 minutes followed by drying in air. Presterilized, purpose-made wide-bore tips can be purchased from a number of commercial companies (e.g., Bio-Rad).

Quantitation of Double-stranded DNA Using Ethidium Bromide

Sometimes there is not sufficient DNA (<250 ng/ml) to assay spectrophotometrically, or the DNA may be heavily contaminated with other substances that absorb UV irradiation and therefore impede accurate analysis. A rapid way to estimate the amount of DNA in such samples is to use the UV-induced fluorescence emitted by ethidium bromide <!> molecules intercalated into the DNA. Because the amount of fluorescence is proportional to the total mass of DNA, the quantity of DNA in the sample can be estimated by comparing the fluorescent yield of the sample with that of a series of standards. As little as 1–5 ng of DNA can be detected by this method. For more information on ethidium bromide, please see Appendix 9 and Chapter 5, Protocol 2.

Saran Wrap Method Using Ethidium Bromide or SYBR Gold

1. Stretch a sheet of Saran Wrap over a UV transilluminator or over a sheet of black paper.
2. Spot 1–5 µl of the DNA sample onto the Saran Wrap.
3. Spot equal volumes of a series of DNA concentration standards (0.1, 2.5, 5, 10, and 20 µg/ml) in an ordered array on the Saran Wrap.

   The standard DNA solutions should contain a single species of DNA approximately the same size as the expected size of the unknown DNA. The DNA standards are stable for many months when stored at –20°C.

4. Add to each spot an equal volume of TE (pH 7.6) containing 2 µg/ml ethidium bromide or an equal volume of a 1:5000 dilution of dimethylsulfoxide <!> (DMSO)/SYBR Gold <!> stock. Mix by pipetting up and down with a micropipette.
5. Photograph the spots using short-wavelength UV illumination for ethidium bromide, or 300-nm transillumination for SYBR Gold (please see Chapter 5, Protocol 2). Estimate the concentration of DNA by comparing the intensity of fluorescence in the sample with that of the standard solutions.

Agarose Plate Method

Contaminants that may be present in the DNA sample can either contribute to or quench the fluorescence. To avoid these problems, the DNA samples and standards can be spotted onto the surface of a 1% agarose slab gel containing ethidium bromide (0.5 µg/ml). Allow the gel to stand at room temperature for a few hours so that small contaminating molecules have a chance to diffuse away. Photograph the gel as described in Chapter 5.

Minigel Method

Electrophoresis through minigels (please see Chapter 5) provides a rapid and convenient way to measure the quantity of DNA and to analyze its physical state at the same time. This is the method of choice if there is a possibility that the samples may contain significant quantities of RNA.

1. Mix 2 µl of the DNA sample with 0.4 µl of Gel-loading buffer IV (bromophenol blue only; please see Table A1-6 in Appendix 1) and load the solution into a slot in a 0.8% agarose minigel containing ethidium bromide (0.5 µg/ml).

   SYBR Gold is too expensive to use routinely in this technique.

2. Mix 2 µl of each of a series of standard DNA solutions (0, 2.5, 5, 10, 20, 30, 40, and 50 µg/ml) with 0.4 µl of Gel-loading buffer IV. Load the samples into the wells of the gel.

   The standard DNA solutions should contain a single species of DNA approximately the same size as the expected size of the unknown DNA. The DNA standards are stable for many months when stored at –20°C.

3. Carry out electrophoresis until the bromophenol blue has migrated ~1–2 cm.
4. Destain the gel by immersing it for 5 minutes in electrophoresis buffer containing 0.01 M $MgCl_2$.
5. Photograph the gel using short-wavelength UV irradiation (please see Chapter 5). Compare the intensity of fluorescence of the unknown DNA with that of the DNA standards and estimate the quantity of DNA in the sample.

MEASUREMENT OF RADIOACTIVITY IN NUCLEIC ACIDS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Radioactive isotopes <!> are used as tracers to monitor the progress of many reactions used to synthesize DNA and RNA. To calculate the efficiency of such reactions, it is necessary to measure accurately the proportion of the radioactive precursor incorporated into the desired product. This goal can be achieved by two methods: (1) differential precipitation of the nucleic acid products with trichloroacetic acid <!> (TCA) and (2) differential adsorption of the products onto positively charged surfaces (e.g., DE-81 paper).

Precipitation of Nucleic Acids with Trichloroacetic Acid

1. Use a soft-lead pencil to label the appropriate number of Whatman GF/C glass fiber filters (2.4-cm diameter). Impale each of the filters on a pin stuck into a polystyrene support.
2. Spot an accurately known volume (up to 5 µl) of each sample to be assayed on the center of each of two labeled filters.

   One of the filters is used to measure the total amount of radioactivity in the reaction (i.e., acid-soluble and acid-precipitable radioactivity). The other filter is used to measure only the acid-precipitable radioactivity. Under the conditions described, DNA and RNA molecules >50 nucleotides long will be precipitated on the surface of the filter.
3. Store the filters at room temperature until all of the fluid has evaporated. This process can be accelerated by using a heat lamp, although this is not usually necessary.
4. Use blunt-end forceps (e.g., Millipore forceps) to transfer one of each pair of filters to a beaker containing 200–300 ml of ice-cold 5% TCA and 20 mM sodium pyrophosphate. Swirl the filters in the acid solution for 2 minutes, and then transfer them to a fresh beaker containing the same volume of the ice-cold 5% TCA/20 mM sodium pyrophosphate mixture. Repeat the washing two more times.

   During washing, the unincorporated nucleotide precursors are eluted from the filters and the radioactive nucleic acids are fixed to them.

   Commercially available, vacuum-driven filtration manifolds that hold up to 24 filters may also be used to wash the filters.
5. Transfer the washed filters to a beaker containing 70% ethanol and allow them to remain there briefly. Then dry the filters either at room temperature or under a heat lamp.
6. Insert each of the filters (washed and unwashed) into a scintillation vial. Measure the amount of radioactivity on each filter.

   ^{32}P can be detected on dry filters by Cerenkov counting (in the ^{3}H channel of a liquid scintillation counter). The efficiency with which Cerenkov radiation can be measured varies from instrument to instrument and also depends on the geometry of the scintillation vials and the amount of H_2O remaining in the filters. With dry filters, the efficiency of Cerenkov counting is ~25% (one radioactive decay in four can be detected). Alternatively, ^{32}P can be measured with 100% efficiency by adding a few milliliters of toluene-based scintillation fluid to the dried filters and counting in the ^{32}P channel of the liquid scintillation counter.

   To measure other isotopes (^{3}H, ^{14}C, ^{35}S, ^{33}P, etc.), it is essential to use toluene-based scintillation fluid and the appropriate channel of a liquid scintillation counter. The efficiency of counting these isotopes varies from counter to counter and should be determined for each instrument.

7. Compare the amount of radioactivity on the unwashed filter with the amount on the washed filter, and then calculate the proportion of the precursor that has been incorporated as described in the panel below.

Adsorption to DE-81 Filters

DE-81 filters are positively charged and strongly adsorb and retain nucleic acids, including oligonucleotides that are too small to be precipitated efficiently with TCA. Unincorporated nucleotides stick less tightly to the filters and can be removed by washing the filter extensively in sodium phosphate. The procedure is essentially identical to that described for precipitation of nucleic acids by TCA, except that the DE-81 filters are washed in 0.5 M Na_2HPO_4 (pH 7.0) instead of TCA/sodium pyrophosphate.

CALCULATION OF THE SPECIFIC ACTIVITY OF A RADIOLABELED PROBE

To calculate the specific activity of a radiolabeled probe in dpm/μg, use the following equation:

$$\text{specific activity} = \frac{L\,(2.2 \times 10^9)\,(PI)}{m + [(1.3 \times 10^3)\,(PI)\,(L/S)]}$$

where

- L = input radioactive label (μCi)
- PI = proportion of the precursor that has been incorporated (cpm in washed filter/cpm in unwashed filter, please see above)
- m = mass of template DNA (ng)
- S = specific activity of input label (μCi/nmole)

The numerator of this equation is the product of three terms: the total dpm in the reaction [$(L)(2.2 \times 10^6$ dpm/μCi)]; the proportion of these dpm which were incorporated (PI); and a factor to convert the final value for specific activity from dpm/ng to dpm/μg (10^3).

The denominator represents the total mass of DNA (in ng) at the end of the reaction, equal to the starting mass (m) plus the mass (in ng) synthesized during the reaction. The latter is calculated from the number of nanomoles of dNMP incorporated [$(PI)(L/S)$] multiplied by four times the average molecular mass of the four dNMPs (4×325 ng/nmole = 1.3×10^3 ng/nmole).

EXAMPLE:

In a random priming reaction in which 50% of the radioactivity has been incorporated into TCA-precipitable material, from a starting reaction containing 25 ng of template DNA and 50 μCi of radiolabeled dNTP with a specific activity of 3000 Ci/mmole: L = 50 μCi, PI = 0.5, m = 25 ng, and S = 3000 Ci/mmole.

$$\text{specific activity of the probe} = \frac{50\,(2.2 \times 10^9)(0.5)}{25 + [(1.3 \times 10^3)(0.5)(50/3000)]}$$

$$= 1.5 \times 10^9 \text{ dpm/μg}$$

DECONTAMINATION OF SOLUTIONS CONTAINING ETHIDIUM BROMIDE

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Removing Ethidium Bromide from DNA

The reaction between ethidium bromide <!> and DNA is reversible (Waring 1965), but the dissociation of the complex is very slow and is measured in days rather than minutes or hours. For practical purposes, dissociation is achieved by passing the complex through a small column packed with a cation-exchange resin such as Dowex AG 50W-X8 (Waring 1965; Radloff et al. 1967) or by extracting with organic solvents such as isopropanol (Cozzarelli et al. 1968) or *n*-butanol <!> (Wang 1969). The former method has been shown to result in the removal of ethidium bromide to a binding ratio below that detectable by fluorescence, a molar ratio of dye:nucleic acid of 1:4000 (Radloff et al. 1967).

Disposing of Ethidium Bromide

Ethidium bromide itself is not highly mutagenic, but it is metabolized by microsomal enzymes to compound(s) that are moderately mutagenic in yeast and *Salmonella typhimurium* (Mahler and Bastos 1974; McCann et al. 1975; MacGregor and Johnson 1977; Singer et al. 1999). A number of methods have been described to decontaminate solutions and surfaces that contain or have been exposed to ethidium bromide. The concentration of ethidium bromide in solution may be reduced to <0.5 µg/ml with activated charcoal, which can then be incinerated (Menozzi et al. 1990). Alternatively, ethidium bromide can be degraded by treatment with sodium nitrite <!> and hypophosphorous acid <!> (Lunn and Sansone 1987).

Decontamination of Concentrated Solutions of Ethidium Bromide
(solutions containing >0.5 mg/ml)

Method 1

This method (Lunn and Sansone 1987) reduces the mutagenic activity of ethidium bromide in the *Salmonella*/microsome assay by ~200-fold.

1. Add sufficient H_2O to reduce the concentration of ethidium bromide to <0.5 mg/ml.
2. To the resulting solution, add 0.2 volume of fresh 5% hypophosphorous acid and 0.12 volume of fresh 0.5 M sodium nitrite. Mix carefully.

   ▲ **IMPORTANT** Make sure that the pH of the solution is <3.0.

   Hypophosphorous acid is usually supplied as a 50% solution, which is corrosive and must be handled with care. Freshly dilute the acid immediately before use.

   Sodium nitrite solution (0.5 M) should be freshly prepared by dissolving 34.5 g of sodium nitrite in H_2O to a final volume of 500 ml.
3. After incubation for 24 hours at room temperature, add a large excess of 1 M sodium bicarbonate. The solution may now be discarded.

Method 2

This method (Quillardet and Hofnung 1988) reduces the mutagenic activity of ethidium bromide in the *Salmonella*/microsome assay by ~3000-fold. However, there are reports (Lunn and Sansone 1987) of mutagenic activity in occasional batches of "blanks" treated with the decontaminating solutions.

1. Add sufficient H_2O to reduce the concentration of ethidium bromide to <0.5 mg/ml.
2. Add 1 volume of 0.5 M $KMnO_4$ <!>. Mix carefully, and then add 1 volume of 2.5 N HCl. Mix carefully, and allow the solution to stand at room temperature for several hours.
3. Add 1 volume of 2.5 N NaOH <!>. Mix carefully, and then discard the solution.

Decontamination of Dilute Solutions of Ethidium Bromide
(e.g., electrophoresis buffer containing 0.5 μg/ml ethidium bromide)

Method 1

The following method is from Lunn and Sansone (1987).

1. Add 2.9 g of Amberlite XAD-16 (Sigma-Aldrich) for each 100 ml of solution. Amberlite XAD-16 is a nonionic, polymeric absorbent.
2. Store the solution for 12 hours at room temperature, shaking it intermittently.
3. Filter the solution through a Whatman No. 1 filter, and discard the filtrate.
4. Seal the filter and Amberlite resin in a plastic bag, and dispose of the bag in the hazardous waste.

Method 2

The following method is from Bensaude (1988).

1. Add 100 mg of powdered activated charcoal for each 100 ml of solution.
2. Store the solution for 1 hour at room temperature, shaking it intermittently.
3. Filter the solution through a Whatman No. 1 filter, and discard the filtrate.
4. Seal the filter and activated charcoal in a plastic bag, and dispose of the bag in the hazardous waste.

NOTES

- Treatment of dilute solutions of ethidium bromide with hypochlorite (bleach) is not recommended as a method of decontamination. Such treatment reduces the mutagenic activity of ethidium bromide in the *Salmonella*/microsome assay by ~1000-fold, but it converts the dye into a compound that is mutagenic in the absence of microsomes (Quillardet and Hofnung 1988).
- Ethidium bromide decomposes at 262°C and is unlikely to be hazardous after incineration under standard conditions.
- Slurries of Amberlite XAD-16 or activated charcoal can be used to decontaminate surfaces that become contaminated by ethidium bromide.

Commercial Decontamination Kits

Several commercial companies sell devices to extract ethidium bromide from solutions with the minimum of fuss and bother. These devices include the EtBr Green Bag (Q•BIOgene) and the Eliminator Dye Removal System (Stratagene).

GEL-FILTRATION CHROMATOGRAPHY

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

This technique, which employs gel filtration to separate high-molecular weight DNA from smaller molecules, is used most often to separate unincorporated labeled dNTPs from DNA that has been radiolabeled. However, it is also used at several stages during the synthesis of double-stranded cDNA, during addition of linkers to blunt-ended DNA, to remove oligonucleotide primers from polymerase chain reaction (PCR), and, in general, whenever it is necessary to change the composition of the buffer in which DNA is dissolved.

Two methods are available: conventional column chromatography, which is used when it is necessary to collect fractions that contain components of different sizes, and centrifugation through gel matrices packed in disposable syringes, which is a rapid method used to free DNA from smaller molecules. The two most commonly used gel matrices are Sephadex and Bio-Gel, both of which are available in several porosities. Sephadex G-50 and Bio-Gel P-60 are ideal for purifying DNA larger than 80 nucleotides in length. Smaller molecules are retained in the pores of the gel, whereas the larger DNA is excluded and passes directly through the column. Bio-Gel P-2 can be used to separate oligonucleotides from phosphate ions or dNTPs. Bio-Gel is supplied in the form of a gel and need only be equilibrated in running buffer before use. Sephadex is supplied as a powder that must be hydrated before use.

Preparation of Sephadex

1. Slowly add Sephadex of the desired grade to distilled sterile H_2O in a 500-ml beaker or bottle (10 g of Sephadex G-50 granules yields 160 ml of slurry). Wash the swollen resin with distilled sterile H_2O several times to remove soluble dextran, which can create problems by precipitating during ethanol precipitation.
2. Equilibrate the resin in TE (pH 7.6), autoclave at 10 psi (0.70 kg/cm^2) for 15 minutes, and store at room temperature.

Column Chromatography

1. Prepare Sephadex or Bio-Gel columns in disposable 5-ml borosilicate glass pipettes or Pasteur pipettes plugged with a small amount of sterile glass wool. Use a long, narrow pipette (e.g., a disposable 1-ml plastic pipette) to push the wool to the bottom of the glass or Pasteur pipette.
2. Use a Pasteur pipette to fill the column with a slurry of the Sephadex or Bio-Gel, taking care to avoid producing bubbles. There is no need to close the bottom of the column. Keep adding gel until it packs to a level 1 cm below the top of the column. Wash the gel with several volumes of 1x TEN buffer (pH 8.0) (please see Appendix 1).
3. Apply the DNA sample (in a volume of 200 µl or less) to the top of the gel. Wash out the sample tube with ~100 µl of 1x TEN buffer, and load the washing on the column as soon as the DNA sample has entered the gel. When the washing has entered the gel, immediately fill the column with 1x TEN buffer.

   ▲ **WARNING** Columns used to separate radiolabeled DNA from radioactive precursors should be run behind Lucite screens to protect against exposure to radioactivity.

4. Immediately start to collect fractions (~200 µl) in microfuge tubes.

   If the DNA is labeled with ^{32}P **<!>**, measure the radioactivity in each of the tubes by using either a hand-held minimonitor or by Cerenkov counting in a liquid scintillation counter. Add more 1x TEN buffer to the top of the gel as required from time to time.

   The DNA will be excluded from the gel and will be found in the void volume (usually 30% of the total column volume). The leading peak of radioactivity therefore consists of nucleotides incorporated into DNA, and the trailing peak consists of unincorporated [^{32}P]dNTPs.

5. Pool the radioactive fractions in the leading peak and store at –20°C.

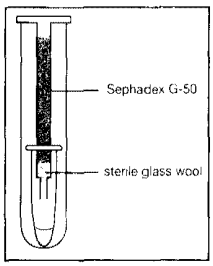
NOTES

Instead of collecting individual fractions, it is possible with practice to follow the progress of the incorporated and unincorporated [^{32}P]dNTPs down the column using a hand-held minimonitor.

- Collect the leading peak into a sterile polypropylene tube as it elutes from the column.
- Clamp the bottom of the column and disconnect the buffer reservoir.
- Discard the column in the radioactive waste.

Spun-column Chromatography

This method is used to separate DNA, which passes through the gel-filtration matrix, from lower-molecular-weight substances that are retained on the column. Spun-column chromatography is particularly useful when separating labeled DNA from radioactive precursors. However, it is also used extensively for other purposes, for example, to remove unwanted nucleotide primers or double-stranded linkers, to change the buffer in which small amounts of DNA are dissolved, or to free crude preparations of minipreparations of plasmid or bacteriophage DNA from inhibitors that prevent cleavage by restriction enzymes. Several samples of DNA can be handled simultaneously. In this respect, spun-column chromatography is much superior to conventional column chromatography.



1. Plug the bottom of a 1-ml disposable syringe with a small amount of sterile glass wool. This is best accomplished by using the barrel of the syringe to tamp the glass wool in place.

2. Fill the syringe with Sephadex G-50 or Bio-Gel P-60, equilibrated in 1x TEN buffer (pH 8.0) (please see Appendix 1). Start the buffer flowing by tapping the side of the syringe barrel several times. Keep adding more resin until the syringe is completely full.

3. Insert the syringe into a 15-ml disposable plastic tube. Centrifuge at 1600*g* for 4 minutes at room temperature in a swinging-bucket rotor in a bench-top centrifuge. Do not be alarmed by the appearance of the column. The resin packs down and becomes partially dehydrated during centrifugation. Continue to add more resin and recentrifuge until the volume of the packed column is ~0.9 ml and remains unchanged after centrifugation.

4. Add 0.1 ml of 1x TEN buffer to the columns, and recentrifuge as in Step 3.

5. Repeat Step 4 twice more.

   Spun columns may be stored at this stage if desired. Several spun columns can be prepared simultaneously and stored at 4°C for periods of 1 month or more before being used. Fill the syringes with 1x TEN buffer and wrap Parafilm around them to prevent evaporation. Store the columns upright at 4°C. Spun columns stored in this way should be washed once with sterile 1x TEN buffer as described in Step 4 just before they are used.

6. Apply the DNA sample to the column in a total volume of 0.1 ml (use 1× TEN buffer to make up the volume). Place the spun column in a fresh disposable tube containing a decapped microfuge tube (please see figure above).

7. Centrifuge again as in Step 3, collecting the effluent from the bottom of the syringe (~100 µl) into the decapped microfuge tube.

8. Remove the syringe, which will contain unincorporated radiolabeled dNTPs or other small components. Using forceps, carefully recover the decapped microfuge tube, which contains the eluted DNA, and transfer its contents to a capped, labeled microfuge tube.

   A rough estimate of the proportion of radioactivity that has been incorporated into nucleic acid may be obtained by holding the syringe and the eluted DNA to a hand-held minimonitor.

9. If the syringe is radioactive **<!>**, carefully discard it in the radioactive waste. Store the eluted DNA at –20°C until needed.

NOTE

- Not all resins are suitable for spun-column centrifugation: DEAE-Sephacel forms an impermeable lump during centrifugation, and the larger grades of Sephadex (G-100 and up) cannot be used because the beads are crushed by centrifugation. If a coarser-sieving resin is required, use Sepharose CL-4B.

SEPARATION OF SINGLE-STRANDED AND DOUBLE-STRANDED DNAs BY HYDROXYAPATITE CHROMATOGRAPHY

Nucleic acids bind to hydroxyapatite by virtue of interactions between the phosphate groups of the polynucleotide backbone and calcium residues in the matrix. Bound nucleic acids can be eluted in phosphate buffers. This step is usually carried out at 60°C, although there is no good reason to do so since the adsorption and elution profiles of nucleic acids are indistinguishable between 25°C and 60°C. The affinity of nucleic acids is determined by the number of phosphate groups that are available to bind to the matrix. Both single- and double-stranded nucleic acids bind to hydroxyapatite in 0.05 M sodium phosphate (pH 6.8). Double-stranded molecules, with their well-ordered and evenly spaced sets of phosphate residues, make many regular contacts with the matrix and therefore require high concentrations of phosphate (0.4 M) for elution. Single-stranded molecules are more disordered and a smaller proportion of their phosphate residues are available for contact with the matrix. Hence, single-stranded DNA is eluted in lower concentrations of phosphate (~0.12 M). Partial duplexes and DNA-RNA hybrid molecules elute at intermediate concentrations.

Nucleic acids are often eluted in such large volumes that they need to be concentrated before they can be used. Ethanol precipitation must be avoided until the phosphate ions have been removed from the solution. This is best achieved by concentrating the eluate by extraction with isobutanol and then removing the salt by chromatography through G-50 Sephadex columns.

Batches of hydroxyapatite vary slightly in their characteristics, and it is therefore important to carry out preliminary experiments to determine the optimal phosphate concentrations for elution of single- and double-stranded nucleic acids. This can be accomplished by setting up two hydroxyapatite columns as described below. One of the columns is loaded with a small amount (~10^5 cpm) of ^{32}P-labeled DNA that has been denatured by boiling for 10 minutes in TE (pH 7.6). The other column receives an equal amount (~10^5 cpm) of ^{32}P-labeled native DNA. Each of the columns is washed with a series of buffers containing increasing concentrations of sodium phosphate (0.01, 0.12, 0.16, 0.20, 0.24, 0.28, 0.32, 0.36, and 0.40 M). The amount of radioactivity eluting at each phosphate concentration is then measured in a liquid scintillation counter (either by Cerenkov counting or in a water-miscible fluor). Usually, single-stranded DNA elutes in 0.14–0.16 M sodium phosphate (pH 6.8), whereas double-stranded DNA is not removed from the column until the phosphate concentration exceeds 0.36 M. In the protocol that follows, SS buffer contains the phosphate concentration that is optimal for elution of single-stranded DNA; DS buffer contains the concentration that is optimal for elution of double-stranded DNA.

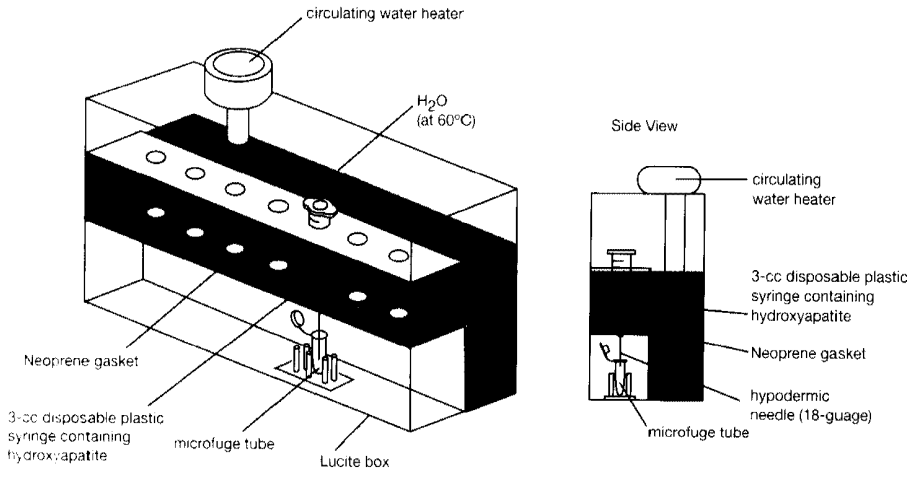


FIGURE A8-4 Apparatus for Hydroxyapatite Chromatography

HYDROXYAPATITE CHROMATOGRAPHY

Hydroxyapatite ($Ca_5[PO_4]_3OH$), the most stable of all calcium phosphates precipitated from aqueous solution, was originally developed as a matrix for protein chromatography by Tiselius et al. (1956). Semenza (1957), working in the Tiselius laboratory, and Main and Cole (1957; Main et al. 1959) seem to have been the first to use hydroxyapatite to fractionate nucleic acids. However, it was work in Bernardi's laboratory (Bernardi 1961, 1965; Bernardi and Timasheff 1961) that led to the widespread and successful use of hydroxyapatite columns to separate single-stranded from double-stranded nucleic acids and to fractionate complex nucleic acids by thermal elution according to their content of (G+C) (Martinson and Wagenaar 1974). During the 1960s and 1970s, hydroxyapatite chromatography became the standard method to investigate the reassociation kinetics of DNAs from many different sources (Britten and Kohne 1966, 1967, 1968; Britten et al. 1974), to construct transcription maps (Sambrook et al. 1972), and to measure the copy number of specific sequences in complex genomes (Gelb et al. 1971; Gallimore et al. 1974). More recently, hydroxyapatite has been used in the preparation of cDNA for subtractive cloning (Davis 1986), as well as for a variety of more menial chores, for example, the isolation of DNA from low-melting-temperature agarose (Wilkie and Cortini 1976) and the removal of contaminants from DNA preparations (Smith 1980). During the decade that led to the development of molecular cloning, hydroxyapatite chromatography was in daily use in many laboratories for a variety of manipulations involving nucleic acids both preparative and analytical. However, there are now better ways to carry out almost all of these tasks, and hydroxyapatite chromatography has all but disappeared from the standard repertoire of laboratory techniques..

1. Determine the concentrations of sodium phosphate that are optimal for elution of single-stranded DNA (SS buffer) and double-stranded DNA (DS buffer) as described above.
2. Prepare SS and DS buffers by diluting 2 M sodium phosphate (pH 6.8).

   2 M sodium phosphate (pH 6.8) is made by mixing equal volumes of 2 M NaH_2PO_4 and 2 M Na_2HPO_4.

3. Suspend the hydroxyapatite powder (Bio-Gel HTP) in 0.01 M sodium phosphate (pH 6.8). Approximately 0.5 ml of packed Bio-Gel HTP is required for each column.

   Bio-Gel HTP has a capacity of 100–200 µg of native DNA/ml of bed volume.

4. Prepare the hydroxyapatite columns in disposable 3-cc plastic syringes as follows:
   a. Remove the barrel from the syringe.
   b. Use the wide end of a Pasteur pipette to push a Whatman GF/C filter to the bottom of the syringe. The filter should completely cover the bottom of the syringe.
   c. Attach an 18-gauge hypodermic needle to the syringe.
   d. Insert the syringe through a Neoprene gasket in the apparatus shown in Figure A8-4.
   e. Use a Pasteur pipette to add enough of the slurry of hydroxyapatite to the syringe to form a column whose packed volume is 0.5–1.0 ml. Wash the column with several volumes of 0.01 M sodium phosphate (pH 6.8). The column will not be harmed if it runs dry; simply rewet before use.
   f. Seal the bottom of the column by sticking a small Neoprene stopper on the end of the hypodermic needle.
5. Load the sample containing the nucleic acid onto the column.

   The concentration of phosphate in the sample should be less than 0.08 M.

6. Remove the Neoprene stopper, and allow the sample to flow through the column.

   There is usually no need to collect and save the loading buffer that elutes from the column.

7. Wash the column with 3 ml of 0.01 M sodium phosphate.

8. Seal the bottom of the column with a Neoprene stopper, and add 1 column volume of SS buffer preheated to 60°C.
9. After 5 minutes, remove the Neoprene stopper and collect the eluate in microfuge tubes. No more than 0.5 ml should be collected in any one microfuge tube. Repeat Steps 8 and 9 two more times.
10. Seal the bottom of the column with a Neoprene stopper, and add 1 column volume of DS buffer preheated to 60°C.
11. After 5 minutes, remove the Neoprene stopper and collect the eluate in microfuge tubes. No more than 0.5 ml should be collected in any one microfuge tube. Repeat Steps 10 and 11 two more times.
12. Allow the eluates to cool to room temperature. DNA can then be extracted as follows:
    a. Add an equal volume of isobutanol to each of the tubes containing the desired nucleic acids.
    b. Mix the two phases by vortexing, and centrifuge the mixture at maximum speed for 20 seconds at room temperature in a microfuge.
    c. Discard the upper (organic) phase.
    d. Repeat the extraction with isobutanol until the volume of the aqueous phase is 100–125 µl.
    e. Remove salts from the DNA by chromatography on, or centrifugation through, a small column of Sephadex G-50 equilibrated in TE (pH 8.0).
    f. Recover the DNA by precipitation with 2 volumes of ethanol at 0°C.

NOTE

- In molecular cloning, nucleic acids fractionated by hydroxyapatite are usually radiolabeled, and the tubes containing the desired fractions can be easily identified by Cerenkov counting in a liquid scintillation counter.

FRAGMENTATION OF DNA

The fragmentation of DNA is often a necessary step preceding library construction or subcloning for DNA sequencing. Fragmentation is typically achieved by physical or enzymatic methods; the most commonly used of these are described in Table A8-6. Although each approach is reasonably successful for generating a range of fragments from a large contiguous segment of DNA, each has its particular limitations. Because they are independent of sequence composition, physical methods for shearing DNA typically result in more uniform and random disruption of the target DNA than enzymatic methods. In particular, methods involving hydrodynamic shearing due to physical stress induced by sonication or nebulization produce collections of appropriately random fragments. The variety in lengths of these fragments is quite large, however, and their use usually requires a subsequent size selection step to narrow the range of fragments that are acceptable for cloning or sequencing.

AUTOMATED SHEARING

During the last few years, a method for hydrodynamic shearing, initially based on the use of high-performance liquid chromatography (HPLC) and called the "point-sink" flow system, has become increasingly refined and, finally, automated (Oefner et al. 1996; Thorstenson et al. 1998). In the point-sink system, an HPLC pump is used to apply pressure to the DNA sample, thereby forcing it through tubing of very small diameter. In the automated process known as HydroShear (commercially available from Gene Machines), a sample of DNA is repeatedly passed through a small hole until the sample is fragmented to products of a certain size. The final size distribution is determined by both the flow rate of the sample and the size of the opening, parameters that are controlled and monitored by the automated system. At any given setting, DNA fragments larger than a certain length are broken, whereas shorter fragments are unaffected by passage through the opening. The resulting sheared products therefore have a narrow size distribution: Typically 90% of the sheared DNA falls within a twofold size range of the target length. It is reasonable to expect that libraries constructed from these DNA fragments are likely to be of higher quality than those made using one of the "old-fashioned" ways. They will certainly contain clones of more uniform size, and possibly may be more comprehensive in their coverage of the genome. However, libraries constructed from sonicated or hydrodynamically sheared DNA, although imperfect, are certainly workable. Perfectionists will feel that the machine is necessary; pragmatists will find it merely desirable.

TABLE A8-6 Hydrodynamic Shearing Methods Used to Fragment DNA

METHOD	PROS AND CONS	REFERENCES
Sonication	Requires relatively large amounts of DNA (10–100 μg); fragments of DNA distributed over a broad range of sizes; only a small fraction of the fragments are of a length suitable for cloning and sequencing; requires ligation of DNA before sonication and end-repair afterward; DNA may be damaged by hydroxyl radicals generated during cavitation (McKee et al. 1977).	Deininger (1983)
Nebulization	Easy and quick; requires only small amounts of DNA (0.5–5 μg) and large volumes of DNA solution; no preference for AT-rich regions; size of fragments easily controlled by altering the pressure of the gas blowing through the nebulizer; fragments of DNA distributed over a narrow range of sizes (700–1330 bp); requires ligation of DNA before nebulization and end-repair afterward.	Bodenteich et al. (1994); Hengen (1997)
Circulation through an HPLC pump	Requires expensive apparatus, ligation of DNA before sonication, and 1–100 μg of DNA; fragments of DNA distributed over a narrow range of sizes that can be adjusted by changing the flow rate; end-repair of fragments before cloning not necessary.	Oefner et al. (1996); Thorstenson et al. (1998)
Passage through the orifice of a 28-gauge hypodermic needle	Cheap, easy, and requires only small amounts of DNA; however, the fragments are a little larger (1.5–2.0 kb) than required for shotgun sequencing; requires ligation of DNA before cleavage and end-repair afterward.	Davidson (1959, 1960); Schriefer et al. (1990)

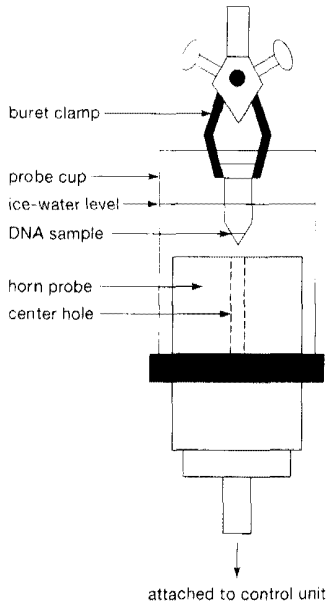


FIGURE 8-5 Cup Horn Sonicator for Random Fragmentation of DNA

The cup horn attachment for a Heat Systems sonicator is depicted with a sample tube in place. The cup horn unit, which contains a large horn probe, is attached to the sonicator control unit and filled with ice water before the sample is sonicated. Typically, the sample tube is held in place from above by using a buret clamp and a ring stand. Alternatively, a tube holder can be fabricated from 1/4-inch plastic and used to hold up to eight tubes for simultaneous processing. (Redrawn, with permission, from Birren et al. 1997.)

Sonication

DNA samples are subjected to hydrodynamic shearing by exposure to brief pulses of sonication. DNA that has been sonicated for excessive periods of time is extremely difficult to clone, perhaps because of damage caused by free radicals generated by cavitation. Most sonicators will not shear DNA to a size smaller than 300–500 bp, and it is tempting to continue sonication until the entire population of DNA fragments has been reduced to this size. However, the yield of subclones is usually greater if sonication is stopped when the fragments of target DNA first reach a size of ~700 bp.

Calibration of the Sonicator

1. Transfer 10 µg of bacteriophage λ DNA (or some other large DNA of defined molecular weight) to a microfuge tube. Add TE (pH 7.6) to a final volume of 150 µl. Distribute 25-µl aliquots of the DNA solution into five microfuge tubes. Store the remaining DNA in an ice bath.

2. Fill the cup horn of the sonicator (Figure A8-5) with a mixture of ice and H_2O. Clamp the five microfuge tubes containing the bacteriophage λ DNA just above the probe.

   If the temperature of the sample rises during sonication, the speed and vigor of fragmentation will increase. It is therefore important to mix the ice and H_2O after each burst of sonication and to add fresh ice when necessary.

3. Sonicate at maximum output and continuous power for bursts of 10 seconds. After each burst, remove one of the microfuge tubes from the sonicator and store it on ice.

4. After sonication is completed, analyze the size of the DNA fragments in each stored sample by electrophoresis through a 1.4% agarose gel. Use suitable standards for molecular-weight markers (please see Appendix 6, Figure A6-4).

5. Stain and photograph the gel, and then estimate the amount of sonication required to produce a reasonable yield of fragments of the desired size (500 bp to 2 kb).

   The times of sonication given in this method are for a cup horn sonicator with a nominal peak output energy of 475 W. Because the actual output of different sonicators varies widely, it is necessary to calibrate each instrument. A probe-type sonicator can be used with a microtip if the volume of the DNA sample is increased to ~250 µl to accommodate the probe. After sonication is completed, concentrate the DNA by precipitation with 2 volumes of ethanol and dissolve in 25 µl of TE (pH 7.6).

Sonication of Target DNA

6. Sonicate the chosen DNA sample for the length of time estimated to produce a reasonable yield of fragments of the desired size (500 bp to 1 kb). Confirm that the fragmentation has gone according to plan by analyzing an aliquot of the sample (~200 ng) by electrophoresis through a 1.4% agarose gel as described above.

Nebulization

Nebulization is performed by collecting the fine mist created by forcing DNA in solution through a small hole in the nebulizer unit. The size of the fragments is determined chiefly by the speed at which the DNA solution passes through the hole, the viscosity of the solution, and the temperature.

Modification of the Nebulizer

1. Modify a nebulizer model number CA-209 (available from CIS-US Inc., Bedford, Massachusetts) by sealing the mouthpiece hole in the top cover with a QS-T plastic cap. Connect a length of Nalgene tubing to the smaller hole. This tubing will be connected to a source of nitrogen gas.

Calibration of the Nebulizer

2. Prepare a sample containing 25 µg of bacteriophage λ DNA or other large DNA of defined size in 500 µl of TE (pH 7.6) containing 25% glycerol. Store the DNA solution in an ice bath for 5 minutes.
3. Place the DNA in the cup of a nebulizer connected to a nitrogen gas source and place the nebulizer in an ice-water bath.

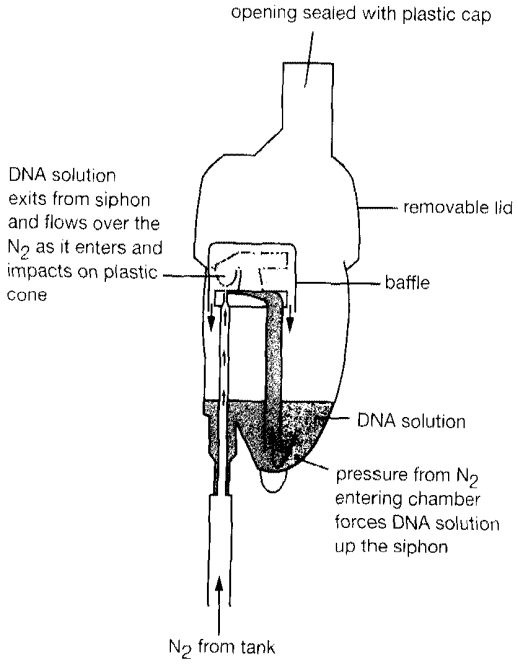


FIGURE A8-6 Nebulizer for Random Fragmentation of DNA

A DNA solution containing glycerol, for viscosity, is placed in the nebulizer. The nebulizer is attached to a nitrogen tank. Pressure from nitrogen entering the chamber siphons the DNA solution from the bottom of the chamber to the top. The solution exits the siphon and impacts on a small plastic cone suspended near the top of the chamber, thus shearing the DNA.

4. Nebulize the DNA sample at 10 psi (69 KPa) for 90 seconds.

   The nebulizer usually leaks a little.

5. After nebulization is completed, collect the DNA by placing the entire unit in a rotor bucket of a bench-top centrifuge fitted with pieces of Styrofoam to cushion. Centrifuge the nebulizer at 2500 rpm for 30 seconds. Analyze the size of the DNA fragments by electrophoresis through a 1.4% agarose gel with appropriate standards, e.g., the fragments generated by digestion of pUC18/19 DNA with *Sau*3AI. For detail of standards, please see Appendix 6, Table A6-4.

6. If the fragments are too large (>1.0 kb) repeat the procedure increasing either the pressure of the nitrogen gas (to 14 psi or 96.5 KPa) or the length of nebulization (2 minutes). Repeat the procedure until conditions are found that produce a reasonable yield of fragments of the desired size (500 bp to 1 kb).

Nebulization and Recovery of the Target DNA

7. Nebulize the target DNA to produce a reasonable yield of fragments of the desired size (500 bp to 1 kb).

8. Confirm that the fragmentation has gone according to plan by analyzing an aliquot of the sample (~200 ng) by electrophoresis through a 1.4% agarose gel as described above.

9. Distribute the remainder of the sample into two microfuge tubes and recover the DNA by precipitation with 3 volumes of ethanol in the presence of 2.5 M ammonium acetate.

10. Centrifuge the DNA sample at maximum speed for 5 minutes in a microfuge, wash the pellet with 0.5 ml of 70% ethanol at room temperature, and centrifuge again. Remove the ethanol and allow the DNA to dry in the air for a few minutes. Dissolve the pellet of DNA in 25 µl of TE (pH 7.6).

CENTRIFUGATION

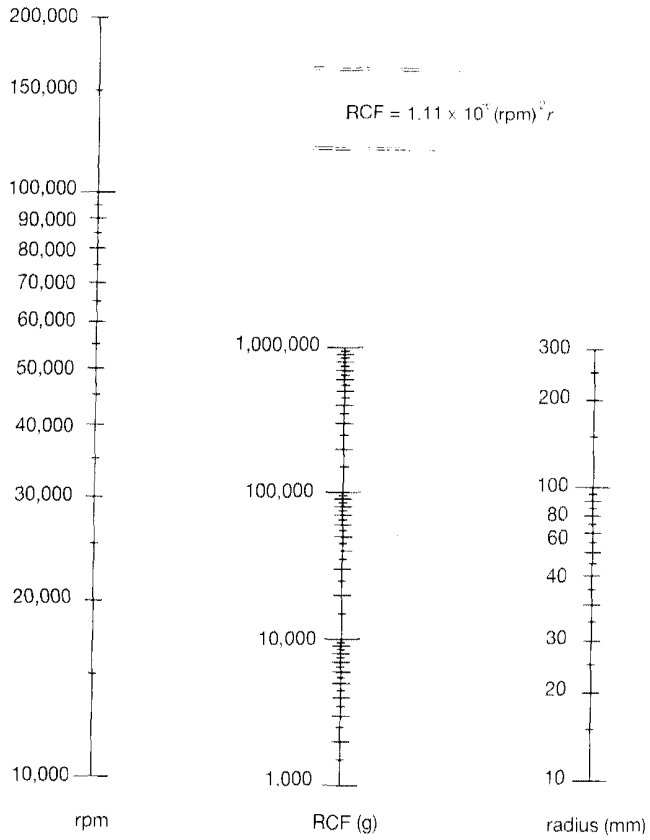


FIGURE A8-7 Nomogram for Conversion of Rotor Speed (rpm) and Relative Centrifugal Force (RCF)

The symbol r represents the radial distance from the center of the rotor to the point for which RCF is required. This is generally equivalent to r_{max} of the rotor. For the example marked, rotor speed = 80,000 rpm and r = 20 mm. RCF can be read as ~145,000g from the middle scale. Using the equation given above, RCF can be calculated as 142,080g. (Figure kindly provided by Siân Curtis.)

TABLE A8-7 Commonly Used Rotors

	R_{MAX}	G_{MAX}	MAX RPM
Sorvall Rotors			
GS3 Sorvall superspeed	151.3	13,700	9,000
GSA Sorvall super superspeed	145.6	27,400	13,000
HB6	146.3	27,617	13,000
AH629 Sorvall ultraspeed	161.0	151,000	29,000
SS-34 Sorvall	107.0	20,500	50,200
Beckman Rotors			
JA17 Beckman	123	17,000	39,800
JA20 Beckman	108	20,000	48,400
SW 28.1 Beckman	171.3	28,000	150,000
Type 50 Beckman	70.1	50,000	196,000
Vti50 Beckman	86.6	50,000	242,000
SW40 Ti Beckman	158.8	40,000	285,000
SW41 Ti Beckman	153.1	41,000	288,000
SW50.1 Beckman	107.3	50,000	300,000
Type 60Ti Beckman	89.9	60,000	362,000
Type 70Ti Beckman	39.5	70,000	504,000

SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS OF PROTEINS

Almost all analytical electrophoreses of proteins are carried out in polyacrylamide gels under conditions that ensure dissociation of the proteins into their individual polypeptide subunits and that minimize aggregation. Most commonly, the strongly anionic detergent SDS is used in combination with a reducing agent and heat to dissociate the proteins before they are loaded onto the gel. The denatured polypeptides bind SDS and become negatively charged. Because the amount of SDS bound is almost always proportional to the molecular weight of the polypeptide and is independent of its sequence, SDS-polypeptide complexes migrate through polyacrylamide gels in accordance with the size of the polypeptide. At saturation, ~1.4 g of detergent is bound per gram of polypeptide. By using markers of known molecular weight, it is possible to estimate the molecular weight of the polypeptide chain(s). Modifications to the polypeptide backbone, such as N- or O-linked glycosylation, however, have a significant impact on the apparent molecular weight. Thus, the apparent molecular weight of glycosylated proteins is not a true reflection of the mass of the polypeptide chain.

In most cases, SDS-polyacrylamide gel electrophoresis is carried out with a discontinuous buffer system in which the buffer in the reservoirs is of a pH and ionic strength different from that of the buffer used to cast the gel. The SDS-polypeptide complexes in the sample that is applied to the gel are swept along by a moving boundary created when an electric current is passed between the electrodes. After migrating through a stacking gel of high porosity, the complexes are deposited in a very thin zone (or stack) on the surface of the resolving gel. The ability of discontinuous buffer systems to concentrate all of the complexes in the sample into a very small volume greatly increases the resolution of SDS-polyacrylamide gels.

The discontinuous buffer system that is most widely used was originally devised by Ornstein (1964) and Davis (1964). The sample and the stacking gel contain Tris-Cl (pH 6.8), the upper and lower buffer reservoirs contain Tris-glycine (pH 8.3), and the resolving gel contains Tris-Cl (pH 8.8). All components of the system contain 0.1% SDS (Laemmli 1970). The chloride ions in the sample and stacking gel form the leading edge of the moving boundary, and the trailing edge is composed of glycine molecules. Between the leading and trailing edges of the moving boundary is a zone of lower conductivity and steeper voltage gradient, which sweeps the polypeptides from the sample and deposits them on the surface of the resolving gel. There the higher pH of the resolving gel favors the ionization of glycine, and the resulting glycine ions migrate through the stacked polypeptides and travel through the resolving gel immediately behind the chloride ions. Freed from the moving boundary, the SDS-polypeptide complexes move through the resolving gel in a zone of uniform voltage and pH and are separated according to size by sieving.

Polyacrylamide gels are composed of chains of polymerized acrylamide that are cross-linked by a bifunctional agent such as *N,N'*-methylene-bis-acrylamide (please see Figure A8-8).

The effective range of separation of SDS-polyacrylamide gels depends on the concentration of polyacrylamide used to cast the gel and on the amount of cross-linking. Polymerization of acrylamide in the absence of cross-linking agents generates viscous solutions that are of no practical use. Cross-links formed from bisacrylamide add rigidity and tensile strength to the gel and form pores through which the SDS-polypeptide complexes must pass. The size of these pores decreases as the bisacrylamide:acrylamide ratio increases, reaching a minimum when the ratio is ~1:20. Most SDS-polyacrylamide gels are cast with a molar ratio of bisacrylamide:acrylamide of 1:29, which has been shown empirically to be capable of resolving polypeptides that differ in size by as little as 3%.

The sieving properties of the gel are determined by the size of the pores, which is a function of the absolute concentrations of acrylamide and bisacrylamide used to cast the gel. Table A8-8 shows the linear range of separation of proteins obtained with gels cast with concentrations of acrylamide that range from 5% to 15%.

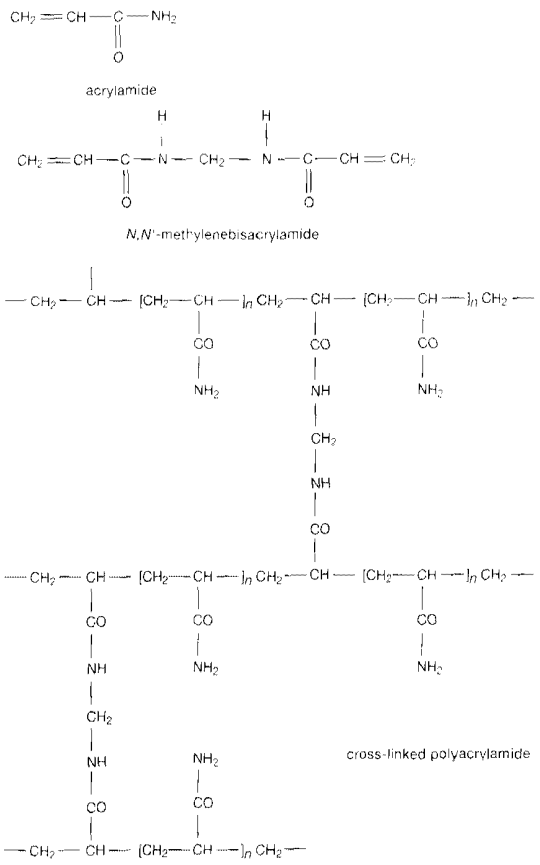


FIGURE A8-8 Chemical Structure of Polyacrylamide

Monomers of acrylamide are polymerized into long chains in a reaction initiated by free radicals. In the presence of *N,N'*-methylenebisacrylamide, these chains become cross-linked to form a gel. The porosity of the resulting gel is determined by the length of chains and degree of cross-linking that occurs during the polymerization reaction.

Reagents

- ***Acrylamide and* N,N′*-methylene-bis-acrylamide:*** Several manufacturers sell electrophoresis-grade acrylamide that is free of contaminating metal ions. A stock solution containing 29% (w/v) acrylamide and 1% (w/v) *N,N'*-methylene-bis-acrylamide should be prepared in deionized warm H_2O (to assist the dissolution of the bisacrylamide). Acrylamide and bisacrylamide are slowly converted during storage to acrylic acid and bisacrylic acid. This deamination reaction is catalyzed by light and alkali. Check that the pH of the solution is 7.0 or less, and store the solution in dark bottles at room temperature. Fresh solutions should be prepared every few months. Note that prepackaged, premixed stock solutions are commercially available (e.g., Life Technologies). These gel systems provide all the components except ammonium persulfate and are certainly convenient but perhaps a bit expensive.

- ***Sodium dodecyl sulfate (SDS):*** Several manufacturers sell special grades of SDS that are sufficiently pure for electrophoresis. Although any one of these will give reproducible results, they are not interchangeable. The pattern of migration of polypeptides may change quite drastically when one manufacturer's SDS is substituted for another's. We recommend exclusive use of one brand of SDS. A 10% (w/v) stock solution should be prepared in deionized H_2O and stored at room temperature. If proteins are to be eluted from the gel for sequencing, electrophoresis-grade SDS should be further purified as described by Hunkapiller et al. (1983).

TABLE A8-8 Effective Range of Separation of SDS-Polyacrylamide Gels

Acrylamide Concentration (%)	Linear Range of Separation (kD)
15	10–43
12	12–60
10	20–80
7.5	36–94
5.0	57–212

Molar ratio of bisacrylamide:acrylamide is 1:29.

- ***Tris buffers for the preparation of resolving and stacking gels:*** It is essential that these buffers be prepared with Tris base. After the Tris base has been dissolved in deionized H_2O, adjust the pH of the solution with HCl as described in Appendix 1. If Tris-Cl or Trizma is used to prepare buffers, the concentration of salt will be too high and polypeptides will migrate anomalously through the gel, yielding extremely diffuse bands.
- ***TEMED (N,N,N′,N′-tetramethylethylenediamine):*** TEMED accelerates the polymerization of acrylamide and bisacrylamide by catalyzing the formation of free radicals from ammonium persulfate. Use the electrophoresis grade sold by several manufacturers. Because TEMED works only as a free base, polymerization is inhibited at low pH.
- ***Ammonium persulfate:*** Ammonium persulfate provides the free radicals that drive polymerization of acrylamide and bisacrylamide. A small amount of a 10% (w/v) stock solution should be prepared in deionized H_2O and stored at 4°C. Ammonium persulfate decomposes slowly, and fresh solutions should be prepared weekly.
- ***Tris-glycine electrophoresis buffer:*** This buffer contains 25 mM Tris base, 250 mM glycine (electrophoresis grade) (pH 8.3), 0.1% SDS.

MATERIALS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Buffers and Solutions

Please see Appendix 1 for components of stock solutions, buffers, and reagents. Dilute stock solutions to the appropriate concentrations.

Acrylamide solutions <!>

Please see Table A8-9 for resolving gel recipes and Table A8-10 for stacking gel recipes.

Protein markers

Standard molecular-weight markers are commercially available (e.g., Life Technologies and Promega).

Protein samples

Samples to be resolved, for example, can be purified protein or cell lysates.

1x SDS gel-loading buffer

50 mM Tris-Cl (pH 6.8)
100 mM dithiothreitol
2% (w/v) SDS (electrophoresis grade)
0.1% bromophenol blue
10% (v/v) glycerol

Store 1x SDS gel-loading buffer lacking dithiothreitol at room temperature. Add dithiothreitol from a 1 M stock just before the buffer is used.

TABLE A8-9 Solutions for Preparing Resolving Gels for Tris-glycine SDS-Polyacrylamide Gel Electrophoresis

⇓ Components / Gel Volume ⇒	Volume (ml) of Components Required to Cast Gels of Indicated Volumes and Concentrations							
	5 ml	10 ml	15 ml	20 ml	25 ml	30 ml	40 ml	50 ml
6% gel								
H_2O	2.6	5.3	7.9	10.6	13.2	15.9	21.2	26.5
30% acrylamide mix <!>	1.0	2.0	3.0	4.0	5.0	6.0	8.0	10.0
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0	6.3	7.5	10.0	12.5
10% SDS	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
10% ammonium persulfate <!>	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
TEMED <!>	0.004	0.008	0.012	0.016	0.02	0.024	0.032	0.04
8% gel								
H_2O	2.3	4.6	6.9	9.3	11.5	13.9	18.5	23.2
30% acrylamide mix <!>	1.3	2.7	4.0	5.3	6.7	8.0	10.7	13.3
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0	6.3	7.5	10.0	12.5
10% SDS	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
10% ammonium persulfate <!>	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
TEMED <!>	0.003	0.006	0.009	0.012	0.015	0.018	0.024	0.03
10% gel								
H_2O	1.9	4.0	5.9	7.9	9.9	11.9	15.9	19.8
30% acrylamide mix <!>	1.7	3.3	5.0	6.7	8.3	10.0	13.3	16.7
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0	6.3	7.5	10.0	12.5
10% SDS	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
10% ammonium persulfate <!>	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
TEMED <!>	0.002	0.004	0.006	0.008	0.01	0.012	0.016	0.02
12% gel								
H_2O	1.6	3.3	4.9	6.6	8.2	9.9	13.2	16.5
30% acrylamide mix <!>	2.0	4.0	6.0	8.0	10.0	12.0	16.0	20.0
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0	6.3	7.5	10.0	12.5
10% SDS	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
10% ammonium persulfate <!>	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
TEMED <!>	0.002	0.004	0.006	0.008	0.01	0.012	0.016	0.02
15% gel								
H_2O	1.1	2.3	3.4	4.6	5.7	6.9	9.2	11.5
30% acrylamide mix <!>	2.5	5.0	7.5	10.0	12.5	15.0	20.0	25.0
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0	6.3	7.5	10.0	12.5
10% SDS	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
10% ammonium persulfate <!>	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
TEMED <!>	0.002	0.004	0.006	0.008	0.01	0.012	0.016	0.02

Modified from Harlow and Lane (1988).

TABLE A8-10 Solutions for Preparing 5% Stacking Gels for Tris-glycine SDS-polyacrylamide Gel Electrophoresis

⇓ Components / Gel Volume ⇒	Volume (ml) of Components Required to Cast Gels of Indicated Volumes							
	1 ml	2 ml	3 ml	4 ml	5 ml	6 ml	8 ml	10 ml
H_2O	0.68	1.4	2.1	2.7	3.4	4.1	5.5	6.8
30% acrylamide mix <!>	0.17	0.33	0.5	0.67	0.83	1.0	1.3	1.7
1.0 M Tris (pH 6.8)	0.13	0.25	0.38	0.5	0.63	0.75	1.0	1.25
10% SDS	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.1
10% ammonium persulfate <!>	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.1
TEMED <!>	0.001	0.002	0.003	0.004	0.005	0.006	0.008	0.01

Modified from Harlow and Lane (1988).

1x Tris-glycine electrophoresis buffer

25 mM Tris
250 mM glycine (electrophoresis grade) (pH 8.3)
0.1% (w/v) SDS

Prepare a 5x stock of electrophoresis buffer by dissolving 15.1 g of Tris base and 94 g of glycine in 900 ml of deionized H_2O. Then add 50 ml of a 10% (w/v) stock solution of electrophoresis-grade SDS and adjust the volume to 1000 ml with H_2O.

Special Equipment

Power supply

A device capable of supplying up to 500 V and 200 mA is needed.

Vertical electrophoresis apparatus

The use of discontinuous buffer systems requires SDS-polyacrylamide electrophoresis to be carried out in vertical gels. Although the basic design of the electrophoresis tanks and plates has changed little since Studier (1973) introduced the system, many small improvements have been incorporated into the apparatuses. Standard vertical as well as minigel vertical electrophoresis systems for separation and blotting of proteins are now sold by many manufacturers (e.g., Life Technologies). Which of these systems to purchase is a matter of personal choice, but it is sensible for a laboratory to use only one type of apparatus. This type of standardization makes it easier to compare the results obtained by different investigators and allows parts of broken apparatuses to be scavenged and reused.

METHOD

Pouring SDS-polyacrylamide Gels

1. Assemble the glass plates according to the manufacturer's instructions.
2. Determine the volume of the gel mold (this information is usually provided by the manufacturer). In an Erlenmeyer flask or disposable plastic tube, prepare the appropriate volume of solution containing the desired concentration of acrylamide for the resolving gel, using the values given in Table A8-9. Mix the components in the order shown. Polymerization will begin as soon as the TEMED has been added. Without delay, swirl the mixture rapidly and proceed to the next step.

   The concentration of ammonium persulfate that we recommend is higher than that used by some investigators. This eliminates the need to rid the acrylamide solution of dissolved oxygen (which retards polymerization) by degassing.

3. Pour the acrylamide solution into the gap between the glass plates. Leave sufficient space for the stacking gel (the length of the teeth of the comb plus 1 cm). Use a Pasteur pipette to carefully overlay the acrylamide solution with 0.1% SDS (for gels containing ~8% acrylamide) or isobutanol (for gels containing ~10% acrylamide). Place the gel in a vertical position at room temperature.

   The overlay prevents oxygen from diffusing into the gel and inhibiting polymerization. Isobutanol dissolves the plastic of some minigel apparatuses.

4. After polymerization is complete (30 minutes), pour off the overlay and wash the top of the gel several times with deionized H_2O to remove any unpolymerized acrylamide. Drain as much fluid as possible from the top of the gel, and then remove any remaining H_2O with the edge of a paper towel.
5. Prepare the stacking gel as follows: In a disposable plastic tube, prepare the appropriate volume of solution containing the desired concentration of acrylamide, using the values given in

Table A8-10. Mix the components in the order shown. Polymerization will begin as soon as the TEMED has been added. Without delay, swirl the mixture rapidly and proceed to the next step.

The concentration of ammonium persulfate is higher than that used by some investigators. This eliminates the need to rid the acrylamide solution of dissolved oxygen (which retards polymerization) by degassing.

6. Pour the stacking gel solution directly onto the surface of the polymerized resolving gel. Immediately insert a clean Teflon comb into the stacking gel solution, being careful to avoid trapping air bubbles. Add more stacking gel solution to fill the spaces of the comb completely. Place the gel in a vertical position at room temperature.

Teflon combs should be cleaned with H_2O and dried with ethanol just before use.

Preparation of Samples and Running the Gel

7. While the stacking gel is polymerizing, prepare the samples in the appropriate volume of 1x SDS gel-loading buffer and heat them to 100°C for 3 minutes to denature the proteins.

Be sure to denature a sample containing marker proteins of known molecular weights. Mixtures of appropriately sized polypeptides are available from commercial sources.

Extremely hydrophobic proteins, such as those containing multiple transmembrane domains, may precipitate or multimerize when boiled for 3 minutes at 100°C. To avoid these pitfalls, heat the samples for 1 hour at a lower temperature (45–55°C) to effect denaturation.

8. After polymerization is complete (30 minutes), remove the Teflon comb carefully. Use a squirt bottle to wash the wells immediately with deionized H_2O to remove any unpolymerized acrylamide. If necessary, straighten the teeth of the stacking gel with a blunt hypodermic needle attached to a syringe. Mount the gel in the electrophoresis apparatus. Add Tris-glycine electrophoresis buffer to the top and bottom reservoirs. Remove any bubbles that become trapped at the bottom of the gel between the glass plates. This is best done with a bent hypodermic needle attached to a syringe.

▲ **IMPORTANT** Do not prerun the gel before loading the samples, since this procedure will destroy the discontinuity of the buffer systems.

9. Load up to 15 µl of each of the samples in a predetermined order into the bottom of the wells. This is best done with a Hamilton microliter syringe or a micropipettor equipped with gel-loading tips that is washed with buffer from the bottom reservoir after each sample is loaded. Load an equal volume of 1x SDS gel-loading buffer into any wells that are unused.

10. Attach the electrophoresis apparatus to an electric power supply (the positive electrode should be connected to the bottom buffer reservoir). Apply a voltage of 8 V/cm to the gel. After the dye front has moved into the resolving gel, increase the voltage to 15 V/cm and run the gel until the bromophenol blue reaches the bottom of the resolving gel (~4 hours). Then turn off the power supply.

11. Remove the glass plates from the electrophoresis apparatus and place them on a paper towel. Use an extra gel spacer to carefully pry the plates apart. Mark the orientation of the gel by cutting a corner from the bottom of the gel that is closest to the leftmost well (slot 1).

▲ **IMPORTANT** Do not cut the corner from gels that are to be used for immunoblotting.

12. At this stage, the gel can be fixed, stained with Coomassie Brilliant Blue or silver salts, fluorographed or autoradiographed, or used to establish an immunoblot, all as described on the following pages.

STAINING SDS-POLYACRYLAMIDE GELS

Unlabeled proteins separated by polyacrylamide gel electrophoresis typically are detected by staining, either with Coomassie Brilliant Blue or with silver salts. In a relatively rapid and straightforward reaction, Coomassie Brilliant Blue binds nonspecifically to proteins but not to the gel, thereby allowing visualization of the proteins as discreet blue bands within the translucent matrix of the gel (Wilson 1983). Silver staining, although somewhat more difficult to perform, is significantly more sensitive. The use of silver staining allows detection of proteins resolved by gel electrophoresis at concentrations nearly 100-fold lower than those detected by Coomassie Brilliant Blue staining (Switzer et al. 1979; Merril et al. 1984). The identification of proteins by silver staining is based on the differential reduction of silver ions, in a reaction similar to that used in photographic processes. Reagents for staining with Coomassie Brilliant Blue as well as kits (e.g., Blue Print Fast PAGE Stain, Life Technologies) are commercially available. Kits for silver staining are commercially available from Pierce and Bio-Rad.

HISTORICAL FOOTNOTE

Coomassie Brilliant Blue R-250 was first used as a laboratory reagent to stain proteins in 1963. Robert Webster, a graduate student in the laboratory of Stephen Fazekas de St. Groth at the Australian National University in Canberra, was searching for a way to locate influenza virus proteins that had been separated by electrophoresis on cellulose acetate strips. At that time, Australia had a thriving wool industry and government laboratories were intensely investigating the mechanism of action of various classes of dyes used for wool dying. Fazekas and Webster reasoned that these dyes must have a high affinity for proteins, and they obtained samples of a great many dyes from the Commonwealth Scientific and Industrial Organization. Included among them was Coomassie Brilliant Blue R-250, which had been used since the turn of the century in the textile dying industry. Webster soon found that Coomassie Brilliant Blue R-250 was a very sensitive stain for proteins, but he was frustrated by extreme day-to-day variation in the intensity of the staining. At home one night, he suddenly realized that the answer to the problem was to fix the protein before staining. He went back to the laboratory and fixed the separated influenza virus proteins with sulfosalicylic acid. After these results were published (Fazekas de St. Groth et al. 1963), the method was rapidly adapted to stain proteins separated by electrophoresis through polyacrylamide gels (Meyer and Lamberts 1965).

Because Coomassie Brilliant Blue R-250 is now a trademark of Imperial Chemical Industries PLC, the dye is generally listed in biochemical catalogs as Brilliant Blue. Two forms of the dye are available: Brilliant Blue G and Brilliant Blue R, which are given different numbers (42655 and 42660) in the Colour Index, a kind of dyer's Bible, in which dyes are classified and arranged according to color. Brilliant Blue G and Brilliant Blue R are, respectively, slightly soluble and insoluble in cold water, and soluble and slightly soluble in hot water and alcohols. Both dyes stain fixed proteins efficiently. They are used at a concentration of 0.05% in methanol:glacial acetic acid:water (50:10:40 v/v).

Staining SDS-Polyacrylamide Gels with Coomassie Brilliant Blue

Coomassie Brilliant Blue is an aminotriarylmethane dye that forms strong but not covalent complexes with proteins, most probably by a combination of van der Waals forces and electrostatic interactions with NH_3^+ groups. Coomassie Brilliant Blue is used to stain proteins after electrophoresis through polyacrylamide gels. The uptake of dye is approximately proportional to the amount of protein, following the Beer-Lambert law.

Polypeptides separated by SDS-polyacrylamide gels can be simultaneously fixed with methanol:glacial acetic acid and stained with Coomassie Brilliant Blue R-250, a triphenylmethane textile dye also known as Acid Blue 83. The gel is immersed for several hours in a concentrated methanol:acetic acid solution of the dye, and excess dye is then allowed to diffuse from the gel during a prolonged period of destaining.

MATERIALS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Coomassie Brilliant Blue R-250

Methanol:acetic acid solution **<!>**

Combine 900 ml of methanol:H_2O (500 ml of methanol and 400 ml of H_2O) and 100 ml of glacial acetic acid.

Step 1 of this protocol requires the reagents listed on page A8.42.

METHOD

1. Separate proteins by electrophoresis through an SDS-polyacrylamide gel as described on page A8.40.
2. Prepare the staining solution by dissolving 0.25 g of Coomassie Brilliant Blue R-250 per 100 ml of methanol:acetic acid solution. Filter the solution through a Whatman No. 1 filter to remove any particulate matter.
3. Immerse the gel in at least 5 volumes of staining solution and place on a slowly rotating platform for a minimum of 4 hours at room temperature.
4. Remove the stain and save it for future use. Destain the gel by soaking it in the methanol:acetic acid solution without the dye on a slowly rocking platform for 4–8 hours, changing the destaining solution three or four times.

   The more thoroughly the gel is destained, the smaller the amount of protein detected by staining with Coomassie Brilliant Blue. Destaining for 24 hours usually allows as little as 0.1 μg of protein to be detected in a single band.

   A more rapid rate of destaining can be achieved by the following methods:

   - Destaining in 30% methanol, 10% acetic acid. If destaining is prolonged, there will be some loss in the intensity of staining of protein bands.
   - Destaining in the normal destaining buffer at higher temperatures (45°C).
   - Including a few grams of an anion exchange resin or a piece of sponge in the normal destaining buffer. These absorb the stain as it leaches from the gel.
   - Destaining electrophoretically in apparatuses that are sold commercially for this purpose.

5. After destaining, store the gels in H_2O in a sealed plastic bag.

   Gels may be stored indefinitely without any diminution in the intensity of staining; however, fixed polyacrylamide gels stored in H_2O will swell and may distort during storage. To avoid this problem, store fixed gels in H_2O containing 20% glycerol. Stained gels should not be stored in destaining buffer, because the stained protein bands will fade.

6. To make a permanent record, either photograph the stained gel (please see Chapter 5, Protocol 2) or dry the gel as described on p. A8.50.

Staining SDS-Polyacrylamide Gels With Silver Salts

A number of methods have been developed to stain polypeptides with silver salts after separation by SDS-polyacrylamide gel electrophoresis. In every case, the process relies on differential reduction of silver ions that are bound to the side chains of amino acids (Switzer et al. 1979; Oakley et al. 1980; Ochs et al. 1981; Sammons et al. 1981; Merril et al. 1984). These methods fall into two major classes: those that use ammoniacal silver solutions and those that use silver nitrate. Although both types of staining are ~100–1000-fold more sensitive than staining with Coomassie Brilliant Blue R-250 and are capable of detecting as little as 0.1–1.0 ng of polypeptide in a single band, silver nitrate solutions are easier to prepare and, by contrast to ammoniacal silver salts, do

not generate potentially explosive by-products. The method given below is a modification of the staining procedure originally devised by Sammons et al. (1981), which has since undergone several improvements (Schoenle et al. 1984). For further information, please see the discussion on silver staining in Appendix 9.

MATERIALS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Acetic acid (1%) <!>

Developing solution

Prepare fresh for each use an aqueous solution of 2.5% sodium carbonate, 0.02% formaldehyde. <!>

Ethanol (30%)

Fixing solution

Ethanol:glacial acetic acid:H_2O (30:10:60) <!>

Photographic reducing solution (Switzer et al. 1979)

Optional, please see Step 10.

Prepare Solution A: Dissolve 37 g of NaCl and 37 g of $CuSO_4$ in 850 ml of deionized H_2O. Add concentrated NH_4OH <!> until a deep blue precipitate forms and then dissolves. Adjust the volume to 1 liter with H_2O.

Prepare Solution B: Dissolve 436 g of sodium thiosulfate in 900 ml of deionized H_2O. Adjust the volume to 1 liter with H_2O.

Mix equal volumes of Solution A and Solution B, dilute the mixture with 3 volumes of H_2O, and use the diluted mixture immediately.

Silver nitrate solution

Prepare fresh for each use, 0.1% solution of $AgNO_3$,<!> diluted from a 20% stock, stored in a tightly closed, brown glass bottle at room temperature.

Step 1 of this protocol requires the reagents listed on page A8.42

METHOD

▲ **IMPORTANT** Wear gloves and handle the gel gently because pressure and fingerprints produce staining artifacts. In addition, it is essential to use clean glassware and deionized H_2O because contaminants greatly reduce the sensitivity of silver staining.

1. Separate proteins by electrophoresis through an SDS-polyacrylamide gel as described on page A8.40.
2. Fix the proteins by incubating the gel for 4–12 hours at room temperature with gentle shaking in at least 5 gel volumes of fixing solution.
3. Discard the fixing solution, and add at least 5 gel volumes of 30% ethanol. Incubate the gel for 30 minutes at room temperature with gentle shaking.
4. Repeat Step 3.
5. Discard the ethanol and add 10 gel volumes of deionized H_2O. Incubate the gel for 10 minutes at room temperature with gentle shaking.
6. Repeat Step 5 twice.

   The gel will swell slightly during rehydration.
7. Discard the last of the H_2O washes, and, wearing gloves, add 5 gel volumes of silver nitrate solution. Incubate the gel for 30 minutes at room temperature with gentle shaking.

8. Discard the silver nitrate solution, and wash both sides of the gel (20 seconds each) under a stream of deionized H_2O.

   Allowing the surface of the gel to dry out will result in staining artifacts.

9. Add 5 gel volumes of fresh developing solution. Incubate the gel at room temperature with gentle agitation. Watch the gel carefully. Stained bands of protein should appear within a few minutes. Continue incubation until the desired contrast is obtained.

   Prolonged incubation leads to a high background of silver staining within the body of the gel.

10. Quench the reaction by washing the gel in 1% acetic acid for a few minutes. Then wash the gel several times with deionized H_2O (10 minutes per wash).

    A shiny gray film of silver sometimes forms on the surface of the gel. This can be removed by washing the gel for 2–3 seconds in a 1:4 dilution of photographic reducing solution. Rinse the treated gel extensively in deionized H_2O.

11. Preserve the gel by drying as described on the following pages.

DRYING SDS-POLYACRYLAMIDE GELS

SDS-polyacrylamide gels containing proteins radiolabeled with ^{35}S-labeled amino acids must be dried before autoradiographic images can be obtained. The major problems encountered when a gel is dried are (1) shrinkage and distortion and (2) cracking of the gel. The first of these problems can be minimized if the gel is attached to a piece of Whatman 3MM paper before it is dehydrated. (For nonradioactive gels, note that the preference may be to dry the gel between acetate sheets to allow transillumination and easy visualization of the dried gel.) However, there is no guaranteed solution to the second problem, which becomes more pronounced with thicker gels containing more polyacrylamide. Cracking generally occurs when the gel is removed from the drying apparatus before it is completely dehydrated. It is therefore essential to keep the drying apparatus in good condition, to use a reliable vacuum line that has few fluctuations in pressure, and to use the thinnest gel possible to achieve the desired purpose. An excellent alternative method is soaking the gel in 3% glycerol, followed by drying in air using a simple apparatus (available from AP Biotech or Owl Scientific).

MATERIALS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Fixing solution

Glacial acetic acid:methanol:H_2O (10:20:70 v/v/v) <!>

Gel dryer

Gel dryers are available from a number of commercial sources (e.g., Life Technologies and Promega). It is best to purchase the dryer from the manufacturer of the SDS-polyacrylamide gel electrophoresis tanks to ensure that the size of the dryer will be tailored to that of the gels and will accommodate several SDS-polyacrylamide gels simultaneously.

Methanol (20%) containing 3% glycerol <!>

Optional, please see Step 1.

Whatman 3MM paper

METHOD

1. Remove the gel from the electrophoresis apparatus and incubate it at room temperature in 5–10 volumes of fixing solution. The bromophenol blue will turn yellow as the acidic fixing solution diffuses into the gel. Continue fixation for 5 minutes after all of the blue color has disappeared, and then wash the gel briefly in deionized H_2O.

   If cracking of polyacrylamide gels during drying is a constant problem, soak the fixed gel in 20% methanol, 3% glycerol overnight before proceeding to Step 2.

2. On a piece of Saran Wrap slightly larger than the gel, arrange the gel with its cut corner on the lower right-hand side.

3. Place a piece of dry Whatman 3MM paper on the damp gel. The paper should be large enough to create a border (1–2 cm) around the gel and small enough to fit on the gel dryer. Do not attempt to move the 3MM paper once contact has been made with the gel.

4. Arrange another piece of dry 3MM paper on the drying surface of the gel dryer. This piece should be large enough to accommodate all of the gels that are to be dried at the same time.

5. Place the sandwich of 3MM paper/gel/Saran Wrap on the piece of 3MM paper on the gel dryer. The Saran Wrap should be uppermost.

6. Close the lid of the gel dryer, and apply suction so that the lid makes a tight seal around the gels. If the dryer is equipped with a heater, apply low heat (50–65°C) to speed up the drying process.
7. Dry the gel for the time recommended by the manufacturer (usually 2 hours for standard 0.75-mm gels). If heat was applied, turn off the heat for a few minutes before releasing the vacuum.
8. Remove the gel, which is now attached to a piece of 3MM paper, from the dryer.
9. Remove the piece of Saran Wrap and establish an autoradiograph as described in Appendix 9, or store the dehydrated gel as a record of the experiment.

IMMUNOBLOTTING

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Immunoblotting is used to identify and measure the size of macromolecular antigens (usually proteins) that react with a specific antibody (Towbin et al. 1979; Burnette 1981; for reviews, please see Towbin and Gordon 1984; Gershoni 1988; Stott 1989; Poxton 1990). The proteins are first separated by electrophoresis through SDS-polyacrylamide gels and then transferred electrophoretically from the gel to a solid support, such as a nitrocellulose, polyvinylidene difluoride (PVDF), or cationic nylon membrane. After the unreacted binding sites of the membrane are blocked to suppress nonspecific adsorption of antibodies, the immobilized proteins are reacted with a specific polyclonal or monoclonal antibody. Antigen-antibody complexes are finally located by radiographic, chromogenic, or chemiluminescent reactions.

Much mumbo-jumbo has been written about ways to avoid the problems that commonly arise in immunoblotting. These problems include inefficient transfer of proteins, loss of antigenic sites, low sensitivity, high background, and nonquantitative detection methods. Although no magic incantation can eliminate all of these undesirable difficulties for every antigen, a small amount of experimentation is usually sufficient to cure all but the most obdurate technical problems. Comprehensive reviews by Bjerrum and Schafer-Nielsen (1986), Bjerrum et al. (1988), and Stott (1989) provide catalogs of potential difficulties in immunoblotting and detailed suggestions for solving them.

Transfer of Proteins from Gel to Filter

Electrophoretic transfer of proteins from polyacrylamide gels to membranes is far more efficient and much quicker than capillary transfer. Transfer is carried out perpendicularly from the direction of travel of proteins through the separating gel, using electrodes and membranes that cover the entire area of the gel. Most commercial electrophoretic transfer devices use large electrodes made of graphite, platinum wire mesh, or stainless steel. In older devices, vertical electrodes were submerged in a tank of transfer buffer in a plastic cradle surrounding the gel and the membrane. The more modern devices use the efficient "semi-dry" method, in which Whatman 3MM paper saturated with transfer buffer is used as a reservoir. For transfer from SDS gels, the membrane is

TABLE A8-11 Buffers for Transfer of Proteins from Polyacrylamide Gels to Membranes

TYPE OF TRANSFER	BUFFER	REFERENCE
Semi-dry	24 mM Tris base 192 mM glycine 20% methanol <!>	Towbin et al. (1979)
Immersion	48 mM Tris base 39 mM glycine 20% methanol <!> 0.0375% SDS	Bjerrum and Schafer-Nielsen (1986)

Methanol minimizes swelling of the gel and increases the efficiency of binding of proteins to nitrocellulose membranes. The efficiency of transfer may be affected by the presence of SDS in the electrophoresis buffer, the pH of the transfer buffer, and whether the proteins were stained in the gel before transfer. To maximize transfer of protein to membranes, the concentration of SDS should be ≤0.1% and the pH should be ≥8.0. CAPS buffer should be used for transfer if the protein is to be sequenced on the membrane (Moos 1992). Glycine interferes with this procedure.

placed on the side of the gel facing the anode. The conditions used for transfer vary according to the design of the apparatus, and it is therefore best to follow the manufacturer's instructions at this stage.

Types of Membranes

Three types of membranes are used for immunoblotting: nitrocellulose, nylon, and polyvinylidene fluoride. Different proteins may bind with different efficiencies to these membranes, and particular antigenic epitope(s) may be better preserved in one case than another. It is therefore worthwhile wherever possible to test the efficiency with which the antigen of interest can be detected on various membranes, using several antibodies.

- ***Nitrocellulose*** (pore size 0.45 µm) remains a standard membrane used for immunoblotting, although membranes with a smaller pore size (0.22 µm or 0.1 µm) are recommended for immunoblotting of small proteins of M_r <14,000 (Burnette 1981; Lin and Kasamatsu 1983). The capacity of nitrocellulose to bind and retain proteins ranges from 80 µg/cm^2 to 250 µg/cm^2, depending on the protein. Proteins bind to nitrocellulose chiefly by hydrophobic interactions (van Oss et al. 1987), although hydrogen bonding between amino acid side chains and the nitro group of the membrane may also be involved. In any event, partial dehydration of the proteins by methanol or salt in the transfer buffer ensures a more lasting bond between the protein and the membrane. Even so, proteins may be lost from the membrane during processing, particularly if buffers containing nonionic detergents are used. Many investigators therefore fix the proteins to nitrocellulose membranes to reduce loss during washing and incubation with antibody (e.g., please see Gershoni and Palade 1982). However, it is important to check that the treatments used for fixation (glutaraldehyde, cross-linking, UV irradiation) do not destroy the antigenic epitope under study. These treatments can also increase the brittleness of nitrocellulose filters that are allowed to dry after transfer.

- ***Nylon and positively charged nylon membranes*** are tougher than nitrocellulose and bind proteins tightly by electrostatic interactions. Their capacity varies from one type of nylon to the next and from one protein to another but is usually in the range of 150 µg/cm^2 to 200 µg/cm^2. The advantage of nylon and charged nylon membranes over nitrocellulose is that they can be probed multiple times with different antibodies. However, nylon membranes have two potential disadvantages. First, as discussed below, no simple and sensitive procedure is available to stain proteins immobilized on nylon and charged nylon membranes. Second, because it is difficult to block all of the unoccupied sites on these membranes, antibodies tend to bind nonspecifically to the filter, resulting in a high background, especially when a highly sensitive detection method such as enhanced chemiluminescence (ECL) is used. In many cases, extended blocking in solutions containing 6% heat-treated casein and 1% polyvinylpyrrolidone (Gillespie and Hudspeth 1991) is required to achieve satisfactory results.

- ***Polyvinylidene fluoride*** (PVDF) (Pluskal et al. 1986) is mechanically strong and manifests a strong interfacial (hydrophobic) interaction with proteins. Before transfer, it is necessary to wet the hydrophobic surface of the membrane with methanol. The capacity of PVDF membranes is approximately equal to that of nylon membranes (~170 µg protein/cm^2). Proteins bind approximately sixfold more tightly to PVDF membranes than to nitrocellulose (van Oss et al. 1987) and are retained more efficiently during the subsequent detection steps. Proteins immobilized on PVDF can be visualized with standard stains such as Amido Black, India Ink, Ponceau S, and Coomassie Brilliant Blue.

Staining of Proteins during Immunoblotting

Separation of proteins in gels and transfer to membranes can be confirmed by staining. This is a simple procedure, but it requires careful choice of a stain that is sufficiently sensitive and appropriate for the type of membrane. Staining can be carried out at several stages in the immunoblotting procedure as outlined below.

- ***Staining gels before transfer to membranes.*** Proteins can be stained in polyacrylamide gels with conventional dyes such as Coomassie Brilliant Blue, destained, and then transferred electrophoretically to nitrocellulose or PVDF filters for immunoblotting (e.g., please see Thompson and Larson 1992). The chief advantage of this method is that proteins remain stained during immunodetection, thereby providing a set of internal markers. However, in some cases, staining of proteins in gels appears to reduce the efficiency of electroelution and/or to interfere with binding of antibody. (The use of prestained protein markers [Life Technologies] provides a set of internal markers during protein transfer without the need to stain the entire gel.)
- ***Staining proteins after transfer to membranes.*** The entire area of nitrocellulose and PVDF membranes can be stained with the removable but insensitive stain Ponceau S (Muilerman et al. 1982; Salinovitch and Montelaro 1986). When more permanent stains are used (e.g., India Ink [Hancock and Tsang 1983], Amido Black [Towbin et al. 1979; Wilson 1979], colloidal gold [Moeremans et al. 1985; Rohringer and Holden 1985], or silver [Yuen et al. 1982]), it is usually necessary to cut a reference lane from the membrane.

Brief exposure to alkali enhances staining with both India Ink and colloidal gold, perhaps by reducing loss of protein from the filter during washing (Sutherland and Skerritt 1986). Under these conditions, it is easily possible to detect a band containing as little as a few nanograms of protein. There is no satisfactory method to stain proteins immobilized on nylon or cationic nylon membranes. The high density of charge on these membranes causes dye molecules to bind indiscriminately to the surface, producing high backgrounds that obscure all but the strongest protein bands.

Blocking Agents

Traditional blocking agents such as 0.5% low-fat dry milk or 5% bovine serum albumin (Johnson et al. 1984; DenHollander and Befus 1989) are suitable for use with chromogenic detection systems based on horseradish peroxidase. However, these solutions are usually rich in residual alkaline phosphatase and should not be used in detection systems that employ this enzyme. This is particularly true with chemiluminescent systems, where the sensitivity is determined not by the strength of the emitted signal but by the efficiency of suppression of background. The best blocking solution for alkaline-phosphatase-based systems contains 6% casein, 1% polyvinylpyrrolidone, 10 mM EDTA in phosphate-buffered saline (Gillespie and Hudspeth 1991). The blocking solution should be heated to 65°C for 1 hour to inactivate residual alkaline phosphatase and then stored at 4°C in the presence of 3 mM sodium azide. For recipes, please see Appendix 1.

Probing and Detection

The antibody that reacts with the epitope of interest can be either polyclonal or monoclonal. In either case, it is not radiolabeled or conjugated to an enzyme but is merely diluted into an appropriate buffer for formation of antibody-antigen complexes. In general, backgrounds in immunoblotting are unacceptably high unless the primary antibody can be diluted at least 1:1000

when enzymatic methods of detection are used and at least 1:5000 when chemiluminescent methods are used. After washing, the bound antibody is detected by a radiolabeled or enzyme-conjugated secondary reagent, which recognizes common features of the primary antibody and carries a reporter enzyme or group. Secondary reagents include:

- ***Radioiodinated antibodies or* Staphylococcal *Protein A,*** which were used in the first immunoblots (e.g., please see Burnette 1981) and for a few years thereafter. However, radiolabeled secondary reagents have now been replaced by nonradioactive detection systems such as enhanced chemiluminscence, which are less hazardous and more sensitive. They remain the most accurate method for semi-quantitative immunoblotting.
- ***Antibodies conjugated to enzymes,*** such as horseradish peroxidase or alkaline phosphatase, for which a variety of chromogenic, fluorescent, and chemiluminescent substrates are available.
- ***Antibodies coupled to biotin,*** which can then be detected by labeled or conjugated streptavidin.

Images of radiolabeled reagents are captured on X-ray film or phosphorimagers, whereas the results of chromogenic and fluorogenic reactions are best recorded by conventional photography. Table A8-12 shows the approximate sensitivity with which the best of these methods can detect a standard antigen using antibodies of high titer and specificity. For more information about these detection methods, please see Appendix 9.

TABLE A8-12 Chromogenic and Chemiluminescent Methods of Detection of Immobilized Antigens

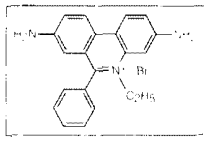
Enzyme	Reagent	Sensitivity	Comments	References
Chromogenic				
Horseradish peroxidase	4-chloro-1 naphthol/ H_2O_2	1 ng	The purple color of oxidized products fades rapidly on exposure to light.	Hawkes et al. (1982); Dresel and Schettler (1984)
	diaminobenzidine <!>/ H_2O_2	250 pg	Potentially carcinogenic. The diaminobenzidine reaction generates a brown precipitate, which is enhanced by the addition of cobalt, silver, and nickel salts.	de Blas and Cherwinski (1983); Gershoni (1988)
	3,3′,5,5′-tetramethylbenzidine	100 pg	Deep purple precipitate.	McKimm-Breschkin (1990)
Alkaline phosphatase	nitro blue tetrazolium/ 5-bromo-4-chloroindolyl phosphate	100 pg	Steel-blue precipitate.	Leary et al. (1983); Blake et al. (1984)
Chemiluminescent				
Horseradish peroxidase	luminol/4-iodo-phenol/ H_2O_2	300 pg	Oxidized luminol emits blue light that is captured on X-ray film. Luminescence generated by intense bands appears within a few seconds,whereas faint bands need at least 30 minutes to develop.	Schneppenheim and Rautenberg (1987); Harper and Murphy (1991); Schneppenheim et al. (1991)
Alkaline phosphatase	AMPPD 3-(4-methoxyspiro[1,2-dioxetane-3′2′-tricyclo-[3.3.1^{3,7}]decan]-4-yl)-phenylphosphate	1 pg	The enzymatically dephosphorylated product emits light. Because of its high turnover number, alkaline phosphatase rapidly generates a strong signal that provides an exquisitely sensitive method of immunodetection.	Gillespie and Hudspeth (1991)

Appendix 9

Detection Systems

STAINING NUCLEIC ACIDS

Ethidium Bromide



Ethidium bromide was synthesized in the 1950s in an effort to develop phenanthridine compounds as effective trypanocidal agents. Ethidium emerged from the screening program with flying colors. It was 10–50-fold more effective against trypanosomes than the parent compound, was no more toxic to mice, and, unlike earlier phenanthridines, did not induce photosensitization in cattle (Watkins and Wolfe 1952). Until recently, ethidium bromide was widely used for the treatment and prophylaxis of trypanomiasis in cattle in tropical and subtropical countries. The chemical structure of ethidium bromide is shown at the right.

Binding of Ethidium Bromide to Nucleic Acids

Ethidium bromide contains a planar tricyclic phenanthridine ring system that is able to intercalate between the stacked base pairs of double-stranded DNA. After insertion into the helix, the drug lies perpendicular to the helical axis and makes van der Waals contacts with the base pairs above and below. Whereas the planar ring system of the drug is buried, its peripheral phenyl and ethyl groups project into the major groove of the DNA helix. At saturation in solutions of high ionic strength, approximately one ethidium molecule is intercalated per 2.5 base pairs, independent of the base composition of the DNA. The geometry of the base pairs and their positioning with respect to the helix are unchanged except for their displacement by 3.4 Å along the helix axis (Waring 1965). This causes a 27% increase in the length of double-stranded DNA (Freifelder 1971) saturated with ethidium bromide.

Ethidium bromide also binds with highly variable stoichiometry to helical regions formed by intrastrand base pairing in RNA and heat-denatured or single-stranded DNA (Waring 1965, 1966; LePecq and Paoletti 1967). The fixed position of the planar group of ethidium bromide and its close proximity to the bases cause the bound dye to display a 20–25-fold increase in fluorescent yield compared to the dye in free solution. UV radiation at 254 nm is absorbed by the DNA and transmitted to the dye; radiation at 302 nm and 366 nm is absorbed by the dye itself. The energy is re-emitted with a quantum yield of 0.3 at 590 nm in the red-orange region of the visible spectrum (LePecq and Paoletti 1967; Tuma et al. 1999).

Most of the commercially available UV light sources emit UV light at 302 nm. The fluorescent yield of ethidium bromide–DNA complexes excited by irradiation is considerably greater at 302 nm than at 366 nm but is slightly less than at shorter wavelength (254 nm). However, the amount of photobleaching of the dye and nicking of the DNA is much less at 302 nm than at 254 nm (Brunk and Simpson 1977).

Staining DNA in Gels

Ethidium bromide is widely used to locate fragments of DNA in agarose gels (Aaij and Borst 1972; Sharp et al. 1973; please see the introduction to Chapter 5 and Protocol 2 in Chapter 5). The dye is usually incorporated into the gel and the electrophoresis buffer at a concentration of 0.5 µg/ml. Although the electrophoretic mobility of linear double-stranded DNA is reduced by ~15% in the presence of ethidium bromide, the ability to examine the gel directly under UV illumination is a great advantage. Since the fluorescent yield of ethidium bromide–DNA complexes is much greater than that of the unbound dye, small amounts of DNA (~10 ng/band) can be detected in

the presence of free ethidium bromide in gels (Sharp et al. 1973). Even smaller quantities of DNA can be detected if the DNA has previously been treated with chloroacetaldehyde, a chemical mutagen that reacts with adenine, cytosine, and guanine (Premaratne et al. 1993). A more practical way to enhance fluorescence is to destain the gel in a solution containing 10 mM Mg^{2+} before examining it under UV illumination (Sambrook et al. 1989).

Quantitating Double-stranded DNA

The formation of complexes between DNA and ethidium bromide can be observed with the naked eye because of the large metachromatic shift in the absorbtion spectrum of the drug that accompanies binding. The original maximum at 480 nm (yellow-orange) is shifted progressively to 520 nm (pink) with a characteristic isosbestic point at 510 nm. This provides a simple way to estimate the concentration of a sample of DNA by quantititive spectrophotometry (Waring 1965).

A faster and more sensitive method utilizes the UV-induced fluoresence emitted by intercalated ethidium bromide molecules. Because the amount of fluorescence is proportional to the total mass of DNA, the quantity of DNA can be estimated by comparing the light emitted by the sample at 590 nm with that of a series of standards (for more information on quantifying DNA with ethidium bromide, please see Appendix 8).

Improved Versions of Ethidium Bromide

Dimers of intercalating dyes bind to DNA with much greater affinity than the parent monomeric compound (Gaugain et al. 1978). Homodimers of ethidium bromide and heterodimers of acridine and ethidium are therefore much more sensitive reagents for detecting DNA than is monomeric ethidium bromide. For example, as little as 30 pg of DNA can be detected on a confocal fluorescence gel scanner (e.g., please see Glazer et al. 1990; Glazer and Rye 1992). However, the price for this increase in sensitivity is very steep; 1 mg of ethidium homodimer costs about ten times more than 1 g of ethidium bromide. Unsymmetric cyanine dyes unrelated to ethidium bromide are more sensitive detectors of DNA, but these dyes are also expensive (please see section on SYBR Dyes below). For information on disposal of ethidium bromide, please see Appendix 8.

Methylene Blue

Also known as Swiss Blue, in recognition of the nationality of Caro who first synthesized the dye in 1876, methylene blue (Fierz-David and Blangey 1949) is sometimes used as a stain for RNA that has been transferred to nitrocellulose filters or to certain types of nylon filters (Herrin and Schmidt 1988) (please see Chapter 7, Protocol 7).

Methylene blue may also be used to stain bands of DNA in agarose gels (please see the panel on **BRIEF PROTOCOL**). The aim is to avoid the use of ethidium bromide and to minimize the exposure of the DNA to UV irradiation, which can generate pyrimidine dimers and lower the biological activity of the DNA. This problem does not seem to be serious, but it is nevertheless a source of concern to some investigators.

Methylene blue has two absorbtion maxima (668 and 609 nm) in the visible spectrum, and is soluble in H_2O. For staining RNA immobilized on nylon or nitrocellulose, the dye is used at a concentration of 0.04% in 0.5 M sodium acetate (pH 5.2). Staining is reversible and can be carried out before hybridization.

BRIEF PROTOCOL

Staining DNA in Gels with Methylene Blue

1. Load and run a gel cast with GTG agarose in 1× TAE buffer.

   The smallest amount of DNA that can be reliably detected as a band by staining with methylene blue is ~40 ng. It is therefore necessary in most cases to load 2–3 times the normal quantity of DNA into the gel.

2. At the end of the electrophoretic run, place the gel in a glass dish containing 5 gel volumes of a solution containing 0.001–0.0025% methylene blue (available from Sigma) in 1 mM Tris-acetate (pH 7.4), 0.1 mM EDTA (pH 8.0).

3. Incubate the gel for 4 hours at room temperature, with gentle agitation on a rotary shaker.

4. Rinse the gel briefly in distilled H_2O and examine it on a light box of the kind used to illuminate X-ray films.

Silver Staining

Count Albert von Bollstadt (1193 or 1206–1280) is thrice famous. He was the teacher of St. Thomas Aquinas, he was an alchemist who described arsenic so clearly that he sometimes receives credit for the discovery of the element, and, of relevance to molecular cloning, he recorded that silver nitrate would stain human skin. Eight centuries later, silver staining has been refined into a highly sensitive technique for postelectrophoretic detection of DNA bands in polyacrylamide and agarose gels. At its best, silver staining can detect bands containing <1.0 ng of DNA. Described below are three general types of silver staining (Merril 1987, 1990; Mitchell et al. 1994).

- ***Photo development,*** like conventional photography and fluorography, uses photonic energy to reduce silver ions to the metallic element. Unfortunately, the simplicity and speed of the method cannot compensate for its lack of sensitivity, which is no better than can be achieved by conventional staining with ethidium bromide.

- ***Diammine staining*** methods use ammonium hydroxide to generate silver diammine complexes, which bind to the nucleic acid (Yuksel and Gracy 1985). Silver ions are then liberated from the complexes by decreasing the concentration of ammonium ions with citric acid. The liberated silver ions are finally reduced to metallic silver by formaldehyde. The basic method established by Johansson and Skoog (1987) is both rapid and reasonably sensitive (0.5–2 ng of DNA/band); greater sensitivity (0.1–1 ng DNA/band) can be achieved using the modifications described by Vari and Bell (1996). However, in our hands, it is difficult to achieve such high sensitivity on a regular basis. In addition, ammoniacal silver salts are potentially explosive and must be handled with great care.

- ***Nondiammine staining*** involves, for example, fixation of the DNA, sensitization of the DNA with glutaraldehyde, impregnation of the gel or membrane with silver nitrate at weakly acidic pH, and reduction of bound silver ions to metallic silver by alkaline formaldehyde. Many different variants of this technique have been published, but most of them suffer to a greater or lesser extent from the same problem: Bands of DNA stain gray or dog-yellow against a variable background of brownish surface staining (Vari and Bell 1996). This problem can be minimized, as in the protocol below, by carefully monitoring the gel during development so as to obtain the greatest discrimination between specific staining of the DNA and background staining of the gel. Differential reduction of silver ions can be improved by adding sodium thiosulfate to the alkaline formaldehyde solution (Bassam et al. 1991); thiosulfate removes silver ions from the gel surface by forming soluble complexes with silver salts. When working well, the nondiammine staining can detect bands of DNA containing 2–5 ng of DNA. A nondiammine silver staining is marketed by Promega as part of the Silver Sequence DNA sequencing kit. The

nondiammine staining method outlined below is simple and is sensitive enough to detect a band containing 2–5 ng of DNA in a polyacrylamide gel. The protocol was kindly provided by Dr. Sue Forrest (Victorian Clinical Genetics Service, Melbourne, Australia).

Silver Staining of DNA in Nondenaturing Polyacrylamide Gels

The volumes in the protocol are appropriate for staining a 150 x 150-mm polyacrylamide gel. Gels of this size are often used to analyze single-stranded DNA by SSCP (please see Chapter 13).

MATERIALS

CAUTION: Please see Appendix 12 for appropriate handling of materials marked with <!>.

Acetic acid (3% v/v) <!>

Developer

Dissolve 22.9 g of sodium carbonate in a final volume of 1 liter of distilled H_2O. Store the solution at room temperature away from direct light (e.g., in a closet).

Ethanol (10% v/v)

Ethanol/Glycerol (10% and 7% v/v, respectively)

Formaldehyde (37% v/v) <!>

Nitric acid (0.7% v/v) <!>

Silver nitrate (0.2% w/v), freshly prepared <!>

METHOD

1. After electrophoresis is complete, place the gel, still attached to one glass plate, in a plastic tray reserved for silver staining. Do not touch the surface of the gel at any time.

   Because pressure on the gel causes background staining, the various solutions used for silver staining are removed from the plastic tray with a pipette attached via a trap to a vacuum line.

2. Rinse the gel twice with distilled H_2O to remove electrophoresis buffer. During rinsing, the gel will float free and the glass plate can then be removed.

3. Fix the gel in 10% ethanol by gentle shaking for 10 minutes on a rocking platform. Remove the 10% ethanol by suction and repeat the process.

   If necessary, the gel can be left for several hours in the second batch of 10% ethanol.

4. Remove the ethanol by suction and add just enough 0.7% nitric acid to cover the gel. Shake the gel gently on a rocking platform for 6 minutes. Remove the nitric acid by suction and rinse the gel with two changes of distilled H_2O.

5. Add just enough 0.2% silver nitrate to cover the gel. Shake the gel gently on a rocking platform for 30 minutes. Rinse the gel and the staining tray three times with distilled H_2O.

   The silver nitrate solution may be reused but staining then may become unreliable.

6. To 100 ml of developer, add 125 µl of formaldehyde solution. Transfer the developer/formaldehyde solution to the staining tray and shake the tray gently in an indirect light (e.g., cover the container with aluminum foil). When the solution turns yellow or when a dark precipitate becomes noticeable, replace the developer/formaldehyde with a second batch of 100 ml of the same solution. Continue to shake the gel in indirect light. Monitor the appearance of bands and background. When the ratio of signal to noise is at its maximum, remove the second batch of developer/formaldehyde.

7. Add 250 ml of 3% acetic acid to the staining tray. Shake the gel gently for 5 minutes.
8. Remove the 3% acetic acid and wash the gel with 10% ethanol. Remove the ethanol and then store the gel for 2 minutes in a fresh batch of 10% ethanol.
9. Photograph the gel by transillumination on a white light box using Polaroid 667 film. The best photographs are obtained when the area around the gel is covered with black paper.
10. For long-term storage, either keep the gel in 10% ethanol/7% glycerol or dry it as follows:
    a. Lay a glass plate 2 cm longer and wider than the gel across the top of a beaker.
    b. Wet a piece of Saran Wrap approximately the same size as the glass plate. Lay the wet wrap on the glass plate, taking care to remove any air bubbles.
    c. Place the gel in the center of the piece of Saran Wrap and then cover it with another piece of pre-wet wrap, the same size as the first. Remove all wrinkles and air bubbles.
    d. Place gel spacers, ~1 cm thick on the Saran Wrap around the four edges of the gel. Use a series of bulldog clips to clamp the gel spacers to the glass plate.
    e. Allow the gel to dry for 24–48 hours, until the Saran Wrap feels crisp.

SYBR Dyes

The SYBR dyes are unsymmetric cyanine compounds, developed by Molecular Probes, that have some advantages as stains for DNA and RNA over phenanthridine dyes such as ethidium bromide. The information about SYBR dyes in the scientific literature is sparse. However, the Web Site of the manufacturer (http://www.probes.com) has much useful information, which is summarized below. For further details on the advantages and use of SYBR dyes, please see Chapter 5, Protocol 2.

Three SYBR dyes are used in molecular cloning: SYBR Green I and II and SYBR Gold. All three dyes are essentially nonfluorescent in free solution but, upon binding to nucleic acids, display greatly enhanced fluorescence and a high quantum yield. SYBR Green I, for example, has a quantum yield of 0.8 upon binding to double-stranded DNA, whereas SYBR Gold has a quantum yield of 0.7 and 1000-fold enhancement of fluorescence (Tuma et al. 1999). Because the SYBR dyes generate strong signals with very little background and have a high affinity for nucleic acids, they can be used in low concentrations and are more sensitive than conventional stains such as ethidium bromide. SYBR Green I and II, however, have some less desirable characteristics:

- These dyes are not optimally stimulated by standard transilluminators that emit UV radiation at 300 nm. The signal strength improves when illumination at 254 nm is used, but at this wavelength, damage to DNA is maximal.
- Both dyes penetrate agarose gels slowly. Postelectrophoretic staining can take 2 hours or more when the gels are thick or contain a high concentration of agarose.
- The dyes are not particularly photostable.
- SYBR Green I is only slightly more sensitive than ethidium bromide in detecting single-stranded DNA in agarose gels.

For these reasons, SYBR Green I and II are not commonly used to stain DNA in agarose gels. However, SYBR Green II detects RNA in denaturing agarose gels with fivefold greater sensitivity than ethidium bromide and does not interfere with northern transfer. The dye is therefore useful when analyzing quantities of RNA that are too small to be detected by ethidium bromide. SYBR

Green I, on the other hand, is used chiefly to quantify DNA in solution, for example, in real time polymerase chain reactions (please see the panel on **REAL TIME PCR** in Chapter 8, Protocol 15).

SYBR Gold, which has come onto the market more recently, is the best of the SYBR dyes. It is tenfold more sensitive than ethidium bromide, and its dynamic range is greater. Unlike SYBR Green I and II, SYBR Gold penetrates gels quickly and can therefore be used to stain DNA and RNA both in conventional neutral polyacrylamide and agarose gels and in gels containing denaturants such as urea, glyoxal, and formaldehyde. Because SYBR Gold most probably binds to the backbone of charged phosphate residues, the electrophoretic mobility of DNA stained with the dye is markedly retarded, and the bands of DNA are sometimes curved. For this reason, gels are stained with SYBR Gold after electrophoresis is complete. The level of background fluorescence is so low that no destaining is required. When excited by standard transillumination at 300 nm, nucleic acids stained with SYBR Gold generate bright gold fluorescent signals that can be captured on conventional black and white Polaroid film (type 667) or on charged couple device (CCD)-based image detection systems. The stained nucleic acids can be transferred directly to membranes for northern or Southern hybridization (Tuma et al. 1999).

Although many enzymatic reactions are not inhibited by SYBR Gold, polymerase chain reactions are sensitive to high concentrations of the dye. Inhibition can, however, be relieved by adjusting the concentration of Mg^{2+} (Tuma et al. 1999) or be avoided by removing SYBR Gold from the template DNA by standard ethanol precipitation.

SYBR Gold is supplied as a 10,000x concentrate in anhydrous dimethylsulfoxide (DMSO). The high cost of the dye precludes its use for routine staining of gels. However, the dye may be cost-effective as an alternative to radiolabeling or silver staining of DNA in techniques such as single-strand conformation polymorphism (SSCP) and denaturing gradient gel electrophoresis (DGGE).

AUTORADIOGRAPHY AND IMAGING

The first autoradiograph was obtained in 1867, when Niepce de St. Victor described the blackening of emulsions of silver chloride and iodide by uranium nitrate and tartrate. The blackening occurred even when the uranium salt was separated from the emulsion by sheets of colored glass. At that time, radioactivity had not been discovered and Niepce struggled unsuccessfully to explain his results in terms of luminescence.

Molecular cloning depends on techniques to map accurately the distribution of radioactive atoms on two-dimensional surfaces. For the last 20 years, autoradiographic images of Southern blots, northern hybridizations, DNA sequencing gels, and library screens have been the icons of the field. In autoradiography, a radioactive specimen emits radiation, generally in the form of β-particles whose image is recorded on photographic emulsion. A diagram of the events that occur during exposure of photographic emulsion to radioactivity is shown in Figure A9-1.

The emulsions used in autoradiography are suspensions of crystals (grains) of silver halide in gelatin. Exposure to radiation activates the halide crystals, producing a latent image that can be converted to a true image by development. Each β-particle emitted by the sample converts a number of silver ions to silver atoms, which are then withdrawn from the crystal lattice. The resulting latent image is unstable since the atoms of silver tend to lose their captured electrons and to resume their places in the lattice. At room temperature, this return reaction has a half-time of ~1 second. At −70°C, the rate of the return reaction is much slower, and there is little fading of the latent image.

During development, the activated nuclei of silver atoms catalyze the conversion of the entire silver halide grain into metallic silver. In most emulsions, it is necessary to activate between 5 and 10 silver atoms per crystal in order to obtain complete conversion of the crystal during development. Crystals with fewer activated silver atoms have a lower chance of development. This means that the intensity of the final image is not proportional to the intensity of the incident radiation. Low levels of radiation will generate developed images that are disproportionately faint. In photography, this phenomenon is known as low-intensity reciprocity failure.

Conversely, exposure to an intense source of light or β-particles can saturate all of the silver bromide crystals so that the emulsion becomes refractory to further radiation. The density of such burnt images is no longer proportional to the intensity of the incident radiation. Developed images whose absorbance at 545 nm exceeds 1.3 on a microdensitometer have saturated the film and cannot be used to quantitate the intensity of the original source of radiation.

The three isotopes most commonly used for autoradiography are ^{35}S, ^{33}P, and ^{32}P, all of which emit β-particles. The energies of these particles are different: ^{35}S emits a particle with a maximum energy of 0.167 MeV that can penetrate film emulsion only to a depth of 0.25 mm (see Figure A9-2). ^{33}P emissions are slightly stronger (0.249 MeV) with a maximum penetration depth of 0.6 mm. Although this depth is sufficient to allow the emitted β-particles to interact productively with silver

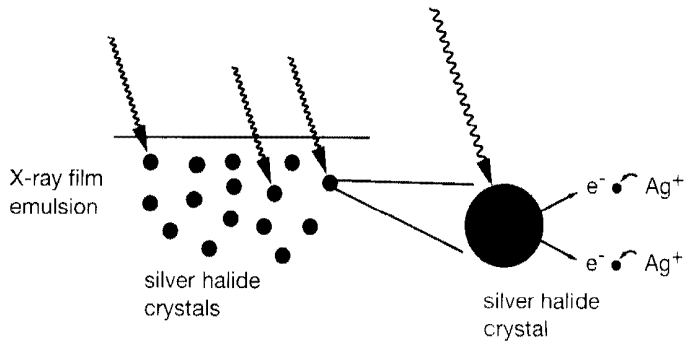


FIGURE A9-1 Events Leading to the Formation of an Autoradiographic Image

The diagram shows that particles entering the autoradiographic image cause ejection of electrons from silver halide crystals. These electrons attract positively charged silver ions, generating precipitates of silver atoms.

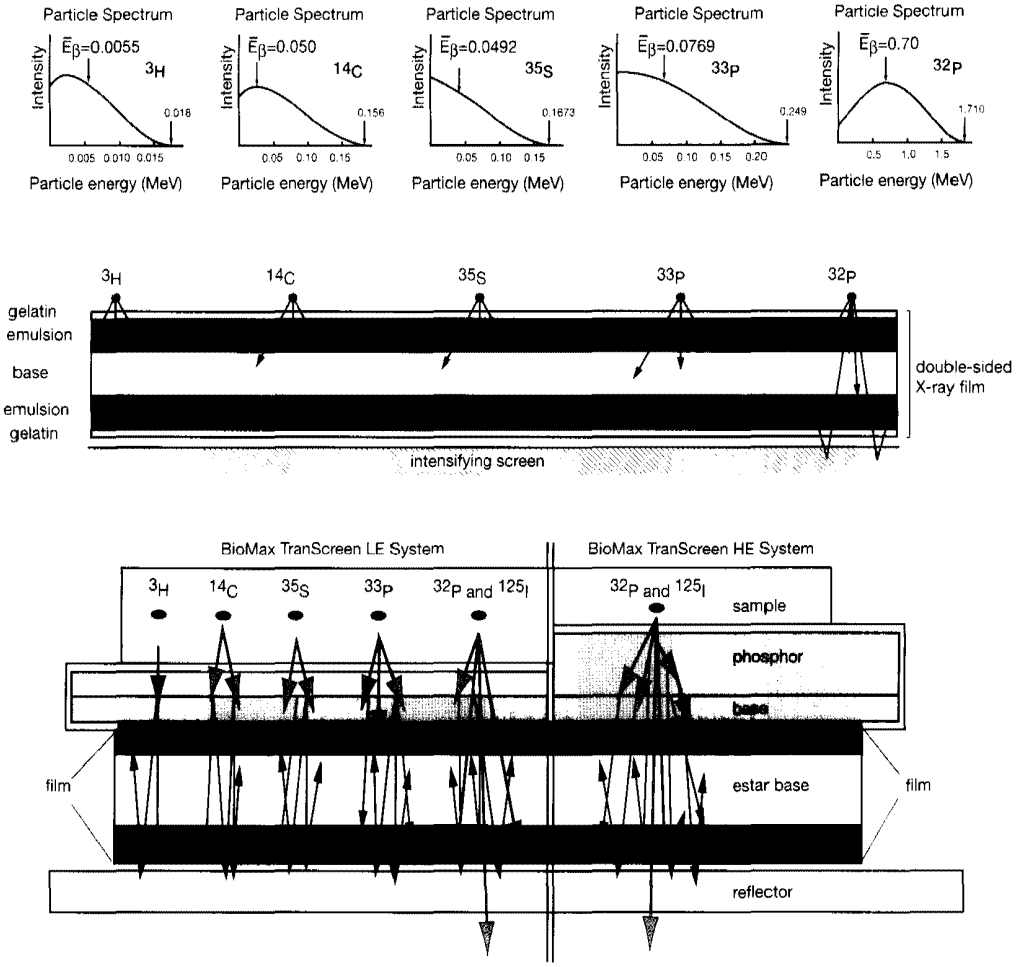


FIGURE A9-2 Energy of Radiation Emitted by Commonly Used Isotopes

(Top) Graphs show the spectra of energies carried by particles emitted by decaying radioactive isotopes. In each case, the arrow marks the average energy per particle. *(Middle)* Diagram showing the depth to which commonly used isotopes penetrate autoradiographic film. *(Bottom)* Principle of Kodak BioMax TranScreen Systems (Modified, with permission, from Eastman Kodak Company; Kodak, BioMax, and TranScreen are trademarks of Eastman Kodak Company.)

halide crystals in the emulsion, it is not enough to allow the particles to pass through barriers (e.g., Saran Wrap) that might be placed between the film and the source of the radiation. Thus, when establishing autoradiographs of ^{35}S- or ^{33}P-labeled material, it is essential that the film and the source of the radiation be directly apposed to one another. To reduce internal absorption of radiation, gels should be as thin as possible and should be fixed and dried before autoradiographs are taken. Nitrocellulose and nylon filters should be thoroughly dried, and care should be taken to ensure that the surface carrying the radioactivity is placed in contact with the film. (*Warning:* Damp gels and membranes stick tightly to the film and usually cannot be removed.)

^{32}P, by contrast to ^{35}S and ^{33}P, emits a β-particle with sufficient energy (1.709 MeV) to penetrate water or plastic to a depth of 6 mm and to pass completely through an X-ray film. Gels and filters therefore need not be completely dried (although the sharpness of the autoradiographic image is much improved if they are) and can be covered with Saran Wrap before they are exposed to the film. Radiation from ^{32}P is strong enough to require shielding by 1-cm Plexiglas, which blocks β-particles while minimizing production of Bremstrahlung.

AUTORADIOGRAPHY AND PHOSPHORIMAGING

Intensifying Screens

Strong β-particles, such as those emitted by decay of ^{32}P, can pass undetected through X-ray film. To increase the efficiency with which high-energy particles are detected, an intensifying screen may be placed behind the X-ray film. Radioactive particles that pass through the film hit the intensifying screen and cause it to emit photons that are captured by silver halide crystals in the emulsion. The efficiency of the intensifying screen is determined largely by the thickness of the phosphor layer. A thicker phosphor layer results in a "faster" screen because the thick layer absorbs more radiation than a thin layer. Thick screens are faster, but generate fuzzier images on film, due primarily to diffusion of light in the phosphor layer.

Most conventional intensifying screens are made of calcium tungstate, which emits blue light after capture of β-particles. Other screens contain rare earths such as lanthanum oxybromide (blue light) or gadolinium oxysulfide, which emits green light. Modern calcium tungstate screens such as Lightning Plus (Dupont, Cronex) enhance the intensity of an autoradiographic image by a factor of ~5 when the film is exposed at low temperature (−70°C) to retard the decay of the latent image (Koren et al. 1970; Swanstrom and Shank 1978). Intensifying screens are usually used in pairs, with double-sided X-ray film and the ^{32}P-labeled sample sandwiched between the two screens in a light-tight cassette.

Conventional intensifying screens do not improve the efficiency with which low-energy β-particles can be captured by X-ray film. It is therefore a waste of time to use these screens with samples that are labeled with ^{35}S, ^{33}P, ^{14}C, or ^{3}H (Laskey and Mills 1977; Sanger et al. 1977). However, specialized screens for use with these isotopes are available from Kodak. These screens are placed between the sample and the film. β-particles from the sample are captured by the screen and converted into photons, which are then detected by the film, as illustrated in Figure A9-2. The signal enhancement provided by TranScreen is generally equivalent to the sensitivity obtained using fluorography.

Preflashing

The spectral emission of some phosphors requires that an appropriately sensitized film be used or much of the light will be wasted. Accordingly, the efficiencies with which low levels of radioactivity can be detected are increased by a further twofold by preexposing the film to a short (~1 msec) flash of light emitted by a stroboscope or a photographic flash unit. This exposure generates stable pairs of silver atoms within each silver halide crystal and therefore increases the probability that an incoming particle of radiation will generate an activated crystal, which will be reduced to silver metal during the developing process. The distance of the light source from the film during preexposure should be determined empirically as follows (Laskey and Mills 1975, 1977). Note that preflashing is not recommended for high-sensitivity films such as BioMax (Kodak). Preflashing these films induces high levels of background.

1. Cover the stroboscope or flash unit with an orange filter (Kodak, Wratten 21 or 22A). This filter reduces the amount of incident blue light, to which X-ray film is very sensitive.
2. Working in total darkness, place the film perpendicular to the light source and at least 50 cm away from it. This prevents uneven illumination. Cover the film with a diffusing screen. If a suitable screen is not available, use a piece of Whatman No. 1 filter paper.

3. Expose a series of test films to the light source for different lengths of time and then develop the films. Cut the films into pieces that will fit neatly into the cuvette holder of a spectrophotometer. Measure the absorbance at 545 nm of the exposed films against a blank consisting of film that was not preexposed. Choose an exposure time that causes the absorbance to increase by 0.15.

Preflashed film has another advantage: The intensity of the image on the film becomes proportional to the amount of radioactivity in the sample (Laskey 1980). The intensity of autoradiographic images on preexposed film can therefore be quantitated by microdensitometry and be used to measure the amount of radioactivity in the original sample. By contrast, the silver halide crystals in film that is not preexposed to light are not fully activated and therefore respond in a sigmoidal fashion to increasing amounts of radioactivity (Laskey and Mills 1975, 1977). This relationship can complicate quantitation of autoradiographic images. The best types of films, for all types of autoradiography except fluorography, are Kodak X-Omat-R and Fuji RX. When preexposed, these films yield images whose absorbances are proportional to the intensity of the source of radioactivity over a range of 0.15–1.0. However, true linear responses are only obtained when (1) the background absorbance of the film at 545 nm is raised to an OD of 0.10–0.20 and (2) the presensitizing flash is brief (~1 msec).

Fluorography

The intensity of autoradiographic images of weak β-emitters such as ^{3}H, ^{14}C, and ^{35}S can be enhanced by impregnating the samples with chemicals that are fluorescent and emit many photons when they encounter a single quantum of radiation (Wilson 1958, 1960). Fluorography increases the sensitivity of detection of ^{14}C and ^{35}S approximately tenfold and permits detection of ^{3}H, which is otherwise virtually invisible to conventional autoradiography. Fluorography is therefore particularly useful for the detection of radiolabeled proteins and nucleic acids in polyacrylamide gels.

In the original methods (Bonner and Laskey 1974; Laskey and Mills 1975), aqueous gels containing the radioactive samples were equilibrated with DMSO, impregnated with the scintillant PPO (2,5-diphenyloxazole), soaked in H_2O to remove the DMSO, dried, and exposed to X-ray film at –70°C. These procedures were costly, tedious (requiring at least 5 hours of work), and irreproducible in inexperienced hands. The most frequent cause of difficulty was the failure to remove DMSO: Complete removal is essential to avoid sticky gels after drying. Because of these problems, a number of alternative solvents have been developed to deliver PPO to the sample. These include ethanol (Laskey 1980), glacial acetic acid (Skinner and Griswold 1983), and several other organic solvents (e.g., please see Shine et al. 1974; Southern 1975).

Despite these improvements, PPO has now largely been replaced as a scintillant by sodium salicylate (Chamberlain 1979) or by commercial scintillants (see below). With sodium salicylate, the level of enhancement is approximately equal to that obtained with organic scintillants, although the bands are slightly more diffuse. Commercially available aqueous scintillants such as En^{3}Hance, Enlightning, or Entensify (NEN Life Science Products) and Amplify (Amersham) are supplied in liquid and spray-on form and, if used in accordance with the manufacturer's instructions, give results every bit as good as those obtained with PPO, with far less work. However, they are exceedingly expensive.

The types of X-ray film used for fluorography should match the fluorescence spectrum of the scintillant. Sodium salicylate emits at 409 nm, whereas PPO emits at 375 nm. Commercial enhancers emit either blue or UV light. Films that are sensitive in this region of the spectrum are called "screen-type" X-ray films and include Kodak BioMax MS, Amersham Hyperfilm-MP, and Fuji RX.

TABLE A9-1 Sensitivity of Autoradiographic Methods for Detection of Radioisotopes

Isotope	Method	Sensitivity (dpm/mm^2)
^{35}S	no enhancement	30–60
	fluorography	2–3
	TranScreen LE (Kodak)	0.8–1.2
^{32}P	direct	2–5
	intensifying screen	0.5
	TranScreen HE (Kodak)	0.05–0.1
^{33}P	direct	15–30
	intensifying screen	1–1.5
	TranScreen LE (Kodak)	0.4–0.6
^{14}C	fluorography	2
	TranScreen LE (Kodak)	0.8–1.2
^{125}I	intensifying screen	1–2
^{3}H	fluorography	10–20
	TranScreen LE (Kodak)	74–110

Sensitivity of Different Autoradiographic Methods

Table A9-1 shows the sensitivities of different autoradiographic methods for the detection of radioisotopes, with and without various enhancements. The amounts of radioactivity shown in the table are those required to obtain a detectable image (A_{545} = 0.02) on preflashed film that is exposed to the sample for 24 hours. Much longer exposure times may be necessary to obtain publishable images.

Setting up Autoradiographs

1. Prepare gels for autoradiography in one of the following ways:

   a. Fix the SDS-polyacrylamide gels containing ^{33}P or ^{35}S, ^{14}C, or ^{3}H as described in Chapter 5, Protocol 11. Use a commercial gel dryer to dry the gels onto Whatman 3MM paper.

   b. Fix the sequencing gels containing ^{33}P or ^{35}S as described in Chapter 12, Protocol 12. Use a commercial gel dryer to dry the gels onto Whatman 3MM paper.

   c. For maximal sensitivity and resolution, fix, dry, and mount the polyacrylamide gels containing ^{32}P on backing paper.

      Satisfactory images of wet, unfixed gels can also be obtained as long as the gels are sealed in a plastic bag or wrapped in Saran Wrap before they are exposed to the film (please see Chapter 5, Protocol 11). ^{32}P-labeled nucleic acids in agarose gels can be detected by exposing the wet gel (wrapped in Saran Wrap) to X-ray film. However, for maximal sensitivity and resolution, transfer the radiolabeled nucleic acids to a solid support (nitrocellulose or nylon membrane) as described in Chapter 6, Protocol 8. Dry the solid supports and cover with Saran Wrap to prevent contamination of intensifying screens and film holders.

2. Place pieces of tape marked with radioactive ink around the edge of the sample on the backing paper or Saran Wrap. Cover the pieces of tape with Scotch Tape. This arrangement prevents contamination of the film holder or intensifying screen with the radioactive ink.

   Alternatively, attach luminescent labels to the paper or Saran Wrap. These may be purchased from several manufacturers (e.g., Stratagene).

3. In a darkroom, place the sample in a light-tight X-ray film holder and cover it with a sheet of X-ray film. If preflashed film is used, the preexposed side should face the sample; if an intensifying screen is used, the preexposed side should face the intensifying screen.
4. Expose the film for an appropriate length of time (see Table A9-1). When conventional intensifying screens or fluorography is used, the film must be exposed at –70°C. The low temperature stabilizes the silver atoms and ions that form the latent image of the radioactive source.
5. Remove the film holder from storage (use gloves to handle holders stored at –70°C). In a darkroom, remove the film as quickly as possible and allow it to warm up to room temperature before developing.

   If it is necessary to obtain another autoradiograph, apply another film immediately and return the film holder and screens to the freezer as rapidly as possible. If condensation forms before the new film can be applied, allow the sample and screens to reach room temperature and wipe away all condensation before applying the new film.
6. Develop the X-ray film either in an automatic X-ray film processor or by hand as follows:

   X-ray developer
   5 minutes

   3% acetic acid stop bath or water bath
   1 minute

   rapid fixer
   5 minutes

   running water
   15 minutes

   The temperatures of all solutions should be 18–20°C.
7. Use the images of the radioactive or luminescent markers to align the autoradiograph with the sample.

Phosphorimaging

Autoradiography has been the mainstay of molecular cloning for many years. Recently, however, two types of phosphorimaging devices have become available that create images of radiation sources on computer screens rather than on conventional photographic film. One type of device (area detector) scans the gel or filter in small windows with a Geiger counter, compiling a contour map of the number of radioactive disintegrations per unit area. The other device uses plates coated with a light-responsive phosphor. The film or filter is directly exposed to the plates, and the energy emitted is stored in a europium-based coating. The plates are then scanned by a laser, releasing photons that are collected to form an image. Both devices present the image on a computer screen. These instruments are more expensive than those required for conventional autoradiography and the images have a lower resolution. However, the images can be detected in ~10–20% of the time required by conventional autoradiography and a darkroom is not required. In addition, the linear range of imaging instruments extends over about five orders of magnitude, an improvement of at least 100-fold over conventional X-ray film. Densitometric analysis of images is therefore more accurate and far simpler. Details of phosphorimaging methodology will vary with the device used and manufacturers' instructions should be followed. The images are captured electronically and can be stored and prepared for publication using programs such as Adobe Photoshop (Adobe Systems Incorporated).

Isotopic Data

TABLE A9-2 Isotopic Data

^{3}H		^{35}S		^{32}P	
Time (Years)	% Activity Remaining	Time (Days)	% Activity Remaining	Time (Days)	% Activity Remaining
1	94.5	2	98.4	1	95.3
2	89.3	5	96.1	2	90.8
3	84.4	10	92.3	3	86.5
4	79.8	15	88.7	4	82.4
5	75.4	20	85.3	5	78.5
6	71.3	25	82.0	6	74.8
7	67.4	31	78.1	7	71.2
8	63.7	37	74.5	8	67.8
9	60.2	43	71.0	9	64.7
10	56.9	50	67.0	10	61.5
11	53.8	57	63.6	11	58.7
12	50.9	65	59.6	12	55.9
12.3	50.0	73	56.0	13	53.2
		81	52.5	14	50.7
		87.1	50.0	14.3	50.0

^{125}I		^{131}I		^{33}P	
Time (Years)	% Activity Remaining	Time (Days)	% Activity Remaining	Time (Days)	% Activity Remaining
4	95.5	0.2	98.3	2	94.7
8	91.2	0.4	96.6	4	89.7
12	87.1	0.6	95.0	6	84.9
16	83.1	1.0	91.8	8	80.4
20	79.4	1.6	87.2	10	76.1
24	75.8	2.3	81.2	12	72.1
28	72.4	3.1	76.7	14	68.3
32	69.1	4.0	71.0	16	64.6
36	66.0	5.0	65.2	18	61.2
40	63.0	6.1	59.3	20	57.9
44	60.2	7.3	53.4	22	54.9
48	57.4	8.1	50.0	24	52.0
52	54.8			25.4	50.0
56	52.4				
60	50.0				

One Curie (Ci) is equivalent to the amount of an isotope undergoing 3.7×10^{10} nuclear disintegrations/second (2.22×10^{12} disintegrations/minute). 1 Ci = 3.7×10^{10} becquerels (Bq).

1 Bq = 2.7×10^{-11} Ci
1 μCi = 37×10^{3} Bq = 37 kBq = 2.22×10^{6} dpm
1 mCi = 37×10^{6} Bq = 37 MBq = 2.22×10^{9} dpm
1 Ci = 37×10^{9} Bq = 37 GBq = 2.22×10^{12} dpm

CHEMILUMINESCENCE

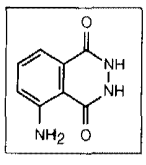
Larry J. Kricka
Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia

Radioactive reagents have been gradually replaced by nonisotopic reagents for some tasks in molecular biology. Concern over laboratory safety and the economic and environmental aspects of radioactive waste disposal have been key factors in this change. Generally, the new nonisotopic systems have improved in terms of analytical sensitivity and the time required to obtain a result. The most prominent nonisotopic analytical methods exploit chemiluminescence. This technique has been particularly effective when used in combination with an enzyme label, so that the amplifying properties of an enzyme label and the high sensitivity of a chemiluminescent detection reaction are combined to produce an ultrasensitive assay (e.g., chemiluminescent detection of peroxidase- and alkaline-phosphatase-labeled proteins and nucleic acid probes). In all of the commonly used applications in molecular biology, the analytical performance of the chemiluminescent systems approaches that of ^{125}I- or ^{32}P-based systems. Chemiluminescent systems also avoid the lengthy signal detection times required with ^{32}P-based methods, yielding results in minutes rather than days. In addition, chemiluminescent probes can be easily stripped from membranes, allowing the membranes to be reprobed many times without significant loss of resolution. Experimental protocols for directly attaching nonisotopic labels to nucleic acids and indirect labeling methods based on biotin, fluorescein, and digoxigenin labels are now well established. The ancillary reagents (e.g., avidin, streptavidin, antidigoxigenin, and antifluorescein enzyme conjugates) required for the indirect methods are widely available. In addition, several companies have developed complete kits of labeling and detection reagents to simplify and facilitate the application of the chemiluminescent assays.

Chemiluminescence is the light emission produced in certain chemical reactions as a result of the decay of chemi-excited intermediates to the electronic ground state. Most chemiluminescent reactions are oxidation reactions because the production of visible light requires highly energetic reactions (63.5 kcal/mole for visible light at 450 nm) (please see the panel on **CHEMILUMINESCENT REACTIONS**).

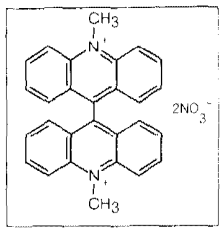
CHEMILUMINESCENT REACTIONS

Acridinium ester + peroxide + base
Adamantyl 1,2-dioxetane aryl phosphate + alkaline phosphatase
Lucigenin + peroxide + base
Luminol + peroxide + base
Nitric oxide + ozone
bis(2,4,6-trichlorophenyl)oxalate + peroxide + fluorescer



Generally, chemiluminescent reactions are inefficient, especially in aqueous environments, and the chemiluminescent quantum yields are typically <10%. Despite such inefficiency, this type of reaction is analytically useful and there are many highly sensitive assays based on compounds with quantum yields of only 1% (e.g., luminol [5-amino-2,3-dihydro-1,4-phthalazinedione]; the chemical structure of luminol is shown at the left).

Chemiluminescence has a long history (for a review, please see Campbell 1988) and some of the compounds in routine use today have been known for a long time. Luminol was first synthesized in 1853 (its chemiluminescent properties were not recognized until 1928), and lucigenin



(bis-*N*-methylacridinium nitrate; the chemical structure of lucigenin is shown at the right) was synthesized in 1935. Chemiluminescent reactions are known in the gas phase (reaction of nitric oxide and ozone), liquid phase (e.g., luminol oxidation reaction), and solid phase (e.g., phosphorous oxidation reaction) (Gundermann and McCapra 1987; Campbell 1988; Van Dyke and Van Dyke 1990).

There are numerous applications for chemiluminescent reactions ranging from the familiar emergency lighting that exploits the fluoresence-sensitized peroxyoxalate oxidation reaction (Cyalume Lightsticks) to the study of phagocytosis using lucigenin or luminol to enhance weak cellular chemiluminescence (Allen and Loose 1976). In molecular biology, chemiluminescent compounds are used as labels in nucleic acid probe and protein blotting applications (e.g., Southern and western blotting), and as reagents to detect enzyme-labeled nucleic acids and proteins (Tables A9-3 and A9-4) (Kricka 1992; Nozaki et al. 1992). Specific advantages of chemiluminescent assays and protocols are improved sensitivity over conventional radiometric, colorimetric, and fluorometric detection systems, hazard-free reagents, rapid results, and versatile assay formats (e.g., solution- and membrane-based assays).

Chemiluminescent Labels

Acridinium Esters and Related Compounds

Light emission from an acridinium-ester-labeled antigen or antibody, prepared using an activated label (2′,6′-dimethyl-4′-[*N*-succinimidyloxycarbonyl]phenyl 10-methylacridinium-9-carboxylate), is triggered by simply adding a mixture of sodium hydroxide and hydrogen peroxide. The light is emitted as a rapid flash lasting <5 seconds, and this time scale imposes certain constraints on the initiation of light production and its measurement (Weeks et al. 1983; Law et al. 1989). Usually, the light emission is measured by injecting the reagents into the assay tube positioned directly in front of a photodetector in the light-tight measuring chamber of the luminometer. Acridinium esters and the acridinium carboxamide analogs (acridinium-9-[*N*-sulfonyl]carboxamide) (Kinkel et al. 1989; Mattingly 1991) are the principal chemiluminescent labels used in immunoassay (available from Assay Designs Inc, Athens, GA; Behringwerke AG, Marburg, Germany; Ciba Corning Diagnostics, Medfield, MA, and Molecular Light Technology Research Ltd, Cardiff, UK). The detection limit for this type of label is ~0.5 attomole (0.5×10^{-18} moles).

Nonseparation DNA probe assays based on hybridization protection have been devised (Arnold et al. 1989). This type of assay does not require the separation of bound from unbound labeled species and so can be conveniently performed in a single step. The hybridization protection assay format exploits the millionfold difference in the hydrolysis rate of an acridinium-ester-

TABLE A9-3 Chemiluminescent Assays for Immunoassay and Nucleic Acid Hybridization Labels

Enzyme	Substrate	Detection Limit (Zeptomoles)
Acridinium ester	NaOH + peroxide	500
Alkaline phosphatase	AMPPD	1
β-D-galactosidase	AMPGD	30
Horseradish peroxidase	luminol + perborate + 4-iodophenol	5,000
Isoluminol	microperoxidase + peroxide	50,000
Xanthine oxidase	luminol + Fe EDTA	3,000

Bronstein and Kricka (1989); Kricka (1991).

TABLE A9-4 Applications of Chemiluminescence in Molecular Biology

TECHNIQUE	EXAMPLE	REFERENCES
Cell surface molecule analysis	CD2	Meier et al. (1992)
Colony screening	*E. coli* transformed with pSP65 containing N-*ras* proto-oncogene	Stone and Durrant (1991)
DNA fingerprinting	Plant and fungal genomes; forensics	Decorte and Cassiman (1991); Bierwerth et al. (1992)
DNA sequencing	Single vector and multiplex	Beck et al. (1989); Creasey et al. (1991); Martin et al. (1991); Karger et al. (1993)
Dot/slot blots	M13 single-stranded DNA	Stone and Durrant (1991)
Gel mobility shift assay	DNA-binding protein complex AP-1 (Jun/Fos)	Ikeda and Oda (1993)
In situ hybridization	Herpes simplex virus I	Bronstein and Voyta (1989)
Northern blotting	LDL receptor	Höltke et al. (1991)
	IL-6, PGDF	Engler-Blum et al. (1993)
PCR product detection	*bcl-2* t(14:18) chromosomal translocation; *Listeria monocytogenes*	Nguyen et al. (1992); Holmstrom et al. (1993)
Plaque screening	M13mp8 containing N-*ras* proto-oncogene	Stone and Durrant (1991)
Reporter gene	β-D-galactosidase *lacZ* gene	Jain and Magrath (1991)
Reverse transcriptase assay	HIV and lentivirus reverse transcriptase	Cook et al. (1992); Suzuki et al. (1993)
RFLP typing	*Clostridium difficile*	Bowman et al. (1991)
Southern blotting	pBR328	Höltke et al. (1991)
	t-PA	Cate et al. (1991)
	HLA class I antigens	Engler-Blum et al. (1993)
Southwestern analysis	Protein: c-*myb* intron DNA interaction	Dooley et al. (1992)
Western blotting	HIV-1 antibodies; transferrin	Bronstein et al. (1992)

labeled probe that is hybridized to complementary target DNA and labeled probe free in solution. Hydrolysis using a pH 7.6 borate buffer destroys the chemiluminescent property of the label, and the light emission produced after the hydrolysis step is due solely to the hybridized labeled probe (available from Gen-Probe, San Diego, CA).

Luminol and Its Analogs

Luminol was the first chemiluminescent compound to be used as an immunoassay label (Schroeder et al. 1978). Light emission is triggered by adding an oxidant (e.g., hydrogen peroxide) in the presence of a suitable catalyst (horseradish peroxidase, microperoxidase, ferricyanide). However, labeling via the 5-amino group of luminol reduced light emission by a factor of 10. Isoluminol, the 6-amino isomer of luminol, is less efficient than luminol (quantum yield 0.1%), but labeling at the 6-position increases the light emission by a factor of 10, and thus this compound, and its amino-substituted analogs, such as *N*-(4-aminobutyl)-*N*-ethylisoluminol (ABEI), have become the favored labels for immunoassay applications (Kohen et al. 1979; Pazzagli et al. 1982).

Pyridopyridazines represent a separate class of chemiluminescent compounds. Early data indicate that these compounds, particularly the 8-amino-5-chloro-7-phenyl and 8-hydroxy-7-phenyl derivatives, will be useful as labels and co-substrates for detection of peroxidase labels. Compared to luminol, these compounds have a much more intense chemiluminescence (~50-fold) (Masuya et al. 1992).

Chemiluminescent Enzyme Assays

Alkaline Phosphatase

Adamantyl 1,2-dioxetane aryl phosphates (e.g., AMPPD; disodium 3-(4-methoxyspiro[1,2-dioxetane-3,2′-tricyclo[3.3.1^{3,7}]decan]-4-yl)-phenylphosphate) and the 5-substituted analogs (e.g., 5-chloro: CSPD; available from Tropix Inc.) have become extremely popular as chemiluminescent substrates for alkaline phosphatase labels (Bronstein et al. 1989, 1990, 1991; Schaap et al. 1989). The detection limit for the enzyme is 1 zeptomole (10^{-21} moles) and the light emission is long-lived (>1 hour), thus making this an ideal system for use with membrane-based assays. The light emission from this reaction can be enhanced by a nylon membrane surface and by certain polymers, for example, polyvinylbenzyl(benzyldimethylammonium) chloride. In the case of nylon, the enhancement is due to sequestering of the dephosphorylated intermediate in hydrophobic domains; these stabilize and minimize nonluminescent decomposition of the intermediate. Chemiluminescent assays for alkaline phosphatase labels are now used widely for blotting and DNA sequencing (Beck and Köster 1990; Tizard et al. 1990).

β-galactosidase

Adamantyl 1,2-dioxetane aryl galactoside substrates (AMPGD) for this enzyme are increasing in popularity. The enzyme cleaves the galactoside group from the 3-position of the aromatic ring to produce a phenoxide intermediate, and this compound decomposes to produce light. The detection limit for the enzyme using this assay is 30 zeptomoles.

Horseradish Peroxidase

Luminol and other cyclic diacylhydrazides serve as chemiluminescent cosubstrates for horseradish peroxidase. The basic isoenzyme of horseradish peroxidase can be assayed in amounts <5 attomoles (5×10^{-18} moles) using an assay reagent comprising luminol, hydrogen peroxide, and an enhancer (e.g., 4-iodophenol or 4-hydroxycinnamic acid) (Whitehead et al. 1983; Thorpe et al. 1985; Thorpe and Kricka 1986). Enhancement of the acidic isoenzymes of peroxidase is much less effective. The role of the enhancer is to increase the intensity of the light emission and reduce background light emission due to oxidation of luminol by peroxide or other oxidants. This dual effect has dramatic impact on the detection of peroxidase activity and increases the ratio of signal to background by several thousandfold (the enhanced chemiluminescent assay reagents are available from Amersham). This sensitive assay for peroxidase (>10-fold more sensitive than a colorimetric assay) has been combined effectively with the catalyzed reporter deposition (CARD) protocol (Wigle et al. 1993). In this amplification scheme, a peroxidase label reacts with a biotin tyramine substrate to produce highly reactive radical products that react with the label and any protein in the immediate vicinity of the label. Next, deposited biotin groups are reacted with streptavidin-peroxidase (in this way, the original peroxidase label is amplified manyfold), and bound peroxidase is detected using the enhanced chemiluminescent assay. Significant improvements in sensitivity were achieved using the combination of CARD and chemiluminescent detection as opposed to colorimetric detection of the deposited peroxidase.

Xanthine Oxidase

This enzyme can be assayed using a mixture of luminol and an iron EDTA complex (Baret et al. 1990; Baret and Fert 1990). The assay is sensitive (detection limit 3 attomoles), and one notable advantage is that the light emission from the xanthine-oxidase-catalyzed chemiluminescent reaction is very long-lived (>96 hours).

Glucose Oxidase

Several chemiluminescent assays for glucose oxidase have been developed. Isoluminol or luminol in the presence of a microperoxidase catalyst can be used to assay peroxide produced by the action of glucose oxidase on glucose (Sekiya et al. 1991); alternatively, the peroxide can be measured using the chemiluminescent fluorophore-sensitized bis(2,4,6-trichlorophenyl) oxalate reaction (Arakawa et al. 1982).

Commercial Reagents, Kits, and Luminometers

Comprehensive surveys of available chemiluminescent reagents and kits and luminometers for the measurement of light emission have been published (please see Stanley 1992, 1993). Also available are a series of compilations of references to current developments in both the fundamental and applied aspects of chemiluminescence (please see Kricka and Stanley 1992; Kricka et al. 1993; Wilkinson 1998). Chemiluminescence can be detected using a range of measuring devices, including a photomultiplier tube (in photon counting or less sensitive photon current mode) or silicon photodiode, or it can be imaged using a CCD camera (Wick 1989) or photographic film (Kricka and Thorpe 1986). CCD cameras are gaining in popularity because they are a convenient and sensitive means of detecting light emission from a two-dimensional source such as a membrane or a 96-well microplate. In addition, the kinetics of light emission are easily monitored, and image enhancement and background subtraction improve the quality of the results.

BIOLUMINESCENCE

Bioluminescent organisms generate light from chemical reactions that are catalyzed by enzymes called luciferases. Many of these reactions are extremely efficient and produce about one photon per luciferase cycle. In the laboratory, light generated by the action of luciferases can be captured, amplified, and measured in a luminometer. A sensitive instrument can detect the light produced by the activity of 2×10^4 luciferase molecules (10^{-20} moles), a level that is several orders of magnitude more sensitive than any non-light-producing enzymatic reaction. Because assays for luciferases are simple to perform, luciferase genes have become widely used as reporters which, when linked to appropriate regulatory elements, provide accurate measurements of the level of gene expression. In the short space of 3 or 4 years, luciferases have replaced chloramphenicol acetyltransferase (CAT) as the reporter system of choice. Assays for luciferase are more sensitive, faster, and less expensive than CAT assays; they do not require the use of radiolabeled compounds and may not involve the destruction of the host cells.

Luciferases are widely distributed in nature and are diverse in structure. The best studied are those of the common North American firefly *Photinus pyralis* (*Photinus* luciferin 4-mono-oxygenase) and the marine microorganism *Vibrio harveyi* (alkanal, reduced-FMN:oxido oxidoreductase). The properties and substrate specificities of these two luciferases are very different.

Firefly Luciferase

Firefly luciferase catalyzes the oxidative decarboxylation of D(−) luciferin in the presence of ATP-Mg^{2+} to generate oxyluciferin and light:

$$\text{luciferin} + \text{ATP-Mg}^{2+} + \text{luciferase} \rightarrow \text{luciferase:luciferyl-AMP} + PP_i$$

$$\text{luciferase:luciferyl-AMP} + O_2 \rightarrow \text{luciferase} + \text{oxyluciferin} + CO_2 + \text{AMP} + h\nu\ (562\ \text{nm})$$

Luciferin is a generic term for substrates that generate light during oxidation catalyzed by luciferases. Firefly luciferin, 6-hydroxybenzothiazole [D-(−)-(6′hydroxy-2′benzothiazolyl)D^2-thiazoline-4-carboxylic acid] was first isolated in pure form from fireflies in 1957 by Bitler and McElroy; 9 mg of pure luciferin was obtained from 15,000 fireflies. Nowadays, luciferin is synthesized chemically.

Properties of Firefly Luciferase

Firefly luciferase is a 521-amino-acid protein with a predicted molecular weight of 57,000 (de Wet et al. 1987) that is targeted to peroxisomes in all organisms in which the protein is expressed (Keller et al. 1987; for review, please see Gould and Subramani 1988). The membrane-bound nature of peroxisomes undoubtedly limits access of substrates (luciferin and ATP) to the enzyme and may account for the low level of light produced when firefly luciferase is assayed in intact mammalian cells (de Wet et al. 1987; Gould and Subramani 1988). Deletion or mutation of the three carboxy-terminal amino acid residues generates cytosolic forms of the enzyme. Several of these cytoplasmic mutants appear to retain full enzymatic activity (Gould and Subramani 1988), but it is not yet known whether they are more accessible to substrates that are added to the extracellular medium.

In addition to tolerating changes at its carboxyl terminus, firefly luciferase will accept alterations to its amino terminus. The enzyme has been expressed in *E. coli* (1) as a fusion protein that lacks the first six amino acids of luciferase and contains eight amino-terminal residues encoded by the expression vector and the synthetic oligonucleotide used for cloning (de Wet et al. 1985) and (2) as a 92-kD fusion with a modified *Staphylococcus aureus* A protein (Subramani and

DeLuca 1988). A series of eukaryotic expression vectors for measuring promoter strength has been developed by Promega. A particularly useful version, the Dual Luciferase Reporter Assay System (DRL), encodes two luciferases, one from the firefly and the other from the sea pansy (*Renilla reniformis*). This dual reporter assay provides a convenient internal standard for normalization of gene expression measurements.

Firefly luciferase, isolated either from fireflies or from *E. coli* expressing a cloned copy of the luciferase gene, is available commercially (Boehringer Mannheim, Sigma, and Promega). A solution of purified firefly luciferase (1 mg/ml) has an absorbance at 280 nm of 0.75 (DeLuca and McElroy 1978).

Assays for Firefly Luciferase

Luminometry. Addition of ATP and luciferin to preparations of firefly luciferase generates a flash of light that peaks 0.3 second later and lasts for a few seconds. Within 1 minute after mixing substrate and enzyme, the intensity of the emitted light falls to ~10% of peak values and then declines more slowly over a period of several minutes. The decay in light emission is caused by slow turnover of the enzyme and product inhibition by pyrophosphate (DeLuca and McElroy 1978). Dissociation of the enzyme-product complex occurs more efficiently in the presence of acetyl–coenzyme A (CoA) (Wood 1991) and nonionic detergents such as Triton X-100 (Kricka and DeLuca 1982). The concentration of ATP in the assay affects the intensity of light produced in the two phases of the reaction. Firefly luciferase has two distinct, catalytically active ATP-binding sites. One site is responsible for the initial flash, whereas the second, which has a higher affinity for ATP, is involved in the continuous production of light of lower intensity (DeLuca and McElroy 1984). Assays for luciferase generally contain concentrations of ATP that are sufficient to saturate both sites of the enzyme. Under these conditions, the amount of light emitted during the initial flash is proportional to the amount of enzyme in the reaction mixture over five orders of magnitude. Because of the short duration of the initial flash, most assays for luciferase require special luminometers that are designed to allow injection and rapid mixing of reagents and immediate analysis of the emitted light. A moderately priced luminometer can detect as little as 0.03–0.10 pg of luciferase.

An improved luciferase assay system is available from Promega. The system generates a burst of light whose intensity remains nearly constant for ~20 seconds and then decays slowly with a half-time of 5–10 minutes. These more favorable kinetics are achieved by including Triton X-100 and acetyl-CoA in the assay system. Luciferin is oxidized more efficiently by luciferyl-CoA than by luciferyl-AMP (Wood 1991). The emitted light can be measured either in a luminometer or in a liquid scintillation counter.

Liquid scintillation spectroscopy. If a luminometer is not available, firefly luciferase can be assayed in a conventional liquid scintillation counter. Reaction conditions have been reported that minimize the intensity of the initial flash and optimize the long-lasting emission of low-intensity light by firefly luciferase (Nguyen et al. 1988). The reactions are carried out in microfuge tubes or 96-well plates (Schwartz et al. 1990), and the emitted light is measured in a channel of a scintillation counter that detects chemiluminescence. Collection of data begins a few seconds after the reagents have been mixed and continues for 3–5 minutes. Under optimal conditions, measurement of long-lasting emission of light in a scintillation counter is as sensitive an assay for firefly luciferase as luminometry.

Some scintillation counters are not equipped with a preset channel for detection of chemiluminescence. However, the standard channels used for counting of atomic disintegration can be used for assay of luciferase provided the coincidence circuit is turned off. The counts per minute are then proportional to the intensity of luminescence. If the coincidence circuit cannot be turned

off, the intensity of luminescence will be proportional to the square root of the counts per minute after background has been subtracted.

Photographic and X-ray film. Firefly luciferase can be detected by exposing the light-emitting reaction to X-ray or photographic films (de Wet et al. 1986; Wood and DeLuca 1987). This method is useful for preliminary screening of samples for the presence or absence of firefly luciferase (e.g., screening tissue extracts of transgenic animals or estimating the relative efficiencies of a series of mutant promoters).

Firefly Luciferase as a Reporter Molecule

The use of luciferase as a reporter gene follows the same basic strategy developed for other reporter gene systems. In brief, a segment of DNA containing putative regulatory *cis*-acting elements is inserted into a plasmid upstream of a cloned copy of luciferase cDNA. The chimeric construct is used to transfect cultured cells of the appropriate type, and luciferase activity is assayed some time later in extracts of the transfected cells. When the assay is carried out in the presence of excess substrates, the amount of luminescence is proportional to the concentration of newly synthesized enzyme. Because extracts are likely to contain ATP, whose concentration is unknown, luciferase is generally measured by adding luciferin as the last substrate. Promoterless expression vectors for luciferase have been described by Nordeen (1988). Another suitable plasmid, pGEM-*luc*, sold by Promega, contains the firefly luciferase gene positioned in the center of the multiple cloning region of pGEM-11Zf(−).

Cultures of mammalian cells (2×10^5 cells) transfected by the calcium phosphate method with a reporter construct in which expression of firefly luciferase is driven by a moderately strong promoter produce ~15–50 ng of luciferase. Luciferase activity is stable for several weeks in extracts of cells stored at 4°C.

Luciferase was first used in 1988 as a reporter for promoter activity in transgenic mice (DiLella et al. 1988). Since then, it has been used sucessfully as a reporter to measure the tissue specificity, developmental expression, and strength of more than a dozen different promoters in transgenic animals. Luciferase is commonly used as a reporter to locate and analyze regulatory elements in mammalian genes (e.g., please see van Zonneveld et al. 1988; Economou et al. 1989; Hudson et al. 1989). In addition, luciferase has been used (1) to identify proteins that influence gene transcription (Waterman et al. 1988; Mellon et al. 1989), (2) to investigate the effects of mRNA structure on protein synthesis (Baughman and Howell 1988; Malone et al. 1989), and (3) to measure rates of intracellular protein recycling (Nguyen et al. 1989). Luminometric measurements can detect ~5×10^5 molecules of luciferase in 10 mg of protein in a crude tissue extract, which is a fivefold increase in sensitivity over assays for CAT activity in tissue extracts (Robinson et al. 1989).

Bacterial Luciferase

Luciferase of *Vibrio harveyi* is a heterodimer whose α and β subunits (355 and 324 amino acids, respectively) are encoded by the bacterial *luxA* and *luxB* genes. These genes, which are part of an operon (Belas et al. 1982; for review, please see Ziegler and Baldwin 1981), have been cloned, sequenced, and expressed, first in *E. coli* (Baldwin et al. 1984; Cohn et al. 1985; Johnston et al. 1986) and then in plants, where the *luxA* and *luxB* cistrons were placed downstream from two separate promoter elements (Koncz et al. 1987). The termination codon of *luxA* and the short intercistronic distance separating *luxA* and *luxB* have been replaced by sequences encoding polypeptide linkers (Boylan et al. 1989; Kirchner et al. 1989). The resulting fused *luxA* and *luxB* genes express highly active luciferase that can be assayed in yeast, *E. coli*, and plant cells. This abil-

ity to express bacterial luciferase as a single polypeptide removes a major obstacle to its use as a reporter molecule in mammalian cells. Bacterial luciferase catalyzes the oxidation of long-chain aldehydes into carboxylic acids using $FMNH_2$ as a cofactor, producing photons as one of the reaction products:

$$RCHO + O_2 + FMNH_2 \rightarrow RCOOH + FMN + H_2O + h\nu\ (490\ nm)$$

where R is an aliphatic moiety containing at least seven carbon atoms; FMN is a flavin mononucleotide; and $FMNH_2$ is a reduced flavin mononucleotide. Bacteria contain enough $FMNH_2$ to drive this reaction, and light emission can therefore be measured directly from intact cells, allowing gene expression to be assayed in real time (e.g., please see Legocki et al. 1986). In yeasts, the intracellular concentration of $FMNH_2$ is much lower, and the intensity of in vivo luminescence is therefore far lower than in intact bacteria. However, the level of light produced in intact yeast cells expressing the bacterial *lux* genes is still 1000-fold over background and significantly above that reported for firefly luciferase (Tatsumi et al. 1988). Intact cells of higher eukaryotes do not contain enough $FMNH_2$ to drive the reaction catalyzed by the fused *luxAB* genes. However, addition to cell extracts of $FMNH_2$ and an $FMNH_2$-generating system (e.g., NAD[P]H:FMN reductase; available from Boehringer Mannheim) generates high levels of luminescence.

Although side-by-side comparisons of efficiency have not been reported, a *luxAB* fusion has two theoretical advantages over firefly luciferase as a reporter molecule: *luxAB* is a cytosolic molecule, and its substrate, the aldehyde decanal, penetrates membranes easily. Although the quantum yield of the photochemical reaction catalyzed by *luxAB* is less than that of firefly luciferase, *luxAB* may turn out to be the more sensitive reporter in intact cells. For a summary of chromogenic and luminescent methods of detection of immobilized antigens, please see Appendix 8, Table A8-12.

Green Fluorescent Protein

The bioluminescent jellyfish, *Aequorea victoria* emits a characteristic green fluorescence, which is due to the interaction of two proteins, the calcium-binding photoprotein aequorin, and the green fluorescent protein (GFP). The emission spectrum of GFP peaks at 508 nm (Johnson et al. 1962), a wavelength close to that of living *Aequorea* tissue, but distinct from the chemiluminescence of pure aequorin, which is blue and peaks near 470 nm. With the initial purification and crystallization of GFP, it was discovered that calcium-activated aequorin could efficiently transfer its luminescent energy to GFP when the two were co-adsorbed onto a cationic support (Morise et al. 1974). Green light is produced when energy is transferred by a Föster-type mechanism from Ca^{2+}-activated aequorin to GFP. Blue light emitted by activated aequorin is captured by a hexapeptide chromophore (beginning at residue 64 of GFP) that contains a cyclic structure (4-[*p*-hydroxybenzylidene]imidazolidin-5-one) attached to the peptide backbone through the 1-position and 2-position of the ring (Shimomura 1979). Interestingly, this structure appears to be conserved among fluorescent proteins, even those from nonbioluminescent organisms (Matz et al. 1999).

The GFP gene was cloned in 1992 (Prasher et al. 1992) and is now used as a reporter in a wide range of organisms (e.g., please see Chalfie et al. 1994; Wang and Hazelrigg 1994; Marshall et al. 1995; Yeh et al. 1995; Chiu et al. 1996; Niedenthal et al. 1996; Misteli et al. 1997). Expression levels of GFP can be measured accurately using a fluorometer. Fluorescence is measured directly in intact living cells in 96-well plates and expression is assessed on an individual cell basis using fluorescence microscopy. Engineered variants of GFPs that fluoresce at different wavelengths and with different intensities have been developed and are available from various commercial sources (e.g., Research Genetics and CLONTECH). For more information on GFP and its applications in molecular biology, please see the information panel in Chapter 17.

ANTIBODIES

Antibodies are used for several major purposes in molecular cloning:

- to screen cDNA libraries for clones that express a specific protein
- to purify fusion proteins that are tagged with a specific epitope
- to detect and quantify foreign protein expressed by recombinant DNA techniques
- to confirm that a cloned gene encodes a protein of interest
- to detect probes labeled with biotin or digoxygenin
- to detect mRNAs by immunohistochemical analysis

Antibodies come in a variety of classes, affinities, and idiotypes. They can be polyclonal or monoclonal, engineered or natural, and raised in animals or generated in vitro. However, as far as molecular cloning is concerned, the specificity of the antibody is far more important than its provenance. Antibodies can be divided into three groups according to how they react with their target proteins.

- ***Antibodies that react with the foreign protein independently of its conformation.*** Antibodies of this type are particularly useful for measuring the total amount of the target protein in crude preparations or in cell extracts. They are usually polyclonal in nature and are prepared by immunizing animals with partially denatured protein or with a peptide whose sequence corresponds to part of the intact protein. However, monoclonal antibodies that are pan-specific are not uncommon.
- ***Antibodies that react only with epitopes specific to the native form of the target protein.*** Antibodies of this type are typically monoclonal, have been raised against native protein, and recognize a given sequence of amino acids only when it occurs in its native three-dimensional configuration. These antibodies are useful for testing whether mutated forms of proteins that have been generated by in vitro mutagenesis are folded correctly or whether a wild-type protein expressed in heterologous cells is assembled into a correct three-dimensional configuration.
- ***Antibodies that react only with denatured forms of the target protein.*** These antibodies are raised against fully denatured antigens and can be either monoclonal or polyclonal. Antibodies of this type are useful for western blotting and for immunological screening of cDNA libraries.

Although there is no way to guarantee the production of particular types of antibodies, it is nevertheless possible to choose an immunization regimen that will favor the production of antibodies with the desired characteristics. However, it is always necessary to screen several independent antisera or a series of monoclonal antibodies to identify those suited to the tasks at hand.

Purification of Antibodies

For many purposes, antisera need not be fractionated before use. However, if the antisera are to be radiolabeled or conjugated to enzymes such as horseradish peroxidase (HRP) or alkaline phosphatase, it is necessary to purify the IgG fraction or, in some cases, to purify the antibody of interest by binding it to its cognate antigen.

Although many techniques have been developed to purify IgG molecules, the method of choice is adsorption to, and elution from, beads coated with protein A, a component of the cell wall of *S. aureus* (Hjelm et al. 1972). For reasons that are not known, this protein (M_r = 42,000)

binds strongly to sites in the second and third constant regions of the Fc portion of the immunoglobulin heavy chain (Deisenhofer 1981). Each IgG molecule therefore contains two binding sites for protein A. Because protein A itself has five potential sites for binding to IgG (Sjödahl 1977), it is possible to form multimeric complexes of the two types of proteins.

Not all immunoglobulins bind to protein A with the same affinity. Antibodies from humans, rabbits, and guinea pigs bind most tightly, followed in decreasing order of affinity by those from pigs, mice, horses, and cows (Kronvall et al. 1970a,b; Goudswaard et al. 1978). Immunoglobulins from goats, rats, chickens, and hamsters bind in a much weaker fashion, and a "bridging" antibody is usually required to purify them by adsorption to protein A.

Within any one species, different classes of immunoglobulins vary in the sequences of their Fc regions and consequently bind to protein A with different affinities. Of the major classes of human IgG, for example, three (IgG1, 2, and 4) bind with high affinity and one (IgG3) binds very weakly, if at all. Similarly, mouse IgG2a binds with high affinity, IgG2b and IgG3 bind tolerably well, and IgG1 binds poorly (Ey et al. 1978). These differences are generally unimportant when dealing with polyclonal sera, where antibodies against the target antigen are distributed throughout all of the major subclasses of IgG. Consequently, purification of polyclonal immunoglobulins raised in rabbits, humans, and mice by binding to protein A may alter the distribution of subclasses of IgG, but it rarely changes the specificity or avidity of the final preparation. However, monoclonal antibodies secreted from hybridomas carry only one subclass of heavy chain. Before attempting to purify a given monoclonal antibody, it is essential to determine the subclass of its heavy chain using commercially available immunological reagents directed against isotypes of the Fc region. If the monoclonal antibody falls into a class that binds poorly to protein A (e.g., by binding to protein G or to human IgG3 or mouse IgG1), it should be purified by another method (e.g., ammonium sulfate precipitation, followed by chromatography on DEAE-cellulose). Alternatively, a bridging antibody can be used to attach the monoclonal antibody to protein A.

Protein A coupled to a solid support by cyanogen bromide is supplied by several manufacturers (e.g., protein A–Sepharose CL-4B; Pharmacia). Each milliliter of swollen gel can bind ~10–20 mg of IgG (equivalent to 1–2 ml of antiserum). Antibodies bind to protein A chiefly by hydrophobic interactions (Deisenhofer 1981) that can be disrupted at low pH. Protein A is remarkably resilient and withstands repeated cycles of exposure to low pH extremely well; it can also be treated with high concentrations of denaturing agents such as urea, guanidine hydrochloride, or potassium isothiocyanate without permanent damage. Most antibodies can withstand transient exposure to low pH, and this treatment is now the standard method to release them in an active form from protein A–Sepharose beads. For additional information on protein A, protein G, and protein L, please see the information panel on **IMMUNOGLOBULIN-BINDING PROTEINS: PROTEINS A, G, AND L** at the end of this appendix.

Although hyperimmune antisera raised in experimental animals contain very high concentrations of immunoglobulin directed against the target antigen, such antisera also always contain antibodies directed against other antigens. In addition, the immunoglobulins in antisera may bind with low avidity to molecules that are not true target antigens. For these and other reasons, antisera can manifest a level of background reactivity that is unacceptably high. There are three ways to deal with this problem:

- ***Use an innocuous blocking agent (e.g., bovine serum albumin, normal serum, or BLOTTO)*** to compete with the immunoglobulin for nonspecific binding sites. Blocking agents are routinely included in solutions used, for example, in immunological screening of expression libraries constructed in plasmid or bacteriophage λ vectors.

- ***Remove antibodies that are directed against specific contaminating antigens by adsorption.*** Methods to remove antibodies directed against bacterial antigens are discussed in Chapter 14,

Protocol 4. Antibodies that cross-react with components in eukaryotic cells can be adsorbed with acetone extracts of a cell line or tissue that are known not to express the true target antigen. If such a line cannot be identified with certainty, use an acetone extract of commercially available dried yeast (Sambrook et al. 1989). Alternatively, unwanted antibodies can be adsorbed to antigens immobilized on nitrocellulose or PVDF membranes.

- ***Separate antibodies directed against the target antigen from contaminating antibodies by affinity purification.*** In some cases (e.g., when the antigen is a protein), the antigen may be coupled to a matrix such as cyanogen-bromide-activated Sepharose. Antibodies directed against epitopes displayed by the protein will be retained by the column; all other immunoglobulins will pass through. The bound antibody is then released from the column by agents that disrupt the antigen-antibody complex (e.g., potassium isothiocyanate and low-pH buffers). Details of the methods used to prepare antibodies by immunochromatography vary from antigen to antigen and from antibody to antibody. However, the general principles are well-described in a number of reviews (e.g., please see Hurn and Chantler 1980; Harlow and Lane 1988, 1999). When using these methods, it is essential to use highly purified antigen and to avoid the batch of antigen that was used to raise the antibody in the experimental animals. Furthermore, it is important to remember that antibodies with different affinities for the antigen will show different patterns of elution from the column; those that bind loosely to the antigen will elute first, and those that bind most tightly will elute last. In fact, antibodies with the highest avidity may be denatured by the elution buffer before they dissociate from the antigen. Thus, there is a tendency during immunopurification to select for antibodies that are specific for the antigen but that bind with low affinity.

Antibodies may be purified on a small scale by adsorption to, and elution from, protein antigens that are immobilized on diazotized paper (Olmsted 1981), or nitrocellulose filters (Burke et al. 1982; Smith and Fisher 1984; Earnshaw and Rothfield 1985), after electrophoresis through SDS-polyacrylamide gels. Antibodies prepared by this method are especially useful for confirming the identity of cDNA clones isolated from expression libraries constructed in bacteriophage λ or plasmid expression vectors. For example, many false-positive clones can be eliminated by purifying antibodies from crude sera by virtue of their ability to bind to a fusion protein partly encoded by the cloned cDNA and testing the ability of these purified antibodies to precipitate the target protein or to react with it on a western blot. However, this method works well only when the antibodies react with epitopes that are displayed on denatured proteins. Typically, ~50 ng of immunopurified antibody is recovered per microgram of target protein loaded on the original SDS-polyacrylamide gel. Because of the idiosyncratic nature of the interactions between antibodies and their target proteins, it is not possible to give conditions for binding and elution that are universally applicable. For example, most antibodies can be eluted from their immobilized antigens with glycine buffer (pH 2.8). However, Earnshaw and Rothfield (1985) found that antibodies to human centromeric proteins could be eluted only with a solution containing 3 M potassium thiocyanate and 0.5 M NH_4OH. Investigators who wish to use this potentially powerful technique should be prepared to invest some effort in defining the optimal conditions for binding and release of their particular antibodies from their target proteins.

Immunological Assays

Antibodies are used in a wide variety of assays, both qualitative and quantitative, to detect and measure the amount of target antigens. These assays include western blotting, immunoprecipitation, and solid-phase radioimmunoassay (RIA).

Western Blotting

Western blotting (Towbin et al. 1979; Burnette 1981; Towbin and Gordon 1984) is to proteins what Southern blotting is to DNA. In both techniques, electrophoretically separated components are transferred from a gel to a solid support and probed with reagents that are specific for particular sequences of amino acids (western blotting) or nucleotides (Southern hybridization). In the case of proteins, the probes usually are antibodies that react specifically with antigenic epitopes displayed by the target protein attached to the solid support. Western blotting is therefore extremely useful for the identification and quantitation of specific proteins in complex mixtures of proteins that are not radiolabeled. The technique is almost as sensitive as standard solid-phase radioimmunoassays and, unlike immunoprecipitation, does not require that the target protein be radiolabeled. Furthermore, because electrophoretic separation of proteins is almost always carried out under denaturing conditions, any problems of solubilization, aggregation, and coprecipitation of the target protein with adventitious proteins are eliminated.

The critical difference between Southern and western blotting lies in the nature of the probes. Whereas nucleic acid probes hybridize with a specificity and rate that can be predicted by simple equations, antibodies behave in a much more idiosyncratic manner. As discussed earlier, an individual immunoglobulin may preferentially recognize a particular conformation of its target epitope (e.g., denatured or native). Consequently, not all monoclonal antibodies are suitable for use as probes in western blots, where the target proteins are denatured. Polyclonal antisera, on the other hand, are undefined mixtures of individual immunoglobulins, whose specificity, affinity, and concentration are often unknown. It is therefore not possible to predict the efficiency with which a given polyclonal antiserum will detect different antigenic epitopes of an immobilized, denatured target protein.

Although there is an obvious danger in using undefined reagents to assay a target protein that may also be poorly characterized, most problems that arise with western blotting in practice can be solved by designing adequate controls. These include the use of (1) antibodies (i.e., preimmune sera, normal sera, or irrelevant monoclonal antibodies) that should not react with the target protein and (2) control preparations that either contain known amounts of target antigen or lack it altogether.

Often, there is little choice of immunological reagents for western blotting — it is simply necessary to work with whatever antibodies are at hand. However, if a choice is available, either a high-titer polyclonal antiserum or a mixture of monoclonal antibodies raised against the denatured protein should be used. Reliance on a single monoclonal antibody is hazardous because of the high frequency of spurious cross-reactions with irrelevant proteins. If, as is usually the case, monoclonal and polyclonal antibodies have been raised against native target protein, it will be necessary to verify that they react with epitopes that either (1) resist denaturation with SDS and reducing agents or (2) are created by such treatment. This test can be done by using denatured target antigen in a solid-phase radioimmunoassay or in western dot blots.

In western blotting, the samples to be assayed are solubilized with detergents and reducing agents, separated by SDS-polyacrylamide gel electrophoresis, and transferred to a solid support (usually a nitrocellulose or PVDF filter), which may then be stained (e.g., with Ponceau S). The filter is subsequently exposed to unlabeled antibodies specific for the target protein. Finally, the bound antibody is detected by one of several secondary immunological reagents (e.g., ^{125}I-labeled or anti-immunoglobulin, or anti-immunoglobulin or protein A coupled to horseradish peroxidase or alkaline phosphatase), followed by autoradiography, enhanced chemiluminescence, or enzymatic production of a colored precipitate. As little as 1–5 ng of an average-sized protein can be detected by western blotting. For more information on western blotting, please see Appendix 8.

Immunoprecipitation

Immunoprecipitation is used to detect and quantitate target antigens in mixtures of proteins. The power of the technique lies in its selectivity: The specificity of the immunoglobulin for its ligand is so high that the resulting antigen-antibody complexes can be purified from contaminating proteins. Furthermore, immunoprecipitation is extremely sensitive and is capable of detecting as little as 100 pg of radiolabeled protein. When coupled with SDS-polyacrylamide gel electrophoresis, the technique is ideal for analysis of the synthesis and processing of foreign antigens expressed in prokaryotic and eukaryotic hosts or in in vitro systems.

The target protein is usually immunoprecipitated from extracts of cells that have been radiolabeled. However, immunoprecipitation can also be used to analyze unlabeled proteins as long as sufficiently sensitive methods are available to detect the target protein after it has been dissociated from the antibody. Such methods include enzymatic activity, binding of radioactive ligands, and western blotting.

Solid-phase Radioimmunoassay

The solid-phase radioimmunoassay (RIA) is a quantitative method that is capable of detecting as little as 1 pg of target antigen. This means that RIAs are sufficiently sensitive to measure, for example, the amount of foreign protein produced by transfected mammalian cell cultures. There are many different kinds of RIAs, which fall into four basic designs:

- ***Competition RIAs:*** In this method, the unlabeled target protein in the test sample competes with a constant amount of radiolabeled protein for binding sites on the antibody. The amount of radioactivity present in the unbound or bound target protein is then measured. This type of assay can be extremely sensitive but requires that target protein be available (preferably in a pure form) to serve both as a competitor and as a standard.
- ***Immobilized antigen RIAs:*** In this method, unlabeled antigen is attached to a solid support and exposed to radiolabeled antibody. Comparison of the amount of radioactivity that binds specifically to the samples under test with the amount that binds to a known amount of immobilized antigen allows the antigen in the test samples to be quantitated. Although used occasionally, this type of assay is not particularly useful for quantitation of small amounts of foreign protein in complex mixtures (e.g., in cell lysates); most of the binding sites on the solid support become occupied by proteins other than the target protein, so that the sensitivity of the assay is comparatively low.
- ***Immobilized antibody RIAs:*** In this method, a single antibody bound to a solid support is exposed to radiolabeled antigen. The amount of antigen in the test sample can be determined by the amount of radioactivity that binds to the antibody. This assay is not useful for quantitating the amount of foreign protein in many different samples, chiefly because of the practical difficulty of radiolabeling the protein either in vivo or in vitro.
- ***Double-antibody RIAs:*** In this method, one antibody bound to a solid support is exposed to the unlabeled target protein. After washing, the target protein bound to the immobilized antibody is quantitated with an excess of a second radiolabeled antibody. This assay is extremely sensitive and specific because the target protein is essentially purified and concentrated by immunoadsorption. Furthermore, many test samples can be processed simultaneously. However, the method requires that the first and second antibodies recognize nonoverlapping epitopes on the target protein. Ideally, the first antibody is monoclonal, whereas the second can be either a polyclonal antibody or a monoclonal antibody of different specificity. However, in

some cases, it may be possible to use the same polyclonal antibody for both parts of the assay. If suitable antibodies are available, this is the method of choice for quantitation of target proteins in complex mixtures.

Radiolabeling of Antibodies

Of the several methods that are available to radioiodinate antibodies, the most commonly used is a reaction in which an electrophilic iodine species is generated by oxidation of $Na^{125}I$ with chloramine-T (*N*-chlorobenzenesulfonamide). The positively charged iodine then reacts with the side chains of tyrosyl, and to a lesser extent, histidyl residues (for review, please see Seevers and Counsell 1982). Iodination by chloramine-T was devised as a method to label small amounts of polypeptide hormones to very high specific activity (Hunter and Greenwood 1962; Greenwood et al. 1963). However, the reaction conditions are so severe as to cause extensive denaturation of proteins that are sensitive to oxidation. The modified procedure devised by McConahey and Dixon (1966, 1980), which uses lower concentrations of chloramine-T and longer reaction times, yields native proteins without a significant loss in the efficiency of radiolabeling. Nowadays, most investigators prefer to use chloramine-T covalently coupled to nonporous polystyrene beads (Iodobeads, Pierce) (Markwell 1982). This has several benefits: It reduces the amount of radiolabeled molecular iodine generated in the reaction and simplifies measurements of the time course and overall efficiency of radioiodination (Cheng and Rudick 1991). In addition, because there is very little contact between the protein and the oxidizing agent, oxidative damage to the protein is minimized. Finally, when iodination is carried out according to standard protocols, no more than one atom of ^{125}I is incorporated per protein molecule. Conformational distortions caused by the introduction of a bulky iodine atom are therefore kept to a minimum.

An alternative to chloramine-T is Iodogen (1,3,4,6-tetrachloro-3α,6α-diphenylglycouril), an oxidizing reagent resembling a fourfold chloramine-T (Fraker and Speck 1978). Iodogen is sparingly soluble in water and forms a thin coating on the wall of the tube used for radioiodination. Solid-phase reagents such as Iodogen and Iodobeads (available from Pierce) give radiochemical yields that are equal to those obtained with free chloramine-T (Woltanski et al. 1990) while significantly reducing the potential for exposure of laboratory workers to volatile forms of radioiodine.

Radioiodination with chloramine-T and similar compounds will not work with proteins and peptides that lack accessible tyrosines or are extremely susceptible to damage by oxidation. These proteins should be labeled by the Bolton-Hunter reagent (Bolton and Hunter 1973). The *N*-succinamidyl group of this reagent (*N*-succinamidyl 3-[4-hydroxy 5-[^{125}I]iodophenyl]propionate) condenses with free amino groups and with imino groups in lysyl side chains to give a derivative in which the radioiodinated phenyl group is linked via an amide bond to the target protein (for more details, please see Langone 1980; Harlow and Lane 1988).

Antipeptide Antibodies

If the sequence of a protein is known or can be deduced from the nucleic acid sequence, specific antisera can be raised by immunizing animals with a synthetic peptide corresponding in sequence to a segment of the native protein (for review, please see Lerner 1984). If information about the primary sequence of the target protein is limited, there may be little or no choice in the peptide sequence used as an immunogen. However, there is a good chance that peptides chosen at random will be at least partially buried in the native protein and they may be too hydrophobic to be

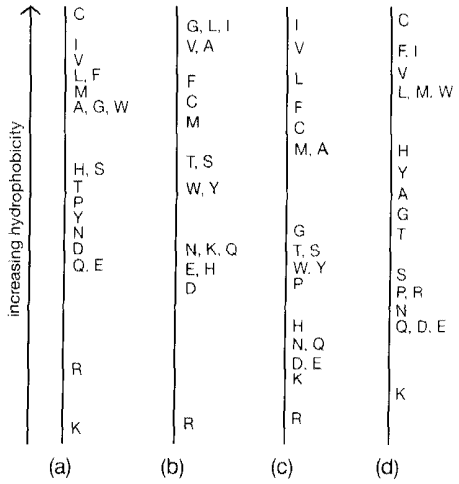


FIGURE A9-3 Hydrophobicity Scales

Several hydrophobicity scales, approximately normalized to each other and increasing from bottom to top. Data from (*a*) Janin (1979); (*b*) Wolfenden et al. (1981); (c) Kyte and Doolittle (1982); and (*d*) Rose et al. (1985). Please also see Figure A7-2 in Appendix 7. (Reprinted, with permission, from Richardson and Richardson 1989 [©Kluwer Academic/Plenum Publishers].)

efficient immunogens. Antibodies directed against these peptides may be of low titer and/or may react only with denatured protein.

Hydrophilic peptides that contain charged residues are much better immunogens and also have a high probability of occupying a surface location on the native protein. Antipeptide antibodies raised against conformationally flexible surface features of proteins such as turns and β-loops are likely to be of high titer and may react efficiently with the native protein. Most of the computing packages that are commonly used to analyze DNA sequences contain programs to search protein sequences for surface peptides that are likely to be good antigens. The goal of these programs is to choose a sequence of 11–15 amino acids that contains a preponderance of polar residues and no more than four adjacent hydrophobic residues. Such peptides are likely to be soluble in aqueous solvents and therefore easy to couple to carrier proteins.

Several scales of hydrophobicity and hydrophilicity have been proposed for amino acids (please see Figure A9-3) (Janin 1979; Wolfenden et al. 1981; Kyte and Doolittle 1982; Rose et al. 1985). Although these scales are in good general agreement, they differ slightly in the order of amino acids and in the relative hydrophobicity values assigned to them. This variability arises because the hydrophobicity of an amino acid residue is the product of several different factors, including electrostatic charge, hydrogen-bonding capability, and surface area. In addition, the hydrophobicity of an amino acid can be assessed experimentally by partitioning into solvents of various types. The variation in hydrophobicity scales of amino acids reflects the particular weightings that different investigators have attached to these and other factors.

Computer programs to predict strongly antigenic sites in proteins rely on hydrophobicity scales alone or in combination with programs that predict secondary structure. The strongest antigenic sites are predicted in regions of the protein surface that are high in charge and low in hydrophobicity. Rarely found in ordered structures such as helices or sheets, such regions usually map to turns and loops that are rich in residues with H-bonding potential.

Highly charged regions also often occur at the amino and carboxyl termini of proteins, which tend to be regions of high flexiblility. However, if the protein of interest is a secretory or transmembrane protein, or is located in organelles such as mitochondria or chloroplasts, it is better not to use peptides derived from the terminal regions. The amino-terminal peptide is likely to be part of a signal or leader sequence that will be cleaved during posttranslational processing. The carboxy-terminal region of transmembrane proteins may be located on the cytoplasmic side of the membrane and may only be accessible to antibodies after cells are permeabilized, fixed, or lysed.

GUIDELINES FOR CHOOSING AN IMMUNOGENIC PEPTIDE

An immunogenic peptide should have:

- a sequence 11–15 amino acids long
- no more than six hydrophobic residues
- no more than four adjacent hydrophobic residues
- no consensus sites for *N*-glycosylation* (N-X-S/T)
- no basic residues if glutaraldehyde is to be used as a cross-linker
- no internal or cysteine residues if m-maleimidobenzoic acid *N*-hydroxysuccinimide ester (MBS) is to be used as a cross-linker
- a tyrosine residue because the peptide can then easily be radiolabeled with ^{125}I

*Oligosaccharide groups added to secretory and plasma membrane proteins can shield the underlying amino acid residues from interacting with antibodies and other ligands. Antibodies generated against a synthetic peptide may therefore not recognize the glycosylated target protein even though the immunogen and the target share amino acid sequences. When raising antibodies against secretory and plasma membrane proteins, it is best to avoid regions that contain consensus sites for glycosylation.

Usually, the synthetic peptide is coupled to a carrier protein such as keyhole limpet hemocyanin through an amino- or carboxy-terminal cysteine residue. Ideally, therefore, the sequence of the peptide should end with a "natural" cysteine that is present in the cognate sequence of the target protein. If this is not possible, choose a peptide that has no cysteine at all. Chemical synthesis of the peptide, which proceeds from the carboxyl terminus to the amino terminus, is then initiated from an "artificial" cysteine that has no counterpart in the natural cognate sequence. This terminal cysteine is used as a linker to couple the synthetic peptide to the carrier protein. Alternatively, multiple copies of a pepide can be linked to each other. Other methods of coupling peptides to carriers are available but are rarely used. These include:

- ***Glutaraldehyde-mediated cross-linking of basic residues in the peptide and a carrier protein.*** Because glutaraldehyde reacts with $-NH_2$ groups to form Schiff's bases, peptides containing lysine residues can be efficiently cross-linked to a protein such as keyhole limpet hemocyanin, whose surface is rich in basic residues. However, lysines, if present in the peptide, are likely to be important contributors to its antigenicity. Modifying them may reduce the antigenicity of the peptide and/or generate antibodies that efficiently recognize the modified residues but do not interact with the unmodified cognate sequence in the target protein.
- ***Use of peptide-resin as the immunogen.*** Novabiochem sells a polymethylacrylamide/Keiselguhr composite derived using ethylene diamine. Antigenic peptides can be synthesized directly on the free amino group of the base resin and remain attached to the support by a covalent bond. After deprotection of the peptide with trifluoroacetic acid and homogenization of the inorganic Keiselguhr, the peptide-polyamide complex can be emulsified with adjuvant and injected into animals.

Immunization with the carrier-peptide complexes will generate a mixture of antibodies of various affinities and specificities. Some antibodies will be directed toward the carrier (protein or resin) or the cross-linking groups; others will react with both peptide and carrier. However, as long as the synthetic protein is at least 11 residues in length, a proportion of the antibodies will be specific for amino acid sequences that are shared by the peptide and its cognate protein. Usually, there is no need to remove antibodies directed against the linking groups and the carrier protein, since these do not interfere with antibodies directed against the target protein.

However, antibodies that cross-react with the carrier protein can, if necessary, be removed by immunoabsorption with resin or by immunoaffinity chromatography on Sepharose columns containing immobilized carrier protein.

Conjugated Antibodies

Antibodies can be conjugated to a wide variety of substances without significant erosion of avidity and specificity. These substances include biotin, fluorochromes of various colors, and enzymes such as alkaline phosphatase, horseradish peroxidase, and urease. Antibodies conjugated to different substances allow simultaneous detection of more than one antigen and generate strong signals that can be further amplified through the use of secondary antibodies and avidin-conjugated enzymes.

- ***Biotinylated antibodies*** are used chiefly to detect antigens by enzyme-linked immunosorbent assay (ELISA), for antibody screening of expression libraries, and for western blotting.

  The labeled avidin-biotin (LAB) technique uses an avidin-enzyme conjugate to detect a biotinylated primary or secondary antibody.

  The bridged avidin-biotin (BRAB) assay, as its name implies, uses avidin as a bridge between a biotinylated primary or secondary antibody and a biotin-enzyme conjugate.

  The avidin-biotin complex (ABC) assay is the most sensitive of these three assays. The biotinylated enzyme is preincubated with avidin, forming large complexes that are then used to detect biotinylated primary or secondary antibodies.

  Biotinylated secondary antibodies are available from several manufacturers (e.g., Pierce and Vector).

- ***Fluorochrome-labeled antibodies*** are used chiefly for immunohistochemistry. Wherever possible, it is best to use fluorochrome-labeled, affinity-purified, anti-immunoglobulins or fluorochromes conjugated to avidin or streptavidin. These reagents, which are commercially available, have lower background and less nonspecific fluorescence than fluorescent antiserum or immunoglobulin fractions. The choice among fluorochromes is usually dictated by the types of filters that are available for the fluorescence microscope. Fluorescein has a lower background than rhodamine, but bleaches more rapidly. Texas Red emits strongly and is resistant to photobleaching. Phycoerythrin, a phycobiliprotein whose fluorescent yield is 30–50 times greater than that of fluorescein and rhodamine, is used when extreme sensitivity is required (Oi et al. 1982). Table A9-5 summarizes the excitation and emission wavelengths of various fluorochromes.

TABLE A9-5 Excitation and Emission Wavelengths of Fluorochromes

FLUOROCHROME	EXCITATION WAVELENGTH (nm)	EMISSION WAVELENGTH (nm)
Fluorescein	495 blue	524 greenish-yellow
Rhodamine	540 green, visible	575 orange-red
Texas Red	595 orange-red, visible	620 red
Oregon Green	496 blue	524 greenish-yellow
R-phycoerythrin	480, 545, 565 green, visible	574 orange-red

- ***Enzyme-conjugated antibodies.*** Antibodies coupled to horseradish peroxidase, alkaline phosphatase, β-galactosidase, and urease are available commercially. These can be used as secondary reagents for the enzymatic detection of primary antibody-antigen complexes. For both immunoscreening and western blotting, enzymatic methods are more sensitive and give lower backgrounds than radiolabeled antibodies or protein A. Enzyme-conjugated antibodies are particularly useful for screening expression libraries because signals develop directly on the nitrocellulose filter (rather than on a sheet of X-ray film or a phosphorimager), so that positive plaques can be located more accurately.

Conjugated enzymes catalyze the formation of insoluble colored precipitates on the surface of a nitrocellulose filter at the site of an antibody-antigen complex (Towbin et al. 1979; Hawkes et al. 1982; Blake et al. 1984; Knecht and Dimond 1984; Towbin and Gordon 1984; Hawkes 1986). Alkaline phosphatase and horseradish peroxidase are the most commonly used conjugated enzymes. Alkaline phosphatase remains active for several hours, during which the end-product of the reaction — a dark blue precipitate of diformazan (McGadey 1970) — continues to accumulate and intensify in color. By contrast, horseradish peroxidase is inactivated within a short period of time by its substrate, peroxide, which may account for the lower sensitivity of horseradish-peroxidase-conjugated antibodies in detecting antibody-antigen complexes on nitrocellulose filters (Mierendorf et al. 1987). Furthermore, some protocols using horseradish peroxidase antibodies employ a carcinogenic chromogen, *o*-dianisidine, which requires special handling and disposal. For a summary of chromogenic and luminescent methods of detection of immobilized antigens, please see Table A8-12 in Appendix 8.

HORSERADISH PEROXIDASE

Horseradish peroxidase (HRP), a heme-containing protein, is generally isolated from wild horseradish roots as a mixture of several isozymes. The classical preparation (Shannon et al. 1966) is predominantly a mixture of two forms (isozymes B and C), each with a M_r ~40,000. HRP catalyzes the transfer of two electrons from a substrate to hydrogen peroxide, to generate H_2O and an oxidized donor. When 3,3′-diaminobenzidine (DAB) is used as the substrate (Graham and Karnovsky 1966), the oxidized product polymerizes to form an intense brown residue that is insoluble in both H_2O and ethanol. In the presence of transition elements, such as cobalt and nickel, the residue is a slatey blue-black. These reactions form the basis of sensitive chromogenic assays for peroxidase that have been used for many years in electron microscopy (Robbins et al. 1971), immunocytochemistry (Nakane and Pierce 1967), various enzyme-linked immunosorbent assays and western blotting (see Figure A9-4). However, diaminobenzidine has some disadvantages as a substrate: It is possibly carcinogenic (Garner 1975; Weisburger et al. 1978), and it is not as sensitive a detector of HRP activity as more recently developed compounds such as 3,3′,5,5′-tetramethylbenzidine (TMB) (e.g., please see Roberts et al. 1991). The oxidation products of TMB are normally soluble but can be "captured" by pretreating blots with dextran sulfate (McKimm-Breschkin 1990) or by using a particulate form of TMB that is available commercially. With these modifications, TMB is suitable for detection of HRP-antibody conjugates that are directed against proteins (including antibodies of another species) or ligands that can be attached to nucleic acid probes.

In addition to DAB and TMB, other HRP substrates that can be oxidized to insoluble chromogenic products include 4-chloro-1-naphthol (purple precipitate) and 3-amino-4-ethylcarbazole (red precipitate). Neither of these substrates is as sensitive a detector of HRP activity as diaminobenzidine. In addition, a number of chromogenic substrates for HRP are available that yield attractively colored, soluble oxidation products. In addition to TMB, these include *O*-phenylenediamine dihydrochloride (OPD) (tangerine orange) (Wolters et al. 1976), and 2,2′-azino-di(3-ethylbenzthiazoline-6-sulphonic acid) (ABTS) (green) (Engvall 1980). The best fluorogenic substrate is 3-(*p*-hydroxyphenyl) propionic acid (HPPA) (violet) (Roberts et al. 1991).

In the presence of hydogen peroxide, HRP can also be used to trigger a cyclical chemiluminescent reaction that results in the oxidation of luminol to an excited form of 3-aminophthalate (Isaacson and Wettermark 1974; Roswell and White 1978; Durrant 1990; Stone and Durrant 1991) (please see Figure A9-5). As this compound returns to ground state, blue light is emitted at 428 nm. However, because the efficiency of this reaction is low, various *para*-substituted phenol-based compounds are used to increase the quantity and duration of the light emitted by the reaction (Whitehead et al. 1983; Hodgson and Jones 1989). In the presence of *p*-iodophenol, for example, the light emitted by the reaction may be increased by >1000-fold and the emission may be sustained for several hours. In addition, enhancers reduce the background by inhibiting chemical oxidation of luminol. Enhancers are thought to achieve these effects by eliminating the rate-limiting step in the reaction cycle by (1) reacting with HRP and (2) forming enhancer radicals that in turn react with luminol to form luminol radicals (please see Figure A9-5). The net result is an increase in the efficiency and speed, but not the characteristics, of the light-producing reaction. Light emitted from the reaction can be captured on X-ray film, phosphorimagers, or CCD cameras.

HRP can be directly coupled to single-stranded nucleic acid probes (Renz and Kurz 1984; Stone and Durrant 1991) or it can be conjugated to an antibody specific for ligands such as biotin or digoxygenin that can be incorporated into nucleic acid probes by standard enzymological techniques. In either case, the HRP/luminol reaction is sufficiently sensitive for nonradioactive detection of nucleic acids in Southern and northern hybridizations (Thorpe and Kricka 1986). Under ideal conditions, the limit of detection of nucleic acid probes is ~5×10^{-17} moles (Urdea et al. 1988). However, the HRP/luminol reaction is considerably less sensitive than chemiluminescent reactions involving cleavage of 1,2-dioxetanes by alkaline phosphatase (please see the information panel on **AMPPD**). In addition, to achieve maximal sensitivity, the HRP/luminol reaction may require extensive optimization of its various components. For these reasons, HRP is now used less than alkaline phosphatase for the chemiluminescent detection of nucleic acids and proteins.

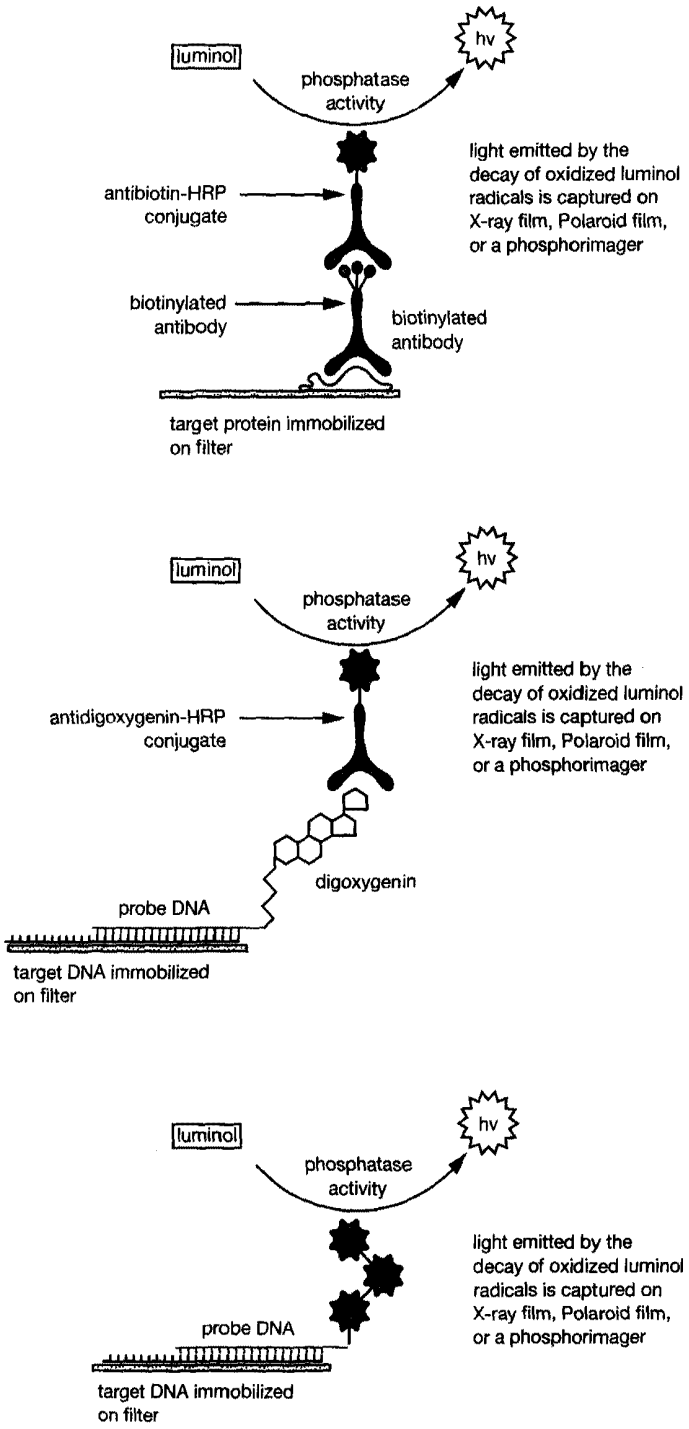
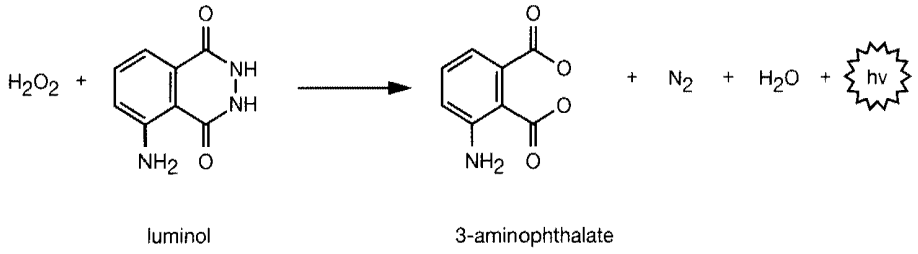


FIGURE A9-4 Experimental Formats for Detection of Immobilized Nucleic Acids and Proteins with Horseradish Peroxidase (HRP)

Light emitted by the decay of oxidized luminol radicals is captured on X-ray film or by a CCD camera.



luminol 3-aminophthalate

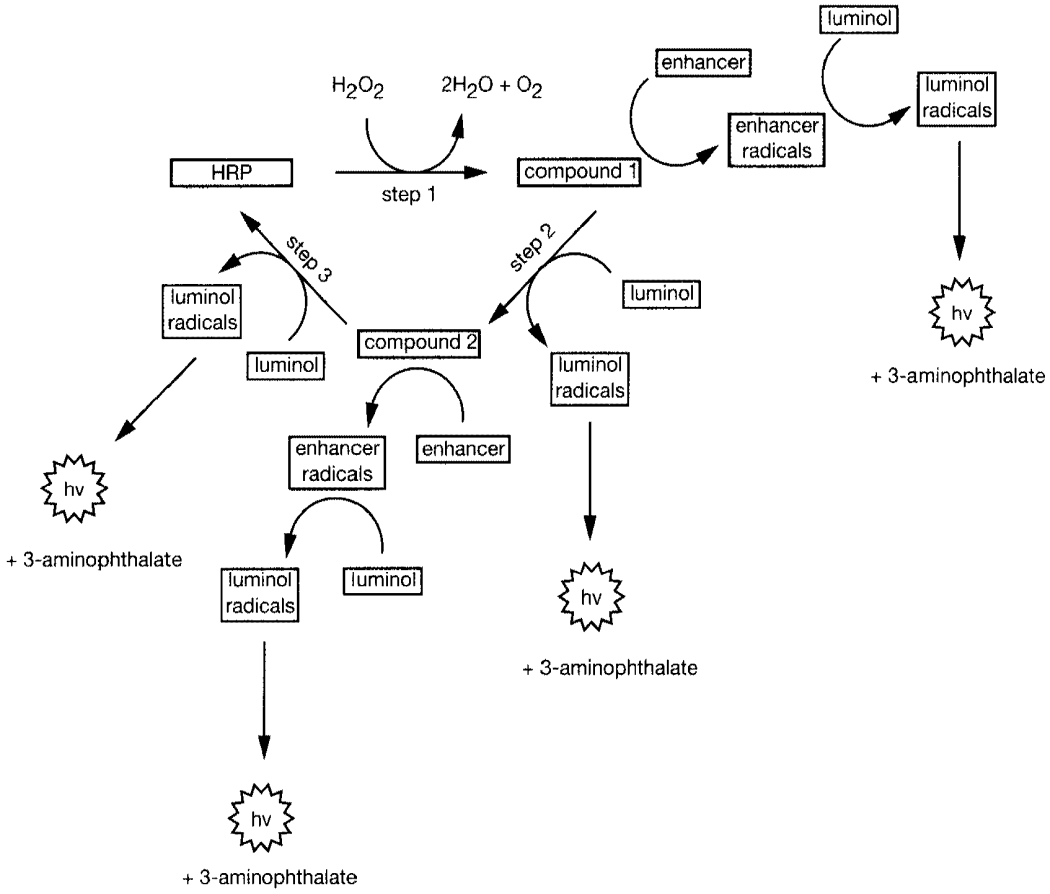
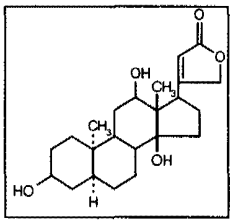


FIGURE A9-5 Proposed Cyclic Reaction for the Generation of Light by Oxidation of Luminol by HRP

DIGOXYGENIN



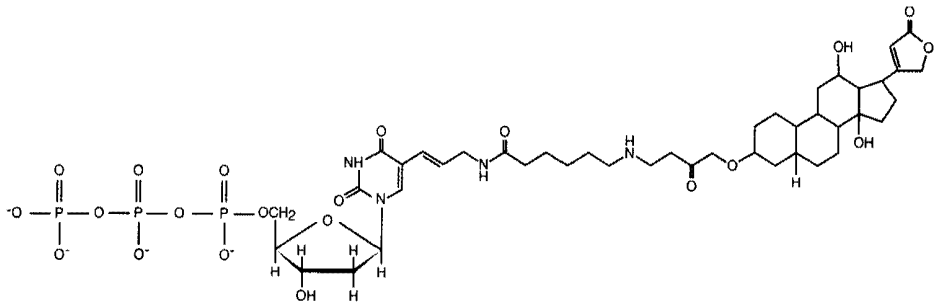
The cardenolide digoxygenin (M_r = 390.53) can be derived by chemical or enzymatic removal of four sugar residues from desacetylanatoside C isolated from *Digitalis purpurea* (Reichstein 1962) or it can be extracted directly from *D. orientalisi* and *D. lanata* L. *Scrophulariaceae* (Maanich and Schneider 1941; Hegenauer 1971). The chemical structure of digoxygenin (Cardwell and Smith 1953, 1954; Pataki et al. 1953) is shown in the box at the right.

In molecular cloning, digoxygenin is used as a ligand that can be incorporated into DNA and RNA probes and detected after hybridization with an anti-digoxygenin-antibody enzyme conjugate (Kessler et al. 1989; Kessler 1991) (please see Figure A9-6). Digoxygenin-labeled probes can therefore be used in Southern, northern, and dot-blot hybridizations. They also have considerable advantages for in situ hybridization, in part because the hapten digoxygenin is not present in animal cells (e.g., please see Tautz and Pfeifle 1989; Hemmati-Brivanlou et al. 1990). This absence eliminates background contributed by endogenous molecules. Such is not always the case with biotin-labeled probes. Digoxygenin-labeled nucleic acid probes can be stripped from nylon filters by heating the filters to 37°C for 20–30 minutes in a large volume of 0.1% SDS, 2 mM EDTA (pH 8.0) (Church and Kieffer-Higgins 1988). After washing, the filters can be reprobed without significant loss of efficiency or sensitivity.

Note that after several cycles of probing, stripping, and reprobing, the background of chemiluminescence may increase significantly due to the accumulation of alkaline phosphatase residue on the filter. This problem can be avoided by treating the filter with proteinase K and formamide between every second and third probing (Dubitsky et al. 1992).

Labeling Nucleic Acids with Digoxygenin

For labeling of nucleic acids, digoxygenin is supplied by the manufacturer (Boehringer Mannheim) in two forms: digoxygenin-11-dUTP (DIG-11-dUTP) and digoxygenin-11-UTP (DIG-11-UTP). In each form, digoxygenin is coupled by an alkali-stable linkage and a spacer arm to deoxyuridine triphosphate and uridine triphosphate, respectively.



- DIG-11-dUTP is a substrate for *E. coli* DNA polymerase, the Klenow fragment of DNA polymerase I, thermostable polymerases such as *Taq*, reverse transcriptase, and terminal transferase. Digoxygenin can therefore be introduced into DNA by a variety of standard reactions including random priming, nick translation, amplification, end-filling, and 3′-tailing.
- DIG-11-UTP is incorporated into RNA during in vitro transcription of DNA templates by bacteriophage-encoded DNA-dependent RNA polymerases (T3, T7 and SP6 RNA polymerases) (please see Table A9-6).

Oligonucleotides that have been 5′-aminated during synthesis can be labeled with digoxygenin by reaction with NHS-digoxygenin (digoxygenin-3-*O*-methylcarbonyl-ε-aminocaproic acid-*N*-hydroxysuccinamide ester; Boehringer Mannheim) (Richterich and Church 1993).

TABLE A9-6 Methods of Labeling Nucleic Acids with Digoxygenin

METHOD OF LABELING	ENZYME	NUMBER OF DIGOXYGENIN MOLECULES INCORPORATED	REFERENCE
Random priming	Klenow fragment	1 per 25–36 nucleotides	Kessler et al. (1990)
Nick translation	*E. coli* DNA polymerase I	1 per 25–36 nucleotides	Höltke et al. (1990)
Tailing	Terminal transferase	1 per 12 nucleotides	Schmitz et al. (1991)
Amplification by PCR	*Taq* and other thermostable polymerases	1 per 25 nucleotides	Seibl et al. (1990)
Transcription	T3, T7, and SP6 RNA polymerases	1 per 25–36 nucleotides	Höltke and Kessler (1990)
cDNA synthesis	Reverse transcriptase	1 per 25–36 nucleotides	McCracken (1989)

For review, please see Kessler (1991).

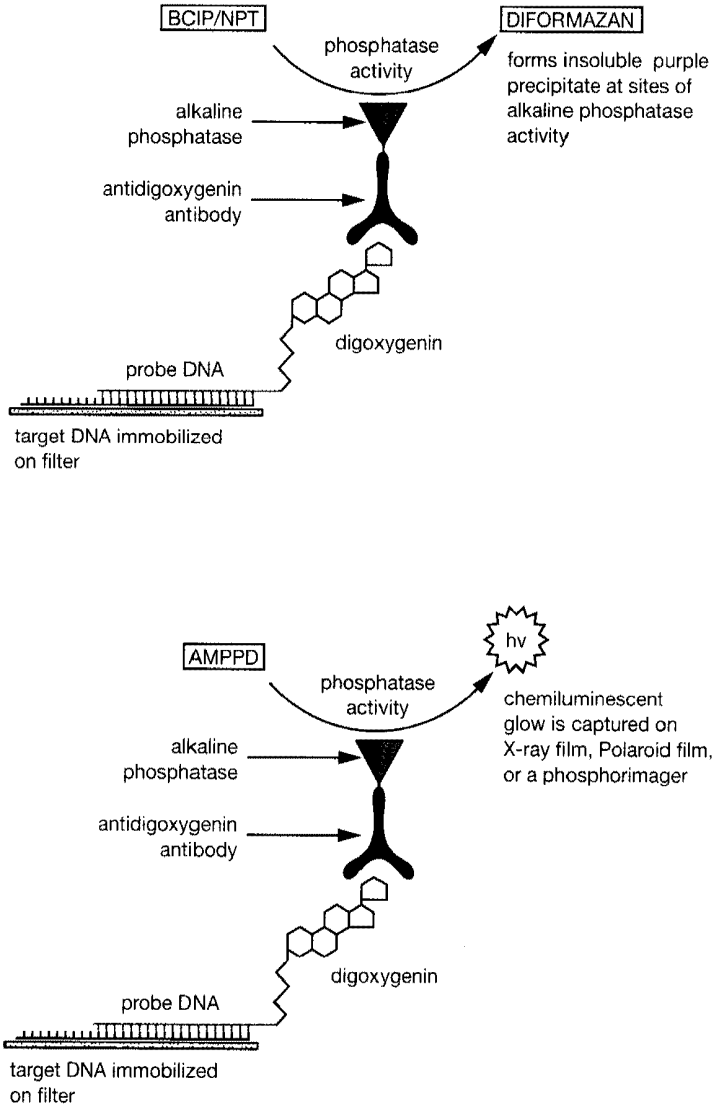


FIGURE A9-6 Detection of Digoxygenin-labeled Nucleic Acid Probes with BCIP/NBT or AMPPD

Digoxygenin-specific Antibodies Coupled to Reporter Enzymes

High-affinity digoxygenin-specific antibody is raised in sheep immunized with the digoxygenin-coated proteins edestin or bovine serum albumin. The most versatile of the immunological reagents available for detection of digoxygenin-labeled probes are FAB fragments of antidigoxygenin immunoglobulin that have been coupled to alkaline phosphatase (Boehringer Mannheim). The best chromogenic substrates for immunolocalized alkaline phosphatase are the binary reagents 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and 4-nitroblue tetrazolium chloride (NBT). For more information on these reagents, please see the information panel on **BCIP**.

Under optimal conditions, where high concentrations (50–100 ng/ml) of digoxygenin-labeled probe are used in the hybridization step and the color reactions are incubated for extended times (16–24 hours), it is possible to detect <1 pg of target DNA in a dot-blot assay (e.g., please see Kerkhof 1992). However, far greater sensitivity and a more linear response are obtained with the soluble chemiluminescent substrate adamantyl 1,2-dioxetane phosphate (AMPPD; Tropix, Inc.; Roche) (Kerkhof 1992; Bronstein et al. 1993). In AMPPD luminescent systems, which have gained rapidly in popularity in recent years, removal of the phosphate residue by alkaline phosphatase stimulates the substrate to emit chemiluminescence at 477 nm, which can be measured using a luminometer. The light occurs as a glow that persists for several hours in solution and for considerably longer on nylon surfaces (Tizard et al. 1990). Images of blots can be captured on X-ray film, on Polaroid instant black and white film (Kricka and Thorpe 1986), on a phosphorimager, or on a CCD camera (Karger et al. 1993). At its best, chemiluminescent detection of digoxygenin-labeled probes is highly sensitive (0.03 pg of target DNA or RNA) and rapid (<30 minutes exposure). The method is therefore ~10-fold more sensitive and 50-fold faster than autoradiographic detection of ^{32}P-labeled nucleic acid probes. In addition, reprobing is simplified since stripping of the filter may not be required (Allefs et al. 1990). As a consequence of these advantages, chemiluminescent assays for alkaline phosphatase labels are used for detection of immobilized nucleic acid sequences in a wide variety of techniques involving blotting and DNA sequencing. For more information on AMPPD systems, please see the information panel on **AMPPD**.

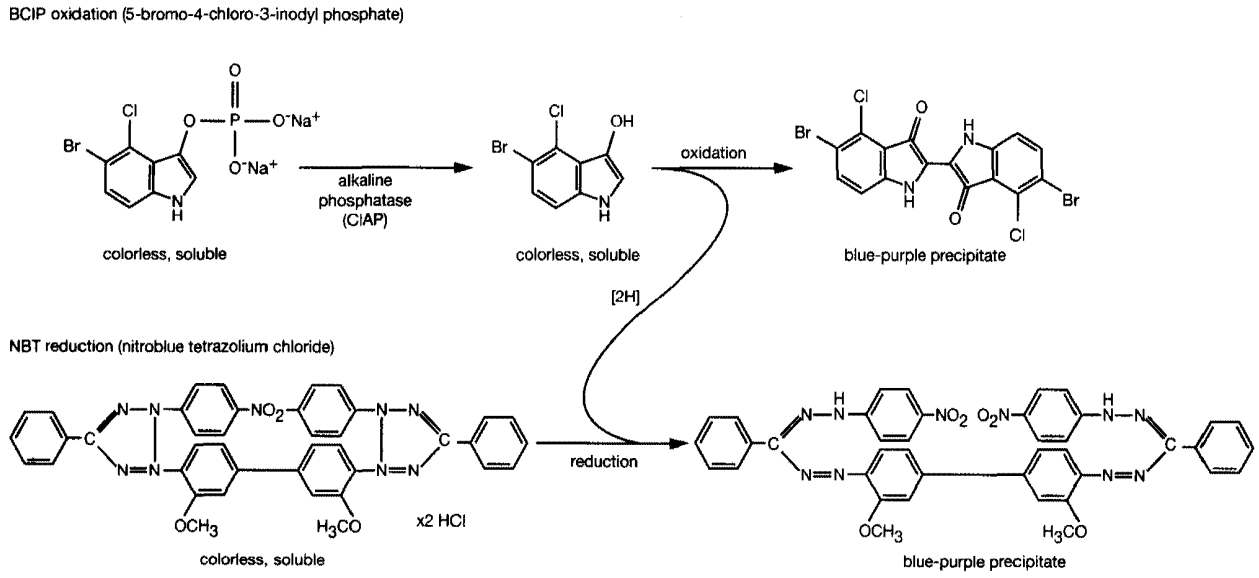
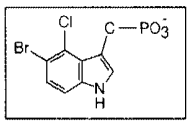


FIGURE A9-7 Oxidation of BCIP and Reduction of NBT in the BCIP/NPT Indicator Reaction

(Reprinted, with permission, from Kessler 1991 [Copyright 1991 ©Academic Press London].)

BCIP



BCIP (5-bromo-4-chloro-3-indolyl phosphate, F.W. 370.4, [disodium salt]; F.W. 433.6 [toluidine salt]) is used in combination with nitroblue-tetrazolium salt (NPT) to detect alkaline phosphatase in situ. The binary reagent forms insoluble precipitates at sites of alkaline phosphatase activity and is the most sensitive indicator system available for chromogenic detection of alkaline phosphatase conjugates. In this reaction (Figure A9-7), alkaline phosphatase catalyzes the removal of the phosphate group from BCIP, generating 5-bromo-4-chloro-3-indolyl hydroxide, which dimerizes to the insoluble blue compound, 5,5′-dibromo-4,4′-dichloroindigo. The two reducing equivalents produced during the dimerization reaction reduce one molecule of nitroblue tetrazolium to the insoluble, intensely purple dye, diformazan (McGadey 1970; Franci and Vidal 1988). Color development occurs over several hours, but the time required to produce maximum signal varies considerably, depending on the amount of antibody bound to the target molecule.

BCIP/NBT is a sensitive detector of alkaline phosphatase activity but is not quantitative. Quantitative assays of the enzyme are carried out using the chromogenic substate *p*-nitrophenyl phosphate (McComb and Bowers 1972; Brickman and Beckwith 1975; Michaelis et al. 1983). This substrate is sold by Sigma under

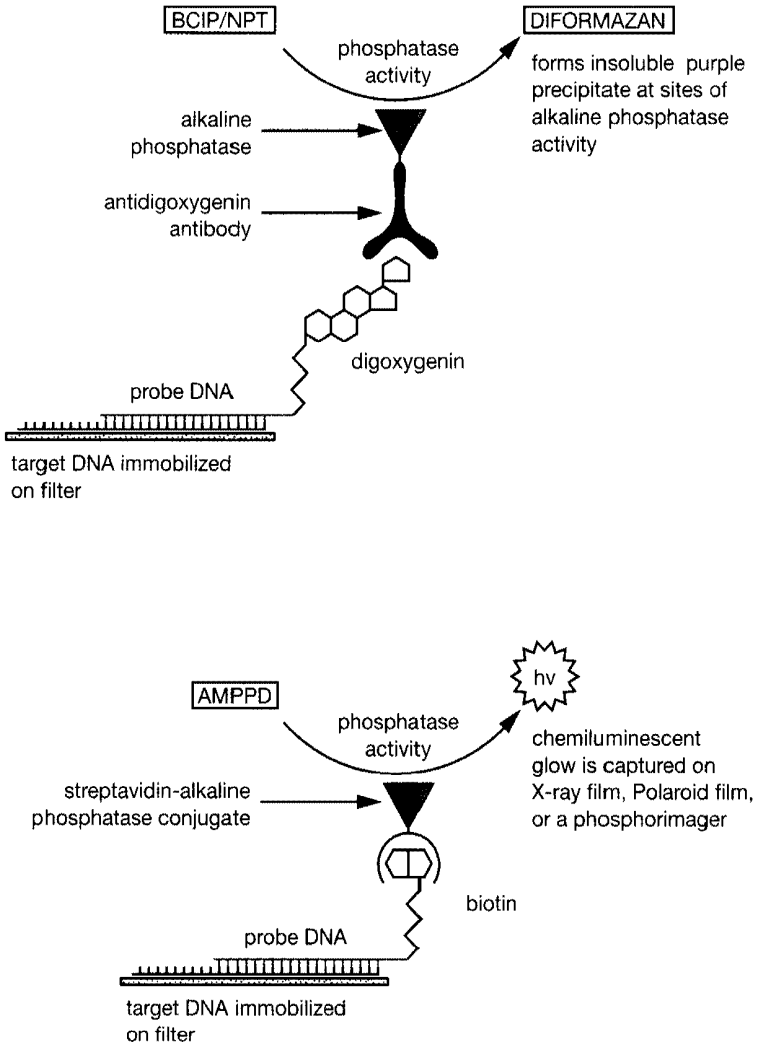


FIGURE A9-8 Detection of Digoxygenin- and Biotin-labeled Nucleic Acid Probes with BCIP/NPT

the name "Sigma No. 104 phosphatase substrate." Hydrolysis of Sigma 104 generates *p*-nitrophenol, which in aqueous solution absorbs strongly at 420 nm. A more sensitive bioluminescence-based assay has also been described (Miska and Geiger 1987) in which D-luciferin-*O*-phosphate is used as a substrate. The amount of luciferin released by hydrolysis of the substrate is titrated using luciferase, and light impulses are then measured in a luminometer.

In molecular cloning, BCIP/NPT is widely used as an indicator in nonradioactive systems to detect nucleic acids or proteins. Almost always, these systems consist of several components that are used in sequence. When working at its best, the BCIP/NPT indicator system is sensitive enough to detect single copies of genes in Southern hybridization of total mammalian DNA. Typical arrangements are shown in Figure A9-8.

Southern, northern, and dot blots can be rehybridized by stripping the precipitate of formazan from charged nylon filters with *N,N*-dimethylformamide. Note that *N,N*-dimethylformamide's vapor is irritating to skin and mucous membranes and should therefore be used only in a well-ventilated chemical fume hood. Digoxygenin-labeled oligonucleotide probes can be removed from membranes by *N,N*-dimethylformamide at room temperature or by heating to 70–75°C in a buffer containing high concentrations of SDS (Richterich and Church 1993). The membranes then can be reprobed.

Chromogenic detection of alkaline phosphatase is far less sensitive than the luminescent method using the soluble 1,2-dioxetane substrate AMPPD (Schaap et al. 1987). In this system, the sensitivity in dot blots is increased from 0.2 pg in 16 hours to 1 fg in 1 hour. In Southern hybridizations, using AMPPD as an indicator, the sensitivity is ~70 fg in 1 hour (e.g., please see Allefs et al. 1990).

In addition to its use as an indicator for alkaline phosphatase conjugates, BCIP/NBT has also been used to detect alkaline phosphatase expressed in sections of vertebrate tissues (Fields-Berry et al. 1992) and to identify bacterial colonies that express BAP (Brickman and Beckwith 1975; for review, please see Manoil et al. 1990).

BCIP is sold as a disodium salt or a toluidine salt. The toluidine salt of BCIP is suitable for use with the commercially available nonradioactive detection systems for nucleic acids and proteins.

AMPPD

AMPPD (adamantyl-1,2-dioxetane phosphate also known as -[2′spiroadamantane]-4-methoxy-3-[3″-(phosphoryl)phenyl]1,2,-dioxetane *or* disodium 3-(4-methoxyspiro[1,2-dioxetane-3,2′-tricyclo-[3.3.1.^{3,7}]decan]-4-yl)-phenylphosphate) is a substrate used in alkaline-phosphatase-triggered chemiluminescence for the detection of biopolymers immobilized on nylon or PVDF membranes (Bronstein and McGrath 1989; Bronstein et al. 1990; Tizard et al. 1990; Gillespie and Hudspeth 1991). Figure A9-9 shows a typical format used for the detection of proteins in western blotting and for hybridization of immobilized nucleic acids.

The sensitivity of chemiluminescent detection of DNA, RNA, and proteins using AMPPD is superior to available colorimetric, bioluminescent, or fluorimetric methods and is at least equal to the traditional autoradiographic techniques that have dominated molecular cloning for the last 20 years (e.g, please see Beck and Köster 1990; Bronstein et al. 1990; Carlson et al. 1990; Pollard-Knight et al. 1990). In solution, the light output from the activity of fewer than 1000 molecules (1 zmole or 10^{-21} moles) of alkaline phosphatase may be measured (Schaap et al. 1989). Less than 0.1 pg of RNA or DNA may be detected in standard Southern and northern hybridizations (e.g, please see Beck et al. 1989), whereas <1 pg of a target protein may be detected in western blots of total cell proteins (Gillespie and Hudspeth 1991).

Figure A9-10 shows the events triggered by dephosphorylation of AMPPD that leads to the production of light. Alkaline phosphatase catalyzes the removal of the single phosphate residue of AMPPD, generating a moderately stable dioxetane anion that fragments into adamantanone and the excited state of a methyl-*meta*oxybenzoate anion. On return to ground state, the anion emits visible, yellow-green light (Bronstein et al. 1989). The dephosphorylation of dioxetanes by alkaline phosphatase is quite efficient with a turnover rate of ~4.0 x 10^3 molecules s^{-1} (Schaap, cited in Beck and Köster 1990). However, the half-life of the excited 1,2 dioxetane anion is comparatively long, varying from 2 minutes to several hours, depending on the

local environment (Bronstein 1990). Thus, when the dephosphorylation reaction is carried out in the presence of excess AMPPD, the dioxetane anion is initially produced more rapidly than it decays. This behavior explains why chemiluminescence is emitted in the form of a "glow" that increases in intensity for several minutes and then persists for several hours (Figure A9-11). On nylon membranes, the kinetics are even slower because the excited dioxetane anion is stabilized by hydrophobic pockets in the membrane (Tizard et al. 1990). The hydrophobic interactions between the nylon and the anion also cause a "blue shift" of ~10 nm in the emitted light, i.e., from 477 nm to 466 nm (Beck and Köster 1990; Bronstein 1990).

In most experimental situations, the extended kinetics of chemiluminescence on nylon filters are advantageous because they allow time to make capture images at several exposures. However, the slow kinetics may be of practical importance when alkaline-phosphatase-triggered chemiluminescence is used to detect extremely low concentrations of DNA, RNA, or protein, for example, when the target band of DNA on a filter is expected to contain $<10^{-18}$ moles. In such cases, CSPD, a halogen-substituted derivative of AMPPD, may be a better choice. The addition of a chlorine atom to the 5-position of the adamantyl group

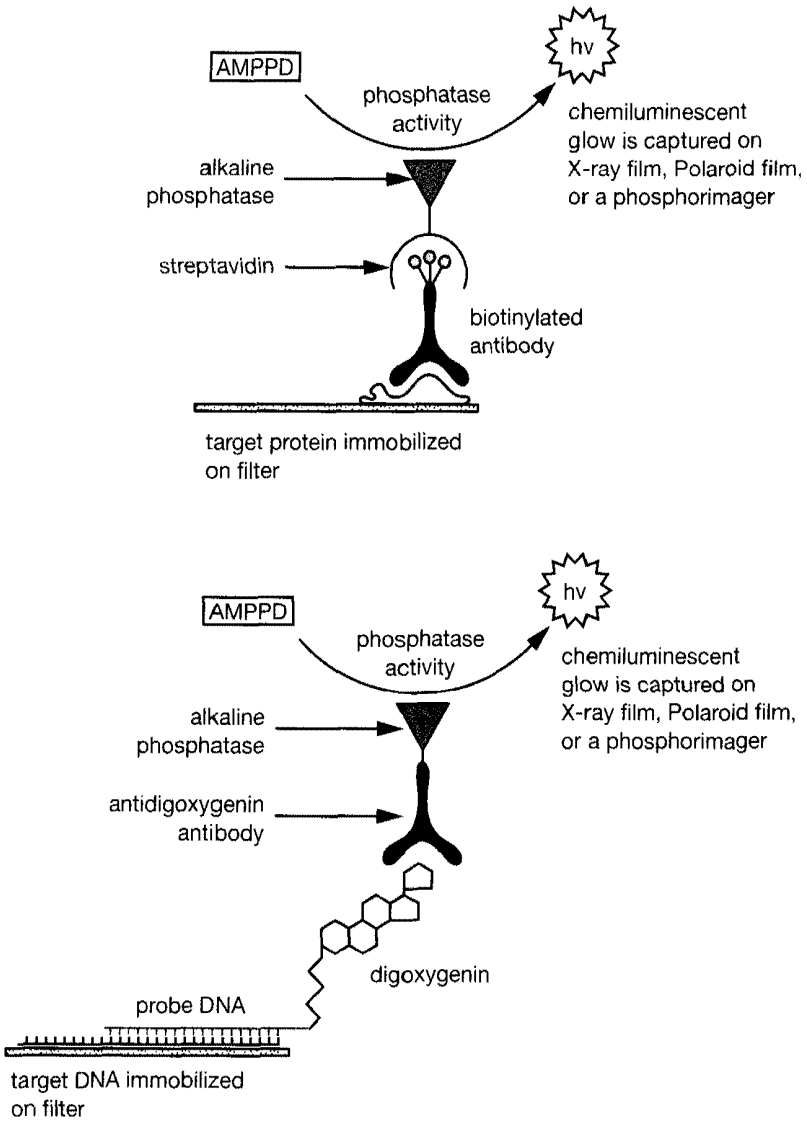


FIGURE A9-9 Detection of Immobilized Nucleic Acids and Proteins with AMPPD

(*Top*) Detection of target protein by western blotting; (*bottom*) detection of nucleic acid sequence in Southern or northern blotting.

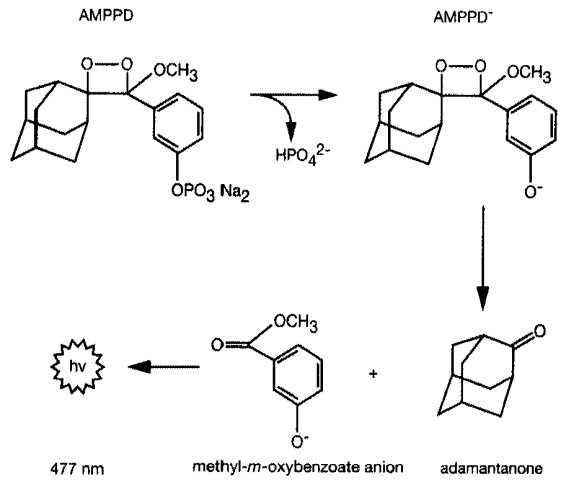


FIGURE A9-10 Chemiluminescent Generation of Light by Dephosphorylation of AMPPD

diminishes the tendency of the 1,2 dioxetane to self-aggregate and restricts its interactions with nylon filters. With this compound, the time to reach maximum light emission is markedly reduced so that very small quantities of target molecules can be detected rapidly (e.g., please see Martin et al. 1991).

CDP-*Star*, a third 1,2-dioxetane substrate, follows a decomposition pathway similar to that of AMPPD and CSPD, but it produces a signal several times brighter. The CDP-*Star* signal also peaks earlier in the reaction and persists for longer (up to several days) than that of its rivals (source: Tropix Web Site www.tropix.com).

Nitrocellulose filters lack the appropriate hydrophobic surfaces and are therefore not recommended for use with either AMPPD, CSPD, or CDP-*Star*; however, PVDF can be used with all three.

AMPPD is an extremely stable compound: The activation energy for its thermal decomposition in H_2O is 21.5 kcal/mole and its half-life at 25°C is ~20 years. Since nonenzymatic hydrolysis is very slow, the background of chemiluminescence on blots is minimal. In fact, the sensitivity of chemiluminescent detection of proteins and nucleic acids is generally limited, not by spontaneous decay of AMPPD, but by the presence of trace amounts in buffers of alkaline phosphatase of bacterial origin (Bronstein et al. 1990).

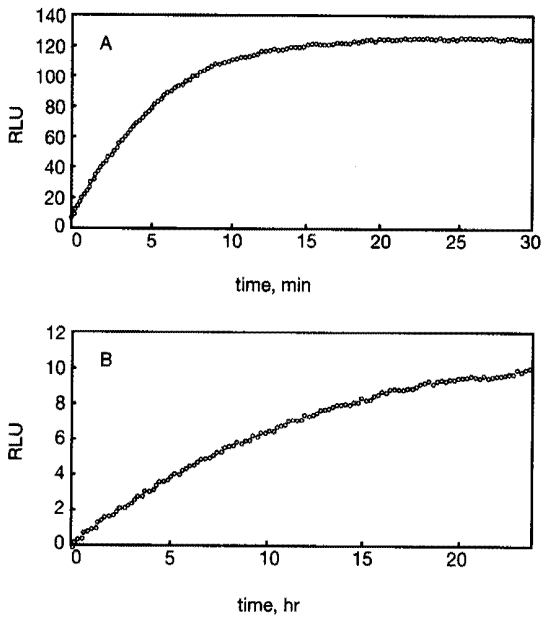
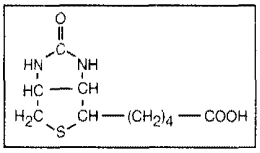


FIGURE A9-11 Kinetics of Chemiluminescence of AMPPD

Shown are the kinetics of chemiluminescence of alkaline-phosphatase-catalyzed decomposition of AMPPD in (*A*) 0.05 M bicarbonate/carbonate/1 mM $MgCl_2$ (pH 9.5) or (*B*) on a Biodye A membrane. Luminescent rates are expressed as relative light units (RLU). (Reprinted, with permission, from Tizard et al. 1990.)

AVIDIN AND BIOTIN



Biotin (vitamin H, coenzyme R; F.W. 244.3) is a water-soluble vitamin that binds with high affinity to avidin (M_r = 66,000), a tetrameric, basic glycoprotein, abundant in raw egg white (for review, please see Green 1975). Because each subunit of avidin can bind one biotin molecule, 1 mg of avidin can bind ~14.8 µg of biotin. The dissociation constant of the complex is ~1.0×10^{-15} M, which corresponds to a free energy of association of 21 kcal/mole. With such a tight association, the off-rate is extremely slow and the half-life of the complex is 200 days at pH 7.0 (Green and Toms 1973). For all practical purposes, therefore, the interaction between avidin and biotin is essentially irreversible. In addition, the avidin-biotin complex is resistant to chaotropic agents (3 M guanidine hydrochloride) and to extremes of pH and ionic strength (Green and Toms 1972). The strength of the interaction between biotin and avidin provides a bridging system to bring molecules with no natural affinity for one another into close contact.

Biotin can be attached to a variety of proteins and nucleic acids, often without altering their properties. Similarly, avidin (or streptavidin, its nonglycosylated prokaryotic equivalent) can be joined to reporter enzymes whose activity can be used to locate and/or quantitate avidin-biotin-target complexes. For example, in enzyme immunoassays, a biotinylated antibody bound to an immobilized antigen or primary antibody is often assayed by an enzyme, such as horseradish peroxidase or alkaline phosphatase, that has been coupled to avidin (Young et al. 1985; French et al. 1986). In addition, in nucleic acid hybridization, biotinylated probes can be detected by avidin-conjugated enzymes or fluorochromes. Derivatives of biotin are used to biotinylate proteins, peptides, and other molecules (for review and references, please see Wilchek and Bayer 1990).

The disadvantage of egg white avidin is the relatively high background of nonspecific binding caused by the presence of oligosaccharide groups. This problem can be reduced by using streptavidin, a tetrameric nonglycosylated 58,000-dalton protein secreted by *Streptomyces avidini* that binds biotin with approximately the same affinity as egg-white avidin.

For more information on biotin, please see the information panel on **BIOTIN** in Chapter 11. Much useful information about avidin-biotin chemistry and avidin-biotin techniques is available in a handbook entitled *Avidin-Biotin Chemistry: A Handbook*, published and sold by Pierce.

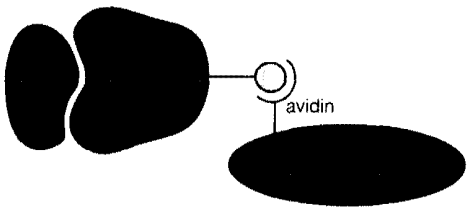


FIGURE A9-12 The Avidin-Biotin Complex

IMMUNOGLOBULIN-BINDING PROTEINS: PROTEINS A, G, AND L

Protein A

Protein A, a cell wall component of *Staphylococcus aureus*, binds to the Fc region of immunoglobulins (IgG) of many mammalian species and thereby helps the bacterium escape the immune response of the host (Forsgren and Sjöquist 1966; for review, please see Langone 1982; Boyle and Reis 1987; Boyle 1990; Bouvet 1994) (please see Table A9-7). The interaction of protein A with the Fc portion of IgG molecules does not block the ability of the antibody to combine with its antigen (Figure A9-13). Protein A has been used extensively for both qualitative and quantitative analysis of immunochemical reactions (Goding 1978; Harlow and Lane 1988, 1999).

TABLE A9-7 Binding of Protein A and Protein G to the Fc Region of Mammalian Immunoglobulins

Immunoglobulin	Protein A (*Staphylococcus aureus*)	Protein G (Streptococci of Groups C and G)
Human IgG1	++	++
Human IgG2	++	++
Human IgG3	–	++
Human IgG4	++	++
Mouse IgG1	+	+
Mouse IgG2a	++	++
Mouse IgG2b	++	++
Mouse IgG3	++	++
Rat IgG1	+	+
Rat IgG2A	–	++
Rat IgG2b	–	+
Rat IgG2c	++	++
Rabbit IgG	++	++
Bovine IgG1	–	++
Bovine IgG2	++	++
Sheep IgG	–	++
Sheep IgG2	++	++
Goat IgG1	+	++
Goat IgG2	++	++
Horse IgG(ab)	+	++
Horse IgG(c)	+	(+)
Chicken	–	(+)
Hamster	(+)	+
Guinea Pig	++	+

Data on human IgG from Forsgren and Sjöquist (1966), Kronvall (1973), and Myhre and Kronvall (1977, 1980a); on mouse from Kronvall et al. (1970a), Chalon et al. (1979), and Myhre and Kronvall (1980b); on rat from Medgyesi et al. (1978) and Nilsson et al. (1982); on rabbit from Kronvall (1973), Myhre and Kronvall (1977), and Forsgren and Sjöquist (1967); on bovine from Lind et al. (1970), Myhre and Kronvall (1981), and Kronvall et al. (1970b); on ovine, equine, and caprine from Kronvall et al. (1970b) and Sjöquist et al. (1972). Data on other animals are from Richman et al. (1982), Björck and Kronvall (1984), Åkerström et al. (1985), and Åkerström and Björck (1986).

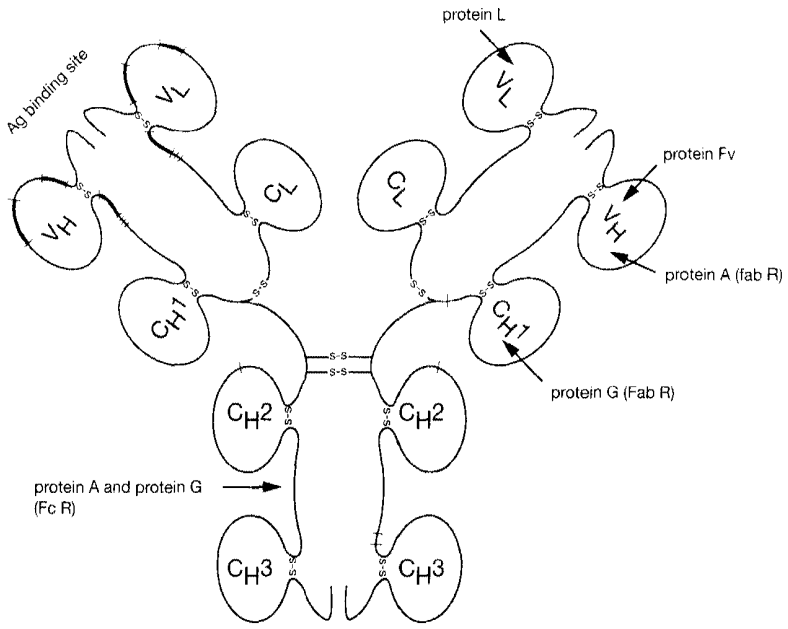


FIGURE A9-13 Immunoglobulin Regions Recognized by the Fab-binding Molecules

(Redrawn, with permission, from Bouvet 1994 [©Elsevier Science].)

TABLE A9-8 Binding of Immunoglobulins to Protein L, Protein A, and Protein G

IMMUNOGLOBULIN	PROTEIN L	PROTEIN LA[a]	PROTEIN A	PROTEIN G
Human				
IgG	++	++	+	+++
IgM	++	++	+	–
IgA	++	++	+	–
IgE	++	++	+	–
IgD	++	++	+	–
Fab	++	++	+	+
F(ab′)2	++	++	–	+
κ light chains	++	++	+	–
scFv	++	++	+	–
Mouse				
IgG1	++	++	+	++
IgG2a	++	++	++	++
IgG2b	++	++	++	++
IgG3	++	++	+	++
IgM	++	++	+	–
IgA	++	++	++	+
Polyclonal				
Mouse	++	++	++	++
Rat	++	++	+	++
Rabbit	+	++	++	+++
Sheep	–	++	++	++
Goat	–	++	+	++
Bovine	–	++	+	++
Porcine	++	++	++	++
Chicken IgY/IgG	++	++	–	–

Modified, with permission, from CLONTECH (www.clontech.com/archive/JUL98UPD/proteinL.html).

[a]Protein LA combines the immunoglobulin-binding domains of protein A and protein L.

- ***When coupled to a radioactive, enzymatic, or fluorescent tag,*** protein A is an excellent reagent to detect and quantitate antibodies with high affinity for the protein. Protein A chemically coupled to particles of colloidal gold can be used to locate IgGs by electron microscopy.
- ***Protein A immobilized on a solid support*** can be used to purify antibodies and to collect immune complexes, antigens, and whole cells (please see Figure A9-14).
- ***Using ELISA-like sandwich techniques, DNA fragments can be purified*** and detected using purified protein A or engineered fusion proteins containing the IgG-binding domains of protein A domains (e.g., please see Lindbladh et al. 1987; Peterhans et al. 1987; Werstuck and Capone 1989; for review, please see Stahl et al. 1993).

Chimeric proteins containing two independent ligand-binding domains have also been explored as adhesive immunological reagents. For example:

- ***Molecules labeled with digoxygenin*** may be labeled with a fusion protein containing an IgG-binding domain of protein A and an antigen-binding site of an antidigoxygenin antibody. The label may then be detected with any antibody that binds to protein A (Tai et al. 1990).
- ***Streptavidin-protein A chimeras*** may be used for indirect labeling of antibodies with enzymes (Sano and Cantor 1991).
- ***Protein A-maltose binding protein chimeras*** may be used as bifunctional reagents for binding antibodies to a solid matrix (Xue et al. 1995).
- ***Protein A*** can be used as an affinity tag to purify fusion proteins synthesized in pro- and eukaryotic cells (e.g., please see Kobatake et al. 1995; Nilsson et al. 1985; for reviews, please see Nilsson and Abrahmsén 1990; Uhlén and Moks 1990; Uhlén et al. 1992; Stahl et al. 1993). The simplest protocol involves purifying fusion proteins containing the Fc-binding domains of protein A by affinity chromatography with an IgG resin (Uhlén et al. 1983; Nilsson et al. 1985).

The gene for protein A (Uhlén et al. 1984) encodes a preprotein of 509 amino acids that consists of a signal sequence, which is removed during secretion: five homologous, independent IgG-binding domains, each of 58 amino acids; and a repetitive carboxy-terminal anchor (region X), of 180 amino acids. Crystallographic analysis of cocrystals of a single IgG-binding domain of protein A bound to human Fc shows that the IgG-binding domain contains two α-helices that form extensive hydrophobic interactions with the second and third constant regions of the Fc domain (Deisenhofer 1981).

Protein G

Protein G was first isolated from streptococcal strains belonging to groups C and G (Björck and Kronvall 1984). Protein G, like staphylococcal protein A, has a high affinity for the Fc region of mammalian antibodies. However, the two proteins are quite different in structure and have different affinities for antibodies (please see Table A9-8) (Åkerström et al. 1985). Whereas protein A is rich in α-helices and forms hydrophobic interactions with immunoglobulins, protein G has a large content of β-sheet structures (Olsson et al. 1987; Gronenborn et al. 1991). Protein G binds to the first constant domain of IgG (Erntell et al. 1985) and the interaction between the two molecules involves alignment of β strands to form a continuous β-sheet across the interacting surfaces (Derrick and Wigley 1992). The structures recognized by protein G and protein A are closely related (Stone et al. 1989). However, many species and subclasses of IgGs that do not bind well to protein A bind with high efficiency to protein G. For example, protein G is able to bind efficiently to human immunoglobulins of the IgG3 subclass, whereas protein A cannot (Sjöbring et al. 1991). Chimeric proteins, containing ligand-binding domains of proteins A and G exhibit the combined specificities of the parental proteins (Eliasson et al. 1988, 1989).

One potential disadvantage of protein G as a reagent to collect immune complexes and to purify IgGs is that it binds strongly to bovine serum albumin (Björck et al. 1987). However, the IgG-binding site and the serum albumin-binding site are structurally distinct (Nygren et al. 1988; Sjöbring et al. 1989) and engineered versions of protein G lacking the albumin-binding site are commercially available. The serum albumin domains of protein G have been used as purification tags in fusion proteins (Nygren et al. 1988, 1991; Sjölander et al. 1993).

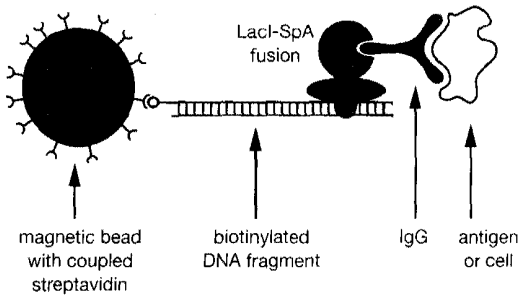


FIGURE A9-14 Isolation of Antigens or Whole Cells Using a LacI-SpA Fusion Protein

The figure illustrates the use of a *lac* repressor (LacI-SpA) fusion protein for reversible recovery of protein antigens or whole cells using magnetic beads with a coupled DNA fragment containing the lac operator (*lac*O) sequence.

Protein L

Protein L (M_r ~76,000) (Åkerström and Björck 1989) is a cell-wall component of the anaerobic bacterium *Peptostreptococcus magnus* (Björck 1988), which binds with high affinity to the κ light chains of immunoglobulins. This interaction does not interfere with the antigen-binding sites of the antibody (Åkerstrom and Björck 1989). Protein L binds to a wide range of Ig subclasses including human, mouse, rat, rabbit, and chicken (please see Table A9-8), but it does not bind to bovine, goat, or sheep Ig. These species-specific binding characteristics make protein L a useful tool for antibody purification, in particular for the purification of monoclonal antibodies from media supplemented with fetal bovine serum or bovine serum albumin and the purification of humanized antibodies from transgenic animals. Protein L is available commercially (e.g., CLONTECH and Pierce).

Appendix 10

DNA Array Technology

Vivek Mittal
Cold Spring Harbor Laboratory

INTRODUCTION

With the growing abundance of sequencing data from such diverse organisms as yeast, bacteria, fruit fly, several plants, and humans comes a pressing need for developing and applying technologies to perform comprehensive functional analyses. DNA microarrays have been developed in response to the need for simultaneous analysis of the patterns of expression of thousands of genes. Microarrays therefore offer tremendous advantages over traditional "single-gene" methods such as northern hybrization, reverse transcriptase–polymerase chain reaction (RT-PCR), and nuclease protection.

The Process

In a typical application, high-density nucleic acid samples, usually cDNAs or oligonucleotides, are delivered (or printed) by a robotic system onto very small, discrete areas of coated substrates (or chips) usually microscopic glass slides or membrane filters, and then immobilized to the substrate. The resulting microarray is then hybridized with a complex mixture of fluorescently labeled nucleic acids (probe) derived from a desired source. Following hybridization, the fluorescent markers are detected using a high-resolution laser scanner. A pattern of gene expression is obtained by analyzing the signal emitted from each spot with digital imaging software. The pattern of gene expression of the experimental sample can then be compared with that of a control for differential analyses. The use of terms varies in the literature; however, in this discussion, probe refers to the mixture of labeled nucleic acids and target denotes the immobilized array of nucleic acid samples.

Applications

Two chief applications of DNA microarray technology are described briefly below. Other applications include gene discovery, genotyping, and pathway analysis.

Analysis of Gene Expression

Gene expression patterns are biologically informative and provide direct clues to function. Correlating changes in gene expression with specific changes in physiology can provide mechanistic insights into the dynamics of various biological processes in an organism. Microarray technology may be used, for example, for simultaneously detecting expression of many genes in different tissues or at different developmental stages; for comparing genes expressed in normal and diseased states; and for analyzing the response of cells exposed to drugs or different physiological conditions. The value of microarrays for identifying patterns of gene expression has been demonstrated clearly in organisms such as yeast, fruit flies, mice, and humans. In yeast, for example, gene expression patterns obtained using cDNA arrays correlate very well with changes in the yeast cell cycle (yeast data sets and analysis tools are available at cmgm.stanford.edu/pbrown/). In explorations of human cancer, expression arrays have aided in the molecular classification of B-cell lymphomas and leukemias. These studies are likely to help us to understand the pathophysiology of cancer and to define targets for therapeutic intervention.

Monitoring Changes in Genomic DNA

Cancer cells typically exhibit genomic instability, including gain-of-function mutations of oncogenes often marked by gene amplifications or translocations and loss-of-function mutations in

tumor suppressor genes often marked by point mutations and subsequent loss of heterozygosity or by homozygous deletions. Three laboratories are now using microarrays for high-resolution analysis of genomic DNA:

- The Brown laboratory (cmgm.stanford.edu/pbrown) uses cDNA microarrays.
- The Gray laboratory (cc.ucsf.edu/gray/) uses microarrays of genomic DNA cloned in BAC vectors.
- The Wigler laboratory (nucleus.cshl.org/wigler/) uses microarrays derived from low-complexity representations of genomic DNA.

Which method or combination of methods will be successful in measuring changes in copy number of nonabundant sequences with high statistical confidence in clinical samples remains to be determined.

Mutations and polymorphisms, in particular single nucleotide polymorphisms (SNPs), can be studied within and among species using high-density oligonucleotide arrays. These so-called mutation detection arrays consist of oligonucleotides representing all known sequence variants of a gene or a collection of genes. Because hybridization to oligonucleotides is sensitive enough to detect single-nucleotide mismatches, an homologous gene carrying an unknown sequence variation can be screened rapidly for a large number of changes. Examples of mutation detection arrays include a p53 gene chip, an HIV gene chip, and a breast cancer BRCA-1 gene chip.

The remainder of this discussion focuses on the technological aspects of DNA microarrays, with little further emphasis on applications. It is imperative to remember that DNA microarrays represent a developing technology and that there remain substantial obstacles in the design and analysis of these microarrays.

OVERVIEW AND SCHEMA OF DNA MICROARRAYS

There are six basic steps in performing a DNA microarray experiment (please see Figure A10-1):

1. Processing cDNA clones to generate print-ready material.
2. Printing cDNA clones (or oligonucleotides) onto a substrate.
3. Sample RNA isolation (either total RNA or mRNA).
4. Preparation of the probe (e.g., cDNA synthesis and labeling).
5. Hybridization of the labeled probe DNA to the DNA arrayed on the substrate.
6. Image acquisition of hybridization results and image analysis.

Clone Sources and Their Description

The initial resources required to design and fabricate gene expression microarrays include genomic or cDNA sequence data, cDNA clones, or both. Traditional approaches to gene discovery, such as cloning, have identified only a small subset of genes in an organism. More recent approaches, for example, large-scale sequencing of expressed sequence tags (ESTs), have greatly enhanced the rate of gene discovery. GenBank now has a collection of >1 million human ESTs. EST sequences are deposited into the databank dbEST (www.ncbi.nlm.nih.gov/dbEST/index.html), a division of GenBank. Sequences stored in dbEST are subjected to an automated

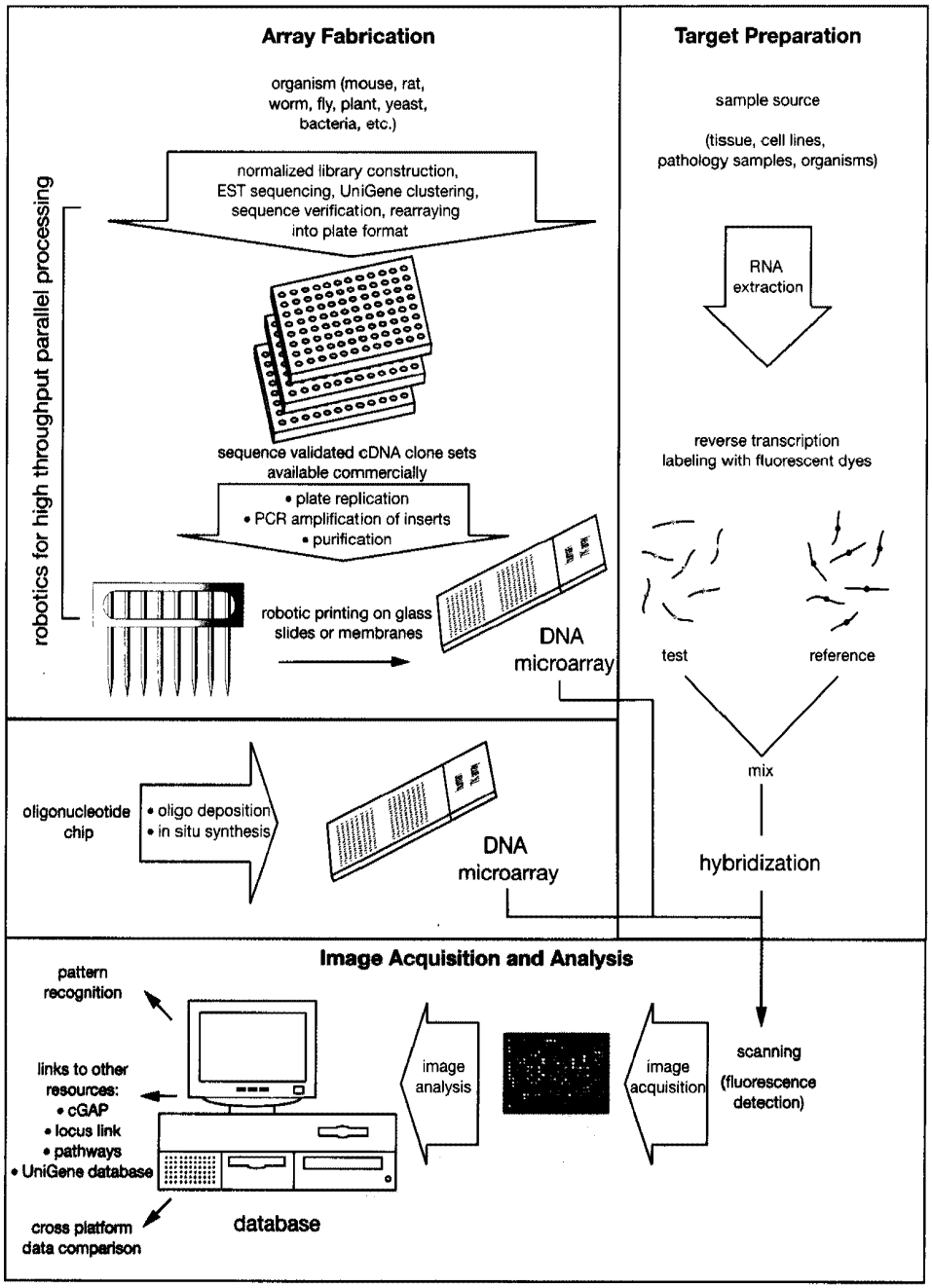


FIGURE A10-1 Flowchart: Performing a Microarray Experiment

process called UniGene that compares ESTs and assembles overlapping sequences into clusters (www.ncbi.nlm.nih.gov/UniGene/index.html). Clone sets, comprising the largest single representative of each cluster, are sequence-verified and then made available to the public through licensed vendors (http://image.llnl.gov/image/html/idistributors.shtml).

The current UniGene collection of human and mouse cDNA sequences (~40,000 and 10,000 sequence-validated clones, respectively) is available for purchase through official vendors such as Research Genetics (www.resgen.com) and Incyte Genomics (gem.incyte.com/gem/products/unigemv2.shtml). The UniGene collection is likely to expand proportionally with progress in

the various genome sequencing projects. Another resource for the human cDNA collection is the German Human Genome Project (http://www.rzpd.de/general/html/glossary/unigene.shtml). A comprehensive, sequence-validated full-length mouse cDNA collection is also available from Riken Genomic Sciences Center, Japan (http://hgp.gsc.riken.go.jp/) and from the National Institute of Aging (http://lgsun.grc.nia.nih.gov/). Table A10-1 presents genomic resources for organisms other than human and mouse that are available for microarray fabrication.

Robotics for High-throughput Processing

Considering the enormous size of clone sets derived from eukaryotic genomes, some degree of automation is typically required to obtain print-ready material with a high degree of accuracy and reproducibility. Two robotic systems, designed for high-throughput parallel processing, encompassing fully automated or semi-automated liquid-dispensing and plate-handling systems are the Robbins Hydra Work Station (www.robsci.com), which uses high-precision Teflon-coated capillaries, and the Beckman Coulter Multimek (www.beckmancoulter.com), which uses disposable tips. In addition to accurate volumetric handling of liquids to and from standard 96-well or 384-well plates, these robotic systems also have reformatting capabilities, allowing the transfer of material between 96- and 384-well plates on the same work surface. Each robotic unit can be integrated with other automated systems. Robots with analogous capabilities are the Qiagen BioRobot (www.qiagen.com/automation/), Tecan AG GENESIS sample processor (www.tecan.com), Packard MultiPROBE (www.packardinst.com), and Tomtec Quadra (www.tomtec.com).

Solid Support and Surface Chemistry

The development of novel solid supports and efficient chemistries for the manufacture of spatially resolved microscopic DNA arrays is essential if the potential of DNA chip technology is to be fully realized. A good substrate must bind DNA efficiently and uniformly, and leave the surface-bound DNA both functional and accessible. The density of DNA attached to the substrate must be high, the bound array must be stable, and the substrate must be chemically inert and have ultra-low intrinsic fluorescence while providing strong signal intensity and a broad dynamic range. Two basic substrates commonly used are glass and membrane filters.

Chemically treated microscope glass slides are the most widely used support. Glass is optically superior, durable under high temperatures of hybridization, and nonporous, thus keeping hybridization volumes to a minimum; it has low autofluorescence and good geometry and is amenable to high-density array fabrication. The major disadvantages of glass are that its planar surface has low loading capacity, the array on a glass support is not reusable, and an efficient chemistry for linking DNA to glass is lacking.

Glass slides coated with amine or aldehyde surface chemistry are available from Corning Microarray Technology (CMT) (http://www.cmt.corning.com/), Cel (www.cel-1.com/bio-molecular_products.htm), and TeleChem International (www.Arrayit.com). It is also possible to attach DNA covalently to glass slides. SurModics' 3D-Link Activated Slides are coated with a three-dimensional matrix that covalently binds amine-terminated DNA (PCR products or oligonucleotides) through endpoint attachment (www.surmodics.com). A common method for making amine-coated glass slides in-house is to treat them with polylysine. Details on the chemistry and instructions for preparing polylysine slides are available on the Brown laboratory Web Site (http://cmgm.stanford.edu/pbrown/index.html).

A major effort is now under way to explore the use of porous materials (such as nitrocellulose, nylon, and acrylamide) as substrates for attachment of nucleic acids. Porous substrates per-

TABLE A10-1 Genomic Resources for Microarrays

ORGANISM	MATERIAL	COMMENTS	RESOURCE
Rat	6000 cDNA clones	1500 sequence-validated clones are available.	Research Genetics (www.resgen.com)
Yeast (*S. cerevisiae*)	ORF-specific primer pairs	These primers are designed to amplify complete coding regions including start and stop codons from genomic DNA (possible because very few yeast genes contain introns). Recently, The Sanger Centre and the Imperial Cancer Research Fund have initiated work on amplification of coding regions from the fission yeast *S. pombe*.	Research Genetics (www.resgen.com)
	intergenic region primer pairs	These primer pairs are designed to amplify regions that lie between the open reading frames (ORFs), from genomic DNA, for use in cloning yeast promoters upstream of reporter genes, assaying deletions or insertions in yeast genes, and mapping transcription-factor-binding sites.	Research Genetics (www.resgen.com)
Arabidopsis thaliana	7900-member clone set	The clones represent genes from all major tissue categories (e.g., roots, rosettes, and inflorescence).	Incyte Genomics (www.incyte.com/products/organisms.html)
	11,500 EST clones	Collection was generated at Michigan State University.	Affymetrix
Caenorhabditis elegans	cDNA arrays representing the complete set of *C. elegans* genes	These microarrays along with support for hybridization and data analysis are available to other *C. elegans* laboratories on a collaborative basis.	Stuart Kim's laboratory at Stanford University (cmgm.stanford.edu/~kimlab/)
	primer pairs	These primer pairs can amplify all or a portion of each of the 19,000 genes.	Research Genetics (www.resgen.com)
Drosophila melanogaster	cDNAs representing 12,000 genes	The cDNA collection is currently undergoing annotation at the Berkley *Drosophila* Genome Project.	Berkeley *Drosophila* Genome Project (www.fruitfly.org) will be available for purchase through Research Genetics (www.resgen.com)
Escherichia coli	4290 ORF sequences	Membrane arrays containing a complete set of PCR-amplified genes.	Genosys Biotechnologies (http://www.genosys.com/expression/frameset.html)
Bacillus subtilis	4107 ORF sequences	PCR-amplified ORFs.	Genosys Biotechnologies
Helicobacter pylori	putative 1590 ORFs of strain 26,695; 91 ORFs unique to strain J99		Genosys Biotechnologies
C. jejuni	1654 genes		Genosys Biotechnologies
In Development:			
Staphylococcus aureus	1900 ORF sequences		Incyte Genomics (www.incyte.com/products/organisms.html)
Candida albicans	ORF sequences		Incyte Genomics (www.incyte.com/products/organisms.html)

mit the immobilization of large amounts of target molecules due to greatly enhanced surface area, providing enhanced sensitivity compared to planar substrates. Porous substrates based on polymer matrices provide a three-dimensional hydrophilic environment similar to free solution for biomolecular interactions. The utility of nitrocellulose-coated glass surfaces for microarrays has been demonstrated. The high binding capacity of nitrocellulose combined with low background

fluorescence makes these membrane-coated slides a superior substrate for microarrays. They are available from Schleicher & Schuell (www.s-und-s.de/).

Microarray Production

Microarrays may be produced using one of the methods described below. Alternatively, prefabricated microarrays from many sources are available commercially. Microarray fabrication technologies are of two kinds: in situ synthesis of nucleic acids and exogenous deposition of prepared materials on solid substrates. Three types of advanced technologies have emerged as early favorites in automated microarray production.

Contact Printing

In this approach, DNA fragments are directly deposited onto a glass support using a precision robot. The DNA sample is loaded into a spotting pin by capillary action and small volumes are transferred to a solid surface, such as a microscope slide, by direct physical contact between the pin and the solid substrate. After the first spotting cycle, the pin is washed and a second sample is then transferred to an adjacent address. A robotic control system and multiplexed print heads allow automated microarray fabrication. One of the crucial factors affecting array fabrication using the contact printing method is the reproducibility and durability of the spotting pins (also called quills or tips). Custom, machine-made high-precision pins can be obtained from Telechem International (www.arrayit.com), Majer Precision Engineering (http://www.majerprecision.com/), and Die-Tech (San Francisco).

The first contact printing arrayer was designed and built in the laboratory of Patrick Brown at Stanford University. An arrayer based largely on the Stanford prototype was subsequently fabricated in the National Human Genome Research Institute (NHGRI, www.nhgri.nih.gov/DIR/LCG/15K/HTML) and at the Albert Einstein College of Medicine (http://sequence.aecom.yu.edu/bioinf/microarray/printer.html). For the "do-it-yourself" approach, detailed instructions for assembling a contact printing arrayer, including a description of the necessary parts, can be found at http://cmgm.stanford.edu/pbrown/mguide/. Commercial robotic systems for making arrays by contact printing are offered by

- Cartesian Technology (www.cartesiantech.com/pixsyspa.htm)
- GeneMachines' OmniGrid (www.genemachines.com)
- Genomic Solutions (www.genomicsolutions.com)
- BioRobotics (www.biorobotics.co.uk)
- Genetix (www.genetix.co.uk/Microarraying.htm)
- Intelligent Bio-Instruments (www.intelligentbio.com)
- Genepak (www.genpakdna.com)
- Amersham Pharmacia Biotech (www.apbiotech.com/application/microarray/)

A less expensive alternative to robotic spotting of samples onto substrates is to use a handheld 384-pin arraying device, available from V & P Scientific, Inc. (www.vp-scientific.com/). This vendor offers pin tools for 96-, 384-, and 1536-well plates, as well as registration accessories that allow the creation of high-density arrays of up to 3456 samples per array.

Photolithography

Developed by Affymetrix (www.affymetrix.com), this method combines photolithography technology from the semiconductor industry with DNA synthetic chemistry to permit the manufacture of high-density oligonucleotide microarrays. A solid support is derivatized with a covalent linker terminated with a photolabile protecting group. A photomask is used to determine which positions react with light. Exposure to light causes dissociation of the protecting group at the exposed site. A chemical coupling reaction is then used to add a specific nucleotide to the new deprotected site, and the process is repeated using a different mask. The end result is a precision-made array, containing many oligonucleotides of known sequence. The Affymetrix method for making arrays demands a technical sophistication that generally cannot be imported to academic laboratory settings.

Pin and Ring

A "pin and ring" device is used for sample spotting by a noncontact mechanism. The ring holds a droplet of solution picked up from the well of a microtiter plate, and the pin punches a smaller droplet from this reservoir onto the substrate. The main advantage of this technology is the ability to generate consistent delivery from a variety of solvents with DNA, proteins, and small molecules on virtually any substrates. The technology as it currently stands has a high degree of precision but lacks the desired robustness. This technology was developed at Genetic Microsystems, now wholly owned by Affymetrix (www.affymetrix.com).

Commercial Arrays and Integrated Array Services

The formidable task of generating microarrays is time-consuming and frequently cost-prohibitive for many laboratories. Some commercial ventures have fabricated microarrays on membranes (nylon and nitrocellulose) and/or glass, and these are available commercially (please see Table A10-2). Commercial arrays are also available from Hyseq (www.hyseq.com) and Stratagene (www.stratagene.com/gc/gene.htm).

Choice of an Array System

The choice of a DNA array for a particular experiment requires consideration of the expense, desired density of DNA on the array, reproducibility among chips, and type of DNA to be immobilized on the surface. The first decision is whether the chips should contain immobilized cDNAs or shorter oligonucleotide sequences. cDNA samples must be spotted onto the chips as complete molecules, whereas oligonucleotides can be either spotted or synthesized on the surface of a chip with high fidelity. Another important decision is whether the user will make the arrays or purchase the chips. The advantages and disadvantages for each of these issues for each approach are presented in Table A10-3.

TABLE A10-2 Commercial Arrays

PRODUCTS	SOURCE	DESCRIPTION	COMMENTS
GeneChip product line (includes high-density DNA arrays for human, murine, yeast, and plant genomes)	Affymetrix www.affymetrix.com	GeneChip expression arrays use ~20 pairs of matched and mismatched oligonucleotides to interrogate each transcript. GeneChips designed for clinical settings include: *p53 GeneChip* to detect single-nucleotide polymorphisms of the p53 tumor-suppressor gene. *HIV GeneChip* to detect mutations in the HIV-1 protease and viral reverse transcriptase genes. *P450 GeneChip*, which focuses on mutations of key liver enzymes that metabolize drugs.	This oligonucleotide-pairing strategy helps identify and minimize the effects of nonspecific hybridization and background signal (a major concern in array experiments) to permit sensitive and accurate recognition of low-intensity hybridization patterns from mRNAs. In addition to expression monitoring, Affymetrix chips are also used for sequence analysis, genotyping, and SNP mapping. An integrated platform comprising an automated hybridization station, array scanner, and software for image analysis is mandatory for Affymetrix DNA chip users. A Pathways analysis software package from the same source is available for image analysis of the GeneFilters.
GeneFilters microarrays (for human, murine, and yeast)	Research Genetics www.resgen.com	The human cDNA arrays consist of 30,000 nonredundant sequence-validated I.M.A.G.E clones (on six different membranes) chosen from the UniGene set. The tissue-specific GeneFilters are gene expression arrays that contain sequences enriched for expression in prostate, ovary, breast, and colon tissues. Membranes carrying 6144 yeast ORFs and 10,000 sequence-validated rat cDNAs are also available.	
Atlas cDNA arrays	CLONTECH www.clontech.com	Arrays on nylon membranes or glass are available for a range of mammalian genes involved in cellular pathways such as oncogenesis, apoptosis, and cell cycle regulation. The arrays contain human, mouse, or rat cDNA and are usually spotted in duplicate. The Atlas human 1.0 cDNA array (on glass) represents 1081 known genes, covering a broad range of biologically significant genes and pathways.	All of the immobilized cDNAs have been tested for optimal performance and minimal cross hybridization. An integrated Expression Array platform, including a total RNA purification/labeling system, image analysis software, and bioinformatics database, is available.
Membrane Expression arrays	Incyte Genomics www.incyte.com/expression/argem1.html	The human array represents 20,000 I.M.A.G.E cDNA clones of human ESTs. The plant array contains clones representing 7900 *Arabidopsis thaliana* genes from all major tissue categories (roots, rosettes, and inflorescence).	Arrays carrying proprietary Incyte sequences are available only to Incyte subscribers. The data analysis is done in-house.
Membrane arrays for human and mouse	Super Array Inc. www.superarray.com	Super Array Incorporated provides pathway-specific gene expression membrane arrays.	Custom arrays are also available.
Expression microarrays	Genometrix Inc. www.genometrix.com	Microarrays are available in membrane format or on microscope slides for analyzing cancer-related human gene expression and mouse gene expression. Arrays for risk and toxicity assessments are also available.	A BioScanner for high-density imaging accompanied by GeneView software for image analysis is available.
Genomic microarray	Vysis Inc. www.vysis.com	The array contains >58 different gene targets, reported to be amplified in various human cancers or cancer cell lines, including HER-2/neu, cyclin D1, and Myc. The array contains large fragments (80–150 kb) of genomic DNA spotted onto chromium-coated glass slides.	The Vysis array is currently the only commercial source of a genomic array for assessing gene copy number. The array can be used to detect amplifications of threefold or greater.
discoveryARRAY Gene Display	Display Systems Biotech www.displaysystems.com	The unique human and mouse discoveryARRAY Gene Display slides are made using the company's proprietary restriction fragment differential display (RFDD-PCR) technology, which identifies and isolates nonredundant expressed cDNA sequences. Tissue-specific sequence arrays for eight different human tissues (liver, brain, kidney, heart, prostrate, lung, mammary, and spleen) are also available.	Display Systems Biotech allows users to discover and characterize completely novel genes by analyzing differential expression patterns of various biological samples. The clones identified by RFDD-PCR are completely independent of expression level or preexisting cDNA libraries or sequence databases.
Arrays carrying covalently bound oligonucleotides	Operon Technologies www.operon.com	Oligonucleotides, 70 nucleotides in length, are each optimized for sequence specificity and melting temperature. The human collage array represents 320 genes from eight functional categories: apoptosis, cancer, cell cycle, transcription factors, neuron/axon guidance, heat shock/stress, blood/inflammation, and aging. The apoptosis array contains 374 genes.	Operon also offers "array-ready" DNA products, normalized for concentration and ready to print. Customized 70 mers are also available.
Micromax	NEN Life Sciences www.nen.com/products/micromax/prod_serv	Two unique arrays are available, one containing a collection of 294 human transcription factors and the other containing a mixture of 196 human kinase and 96 phosphatase genes.	

TABLE A10-3 Choice of an Array System

Array Type	Advantages	Disadvantages
cDNA arrays	• Availability of large nonredundant, sequence-validated clone sets from different organisms. • Because prior sequence information is not required, cDNA arrays are an excellent choice for gene discovery. • cDNA size is optimal for specific hybridization. • Technology for spotting DNA on glass and membranes is readily available.	• Processing clones to generate print-ready material is cumbersome. • Density of printed cDNA is lower than for in-situ-synthesized oligonucleotides. • cDNA sequences may contain repetitive sequences such as *Alu* in human 3′ UTRs or B elements in rodent 3′ UTRs, resulting in cross-hybridization. • Clone authentication can be a problem.
Oligonucleotide chips	• High-density chips can be fabricated. • In situ synthesis on the chip generates consistent and uniform geometry of spots. • Oligonucleotide arrays may be used for genotyping as well as for gene expression analysis. • Maintaining large collections of cloned DNA molecules is not essential. • Because oligonucleotides are relatively short and can be designed for any gene region, oligonucleotide chips can be used for sequencing, identification of polymorphisms, and potentially for identification of different transcript splice variants.	• There is an absolute sequence data requirement for designing oligonucleotides. • Oligonucleotide selection rules are not well defined. • Short sequences are not the best targets for hybridization and, therefore, appropriate controls must be included. • Oligonucleotide chips are extremely expensive and the current manufacturing process is inherently inflexible.
Homemade systems	• Technology for array fabrication is readily available and well suited to laboratory setting. • Once the infrastructure is established, it is a less expensive way of making arrays. • Unlimited number of arrays can be generated as required (quite useful considering microarray experiments are performed with a large number of arrays). • Most flexible and versatile way to study any available genes. New genes can be added easily. • User has complete control over the research and is less dependent on commercial companies.	• User must learn to design and build chips. • Not the system of choice if only a small number of arrays are needed or a limited number of experiments are performed because capital costs run high. (This cost can be circumvented by establishing core facilities and shared resources.) • Tracking of clones and print material can be cumbersome and error-prone.
Commercial systems	• Uniformity in chip quality. • Burden of production and quality control is on the manufacturers. • User has access to proprietary technologies and private clone resources. • Integrated packages are available for array hybridization and downstream informatics.	• Limited flexibility. • Chip costs are extremely high. • User is totally dependent on outsourced production of chips.

IMAGING AND ANALYSIS OF THE ARRAY

Sample Processing and Hybridization

Once the array has been printed or purchased, the next step is to hybridize a labeled probe to the immobilized DNAs on the array. Typically, the probes for arrays are labeled representations of cellular mRNA pools isolated from various biological resources, such as cell cultures, tissues of model organisms, clinical biopsies, and histological samples. Detailed protocols for preparing the probe samples (including isolation of RNA, synthesis of cDNA, and incorporation of fluorescent dyes) and hybridizing the probe molecules to the immobilized DNA arrays are available on the following Web Sites:

Stanford University	www.cmgm.stanford.edu/pbrown/
Albert Einstein College of Medicine	www.sequence.aecom.yu.edu/bioinf/microarray/protocol.html
NHGRI	www.nhgri.nih.gov/DIR/LCG/15K/HTML/protocol.html
Cold Spring Harbor Laboratory	www.nucleus.cshl.org/wigler/
Collection of protocols	www.protocol-online.net/molbio/DNA/dna_microarray.htm
TIGR protocols	www.tigr.org/tdb/microarray

Detection of the Hybridization Signal

After hybridization, the DNA microarray is scanned to monitor the fluorescence of each probe that was successfully hybridized to the target. Most microarrays utilize two fluorophores, typically, the most commonly used fluorophores are Cy3 (green channel excitation) and Cy5 (red channel excitation). To generate a complete microarray image, it is necessary to acquire an image for each of these fluorophores. In general, two different scanning approaches are used: (1) sequential scanning acquires one image at a time and then builds the ratio image after acquisition is completed, and (2) simultaneous scanning acquires both images at the same time. Although a range of scanners is available, it is difficult to judge their relative performances at this time. A listing of scanners and their notable features is presented in Table A10-4.

TABLE A10-4 Scanning Systems

Scanning System	Source	Description
GenePix 4000	Axon Instruments, Inc. www.axon.com	This highly compact scanner has 10-μm resolution and simultaneous dual wave length scanning. Its main strengths are a very short scan time and perfect pixel-to-pixel registration of the images in both channels. (Keeping the pixels in register between channels is a problem often associated with dual-wavelength scanners.) It also includes elaborate image analysis software. The Axon scanner is currently limited to using only two fluorophores.
ScanArray 5000	General Scanning, Inc. www.gsilumonics.com/products_frame/datashts/scanarray/sa5000.htm	This microarray scanner utilizes a confocal laser with a resolution of 10 μm and includes advanced software for image acquisition, image analysis, and data mining. The distinguishing feature is the use of 4 different excitation lasers and 10 emission filters, providing users with the flexibility of many different fluorescent dyes (emission spectra between 500 and 700 nm) for sample labeling. This setup permits more than two differently labeled samples to be used for probing each microarray.
GeneTAC 1000	Genomic Solutions www.genomicsolutions.com/products/bio/img.htm	GeneTAC 1000 uses a CCD camera and a highly sensitive detector capable of detecting up to four fluorophores per experiment. It also has an automated sample holder, with provisions for loading and scanning 24 slides at one time.
HP GeneArray Scanner	Hewlett-Packard www.affymetrix.com/products/ins_array.html	The HP GeneArray Scanner is a high-resolution (3-μm) scanner designed to read Affymetrix GeneChip probe arrays.
The Storm system	Molecular Dynamics www.mdyn.com/arrays/gen.htm	This system combines phosphorimager autoradiography technology with two nonradioactive fluorescent-labeling techniques: direct fluorescence and chemifluorescence. The Storm system scans storage phosphor screens plus fluorescent gels and chemifluorescent blots in a 35 x 43-cm scanning area and is thus well-suited for scanning membrane arrays.
GMS 418 Array Scanner	Genetic Microsystems www.affymetrix.com	This system is a scanning laser confocal imaging epifluorescence microscope with a 10-μm resolution.

Image Analysis

The objective of microarray image analysis is to extract hybridization signals from each probe. Signals are measured as absolute intensities for a given target (essentially for arrays hybridized to a single sample) or as ratios of two probes with different fluorescent labels, representing two separate treatments to be compared with one probe as an internal control. The ratio of two signals provides relative response ratios rather than an estimate of an absolute signal. A typical color image is shown in Figure A10-2.

Once images are obtained in digitized form, they are subjected to further analyses using a variety of software programs. These programs provide a more accurate quantification of the intensity ratio. Undesired features of the data such as uneven spots, dust on the slides, and nonspecific hybridizations are flagged as inadequate and are not considered for further analysis. Background fluorescence, such as autofluorescence of the solid support or nonspecific binding of sample to the array, can also be subtracted from the intensity of a feature. Subsequently, the mean,

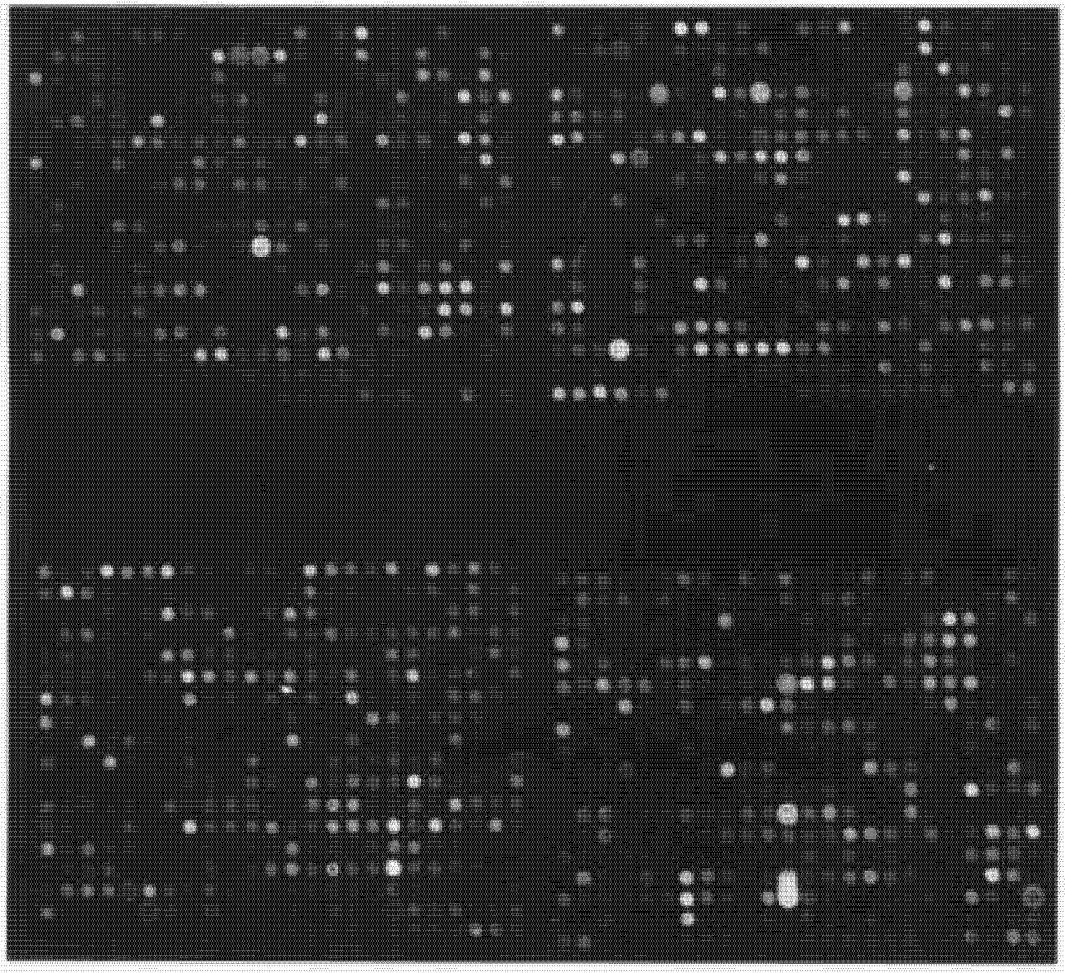


FIGURE A10-2 Expression Patterns in a Human cDNA Array (from cells induced with the human tumor suppressor PTEN, protein tyrosine phosphatase, and tensin homolog)

A portion of a human cDNA array hybridized with a red fluor-tagged experimental probe and a green fluor-tagged reference (uninduced) probe. The measured intensities from the two fluorescent reporters have been false-colored red and green and overlaid. Yellow signals (red plus green) indicate roughly equal amounts of bound cDNA from each sample and therefore have equal intensities in the red and green channels. Spots containing mRNA present at a higher level in one or the other sample show up as predominantly red or green. (Image provided by Vivek Mittal and Michael Wigler, Cold Spring Harbor Laboratory.)

median, and standard deviation of pixel intensities in each feature is determined and subjected to further detailed analysis. The following image analysis programs are commonly used:

- ***GenePix Pro*** is a program for scanning and analyzing microarray slides using the GenePix 4000 Scanner. It can also be used to analyze images acquired with other scanners. GenePix Pro automatically aligns the blocks of feature indicators on an image and aligns feature indicators within the blocks with features on the image. It is fully integrated with the Web in such a way that the gene names or their accession numbers can be linked to any database. It can be purchased directly from Axon (www.axon.com).
- ***ScanAlyze,*** a program similar to GenePix Pro, was developed by Michael Eisen at Stanford University and is freely available from http://rana.stanford.edu/software/.

Other image analysis programs include ArrayVision (from Imaging Research Inc.: http://imaging.brocku.ca/products/); deArray (from NHGRI: www.nhgri.nih.gov/DIR/LCG/15K/HTML); ImaGene (from BioDiscovery: http://www.biodiscovery.com/products/ImaGene/imagene.html); TIGR Spotfinder (from TIGR: www.tigr.org/softlab/); MicroArray Suite (from Scanalytics: www.scanalytics.com/sos/product/hts/microarray.html); GenExplore (from Applied Maths: www.applied-maths.com/ge/ge.htm); GeneData AG (from Basel: www.genedata.com/prod-exp.html); Partek Pro 2000 (from Partek Inc.: www.partek.com/); and Spotfire.net (from Spotfire: www.spotfire.com).

Some Considerations in Performing a Microarray Experiment

Because microarray experiments are miniaturized and of a high-throughput nature, they are sensitive to both external and internal fluctuations, for example, during delivery of the target DNAs onto the substrate and during hybridization. Approximately 7–8% of the variation in the data is estimated to be due to background noise. It is therefore imperative to follow proper averaging and normalization procedures, both to minimize these variables and to ensure that the extracted data are reliable and meaningful. It is also important to note that, as of this writing, there are no globally accepted rules or standards for performing controlled microarray experiments. Furthermore, the platforms as well as the methods for data extraction and analysis are subject to great variability. The following are general recommendations for performing experiments on homemade arrays and should be used as guidelines only.

- ***Replicate each experiment on multiple arrays.*** Averaging the normalized ratios for the same controls to compensate for array-to-array variation.
- ***Perform experiments using color reversal to account for variations in dye incorporation.*** The purpose of reversing the label for the two replicates is to reduce variability in signals that can occur due to differential incorporation of fluorescent labels by reverse transcriptase or other associated technical variability. False positive results can be substantially reduced using this strategy.
- ***Randomize the address of arrayed DNAs on the surface of the substrate*** to control for variation in hybridization that may occur at different locations on the array.
- ***Include a large number of controls on every array*** to verify the success of the probe synthesis, labeling, and hybridization. Some recommended controls include:

  Doping/spiking RNA controls, where mRNAs synthesized from cDNA clones by in vitro transcription are added to each labeling reaction. Ideally, these mRNAs are derived from genes that are not homologous to the organism under study (e.g., nonhomologous plant mRNAs are used on the human expression array). The cognate genes for these mRNAs

are represented on the array. These controls are useful for determining variations in differential labeling of samples, for assaying the stringency of the hybridization, and for calibrating the ratios for comparisons among experiments.

Intensity series clones, where clones from the organism under study that represent high, medium, and low expression are represented on the array. These are useful for estimating sensitivity of detection and measuring relative saturation and scanning effects.

Blocking controls. Most array hybridization experiments require the addition of cold DNA (e.g., human C_0t1 DNA and polyadenylic acid for human cDNA expression arrays) to block nonspecific annealing including hybridization due to repetitive elements. Inclusion of human C_0t1 DNA or oligo(dA) spots on human expression arrays should have little or no hybridization signal if repetitive elements have been successfully blocked.

Background hybridization controls. It is often useful to include blank salt spots and spots of DNA from another organism whose mRNA is not represented in the sample. These controls give an estimate of the nonspecific component of the hybridization and are useful for subtracting background from feature values.

Contamination controls. DNA samples that are used multiple times for printing arrays may become contaminated with bacteria or yeast. Total genomic DNA or cDNA clones of common contaminants such as *E. coli* and yeast are represented in the array to monitor for spurious hybridizations.

cDNA synthesis and RNA label quality control. Hybridization intensity is often dictated by the size of the labeled sample. This variable can be monitored with appropriate quality control elements. For example, DNA fragments, encompassing 100–200 bp tiled across an entire gene, can be spotted on arrays. Labeled cDNA probes of high integrity will hybridize efficiently to all spots, whereas cDNAs of lesser quality will generate proportionately weaker signals with fragments that map to the 5′ end of the gene.

- ***Control of geometric artifacts.*** The use of standard floating coverslips for DNA array hybridization sometimes results in uneven distribution of the probe DNA on the array, which compromises the hybridization. This situation occurs because the hybridization volume under a coverslip is small, which results in inefficient mixing of the hybridization fluid and generation of localized temperature gradients. The compromise is manifest by geometrical distortions in the image due to variations in hybridization at different locations within the array. One way to circumvent this variable is to use custom-made raised-edge coverslips (also referred to as "lifters"; available from www.eriesci.com/). These coverslips provide separation and ensure even dispersal of hybridization solution between the array and the coverslip. Another solution is to use larger hybridization volumes within specially designed hybridization chambers. These ensure efficient mixing of the hybridization fluid, which brings more probe molecules into contact with the cognate target in the array, thus increasing the number of productive events. Hybridization chambers are available from Schleicher & Schuell (www.s-und-s.de/) and CLONTECH (www.clontech.com).

MICROARRAY DATA MINING, ANALYSIS, AND MANAGEMENT

Because microarray analysis is a high-throughput technology, the amount of data being generated is expanding at a tremendous rate. The handling and analyses of these data therefore require elaborate databases, query tools, and data-visualization software. A brief description of some of these databases and analysis software follows:

- ***Another MicroArray Database.*** AMAD is a flat file, Web-driven database system written entirely in PERL and javascript, and intended for use with microarray data. AMAD supports both ScanAlyze and GenePix image analysis software. Further multivariate statistical analyses are performed on the values obtained from the data analysis program by using Mike Eisen's Cluster analysis and TreeView; both are freely available from www.rana.Stanford.edu/software/. AMAD, developed by a group of experienced scientists originally from Stanford University, is the only free database (available from www.microarrays.org) and provides a reliable analysis and data storage tool for investigators with little or no experience in handling enormous amounts of microarray data.
- ***Resolver (Rosetta Inpharmatics).*** Resolver is a comprehensive storage, visualization, and analysis tool for high-volume gene expression data obtained using cDNA microarrays, oligonucleotide arrays, and other technologies. It is interlaced with a powerful discovery tool for pathway interrogation, functional gene assignment, and compound analysis. Multi-experiment analyses are possible through correlation plots, cluster trees, and BLAST-like searches that can be conducted within the database. This feature enables the user to perform ranked similarity searches across entire data sets containing tens of thousands of gene expression profiles. More information about Resolver is available at www.rii.com/prodserv/resolver/index.htm.
- ***GeneSpring (Silicon Genetics)*** is a powerful analysis and visualization software suite available for genomic expression experiments. It is capable of handling and analyzing enormous data sets from any organism. Clustering, pathway determination, regulatory sequence detection, Eigenmode analysis, and a variety of data visualization tools are principal features of GeneSpring. A copy of a free demonstration version and software for purchase are available at www.sigenetics.com.
- ***IPLab MicroArray Suite for Macintosh (Scanalytics)*** is image analysis software capable of performing basic routines for extracting and visualizing microarray data generated from virtually any scanning device. Originally developed by researchers at the National Institute for Human Genome Research (NHGRI), this software is now a completely supported product available from Scanalytics Inc. (www.scanalytics.com/sos/product/hts/microarray.html).

A DNA microarray database can be linked to a reference database, for example, the UniGene database, to allow access to information about particular genes. Each UniGene entry corresponds to a single human or rodent gene and provides a direct link to GenBank gene and EST entries, to SWISS-PROT entries, and to literature entries through PubMed. Several groups have begun to assemble comprehensively curated gene databases for mammals. Such databases will facilitate rapid retrieval of information concerning functional and biochemical pathways. Some of the databases under development are:

Kyoto Encyclopedia of Genes and Genomes (KEGG)	http://www.genome.ad.jp/kegg/
Locus Link	www.ncbi.nlm.nih.gov/LocusLink/
WIT	wit.mcs.anl.gov/WIT2/
SPAD Signaling pathway Database	www.grt.kyushu-u.ac.jp/spad/
Genecards	http://bioinformatics.weizmann.ac.il/cards/
BRITE	www.genome.ad.jp/brite_old/
EGAD	www.tigr.org/docs/tigr-scripts/egad_scripts/role_report.spl
Stanford genome resources	http://genome-www4.stanford.edu/cgi-bin/SMD/source/sourceSearch
Database of transcribed sequences	www.cbil.upenn.edu/DOTS
South African National Bioinformatics STACK database	www.sanbi.ac.za.Dbases.html
TIGR Gene Indices	www.tigr.org/tdb/tgi.shtml

The large-scale data sets normally obtained from related microarray experiments are subjected to pattern recognition analysis to identify groups of genes that are regulated in a similar manner across many experiments. Pattern recognition is performed using a variety of multivariate analytical algorithms, such as hierarchial clustering, k-means clustering, and self-organizing maps. Typically, the interpreted array analysis highlights a relatively smaller number of spots representing differentially expressed mRNAs whose cognate genes are further validated by resequencing and whose patterns of expression are confirmed by other more reliable but low-throughput methods such as northern blotting, nuclease protection, or RT-PCR (for details of these protocols, please see Chapter 7, Protocols 8 and 10, and Chapter 8, Protocol 8).

As large-scale gene expression data accumulate over time, public access to these data becomes an important issue. In its current format, the data are widely dispersed and lack uniform structure and retrieval modalities. Efforts to establish some type of standardized data format for storing and communicating microarray-based gene expression data are in progress, largely due to efforts of the European Bioinformatics Institute (www.ebi.ac.uk/microarray/MGED/). The main objectives of this effort are (1) to ensure reproducibility and verifiability of results, (2) to identify controlled vocabularies for annotating the samples and experiments, and (3) to define the data communication standards. The development of standards for cross-platform data comparison and normalization is also under way.

EMERGING TECHNOLOGIES

The field of microarray technology presents a tremendous technical challenge for both academic institutions and industry. Commercial companies are vying with one another to establish market dominance. Novel platforms are being developed that promise higher throughput and better reproducibility; but which technology will emerge as the system of choice remains to be determined. This discussion surveys some of the emerging technologies that are likely to have a significant impact on the future of microarray research.

Novel Microarray Fabrication Methods

- ***Piezoelectric Printing*** uses technology analogous to that employed in ink-jet printers. The sample is loaded into a miniature nozzle equipped with a piezoelectric element around the nozzle. An electric current is applied to the piezoelectric element, which causes the nozzle to expel a precise amount of liquid from the jet onto the substrate. After the first step, the jet is washed and the second sample is loaded and deposited to an adjacent address. This method allows high-density gridding of virtually any molecule of interest, including DNA, antibodies, and small molecules, onto the substrate of choice. Arrays can be produced on nonporous (glass), porous (filters), and three-dimensional (Hydro-gel) substrates. The piezoelectric technology is not currently as robust as photolithography or microspotting. The BioChip Arrayer from Packard Instrument Co. uses piezoelectric "drop-on-demand" tips to provide noncontact dispensing (www.packardinst.com).
- ***Bubble Jet Technology*** uses the ink-ejecting mechanism used in printing devices. A Bubble Jet ink-jet device is used to eject 5′-terminal-thiolated oligonucleotides onto glass surfaces. The printing head (which is similar to conventional piezoelectric devices) consists of a heater that generates a bubble of fluid in a nozzle. The bubble pressure then ejects a microdroplet of material from the aperture.

Alternate approaches for fabrication of high-density in-situ-synthesized oligonucleotide arrays are under way. The Digital Optical Chemistry (DOC) system developed by Skip Garner and colleagues at the University of Texas Southwestern Medical Center (www.pompous.swmed.edu/) consists of three parts: a Digital Micromirror Device that selectively focuses ultraviolet light onto a glass substrate where chemistry is done, a fluidics system that delivers the chemical reagents in the proper sequence, and a computer program that controls both the Digital Micromirror Device and fluidics system according to the desired sequence pattern. This system alleviates the need for expensive photolithographic masks (using instead a "virtual" digital mask) and generates unique chips very rapidly. Another benefit of using DOC is that the machine is a benchtop unit that can be manufactured for use in any laboratory. In a similar technological direction, Xiaolian Gao at the University of Houston (www.zeiram.chem.uh.edu/gao/) uses a digital photolithographic system and new oligonucleotide deprotection chemistry using photogenerated acids and newly developed synthesis microreactors. These technological improvements should in the future permit chips of any design to be made in regular research laboratories at an affordable cost and with improved quality.

Resequencing

Also known as sequencing by hybridization (SBH), resequencing uses a set of oligonucleotides comprising all possible combinations of sequences in a given length that are synthesized and immobilized on a chip. The DNA fragment to be sequenced is broken down into smaller pieces, fluorescently labeled, and hybridized to the chip. The sequence of the DNA emerges from the pattern of fluorescence bound to the nested sequence. Since the chip contains all possible sequences, this system has the distinct advantage of being able to sequence DNA from any source. Fragments containing nonrandom sequences such as direct and inverted repeats presently limit this technique. The technology is being developed by Hyseq Inc., in collaboration with PE Corporation (www.hyseq.com).

Specific, High-throughput Genotyping by Primer Extension on Microarrays

A system to genotype SNPs and point mutations by a DNA-polymerase-assisted primer extension reaction using a microarray format is being developed by Ann-Christine Syvanen at Uppsala University in Sweden. Miniaturized reaction chambers formed on microscope slides allow analysis of 80 individual samples for hundreds of SNPs, potentially allowing the generation of >20,000 genotypes per slide. Minisequencing single-base extensions from one primer per SNP with dideoxynucleotides labeled with four fluorophores, or alternatively, extension of two allele-specific primers per SNP with one fluorescent deoxynucleotide, may be used. The genotypes of the samples at each SNP are determined by measuring the slides in a fluorescence array scanner.

Direct Identical-by-descent Mapping Using DNA Arrays

This mapping method is being developed in both the Brown (www.cmgm.stanford.edu/pbrown/) and the Cheung (www.w95vcl.neuro.chop.edu/vcheung/) laboratories and facilitates the isolation and physical mapping of DNA fragments shared identical-by-descent (IBD) between individuals. IBD mapping is a combination of two techniques: genomic mismatch scanning (GMS), a method for genetic linkage mapping, and DNA microarray technology. In this method, specialized enzymes are used to isolate DNA fragments shared IBD between two individuals. The isolated DNA fragments are mapped at high resolution by hybridization to a DNA array representing physically ordered genomic segments. Compared to traditional genotyping methods, direct IBD

mapping allows the entire genome to be analyzed in one step without laborious locus-by-locus genotyping. Microarray-based genetic linkage mapping will be useful in mapping genes that underlie complex genetic traits in yeast and in the distribution of meiotic recombination events across the entire genome. Parallel work is directed at adapting the GMS methodology to the human genome. In a demonstration of the feasibility of this technique, a gene for congenital hyperinsulinemia has been mapped by the Cheung lab.

Protein Microarrays

As a complementary approach to gene expression profiling on cDNA microarrays, microarrays of specific antibodies are being developed to measure the abundance of thousands of different proteins in samples from cells, or in biological fluids, such as serum or urine. Preactivated surfaces are used for the covalent immobilization of antibodies. Proteins from experimental and reference samples are differentially labeled with fluorescent dyes and hybridized to the array. Evaluation of the various applications of protein microarrays in detecting and diagnosing disease is under way. Because this application would not be restricted to antigen-antibody systems, protein microarrays should provide a general resource for high-throughput screens of gene expression, receptor-ligand interactions, and protein-protein interactions. Information on the development of protein arrays is available at www.molgen.mpg.de/~proteingroup/LuekingetalPub.html.

Tissue Microarrays

This array-based high-throughput technique facilitates gene expression and copy-number surveys of very large numbers of tissue specimens. For example, as many as 1000 cylindrical tissue biopsies from individual tumors can be distributed in a single tumor tissue microarray. Sections of the microarray provide targets for parallel in situ detection of DNA, RNA, and protein targets in each specimen on the array, and consecutive sections allow the rapid analysis of hundreds of molecular markers in the same set of specimens. Tissue microarrays were developed in the Olli-P. Kallioniemi laboratory at the National Human Genome Research Institute (www.nhgri.nih.gov/DIR/CGB/TMA/about.html). A machine to make tissue microarrays is available from Beecher Instruments (www.beecherinstruments.com).

Use of Arrays to Detect DNA-Protein Interactions

The potential of DNA microarrays for high-throughput screening of DNA-protein interactions is being exploited in several laboratories. One of the major interests is the detection of transcription-factor-binding sites in the genome under different physiological conditions. The general approach uses construction of an array containing both the intergenic and the coding region of the genome. Next, the total DNA and protein content of a cell is cross-linked in vivo, and a chromatin immunoprecipitation (ChIP) experiment using antibody specific to a transcription factor is performed. The DNA component of the immunoprecipitate (DNA–protein–antibody complex) is labeled with fluorescent dyes by PCR and hybridized to the array. A positive spot on the array is likely to be the target of the protein. This approach is feasible for organisms whose complete genomic sequence is known. Methods to analyze genome-wide protein-DNA interactions in yeast are being developed by Iyer and Brown at Stanford University.

Using a more general approach, the Church laboratory (http://arep.med.harvard.edu/gmc/) has generated arrays of single-stranded DNA oligonucleotides (all possible sequence combinations) carrying a constant sequence region at one end. The single-stranded arrays are converted to a double-stranded array by annealing and enzymatically extending a complementary primer.

Arrays containing all possible permutations of a site at each position will facilitate exploration of the spectrum of sequence-specific protein-binding sites in genomes.

Barcode Chip

In an effort to assign a function to every gene in the yeast genome, a novel high-throughput method is under development at Stanford University. The project is described on the Web Site http://sequence-www.stanford.edu/group/yeast_deletion_project/deletions3.html.

A library of yeast strains, harboring a deletion in each gene, is generated. Each of the deletions is tagged with a unique 20-mer DNA sequence that acts as a molecular barcode for individual deletions. The mixture of all such tag strains then allows for the analysis of the entire genome with the manipulation of a single sample. During growth under a variety of conditions, the loss of a tag indicates the loss of a corresponding deletion from the population. The concentration of each tag is determined by PCR amplification of the mixture and hybridization to a chip that contains tag complements at defined positions. Thus, phenotypes of individual strains can be analyzed in parallel. This approach has prospects for use in other organisms as well.

Bioelectronic Chips

A number of companies are focusing on creating microchips using microelectronics. Nanogen-chips (www.nanogen.com/) contain probes that can be electrically activated individually. A sample is applied to the chip, and the spot is electrically activated, which allows the sample to be concentrated and bound efficiently. The chip is then washed and another sample applied. Thus, multiplex hybridizations can be performed to the same chip. This technology has potential applications in a number of other analyses, including antigen-antibody, enzyme-substrate, cell-receptor, and cell separation techniques.

Orchid Biocomputer (Princeton, NJ; www.orchid.com) is designing a microfluidics chip that has precise control over the flow of process chemicals and temperature integration. This chip is used for screening SNPs. Both of these technologies have implications in drug screening and clinical diagnostics.

SUMMARY

From providing detailed profiles of the differential expression of literally thousands of genes to exploiting protein microarrays to build a resource of protein interactions, the applications of array technology are likely to produce an explosion of information on many biological fronts. The final challenge, beyond the emerging technological developments, may well lie in developing effective means for storing, sharing, and analyzing the information.

Appendix 11

Bioinformatics

David Mount, *University of Arizona, Tucson*
Judy Cuddihy, *Cold Spring Harbor Laboratory*

The following tables present widely used software for analyzing DNA, RNA, and protein sequences, as well as databases and types of searches that can be performed for DNA, RNA, and protein structure and function. The organization for Table A11-1 is outlined below:

- DNA
  - Sequence Submission
  - Sequence Alignments (Pairwise and Multiple)
  - Scoring Matrices
  - Motifs and Patterns
  - Genes, Exons, and Introns
  - Promoters, Transcription-factor-binding Sites
  - Other Regulatory Sites
- RNA
  - Secondary Structure
  - RNA-specifying Genes, Motifs
- Protein
  - Sequence Alignment
  - Motifs, Patterns, and Profiles

Because of the ever-changing nature of software and World Wide Web sites, it is worth the effort to consult the relevant Web or FTP sites for more information and the latest versions of software and databases. In most instances, a license is needed if the programs or databases are used for commercial purposes; see each site for further information. Commercial programs have not been included in this survey. For more information on these options, refer to the Web Sites of the relevant companies.

In addition to the specific sites for software and databases, there are several "gateway" Web pages that offer access to many of the sites listed in the tables, as well as to additional sites. The following Web pages offer a good starting place for many searches:

National Center for Biotechnology Information (NCBI)
Web Site: http://www.ncbi.nlm.nih.gov

This site offers a wide range of databases and searches via Web-based interactive forms, including GenBank, PubMed, BLAST, genome biology resources and databases, cancer resources and databases, and many other resources. These resources are integrated by Entrez, a search and retrieval system that also includes cross-referenced information. Submission to the GenBank database can be made here.

Baylor College of Medicine Search Launcher
Web Site: http://www.hgsc.bcm.tmc.edu/SearchLauncher

Web-based interactive forms and links are provided for many types of searches, including nucleic acid, protein sequence and pattern, species-specific protein, sequence alignments, gene features, and protein secondary structure, as well as for many sequence utilities.

The Sanger Centre
Web Site: http://www.sanger.ac.uk

Web-based interactive forms and links are provided to a wide range of databases and software utilities, including the many genome sequencing projects.

EBI-European Bioinformatics Institute, EMBL Outstation
Web Site: http://www.ebi.ac.uk

Web-based interactive forms and links to databases and software utilities are given. Submissions to EMBL Data Libraries can be made here.

ExPASy Molecular Biology Server-Expert Protein Analysis System, Swiss Institute of Bioinformatics
Web Site: http://www.expasy.ch

This protein-specific site provides a wide range of links to database and software utility resouces for analysis of protein structure and function.

For more extensive information on the mathematical and statistical underpinnings of bioinformatics software, as well as methods for database searching and analysis of DNA, RNA, and protein sequence, please see Mount (2001).

TABLE A11-1 Bioinformatics Software

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
Sequence Submission				
BankIt	http://www.ncbi.nlm.nih.gov/BankIt	Web-based sequence submission tool for direct submission to GenBank; submitted sequence must be DNA/RNA sequence that has been sequenced by the submitter. This tool is for one or a few submissions with uncomplicated sequence annotation.	FASTA	Returned by e-mail are: preliminary GenBank flat file; GenBank accession number; completed GenBank flatfile.
Sequin (formerly called Authorin)	http://www.ncbi.nlm.nih.gov/Sequin/index.html *FTP:* www.ncbi.nlm.nih.gov/sequin	Submission and update of sequence submission to GenBank, EMBL, and DDJB. Handles simple and complex submissions; automatically performs functions that need to be done manually in other submission programs. Has powerful sequence annotation tools, built-in validation functions; graphical and sequence views. Runs on UNIX and PC; has NetworkAware mode to exchange information with NCBI.	FASTA for sequence files; also PHYLIP, NEXUS, MACAW, FASTAGAP for population, phylogenetic, and mutation studies.	GenBank flat file; GenBank accession number.
Webin	http://www.ebi.ac.uk/Tools/index.html	Web-based interface preferred for EMBL sequence submission. Data can be modified and viewed before submission; has bulk submission procedure for 25 or more related sequences; also EBI vector screening software.	EMBL Data Libraries accession number.	Local alignment.
Sequence Alignments (methods for aligning two sequences at a time; for database searches that give pairwise alignments, see Table A11-2)				
BLAST (*B*asic *A*lignment *S*earch *T*ool) (Altschul et al. 1990; Karlin and Altschul 1990, 1993; Tatusova and Madden 1999)	http://www.ncbi.nlm.nih.gov/gorf/bl2.html	The BLAST algorithm is primarily used for similarity searches in sequence databases, but it may also be used for aligning a pair of sequences. The BLAST Web Site also provides a page for aligning two sequences by first searching for conserved words and using these to seed an alignment.	FASTA or Accession/Genbank index number (web site)	Local alignment of sequences based on word "seeds" and graphical representation of alignment.
FASTA programs				
FASTA (Pearson and Lipman 1988)	http://fasta.bioch.virginia.edu *FTP:* ftp.virginia.edu, dir /pub/fasta	FASTA is used primarily for similarity searches in sequence databases, but it may also be used for aligning a pair of DNA sequences or a pair of protein sequences. Web-based interactive form or downloadable software. Permission needed for commercial use. FASTA may be established on a variety of computer platforms, including Macintosh and PC Windows, but without a graphical interface.	FASTA or Accession/Genbank index number (web site)	Single, local (Smith-Waterman) alignment between two sequences.

(Continued on following pages.)

TABLE A11-1 (*Continued*)

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
LFASTA	http://fasta.bioch.virginia.edu	Compares two protein or DNA sequences for local similarity; uses the FASTA word search method to locate regions of similarity and then performs and scores a local (Smith Waterman) alignment. BLOSUM50 scoring matrix used by default but other matrices can also be specified; gap penalties can also be designated. Use LALIGN or SIM instead.	FASTA	Local sequence alignment.
FASTX/FASTY	http://fasta.bioch.virginia.edu	Compares translated DNA sequence to protein sequence database.	FASTA	Translates DNA sequence in 3 forward and 3 reverse frames; allows frameshifts.
LALIGN, LALIGN0, PLALIGN (see Huang et al. 1990; Huang and Miller 1991)	http://fasta.bioch.virginia.edu	Compares 2 DNA or 2 protein sequences to identify regions of sequence similarity; default scoring matrix is BLOSUM50 but others can be specified; optional gap penalties; uses sim algorithm to provide a local (Smith-Waterman) alignment; provides *n* different alignments with the same two residues never aligned more than once using the SIM method; can identify internal repeats. LALIGN scores penalties for gaps at the ends of the alignment; LALIGN0 does not penalize end gaps. PLALIGN produces a graphical plot of the alignment much like a dot matrix.	FASTA	Reports several sequence alignments if there are similar regions as well as similarity scores. Web site reports significance of alignment score based on PRSS analysis (see next entry).
PRSS (see Huang et al. 1990; Huang and Miller 1991)	http://fasta.bioch.virginia.edu	Evaluates the significance of pairwise similarity scores of 2 DNA or 2 protein sequences; uses BLOSUM50 matrix and standard gap penalties as default for proteins, but others may be used; second sequence is scrambled (at level of individual characters or of words) many times and each scrambled sequence is aligned with first sequence to give a range of unrelated sequence scores. From these scores, an Expect value for the original alignment score is calculated, i.e., the number of alignments between unrelated sequences that can achieve such a score.	FASTA	Reports similarity scores for 2 sequences; the second sequence is then shuffled a specified number of times compared with the first sequence.
Bayes Block Aligner (primarily used for sequence alignments but also has been used for similarity searching in sequence databases) (Zhu et al. 1997, 1998)	www.wadsworth.org/resnres/bioinfo	Finds all possible blocks (short ungapped alignments between a pair of sequences up to a specified number of blocks), then generates all possible alignments of the sequences that includes compatible sets of these blocks. There is no gap penalty; unmatched sequence regions between blocks are neither aligned nor scored. Scores alignments with a series of log odds scoring matrices to provide an odds score for	FASTA	Most probable sequence alignments according to different sampling criteria; probability that a given pair of residues is aligned; probability of each scoring matrix and each choice for number of blocks, and Bayesian probability that sequences are related.

		each alignment and matrix combination. The sum of all of these scores is calculated and the contribution of each alignment, block number, and scoring matrix to this sum is used to calculate posterior probabilities for the alignment, scoring matrix, etc., using conditional probability calculations (Bayesian statistics). Sequence characters aligned most often regardless of other variables may be determined. Download and compile software for Sun Solaris or SGI Irix or as C source code. PC Windows version also available with no graphic interface. Licensing agreement needed.		
SIM (Huang et al. 1990; Huang and Miller 1991)	http://www.expasy.ch/tools/sim-prot.html	Like LALIGN, finds *n* alignments by a local (Smith-Waterman) alignment algorithm with much increased speed of calculation.	FASTA	Reports *or* local alignments in order of similarity score.
Scoring Matrices				
BLOSUM scoring matrices (see also BLOCKS database) (Henikoff and Henikoff 1992, 1993, 1994; Henikoff et al. 1995)	http://www.blocks.fhcrc.org	Examines multiple alignments of related protein regions, in order to produce a matrix of amino acid substitutions scores, rather than extrapolating from scores of divergent but closely related sequences (as do the PAM scoring matrices). Most alike proteins in the alignment may be clustered to reduce the frequency of changes among the more commonly represented amino acids (e.g., for BLOSUM62, sequences that are 62% or more identical are clustered). There is no specific evolutionary model for these matrices as there is for PAM matrices; all sequences are considered just as likely to be the ancestor of the others (a star phylogeny).		Detection of similarity between protein sequences as evidence of evolutionary homology.
DNA PAM matrices	http://blast.wustl.edu/, see "Improved sensitivity of nucleic acid database searches using application-specific scoring matrices."	Assumes Markov model of change in DNA sequences (see PAM matrices below).		Prediction of evolutionary distance between DNA sequences.
PAM (*P*ercent *A*ccepted *M*utation or mutations that survive natural selection) (Dayhoff et al. 1978; Schwartz and Dayhoff 1978; for more recent versions, please see Gonnett et al. 1992; Jones et al. 1992)	FTP: ftp://ncbi.nlm.nih.gov, directory/blast/matrices	Predicts the expected amount of substitution in protein sequences that have had a given amount of evolutionary time to diverge from a common ancestor sequence. Assumes a Markov model of change (no site-to-site variation and no change in composition) and forward and reverse changes equally as likely. Matrix values give log odds scores for matches or mismatches. Best odds score for alignment is found when scoring matrix that matches number of mismatches in the alignment is used.		When used with alignment program, provides measure of similarity between 2 sequences being characterized by evolutionary distance. 1 PAM corresponds to the average change in 1% of all amino acid positions.

(Continued on following pages.)

TABLE A11-1 *(Continued)*

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
DNA Sequence Alignment: Multiple Sequence Alignment				
MSA (Lipman et al. 1989)	http://www.psc.edu/general/software/packages/msa/msa.html *Interactive Web forms at:* http://www.ibc.wustl.edu/service/msa/index.html (other programs also available on this site)	Aligns several nucleic acid or protein sequences globally; uses dynamic programming to produce optimal alignment of all sequences at the same time; sequences are weighted in accord with their similarity to other sequences in the group. Uses sum of pairs score for evaluating columns in the alignment; normally limited to 3 sequences because it requires large amounts of computer time and memory. Three versions allow alignment of >3 sequences; up to 50 sequences where each sequence has <150 residues; up to 25 sequences where each sequence has <500 residues; up to 10 sequences where each sequence has <1000 residues. Interactive Web-based form; UNIX.	FASTA, GCG, PIR and other formats recognized by the READSEQ program.	Initially shows heuristic alignment based on a progressive pairwise alignment; this alignment is used to limit the search space for an optimal alignment; an optimal alignment is then shown; calculation of alignment costs; divergence values for each pairwise alignment; maximum score of optimal alignment.
CLUSTALW (Higgins and Sharp 1988; Higgins et al. 1992; Thompson et al. 1994)	*Interactive Web-based form at:* http://www.ibc.wustl.edu/ibc/msa.html http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html http://www.infobiogen.fr/ docs/ClustalW/clustalw.html *FTP:* ftp.ebi.ac.ukdir/pub/software/unix/clustalw.tar.Z	Performs multiple alignments on a set of DNA or protein sequences or adds new sequences to an existing alignment; uses a progressive alignment of sequence pairs and groups using as a guide of similarity a neighbor-joining phylogenetic tree initially made by pairwise alignments of the sequences. The contribution of individual sequences to the alignment is weighted in accord with the amount of divergence from the other sequences (different from MSA). Uses position-dependent gap penalties depending on the estimated divergence in each region of the aligned sequences. Excellent method for sequences that are not very divergent; can make errors in the initial alignments of most-alike sequences that becomes more possible with increasing divergence. Matrices that may be used are BLOSUM80,62, 45,30; Gonnett80,120,160,250,350; PAM20,60, 120,350. CLUSTALW (weighted alignments) is a major update and rewrite of CLUSTALV. Interactive Web-based form; ANSI-C version for UNIX and VMS or any machine with ANSI-C compiler; executables for major platforms MAC, PC. Clustal X provides a graphic interface for an X windows environment.	FASTA; also GCG/MSF, NBRF/PIR, EMBL/Swiss Prot, GDE flat file, CLUSTAL, and GCG9-RSF formats. All sequences to be aligned must be in one file. Formats recognized by REASEQ program may be used.	Output formatted in CLUSTAL (shows sequence aligned in blocks), NBRF/PIR, GCG/MSF for input into GCG software; PHYLIP for input into Phylip, GDE flatfiles. For graphical views of the phylogenetic trees calculated after sequence alignment, other software, such as Phylip, must be used.
PIMA (Pattern-induced Multiple Alignment) (Smith and Smith 1990, 1992)	*Interactive Web-based form at:* http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html *FTP:* ftp.ebi.ac.ukdir/pub/software/unix	Constructs multiple sequence alignments by first performing pairwise alignments between the sequences and clustering the sequences into one or more families based on the scores.	See previous entry.	Multiple sequence alignment.

		Common patterns are located in each family and the set of sequences is then aligned using these patterns as starting points; uses constraints to determine gap locations. Interactive Web-based form; UNIX.		
MACAW (Multiple Alignment Construction and Analysis Workbench) (Schuler et al. 1991)	*FTP:* ncbi.nlm.nih.gov:dir/pub/macaw	Locates, analyzes, and edits blocks of localized sequence similarity among multiple sequences; constructs multiple alignments by analyzing, editing, and combining blocks of aligned sequence segments. Sequence blocks are located using a sum of pairs scoring system to produce a local multiple sequence alignment or the Gibbs sampler (see Gibbs sampler, below), a statistical method, is used. The user interface allows the editing and linking of blocks. PC, NT, Mac.	FASTA	Multiple sequence alignment based on blocks; statistical evaluation of block similarity
SAM (Sequence Alignment and Modeling System) (Haussler et al. 1992; Hughey and Krogh 1995, 1996)	http://www.cse.ucsd.edu/research/compbio/sam.html *FTP:* ftp.cse.ucsc.edu:dir/pub/protein/sam1.01.tar.Z.crypt	Uses a hidden Markov model to represent a multiple sequence alignment of a set of full sequences or partial sequences. The HMM allows matches, insertions, and deletions in the alignment. A model is produced for each sequence set by training the model with a family of related sequences. A large number of sequences (>50) are required to produce a representative model, and the larger the set, the better the model produced. A multiple sequence alignment may be used as input. The trained HMM will produce a multiple sequence alignment of the training sequences, and may also be used to search for new family members in sequence databases. HMMs have the advantage of offering a well-defined probabilistic model of sequence alignments. Conversion programs for HMMER formats are included. Encrypted source code; UNIX, Mas-Par.	DNA, RNA, and protein alphabets; user-defined alphabets, most common formats such as FASTA, GenBank, NBRF, EMBL, GCG, DNA Strider, Phylip, GCG/MSF, PAUP NEXUS, PIR CODATA; can be given a predetermined multiple sequence alignment. Uses READSEQ program to recognize sequence formats.	Hidden Markov model-based multiple sequence alignment; database search for new family members.
HMMER (Haussler et al. 1992; Hughey and Krogh 1995, 1996)	http://genome.wustl.edu/eddy/hmmer.html *FTP:* genome.wustl.edu:dir/pub/eddy	Uses hidden Markov models to produce multiple sequence alignment of DNA, RNA, or protein sequences (see also SAM entry). Uses hidden Markov models to search databases for homologs of the query sequence family. UNIX and ANSI-C versions.	FASTA, GenBank, EMBL, SwissProt, GCG/MSF, SELEX	Hidden Markov model-based multiple sequence alignment. Smith-Waterman local alignments found by database searching. Global Needleman Wunsch alignments found by database searching.
SAGA (Sequence Alignment by Genetic Algorithm) (Notredame and Higgins 1996)	http://www.ebi.ac.uk/~cedric	Creates multiple sequence alignments using genetic algorithms (developed as computer science tool) and user-defined objective functions (descriptors of multiple alignment quality). Software evolves population of alignments in	See program notes.	Globally optimal multiple sequence alignment.

(Continued on following pages.)

TABLE A11-1 *(Continued)*

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
		quasi-evolutionary manner; improves the fitness of the population using the objective functions. The software can align <20 sequences that are <400 residues long. A related program, COFFEE (Notredame et al. 1998; available at same site), can be used to optimize the multiple sequence alignment obtained. ANSI-C and UNIX versions.		
DIALIGN segment alignment (Morgenstern et al. 1996)	http://www.gsf.de/biodv/dialign.html	Compares whole segments of sequences instead of single residues. Constructs pairwise and multiple sequence alignments from gap-free pairs of equal-length segments, called diagonals. Especially useful for detecting local similarities in otherwise unrelated sequence. Download executables for DEC, SUN, HP, SG, CONVEX, and LINUS. (For commercial use, contact genomatix@gsf.de. DIALIGN2 offers a Web-based interactive form.	EMBL, FASTA, GCG/RSF, GenBank.DNA/protein sequences <1 Mb.	Alignment of input sequence; graphical representation of degree of local similarity between aligned sequences; sequence tree showing degrees of similarity.
MultAlin (Corpet 1988)	http://protein.toulouse.inra.fr/multalin.html	Web-based interactive form for multiple sequence alignment. Alignment parameters can be set; consensus level options can be chosen for output.	Multi-Alin FASTA; EMBL-SwissProt; GenBank	Multiple sequence alignment output as colored image, plain text, or HTML.
PRRP progressive global alignment (randomly or doubly-nested) (Gotoh 1995)	ftp.genome.ad.jp/pub/genome/saitama-cc	Iterative method of multiple sequence alignment based on a profile method in which groups of sequences are repeatedly aligned, guided by a tree until best score is obtained. Avoids errors made by progressive alignment methods when sequences are divergent.	Special format (see program notes).	Best alignment after repeated trials.
DNA: Motifs and Patterns				
MEME (*M*ultiple *E*xpectation *M*aximization for *M*otif *E*licitation) (Bailey and Elkan 1994, 1995; Bailey 1995)	http://www.sdsc.edu/MEME/meme/website/intro.html *Interactive Web-based form at:* www.sdsc.edu/MEME/website/meme.html	Detects motifs (conserved sequence patterns of the same length; no gaps; no covariations between positions) in groups of related DNA or protein sequences. The software uses a statistical technique (the expectation maximization or EM algorithm) to locate motifs of a given length range, of which there may be 0, 1, or more in an individual sequence, as specified by the user. The output is a PSSM (a position-specific-scoring-matrix) of each motif found with columns that represent motif positions and rows that represent the distribution of residues in each column in a log-odds format. The matrix may be used to search sequence		Results are e-mailed to user, including: PSSMs, which represent the distribution of residues in each column of the motifs found; diagram showing the information content of each column (represents degree of conservation at each motif position), multilevel consensus sequence, examples of motifs found, and location in input sequences; motif also output in BLOCKS format. PSSM may be subsequently

		databases using MAST. License needed for commercial use. Interactive Web-based form; ANSI source code.		used for database searches using MAST.
Meta-MEME (Grundy et al. 1997)	*FTP:* metameme.sdsc.edu *Interactive Web-based form at:* http://www.sdsc.edu/MEME/meme/website/intro.html	Combines DNA or protein motif patterns from MEME into a hidden Markov model (HMM) that represents an alignment of the sequences through alignments of the conserved motifs. The HMM may then be used to search sequence databases for other family members. Interactive Web-based form; ANSI-C source code and program binaries for selected platform.	Set of similar DNA or protein sequences in variety of formats (e.g., FASTA) or a set of motif models previously determined using MEME.	Motif-based multiple sequence alignment of query set of sequences for use in searching sequence databases for additional family members. Computes Viterbi score or total probability score for homology detection reported as log-odds scores
ASSET (Aligned Segment Statistical Evaluation Tool) (Neuwald and Green 1994)	*FTP:* ncbi.nlm.nih.gov Dir:/pub/asset	Performs a depth-first search for conserved sequence patterns to create a multiple sequence alignment by maximizing the log likelihood ratio statistic. Available as C code for UNIX system.	FASTA	Patterns found are reported as motifs with score and locations in sequences.
BLOCKS server (see also Table A11-2) (Smith et al. 1990; Henikoff and Henikoff 1991; Henikoff et al. 1999a,b)	http://www.blocks.fhcrc.org	Blocks are conserved patterns of amino acid sequences of the same length (no gaps) found in members of a protein family. The BLOCKS resource will create blocks from a multiple sequence alignment or locate them in unaligned sequences by either a pattern-searching method or statistical method (Gibbs sampling). BLOCKS server is used to produce position-specific scoring matrices (PSSMs) and a consensus (cobbler) sequence that may in turn be used to scan additional proteins in sequence databases for presence of block. An input sequence may also be scanned against blocks databases. Blocks have been calibrated for specificity against the Prosite and Swiss-Prot databases.	Input sequences in FASTA format on alignment (see description).	Blocks found are reported in their own (blocks) format. PSSMs and a cobbler (consensus) are also given for further database searching. Output can be used to predict PCR primers using CODEHOPS.
BLIMPS–BLOCKS Improved Searcher (Henikoff and Henikoff 1991; Wallace and Henikoff 1992)	http://blocks.fhcrc.org *FTP:* ncbi.nlm.nih.gov Dir:/repository/blocks/unix/blimps *FTP:* howard.fhcrc.org Dir:/blimps/	Alignment search software that scores sequence against BLOCKS database or blocks against sequence. (See also BLOCKS database in DNA: Motifs and Patterns.)	See BLOCKS server.	See BLOCKS server.
emotif (Exploring the *Motif* Universe (Nevill-Manning et al. 1998)	http://dna.stanford.edu/emotif/simple.html	Forms motifs or subsets of aligned sequences. Interactive Web-based form.	Aligned sequence.	List of retrieved motifs.
MAST (Motif Alignment and Search Tool) (Bailey and Gribskov 1998)	http://www.sdsc.edu/MEME/meme/website/mast.html	Searches databases for sequences that contain one or more motifs. (Motifs are defined on this site as conserved sequence patterns of the same length that occur one or more times in a related set of sequences.) Motifs are identified by searching through a set of related sequences with the MEME or	One or more motifs in MEME output or any file an an appropriate format (GCG profile file, concatenated GCG profile files, motif file format). The individual motifs cannot contain gaps.	Retrieved motifs; maximum of 50 matches returned. Output consists of names of high-scoring sequences, motif diagrams showing order and spacing of motifs (PSSM that describes score of each possi-

(Continued on following pages.)

TABLE A11-1 (*Continued*)

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
		BLOCKS servers; each are represented by a PSSM (position-specific scoring matrix) showing the variation in each column of the aligned motifs as a log odds score. Each PSSM is moved across each database sequence to identify high-scoring matches. Sequences with highest-scoring matches are identified. Web-based interactive form.		ble letter at each position in the pattern), detailed annotation of sequence; sequences shown with the location and scores for each motif found, and a combined probability score that these matches would have occurred by chance is also shown. HTML or ASCII format.
Gibbs sampler (Gibbs Sampling Strategy for Multiple Alignment) (Lawrence et al. 1993)	*FTP:* ncbi.nlm.nih.gov Dir:/pub/gibbs	Finds conserved patterns of sequence of the same length in a set of related sequences (no gaps, no correlations between positions) by a Gibbs sampling strategy. This method repeatedly rescans individual sequences in the set with a trial PSSM (position-specific scoring matrix) to find conserved patterns, using a probabilistic approach for choosing a matching location in the sequence, and then updates the matrix. Eventually, common patterns will be recruited to the matrix. May require initial estimate of pattern length (see also BLOCKS Web Site and MACAW).	FASTA	Location of conserved pattern in the sequences.
DNA: Genes, Exons, Introns				
Genie (Kulp et al. 1996; Reese et al. 1997)	http://www.fruitfly.org/seq_tools/genie.html	Finds potential genes in *Drosophila* genomic sequences using a hidden Markov model trained on known *Drosophila* genes, including a neural network trained on *Drosophila* splice sites. Tries to find one gene match on each strand of each sequence submitted; multiple sequences can be submitted. Software has been trained on human genes, but claims to give good results for other vertebrate sequences, *Drosophila*, other invertebrates.	DNA sequence in FASTA or multiple-FASTA; maximum DNA sequence length is 90,000 bases.	Highest-scoring gene location on each input DNA strand.
GeneMark (Borodovsky and McInich 1993)	http://genemark.biology.gatech.edu/GeneMark/gm_info.html *e-mail service:* genemark@amberbiology.gatech.edu ***OR*** genemark@embl-ebi.ac.uk	Identifies protein-coding regions in prokaryotic or eukaryotic genomic DNA sequences. Uses inhomogeneous Markov chain models, each trained on known genes of each organism. Analyzes EST sequences; predicts long exons; can be used for designed RT-PCR primers. Web-based interactive form; e-mail server. License required for use of local version. GeneMark.hmm algorithm (Lukashin and Borodovsky 1998) is an extension of GeneMark used for more accurate finding of 5′ ends in bacterial, human, and *C. elegans* genes.	DNA sequence in GCG, FASTA, EMBL, GenBank, PIR, NBRF, Phylip, text formats.	Coding signals for putative genes shown in graphical format (PSGRAPH). Text output for ORF predictions; possible frameshifts; right and left end of ORFs; orientation and frame of region; mean probability function; internal exon predictions.

Grail II: Gene Recognition and Analysis Internet Link (Uberbacher and Mural 1991)	http://compbio.ornl.gov/Grail-1.3/	Predicts exons in human genomic sequences. Uses neural networks trained on intron length, GC-composition, codon usage, 6-mer preference scores, splice junction scores, and other features. These scores are weighted in the neural network so as to give the best overall prediction of the training sequences. For new sequences, a sliding sequence window is moved across the sequence and a prediction (exon score) is made for each window.	Several formats.	List of best candidate exons and most probable gene structure.
GeneParser (Snyder and Stormo 1993, 1995)	*FTP:* beagle.colorado.edu/pub/GeneParser	Predicts gene structure in genomic DNA sequences. Uses sequence indicators found in known exons, introns, and splice junctions (see Grail) to evaluate likelihood that a sequence interval represents an exon or intron. Applies dynamic programming algorithm to find the exon/intron combination that maximizes the likelihood function for predicting a gene structure. Neural network used to adjust weights to sequence indicators for best fit to known genes.	See documentation.	List of best candidate exons and most probable gene structures.
GENSCAN (Burge and Karlin 1997)	http://CCR-081.mit.edu/GENSCAN.html Web servers also at Stanford University, DKFZ/EMBnet. Heidelberg, Pasteur Institute e-mail server at Pasteur Institute	Predicts genes in genomic DNA based on probabilistic models in gene structure in various organisms. Executables available for UNIX; license needed for commercial use.	DNA sequence, one-letter code.	Most probable gene structures.
NetGene (Brunak et al. 1991)	http://130.225.67.199/services/NetGene/index.html *e-mail service:* netgene@cbs.dtu.dk	Predicts splice sites in vertebrate genes. Uses neural networks trained to recognize differences in sequence features between exon and nonexon sequences (see Grail).	FASTA	Produces list of most strongly predicted exons.
FGENES (V.V. Solovyev, unpubl.)	http://genomic.sanger.ac.uk/gf/gf.html	Predicts genes and exons by using pattern-based structure prediction. Algorithm based on pattern recognition of different types of exons, promoters, and poly(A) signals; dynamic programming finds optimal combinations and then constructs set of gene models. Interactive Web-based form. FGENESH+ (A.A. Salamov and V.V. Solovyev, unpubl.) predicts multiple genes in genomic DNA sequence using a hidden Markov model of genes and similarity with known protein. FGENES-M produces pattern-based human multiple variants of gene structure prediction.	FASTA, plain sequence.	Text output: number of predicted genes, number of predicted exons, positions of predicted genes and exons.
MZEF (Michael Zhang's Exon Finder) (Zhang 1997)	http://sciclio.cshl.org/genefinder *FTP:* phage.cshl.org Dir:/pub/science/mzef	Predicts internal coding exons. Based on prediction algorithm using quadratic discriminant function for multivariate statistical pattern recognition (best discriminates between exon and nonexon sequences using sequence features similar to those described for Grail). Software for SUN, DEC Alpha. Need site license for commercial use.	FASTA; maximum of 200 kb.	Exon boundaries in base pairs, posterior probability, frame preference score, ORF ordinator, acceptor score, coding preference score, donor score.

(Continued on following pages.)

TABLE A11-1 (*Continued*)

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
DNA: Promoters, Transcription-factor Binding Sites				
Transfac (Heinemeyer et al. 1999)	http://transfac.gbf.de/TRANSFAC *FTP:* transfac.gbf.de	Transfac is a relational database of transcription factor *cis*-acting binding sites and transcription factors from many organisms, and also includes PSSMs (position-specific scoring matrices) representing the observed variation in each column of the aligned binding sites for one factor. (Some matrices may not be specific so that false positive matches are common.) Download is flat ASCII files. Software is for noncommercial use only. Programs include Coinspector, MatInspector, MatInd.	See web site.	Predicted location of TF binding sites.
CorePromoter (Zhang 1998)	http://sciclio.cshl.org/ genefinder/ CPROMOTER	Predicts transcriptional start sites and localizes them into 50–100-bp core-promoter regions. Uses quadratic discrimination analysis to distinguish sequence features (e.g., conserved sequence words) characteristic of promoter sequences. Web-based interactive form. License needed for commercial use.	FASTA; DNA sequence must be <2 kb.	N-profile position scores; strand; prior probability; QDA data; maximum number of predictions.
NNPP (*P*romoter *P*rediction by *N*eural *N*etwork) (Reese 1994; Reese and Eeckman 1995; Reese et al. 1996)	http://www.fruitfly.org/ seq_tools/ promoter-instrucs.html	Finds eukaryotic and prokaryotic promoters in DNA sequence. Uses time-delay neural networks (TDNN) that recognize features such as the TATA box, the GC box, the CAAT box, and the initiator region that spans the transcription start site (TSS). TDNN is repeatedly trained and re-trained until it recognizes the most important features of each element as a scoring matrix. Predictions for each element are then combined so that elements can be found even when their spacing varies.	FASTA	Transcription start sites most likely to be promoters. For eukaryotes, list of the 51 bases spanning −40 to +11 where +1 is the predicted transcription start site. For prokaryotes, 46 bases spanning from −41 to + 5 where +1 is the predicted transcription start site.
TESS (*T*ranscription *E*lement *S*earch *S*ystem) (Schug and Overton 1997)	http://www.cbil.upenn.edu/tess	Finds potential transcription-factor-binding sites in DNA sequence (see Transfac). Has Web-based interactive forms for combined search query, browsing Transfac, string-based searching, filtered string-based searching. Options for filtering, setting thresholds, attributes, and scoring schemes.	DNA sequence	Predicted sites.
TFBIND (Tsunoda and Takagi 1999)	http://tfbind.ims.u-tokyo.ac.jp	Searches for transcription-factor-binding sites, including TATA boxes, GC boxes, CCAAT boxes, and transcription start sites.	FASTA	Predicted sites.

		Uses weight matrix in Transfac database and cut-off values (see Transfac). Web-based interactive form.		
TSSG (V.V. Solovyev et al., unpubl. [see http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html])	http://genomic.sanger.ac.uk/gf/Help/tssg.html for information *e-mail service:* service@theory.bchs.uh.edu Put tssg in subject line of message.	Recognizes human PolII promoter region and transcription start site. Uses a linear discriminant function that combines characteristics of functional motifs and oligonucleotide composition of transcription start sites.	DNA sequence	Promoter region, including sequence name, LDF threshold and length of sequence, number of predicted promoter regions, positions of predicted sites, TATA box positions, transcription start site positions, functonal motifs for each predicted region.
TSSW (V.V. Solovyev et al., unpubl. [see http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html])	http://genomic.sanger.ac.uk/gf/Help/tssw.html-for information *e-mail service:* service@bchs.uh.edu *OR* services@bioinformatics.weizmann.ac.il Put tssw in subject line of message.	Recognizes human PolII promoter region and transcription start site. Uses linear discriminant function for prediction; combines characteristics describing functional motifs and oligonucleotide composition of the sites.	Name of sequence; sequence letters. Maximum line length is 79 characters.	Name of sequence; linear discriminant function threshold and length of sequence; number of predicted promoter regions; positions of predicted sites, TATA box, and transcription start site positions; functional motifs for each predicted region.
MAR-Finder (Kramer et al. 1996; Singh et al. 1997)	http://www.ncgr.org/ MarFinder/about.html	Deduces presence of matrix-associated regions (MARs) in DNA sequence. (Matrix-associated regions are the regions at which the chromatin fiber attaches to the nuclear matrix and are necessary for transcriptional regulation.) Method used is statistical inference and a Boolean logic network; 18 motifs known to occur in the neighborhood of MARs are used in the analysis. Web-based interactive form; registration necessary.	FASTA	High-scoring regions in base pairs, including average and integrated strengths; graphical output showing output of relative positions of potential MARs.
GenomeInspector (Quandt et al. 1996a,b)	http://www.gsf.de/biodv/genomeinspector.html	Detects distance correlation between open reading frames and transcription binding sites on megabases of nucleotide sequence; uses large-scale correlation analysis. Download executables for DEC Alpha, SUN, HP, Silicon Graphics, CONVEX; UNIX and X-windows. License needed for commercial use.	Many formats including database annotation, GCG, ConsInspector, MatInspector.	Graphical display showing correlation of ORFs and ARS-binding factor 1 binding sites.
DNA: Other Regulatory Sites				
consensus and wconsensus (Hertz and Stormo 1999)	FTP:beagle.colorado.edu/pub/consensus	A sliding window from each sequence is scanned against windows of the remaining sequences to find the most alike regions. The combined alignments of two or more windows are then used to recruit additional matching regions from other sequences. The algorithm is greedy in the sense that the first sequences to be matched determine the ultimate alignment of all sites found. Thus,	Set of input sequences.	Aligned sites in input sequences and position-specific scoring matrices for these sites.

(*Continued on following pages.*)

TABLE A11-1 (*Continued*)

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
		multiple cycles of matching are necessary. In consensus, window width is user-specified, and in wconsensus, it is not specified since the method can recruit flanking sequence if informative. The object is to maximize the information content (highest conservation of bases in each column) of a PSSM (position-specific scoring matrix). Includes important statistical evaluation of the PSSM. Note that there is no accounting for covariation within the sites as in palindromic regulatory sites, but methods to locate such regions have been described (Hertz and Stormo 1995). Available for implementation on UNIX platforms.		
polyadq (Zhang 1997)	http://sciclio.cshl.org/mzhanglab/tabaska/polyadq/polyadq_form.html	Decides if a given AATAA or ATTAAA hexamer is a true poly(A) signal. Algorithm uses quadratic discriminant analysis (see MZEF) Web-based interactive form.	FASTA. Can set cut-off score by four methods.	Prediction for each AATAA or ATTAAA found in the query sequence. Output is head, plus or minus prediction, site, sequence, and score.
Pol3scan (Pavesi et al. 1994)	http://irisbioc.bio.unipr.it/pol3scan.html	Searches eukaryotic DNA sequence for PolIII intergenic control regions; can discriminate between tRNA genes and related class III elements. Algorithmic method based on analysis of 231 eukaryotic tRNA promoter regions; weight (position-specific scoring) matrices and weight vectors used for scoring. Web-based interactive form.	ReadSeq; also all common formats including FASTA, GenBank, EMBL, GCG, etc.	Predictions of class III intragenic region and presence of transcription termination sites.
RNA: Secondary Structure				
MFOLD (Jaeger et al. 1989; Zuker 1989; Walter et al. 1994)	http://bioweb.pasteur.fr/seqanal/interfaces/mfold.html	Predicts RNA secondary structure; computes number of folds, 5′ base number (start), and 3′ base number (stop). Uses energy minimization method of Zuker. A maximum of 1400 bases can be folded. Web-based interactive form. Can be implemented on a local UNIX host with graphic interface for exploring alternative structural models in given energy range.	RNA and DNA 1-nucleotide sequence in most common formats, including FASTA, IG, GenBank, NBRF, EMBL, etc.	Energy matrices that determine the optimal and suboptimal secondary structure; computed folding has energy within given range of computed minimum free energy. Matrices are written to an output file. Plotfold software is used to read the file and display a representative set of optimal and suboptimal secondary structures.
Vienna RNA package (Shapiro 1988; McCaskill 1990; Shapiro and Zhang	http://www.tbi.univie.ac.at/~ivo/RNA	Calculates predictions of RNA structures with minimum free energies as well as equilibrium partition functions and base-pairing probabil-	RNA sequence; HIT representation for structure representation	RNAfold produces plots and dot plot of minimum free energy structure and base pair

1990; Bonhoeffer et al. 1993; Fontana et al. 1993; Hofacker et al. 1994; Wuchty et al. 1999)		ities. Web-based interactive form for moderate sized RNAs. RNAfold calculates secondary RNA structure. RNAeval calculates energy of RNA sequence on given structure. RNAheat calculates specific heat of RNA. RNAinverse finds an RNA sequence with a given secondary structure. RNAdistance calculates the distance of RNA secondary structure. RNApbdist calculates the distance of thermodynamic RNA secondary structure groups. RNAsubopt calculates suboptimal secondary structure. Download source codes (C code library) or stand-alone programs for UNIX, LINUX.		probabilities. RNAeval gives energy in Kcal/mole; RNA heat, temperature in degrees C, specific heat in Kcal/(mole*K). RNAinverse, found sequence, Hamming distance to start sequence. RNAdistance, base pair distribution. RNApbdist, partition function and matrix of base-pairing probabilities, dot plots. RNAsubopt, structure in bracket notation and energy in Kcal/mole.

RNA: RNA-specifying Genes, Motifs

tRNAscan-SE (Fichant and Burks 1991; Eddy and Durbin 1994; Pavesi et al. 1994; Lowe and Eddy 1997)	http://www.genetics.wustl.edu/eddy/tRNAscan-SE	Searches for tRNA genes in genomic DNA or RNA sequence. Candidate tRNAs are identified and analyzed by a highly selective tRNA covariance model (sequence variation that occurs at two separated sites to conserve base pairing in regions of RNA secondary structure); searches are done at 30,000 bp/sec. The method also searches for eukaryotic PolIII sites and RNA promoters. Web-based interactive form.UNIX source code can also be accessed from the Web page.	Raw sequence, FASTA, GenBank, EMBL, GCG, IG. Searches are best as one sequence at a time, each sequence under 100,000 nucleotides.	Candidate tRNA genes.
snoRNA (Lowe and Eddy 1999)	http://rna.wustl.edu/snoRNAdb	Searches prescreened sections of genomic sequence with characteristic sequence patterns for 2′-*O*-ribose methylation guide snoRNA genes (sno, small nucleolar). The algorithm uses probabilistic methods similar to those used in speech recognition and computational liguistics. SCFG models (stochastic, context-free grammars) are used to recognize regions with a particular pattern of sequence covariation in separated regions of sequence that corresponds to sequences that base pair to form dsRNA secondary structure.		

Protein: Sequence Alignment

FASTA, SSEARCH, Baves block aligner	Please see DNA sequence alignment section.			

(Continued on following pages.)

TABLE A11-1 (*Continued*)

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
Proteins: Motifs, Patterns, Profiles				
emotif, ProfileMaker, Psi-BLAST, Phi-BLAST, PowerBLAST, Blast 2, Bayes block aligner, LAMA, BLOCKS	Please see DNA motifs, patterns, profiles section.			
PowerBLAST	http://www.ncbi.nlm.nih.gov/Kuehl/prefinished/powblast.html	Provides added functionality over the traditional BLAST programs. Allows masking of a sequence against human repetitive elements, comparison of masked sequence against multiple databases simultaneously. Allows rapid examination of ESTs, STSs, and gene hits in the context of their spatial relationships to one another. User can set additional parameters, filter BLAST hits by organism, and choose multiple output options. Web-based interactive form.	FASTA	Text, HTML, or ASN.1 output. A file is output for each sequence in the query. The out put can be viewed with a mask. Sequin can be used to annotate the BLAST hits.
PfScan (ProfileScan) (Swiss Institute of Bioinformatics [see www site])	http://www.isrec.isb-sib.ch/software/PFSCAN_form.html	Scans protein or DNA sequence against profiles from the PROSITE catalog. PROSITE contains amino acid patterns found in families of functionally related proteins, e.g., representing active site of protein kinases. The patterns are given as regular expressions that provide a list of the sequential amino acids found in the sites showing conserved, variable, amino acids not found and skipped amino acid positions. Also scans profiles against Swiss-Prot protein sequence database. Can search against the Pfam database. Web-based interactive form.	FASTA	Optimal alignment scores for profiles—PROSITE profiles N score; Pfam-A Nscore; Gribskov Z score; PROSITE patterns.
Pfam (Sonnhammer et al. 1997, 1998; Bateman et al. 1999, 2000)	http://www.sanger.ac.uk/Software/Pfam *Sweden:* www.cgr.ki.se/Pfam *USA:* pfam.wustl.edu	Enables rapid and automatic classification of predicted proteins into protein domain families; considers domain organization of proteins. Annotates protein using hidden Markov model software (see HMMER) that represents highly suitable probabilistic models of protein families, predicts genes, and annotates DNA sequence using the Wise2 package. The database contains 2128 families matching 65% of the proteins in Swiss-Prot. Genomic DNA can be directly searched against the Pfam library. This software was used to annotate the *C. elegans* genome.	FASTA	Graphical representation of domain structure. Name of domain in Pfam; start and end points of domain; statistical score of match in bits and *E* values; alignment of matching section of query sequence to relevant hidden Markov model. Provides functional annotation, literature references, and database links for each family. The output can be linked to the SCOP database to determine if the domain's structure has been solved.

ScanProsite (Swiss Institute of Bio informatics; see WWW site)	http://www.expasy.ch/tools/scnpsite.html	Scans sequence against patterns in PROSITE or a pattern against Swiss Prot. Web based interactive form.			
PPSEARCH (European Bioinformatics Institute, Hinxton, UK)	http://www2.ebi.ac.uk/ppsearch	Scans sequence against patterns in PROSITE.			Graphical output.
PROSITE (Klaus Hofmann, University of Vienna; Manfred D. Zorn, Lawrence Berkeley Laboratory)	http://pbilabo.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_prosite.html *FTP:* genome.lbl.gov Dir: /pub/prosite File: ProSite.shar.Z	Scans sequence against patterns from PROSITE; can define level of mismatches. Web-based interactive form. Download UNIX, C code. Also available commercially. Many other programs that use PROSITE are listed at www.expasy.ch/cgi-bin/lists/prosite.org	Swiss Prot, FASTA, ASCII		ASCII. Positions of patterns shown under sequence; complete PROSITE pattern; position pattern of pattern occurrences for each pattern.
BEAUTY (BLAST Enhanced Alignment Utility) (Worley et al. 1995, 1998)	http://gcb.icm.edu:8088/searchlauncher/launcher.html	An enhanced version of NCBI BLAST that facilitates identification of functions of matched sequences. Integrates protein families, conserved regions, annotated regions, alignment displays, and WWW resources. Includes new databases of conserved regions and functional domains for protein sequences in NCBI Entrez. An updated version of BEAUTY-X provides enhanced output for DNA queries, ability to search any protein database, and more domain information. Web-based interactive form.	FASTA and other formats.		Schematic display compares relative locations of conserved regions, annotated domains and sites, locally aligned regions matched in a BLAST search, and WWW links to external databases for additional functional information.

TABLE A11-2 Database Similarity Search Software

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
BLAST (*B*asic Alignment Search *T*ool) (Altschul et al. 1990; Karlin and Altschul 1990, 1993; Tatusova and Madden 1999)	http://www.ncbi.nlm.nih.gov/BLAST	Sequence similarity search tool for analysis of nucleotide (DNA) and protein databases; uses heuristic algorithm that seeks high-scoring words and then uses these to produce gapped local alignments; filters out low-complexity regions from the search (optional); assigns scores with well-defined statistical interpretation so that real matches can be distinguished from random background hits. Default scoring matrix is BLOSUM62. The significances of the matches are given an Expect (*E*) score, the expected number of alignments between a random query sequence and a database of random sequences of the same "effective" length and number that will score as well. Available as interactive Web-based interface, stable URL, e-mail server, network BLAST with downloadable client for AIX, UNIX (see Web Site for details). Subprograms include: *blastn*–compares nucleotide query sequence against nucleotide sequence database *blastp*–compares amino acid query sequence against protein sequence database *blastx*–compares 6-frame conceptual translation products of nucleotide query sequence (both strands) against protein sequence database *tblastn*–compares protein query sequence against nucleotide sequence database dynamically translating all 6 reading frames *tblastx*–compares 6-frame translations of nucleotide query sequence against 6-frame translations of nucleotide sequence database A stand-alone BLAST system can be established locally on many computer platforms, including PC Windows 32, by FTP from ncbi.nlm.nih.gov/blast/executables.	FASTA or Accession/ GenBank index no.	Mouse-clickable histogram of matches; list of high-scoring database sequences; gapped, local alignment of query sequence with matching data base sequences with statistical evaluation; list of parameters used and statistical calculations made during search.
PHI-BLAST (*P*osition *H*it *I*nitiated *BLAST*) (Zhang et al. 1998)	http://www.ncbi.nlm.nih.gov/BLAST	Finds protein sequences that share an amino acid pattern, entered using PROSITE syntax (a regular expression). Program is integrated with PSI-BLAST so that query results can be used to start one or more rounds of PSI-BLAST searching using Web-based interactive form.	FASTA or Accession/ GenBank index no.	Statistical significance of sequences that contain patterns are reported using *E* values. Multiple sequence alignment (a position-specific scoring matrix of the matching sequences) can then be used for additional rounds (iterations) of database searching using PSI-BLAST.
PSI-BLAST (*P*osition-specific *I*nitiated *BLAST*) (Altschul et al. 1997)	http://www.ncbi.nlm.nih.gov/BLAST	Search for similar protein sequences using a multiple sequence alignment as the query, thus allowing searches for new combinations of residues found in the alignment. Starts with a regular BLAST search, but as new matching sequences are found, these are added to the alignment and an iteration of the search can then be performed; multiple iterations are possible until no more matching sequences are found. Web-based interactive form.	FASTA or Accession/ GenBank index no.	A multiple sequence alignment (a position-specific scoring matrix of the matching sequences) replaces the initial pattern and may be used for additional rounds (iterations) of database searching until no more matching sequences are found.

BLAST2 (Gapped BLAST) (Altschul et al. 1997)	http://www.ncbi.nlm.nih.gov/BLAST	The newer version 2 of BLAST returns a local (Smith Waterman) alignment from BLAST searches and uses the statistical significance of the alignment scores to identify the most alike sequences; this is the default condition for the BLAST2 program. Earlier versions of BLAST identified high-scoring regions, and used the sum of these scores to evaluate the signifi cance of the sequence similarity.	FASTA or Accession/ GenBank index no.	
BLAST-Genome Sequences	http://www.ncbi.nlm.nih.gov/genome/seq	Searches a query sequence against the human genome data base (or any other genomic sequence database). Interactive Web-based form.	FASTA or Accession/ GenBank index no.	Match of sequence by chromosome number.
WU-BLAST (Washington University *BLAST*) (Altschul et al. 1990; Gish and States 1993; Karlin and Altschul 1993; Altschul and Gish 1996; W. Gish *unpubl.*)	http://blast.wustl.edu	Performs similarity searches of proteins and nucleotide sequence databases; similar to NCBI BLAST but developed independently and can produce different results. WU-BLAST2 uses gapped alignments to identify matching sequences; produces a gapped local (Smith-Waterman) alignment; calculates statistical significance based on the combined scores of high-scoring regions (the sum statistics method as in earlier pre-version 2 versions of NCBI BLAST). BLASTN, BLASTP, TBLASTN, BLASTX and TBLASTX programs are available (see NCBI BLAST); includes improvements for identifying coding regions and uses DNA scoring matrices for more significant alignment of nucleic acids. Contains enhancements for optimized execution, reduced virtual memory requirements, support for parallel processing, and eXtended Database Format (XDF). Software is downloadable from Web Site: License is free for academic and nonprofit use; fee for commercial use.	FASTA or Accession/ GenBank index no.	Mouse-clickable map of matching sequences, lists of matching sequences with sum-statistics scores, gapped alignments with potentially multiple regions of similarity; can be used to show all exons in multi-exon sequence; all complete on par-tial copies of repetitive ele ment in genomic sequence.
FASTA Programs				
FASTA (Pearson and Lipman 1988)	http://fasta.bioch.virginia.edu *FTP:* ftp.virginia.edudir /pub/.fasta	The FASTA package includes a suite of programs for data base similarity searching and for pair-wise sequence align ment. Compares DNA sequence to another DNA sequence or to a DNA sequence database (designated library by the programs), or a protein sequence to another protein sequence or to a protein sequence database; also compares protein sequence to translated DNA sequence or DNA sequence database (library). FASTA3 (version 3) returns matches based on a normalized *Z* score,the number of standard deviations above a mean of 50 and standard deviation of 10 for matching all unrelated sequences of the same length in the database. The *E* (expect) value of the *Z* score, the number of unrelated sequences in the database that are expected to score as well, is given. Can search individual genomes; databases listed on Web form. Web-based interactive form, or down-loadable software. Permission needed for commercial use. FASTA may be established on a variety computer platforms, including Macintosh and PC Windows, but without a graphical interface.	FASTA or Accession/ GenBank index no.	Provides a graphical display of the range of alignment scores found between the query sequence and each database sequence, with a fit of the scores with unrelated proteins to the extreme value distribution. A list of high-scoring sequences is then given followed by gapped alignments of the query and matched data base sequences. Normalized *Z* scores are used to describe the relationship of the alignment score between related sequences to that found for unrelated sequences in the same length range (mean 50, standard deviation 10). The *E* value of these high *Z* scores is the expected number of unrelated sequence alignments with *Z* scores as high as that found with the matched sequence.

(Continued on following pages.)

TABLE A11-2 (*Continued*)

Program Name	Web Site/FTP Site	Description/Platforms	Input Format	Interpretation/Results
TFASTX/TFASTY FASTX, FASTY (Chao et al. 1992)	http://fasta.bioch.virginia.edu	Compares protein sequence to translated DNA sequence or DNA sequence database (TFASTX/TFASTY) or translated DNA sequence against protein sequence database (FASTX/FASTY); designed to accommodate high error rate in EST sequences by allowing frameshifts and substitutions.	FASTA or Accession/Genbank index no.	Single alignment between 2 sequences; DNA sequence is translated in 3 forward and 3 reverse frames; protein query is compared to each of the 6 derived protein sequences. DNA sequence is translated from one end to the other and intervening sequences are not edited out; termination codons are translated into unknown amino acids.
FASTS/TFASTS FASTF/TFASTF	http://fasta.bioch.virginia.edu	Compares sequence of peptide fragments (mass-spectrophotometric analysis) (FASTS) or an ordered peptide mixture (FASTF) against a protein (FASTS) or DNA (tFASTS) data base.	See Web site	Similar to FASTA, includes list of matching sequences in the database.

TABLE A11-3 Bioinformatics Databases

Database	WWW Address	Description
DNA		
GenBank	www.ncbi.nlm.nih.gov/Genbank/GenbankSearch.html	DNA sequence database maintained by National Center for Biotechnology Information (NCBI); member of International DNA Databases.
EMBL (Baker et al. 2000)	www.ebi.ac.uk/embl/	DNA sequence database maintained by European Bioinformatics Institute (EBI); member of International DNA databases.
DDBJ	www.ddbj.nig.ac.uk	DNA sequence database maintained by DNA Data Bank of Japan; member of International DNA Databases.
DNA: Motifs, Patterns		
Codon use database	http://www.kazusa.or.jp/codon/	
Arabidopsis	http://genome-www.stanford.edu/Arabidopsis/splice_site.html	Consensus splice sites.
Yeast splice sites by M. Ares, Jr. laboratory (Spingola et al. 1999)	http://www.cse.ucsc.edu/research/compbio/yeast_introns.html	
DNA: Promoters and Regulatory Sequences		
EPD (*E*ucaryotic *p*romoter *D*atabase) (Bucher 1990; Perier et al. 1999)	http://www.epd.isb-sib.ch/; http://www.epd.isb-sib.ch/promoter_elements/	
TRRD (*T*ranscriptional *R*egulatory *R*egion *D*atabase;) (Kolchanov et al. 1999)	http://wwwmgs.bionet.nsc.ru/mgs/systems/geneexpress/	
TSSW (Recognition of human PolII promoter region and start of transcription by linear discriminant function analysis)	http://genomic.sanger.ac.uk/gf/gf.shtml; http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html	
OOTFD (Object-Oriented *T*ranscription *F*actor *D*atabase) (Ghosh 1998)	http://www.isbi.net/	
PLACE (plant *cis*-acting regulatory elements) (Higo et al. 1999)	http://www.dna.affrc.go.jp/htdocs/PLACE/	
PlantCARE (plant *cis*-acting regulatory elements) (Rombauts et al. 1999)	http://sphinx.rug.ac.be:8080/PlantCARE/index.htm	
Thyroid receptor resource	http://xanadu.mgh.harvard.edu/receptor/trrfront.html	

(Continued on following pages.)

TABLE A11-3 (*Continued*)

DATABASE	WWW ADDRESS	DESCRIPTION
RNA		
5S ribosomal RNA data bank (Szymanski et al. 1999)	http://rose.man.poznan.pl/5SData/5SRNA.html and mirrored at http://userpage.chemie.fu-berlin.de/fb_chemie/ibc/agerdmann/5S_rRNA.html	
Gobase for Mitochondrial Sequences (Korab-Laskowska et al. 1998)	http://alice.bch.umontreal.ca/genera/gobase/gobase.html	
Guide RNA (gRNA) database (Souza and Göringer 1998)	http://www.biochem.mpg.de/~goeringe/	
tRNA Genes, higher-plant mitochondria (Ceci et al. 1999)	ftp://ftp.ebi.ac.uk/pub/databases/plmitrna/	
Nucleic acid database and structure resource (Berman et al. 1998)	http://ndbserver.rutgers.edu/	
Pseudobase (Pseudoknot database maintained by E van Batenburg, Leiden University)	http://wwwbio.leidenuniv.nl/~batenburg/pkb.html	
Ribosomal RNA mutation databases (Triman and Adams 1997)	http://www.fandm.edu/Departments/Biology/Databases/RNA.html	
RNA modification database (Limbach et al. 1994; Rozenski et al. 1999)	http://medlib.med.utah.edu/RNAmods/	
RNA secondary structures (Group I introns–16S rRNA–23S rRNA) (Gutell 1994; Schnare et al. 1996 and references therein)	http://pundit.icmb.utexas.edu	
RNA structure database	http://grserv.med.jhmi.edu/~venk/rna/	
RNA World at IMB Jena (Sühnel 1997)	http://www.imb-jena.de/RNA.html	
rRNA (database of ribosomal subunit sequences) (De Rijk et al. 1992, 1999)	http://rrna.uia.ac.be/	
Small RNA database	http://mbcr.bcm.tmc.edu/smallRNA/smallrna.html)	
snoRNA database for *S. cerevisiae* (Lowe and Eddy 1999)	http://rna.wustl.edu/snoRNAdb/	
tmRNA database (Wower and Zwieb 1999)	http://psyche.uthct.edu/dbs/tmRDB/tmRDB.html	

uRNA database (Zwieb 1997)	http://psyche.uthct.edu/dbs/uRNADB/uRNADB.html	
Viroid and viroid like RNA sequences (Lafontaine et al. 1999)	http://www.callisto.si.usherb.ca/~jpperra	

Protein: Motifs Based on Sequence Alignments

HSSP (*H*omology-derived Secondary Structure of *P*roteins) (Sander and Schneider 1991)	http://www.sander.embl-ebi.ac.uk/hssp	Derived database that merges two- and three-dimensional structure and sequence information.
BLOCKS emotif, pfam,prints, prodom	See Table A11-1.	
SYSTERS (Krause et al. 1999)	http://www.dkfz-heidelberg.de/tbi/services/cluster/systerform	Based on clustering of all similar sequences in Swiss-Prot.
Protomap (Yona et al. 1998)	http://www.protomap.cs.huji.ac.il	Automatic hierarchical classification of all Swiss-Prot proteins.
Prodom (Corpet et al. 1998)	http://protein.tolouse.inra.fr/prodom.html	Groups of sequence segments or domains from similar sequences found in Swiss-Prot database by Blastp algorithm; aligned by multiple sequence alignment.

Protein: Structural Alignment

SCOP (Structural Classification of *P*roteins) (Murzin et al. 1995)	http://scop.mrc-imb.cam.ac.uk/scop/	Classification of ll proteins whose structure is known based on expert analysis; includes all PDB entries.
FSSP (*F*old classification based on Structure-Structure alignment of *P*roteins) (Holm and Sander 1996)	http://www2.ebi.ac.uk/dali/fssp/fssp.html	Three-dimensional structure comparison of protein structures in PDB; structural alignments are performed by Dali program.
3D-Ali (Pascarella and Argos 1992)	http://embl-heidelberg.de/argos/ali/ali_info.html	Aligned protein structures and related sequences using only secondary structures assigned by the author of the structures.
National Center for Biotechnology Information Structure Group	http://www.ncbi.nlm.nih.gov/Structure/	Molecular modeling database (MMDB), vector align ment search tool (VAST) for structural comparisons, viewers, threader software.
Biomolecular Structure and Modelling group at the University College, London	http://www.biochem.ucl.ac.uk/bsm/	CATH database, a hierarchical domain classification of protein structures by class, architecture, fold family, and superfamily, other databases and structural analyses, threader software.
European Bioinformatics Institute, Hinxton, Cambridge	http://www2.ebi.ac.uk/	Databases, TOPS protein structural topology cartoons, Dali domain server, and FSSP database.

(Continued on following pages.)

TABLE A11-3 (*Continued*)

DATABASE	WWW ADDRESS	DESCRIPTION
Protein: Structural Coordinates		
PDB (*Protein Data Bank*) (Berman et al. 2000)	http://www.rcsb.org/pdb	Three-dimensional macromolecular structure data determined mainly by X-ray crystallography and nuclear magnetic resonance; also contains atomic coordinates, reference citations, and primary and secondary structure information. Operated by the Research Collaboration for Structural Bioinformatics, Rutgers.

This survey of databases lists the principal databases as well as some of the more specialized databases for DNA, protein, and RNA. There are several Web pages that list and provide links to the large array of databases. Among these are:

Gabriel's Hot List
http://www.bmm.icnet.uk/useful/usefulz.html
Extensive lists of links to sequence and structure databases and to specialized databases.

A List of Databases, Rockefeller University
http://linkage.rockefeller.edu/wli/gene/databases.html
Links to DNA, protein, and genome databases.

Bioinformatics Resouces, Bioinformatics Group, University of Waterloo, Canada
http://wh.math.umaterloo.ca/bioinfo_res.html
Links to DNA, protein, and genome databases.

The RNA World
http://www.imb-jena.de/RNA.html
Links to RNA-specific databases.

Amos' WWW Links
http://www.expasy.ch/alinks.html
Extensive lists of links to protein databases and resources.

and also see
http://www.bioinformaticsonline.com, a Web Site for Mount D. *Bioinformatics: Sequence and genome analysis* (Cold Spring Harbor Laboratory Press 2001)

Appendix 12

Cautions

GENERAL CAUTIONS

The following general cautions should always be observed.

- Become **completely familiar** with the properties of all substances used before beginning the procedure.
- **The absence of a warning** does not necessarily mean that the material is safe, since information may not always be complete or available.
- If **exposed** to toxic substances, contact the local safety office immediately for instructions.
- **Use proper disposal procedures** for all chemical, biological, and radioactive waste.
- For specific guidelines on **appropriate gloves**, consult the local safety office.
- Handle **concentrated acids and bases** with great care. Wear goggles and appropriate gloves, as well as a face shield if handling large quantities.

  Do not mix strong acids with organic solvents as they may react. Sulfuric acid and nitric acid especially may react highly exothermically and cause fires and explosions.

  Do not mix strong bases with halogenated solvent as they may form reactive carbenes which can lead to explosions.

  When preparing diluted solutions of acids from concentrated stocks, add acid to water ("If you do what you oughta, add acid to wata").

- Never **pipette** solutions using mouth suction. This method is not sterile and can be dangerous. Always use a pipette aid or bulb.
- Keep **halogenated and nonhalogenated** solvents separately (e.g., mixing chloroform and acetone can cause unexpected reactions in the presence of bases). Halogenated solvents are organic solvents such as chloroform, dichloromethane, trichlorotrifluoroethane, and dichloroethane. Some nonhalogenated solvents are pentane, heptane, ethanol, methanol, benzene, toluene, *N*,*N*-dimethylformamide (DMF), dimethylsulfoxide (DMSO), and acetonitrile.
- **Laser radiation**, visible or invisible, can cause severe damage to the eyes and skin. Take proper precautions to prevent exposure to direct and reflected beams. Always follow manufacturers safety guidelines and consult the local safety office. For more detailed information, see caution below.
- **Flash lamps**, due to their light intensity, can be harmful to the eyes and may explode on occasion. Wear appropriate eye protection and follow the manufacturer's guidelines.
- **Photographic fixatives and developers** contain harmful chemicals. Handle them with care and follow manufacturer's directions.

- **Power supplies and electrophoresis equipment** pose serious fire hazard and electrical shock hazards if not used properly.
- **Microwave ovens and autoclaves** in the lab require certain precautions. If the screw top on the bottle is not loose enough, and there is not enough space for the steam to vent, the bottle can explode when the containers are removed from the microwave or autoclave. Always loosen bottle caps before microwaving or autoclaving.
- Use extreme caution when handling **cutting devices** such as microtome blades, scalpels, razor blades, or needles. Microtome blades are extremely sharp! If unfamiliar with their use, have an experienced person demonstrate proper procedures. For proper disposal, use a "sharps" disposal container in the lab. Discard used needles unshielded, with the syringe still attached. This method prevents injuries (and possible infections) while manipulating used needles since many accidents occur while trying to replace the needle shield. Injuries may also be caused by broken Pasteur pipettes, coverslips, or slides.

GENERAL PROPERTIES OF COMMON CHEMICALS

The hazardous materials list can be summarized in the following categories:

- Inorganic acids, such as hydrochloric, sulfuric, nitric, or phosphoric, are colorless liquids with stinging vapors. Avoid spills on skin or clothing. Dilute spills with large amounts of water. The concentrated forms of these acids can destroy paper, textiles, and skin as well as cause serious injury to the eyes.
- Salts of heavy metals are usually colored powdered solids that dissolve in water. Many of them are potent enzyme inhibitors and therefore toxic to humans and to the environment (e.g., fish and algae).
- Most organic solvents are flammable volatile liquids. Breathing their vapors can cause nausea or dizziness. Also avoid skin contact.
- Other organic compounds, including organosulphur compounds such as mercaptoethanol or organic amines, have very unpleasant odors. Others are highly reactive and must be handled with appropriate care.
- If improperly handled, dyes and their solutions can stain not only the sample, but also skin and clothing. Some of them are also mutagenic (e.g., ethidium bromide), carcinogenic, and toxic.
- Nearly all names ending with "ase" (e.g., catalase, β-glucuronidase, or zymolase) refer to enzymes. There are also other enzymes with nonsystematic names like pepsin. Many of them are provided by manufacturers in preparations containing buffering substances, etc. Be aware of the individual properties of materials contained in these substances.
- Toxic compounds often used to manipulate cells (e.g., cycloheximide, actinomycin D, and rifampicin) can be dangerous and should be handled appropriately.

HAZARDOUS MATERIALS

Acetic acid (concentrated) must be handled with great care. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and goggles. Use in a chemical fume hood.

Acetonitrile is very volatile and extremely flammable. It is an irritant and a chemical asphyxiant that can exert its effects by inhalation, ingestion, or skin absorption. Treat cases of severe exposure as cyanide poisoning. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood. Keep away from heat, sparks, and open flame.

Acrylamide (unpolymerized) is a potent neurotoxin and is absorbed through the skin (effects are cumulative). Avoid breathing the dust. Wear appropriate gloves and a face mask when weighing powdered acrylamide and methylene-bisacrylamide. Use in a chemical fume hood. Polyacrylamide is considered to be nontoxic, but it should be handled with care because it might contain small quantities of unpolymerized acrylamide.

Actinomycin D is a teratogen and a carcinogen. It is highly toxic and may be fatal if inhaled, ingested, or absorbed through the skin. It may also cause irritation. Avoid breathing the dust. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood. Solutions of actinomycin D are light-sensitive.

$AgNO_3$, *see* **Silver nitrate**

α-Amanitin is highly toxic and may be fatal by inhalation, ingestion, or skin absorption. Symptoms may be delayed for as long as 6–24 hours. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood.

Aminobenzoic acid may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Ammonium acetate, $H_3CCOONH_4$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Ammonium chloride, NH_4Cl, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Ammonium formate, *see* **Formic acid**

Ammonium hydroxide, NH_4OH, is a solution of ammonia in water. It is caustic and should be handled with great care. As ammonia vapors escape from the solution, they are corrosive, toxic, and can be explosive. Use only with mechanical exhaust. Wear appropriate gloves. Use only in a chemical fume hood.

Ammonium molybdate, $(NH_4)_6Mo_7O_{24}{\cdot}4H_2$) (or its **tetrahydrate**) may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Ammonium persulfate, $(NH_4)_2S_2O_8$, is extremely destructive to tissue of the mucous membranes and upper respiratory tract, eyes, and skin. Inhalation may be fatal. Wear appropriate gloves, safety glasses, and protective clothing. Always use in a chemical fume hood. Wash thoroughly after handling.

Ammonium sulfate, $(NH_4)_2SO_4$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Ampicillin may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Aprotinin may be harmful by inhalation, ingestion, or skin absorption. It may also cause allergic reactions. Exposure may cause gastrointestinal effects, muscle pain, blood pressure changes, or bronchospasm. Wear appropriate gloves and safety glasses. Do not breathe the dust. Use only in a chemical fume hood.

Arc lamps are potentially explosive. Follow manufacturer's guidelines. When turning on arc lamps, make sure nearby computers are turned off to avoid damage from electromagnetic wave components. Computers may be restarted once the arc lamps are in operation.

Aspartic acid is a possible mutagen and poses a risk of irreversible effects. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood. Do not breathe the dust.

Bacterial strains (shipping of): The Department of Health, Education, and Welfare (HEW) has classified various bacteria into different categories with regard to shipping requirements (please see Sanderson and Zeigler, *Methods Enzymol. 204:* 248–264 [1991] or the instruction brochure by Alexander and Brandon (*Packaging and Shipping of Biological Materials at ATCC* [1986] available from the American Type Culture Collection [ATCC], Rockville, Maryland). Nonpathogenic strains of *Escherichia coli* (such as K-12) and *Bacillus subtilis* are in Class 1 and are considered to present no or minimal hazard under normal shipping conditions. However, *Salmonella, Haemophilus*, and certain strains of *Streptomyces* and *Pseudomonas* are in Class 2. Class 2 bacteria are "Agents of ordinary potential hazard: agents which produce disease of varying degrees of severity...but which are contained by ordinary laboratory techniques."

BCIG, *see* **5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside**

Biotin may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Bisacrylamide is a potent neurotoxin and is absorbed through the skin (the effects are cumulative). Avoid breathing the dust. Wear appropriate gloves and a face mask when weighing powdered acrylamide and methylene-bisacrylamide.

Blood (human) and blood products and Epstein-Barr virus. Human blood, blood products, and tissues may contain occult infectious materials such as hepatitis B virus and HIV that may result in laboratory-acquired infections. Investigators working with EBV-transformed lymphoblast cell lines are also at risk of EBV infection. Any human blood, blood products, or tissues should be considered a biohazard and be handled accordingly. Wear disposable appropriate gloves, use mechanical pipetting devices, work in a biological safety cabinet, protect against aerosol generation, and disinfect all waste materials before disposal. Autoclave contaminated plasticware before disposal; autoclave contaminated liquids or treat with bleach (10% [v/v] final concentration) for at least 30 minutes before disposal. Consult the local institutional safety officer for specific handling and disposal procedures.

Boric acid, H_3BO_3, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and goggles.

5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (BCIG; X-gal) is toxic to the eyes and skin and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety goggles.

Bromophenol blue may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

***n*-Butanol** is irritating to the mucous membranes, upper respiratory tract, skin, and especially the eyes. Avoid breathing the vapors. Wear appropriate gloves and safety glasses. Use in a chemical fume hood. *n*-Butanol is also highly flammable. Keep away from heat, sparks, and open flame.

Cacodylate contains arsenic, is highly toxic, and may be fatal if inhaled, ingested, or absorbed through the skin. It is a possible carcinogen and may be mutagenic. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Cacodylic acid is toxic and a possible carcinogen. It may be mutagenic and is harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood. Do not breathe the dust.

Carbenicillin may cause sensitization by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Carbon dioxide, CO_2, in all forms may be fatal by inhalation, ingestion, or skin absorption. In high concentrations, it can paralyze the respiratory center and cause suffocation. Use only in well-ventilated areas. In the form of dry ice, contact with carbon dioxide can also cause frostbite. Do not place large quantities of dry ice in enclosed areas such as cold rooms. Wear appropriate gloves and safety goggles.

Cesium chloride, CsCl, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Cetylpyridinium bromide (CPB) causes severe irritation to the eyes, skin, and respiratory tract. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Cetyltrimethylammonium bromide (CTAB) is toxic and an irritant and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Avoid breathing the dust.

CH_3CH_2OH, *see* **Ethanol**

$C_6H_5CH_2SO_2F$, *see* **Phenylmethylsulfonyl fluoride**

$CHCl_3$, *see* **Chloroform**

$C_7H_7FO_2S$, *see* **Phenylmethylsulfonyl fluoride**

Chloramphenicol may be harmful by inhalation, ingestion, or skin absorption and is a carcinogen. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Chloroform, $CHCl_3$, is irritating to the skin, eyes, mucous membranes, and respiratory tract. It is a carcinogen and may damage the liver and kidneys. It is also volatile. Avoid breathing the vapors. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood.

Citric acid is an irritant and may be harmful by inhalation, ingestion, or skin absorption. It poses a risk of serious damage to the eyes. Wear appropriate gloves and safety goggles. Do not breathe the dust.

CO_2, *see* **Carbon dioxide**

Cobalt chloride, $CoCl_2$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

$CoCl_2$, *see* **Cobalt chloride**

Coomassie Brilliant Blue may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Copper sulfate, $CuSO_4$, may be harmful by inhalation or ingestion. Wear appropriate gloves and safety glasses.

CPB, *see* **Cetylpyridinium bromide**

***m*-Cresol** may be fatal if inhaled, ingested, or absorbed through the skin. It may also cause burns and is extremely destructive to the eyes, skin, mucus membranes, and upper respiratory tract. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

CsCl, *see* **Cesium chloride**

CTAB, *see* **Cetyltrimethylammonium bromide**

$CuSO_4$, *see* **Copper sulfate**

Cysteine is an irritant to the eyes, skin, and respiratory tract. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Do not breathe the dust.

DEAE, *see* **Diethylaminoethanol**

DEPC, *see* **Diethyl pyrocarbonate**

Dichloromethylsilane, *see* **Dichlorosilane**

Dichlorosilane is highly flammable and toxic and may be fatal if inhaled. It is harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety goggles. Use only in a chemical fume hood. It reacts violently with water. Keep away from heat, sparks, and open flame. Take precautionary measures against static discharges.

Diethylamine, $NH(C_2H_5)_2$, is corrosive, toxic, and extremely flammable. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood. Keep away from heat, sparks, and open flame.

Diethylaminoethanol (DEAE) may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Diethyl ether, Et_2O or **$(C_2H_5)_2O$,** is extremely volatile and flammable. It is irritating to the eyes, mucous membranes, and skin. It is also a CNS depressant with anesthetic effects. It may be harmful by inhalation, ingestion, or skin absorption. Avoid breathing the vapors. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood. Explosive peroxides can form during storage or on exposure to air or direct sunlight. Keep away from heat, sparks, and open flame.

Diethyl pyrocarbonate (DEPC) is a potent protein denaturant and is a suspected carcinogen. Aim bottle away from you when opening it; internal pressure can lead to splattering. Wear appropriate gloves and lab coat. Use in a chemical fume hood.

Diethyl sulfate (DES), $(C_2H_5)_2SO_4$, is a mutagen and suspected carcinogen. It is also volatile. Avoid breathing the vapors. Wear appropriate gloves. Use in a chemical fume hood. Use screw-cap tubes for all DES-treated cultures and mechanical pipettors to manipulate DES solutions. Dispose of all DES-treated cultures in bleach.

***N,N*-Dimethylformamide (DMF), $HCON(CH_3)_2$,** is irritating to the eyes, skin, and mucous membranes. It can exert its toxic effects through inhalation, ingestion, or skin absorption. Chronic inhalation can cause liver and kidney damage. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Dimethylsulfate (DMS), $(CH_3)_2\ SO_4$, is extremely toxic and is a carcinogen. Avoid breathing the vapors. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood. Dispose of solutions containing dimethylsulfate by pouring them slowly into a solution of sodium hydroxide or ammonium hydroxide and allowing them to sit overnight in the chemical fume hood. Contact the local safety office before re-entering the lab to clean up a spill.

Dimethylsulfoxide (DMSO) may be harmful by inhalation or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood. DMSO is also combustible. Store in a tightly closed container. Keep away from heat, sparks, and open flame.

Dinitrophenol (DNP) may be fatal by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood.

Diphenyloxazole (PPO) may be carcinogenic. It may be harmful by inhalation, ingestion, or skin absorption.Wear appropriate gloves and safety glasses. Consult the local institutional safety officer for specific handling and disposal procedures.

Disodium citrate, *see* **Citric acid**

Dithiothreitol (DTT) is a strong reducing agent that emits a foul odor. It may be harmful by inhalation, ingestion, or skin absorption. When working with the solid form or highly concentrated stocks, wear appropriate gloves and safety glasses. Use in a chemical fume hood.

DMF, *see* ***N,N*-Dimethylformamide**

DMS, *see* **Dimethylsulfate**

DMSO, *see* **Dimethylsulfoxide**

DNP, *see* **Dinitrophenol**

Dry ice, *see* **Carbon dioxide**

DTT, *see* **Dithiothreitol**

EDC, *see* ***N*-Ethyl-*N*′-(dimethylaminopropyl)-carbodiimide**

EMS, *see* **Ethyl methane sulfonate**

Ethanol (EtOH), CH_3CH_2OH, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Ethanolamine, $HOCH_2CH_2NH_2$, is toxic and harmful by inhalation, ingestion, or skin absorption. Handle with care and avoid any contact with the skin. Wear appropriate gloves and goggles. Use in a chemical fume hood. Ethanolamine is highly corrosive and reacts violently with acids.

Ether, *see* **Diethyl ether**

Ethidium bromide is a powerful mutagen and is toxic. Consult the local institutional safety officer for specific handling and disposal procedures. Avoid breathing the dust. Wear appropriate gloves when working with solutions that contain this dye.

Ethyl acetate may be fatal by ingestion and harmful by inhalation or skin absorption. Wear appropriate gloves and safety goggles. Do not breathe the dust. Use in a well-ventilated area.

***N*-Ethyl-*N*′-(dimethylaminopropyl)-carbodiimide (EDC)** is irritating to the mucus membranes and upper respiratory tract. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Handle with care.

Ethyl methane sulfonate (EMS) is a volatile organic solvent that is a mutagen and carcinogen. It is harmful if inhaled, ingested, or absorbed through the skin. Discard supernatants and washes containing EMS in a beaker containing 50% sodium thiosulfate. Decontaminate all material that has come in contact with EMS by treatment in a large volume of 10% (w/v) sodium thiosulfate. Use extreme caution when handling. When using undiluted EMS, wear protective appropriate gloves and use in a chemical fume hood. Store EMS in the cold. DO NOT mouth pipette EMS. Pipettes used with undiluted EMS should not be too warm; chill them in the refrigerator before use to minimize the volatility of EMS. All glassware coming in contact with EMS should be immersed in a large beaker of 1 N NaOH or laboratory bleach before recycling or disposal.

EtOH, *see* **Ethanol**

$FeCl_3$, *see* **Ferric chloride**

Ferric chloride, $FeCl_3$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood.

Formaldehyde, HCOH, is highly toxic and volatile. It is also a carcinogen. It is readily absorbed through the skin and is irritating or destructive to the skin, eyes, mucous membranes, and upper respiratory tract. Avoid breathing the vapors. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood. Keep away from heat, sparks, and open flame.

Formamide is teratogenic. The vapor is irritating to the eyes, skin, mucous membranes, and upper respiratory tract. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood when working with concentrated solutions of formamide. Keep working solutions covered as much as possible.

Formic acid, HCOOH, is highly toxic and extremely destructive to tissue of the mucous membranes, upper respiratory tract, eyes, and skin. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses (or face shield) and use in a chemical fume hood.

β-Galactosidase is an irritant and may cause allergic reactions. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Giemsa may be fatal or cause blindness by ingestion and is toxic by inhalation and skin absorption. There is a possible risk of irreversible effects. Wear appropriate gloves and safety goggles. Use only in a chemical fume hood. Do not breathe the dust.

Glassware, pressurized must be used with extreme caution. Handle glassware under vacuum, such as desiccators, vacuum traps, drying equipment, or a reactor for working under argon atmosphere, with appropriate caution. Always wear safety glasses.

Glass wool may be harmful by inhalation and may cause skin irritation. Wear appropriate gloves and mask.

Glutaraldehyde is toxic. It is readily absorbed through the skin and is irritating or destructive to the skin, eyes, mucous membranes, and upper respiratory tract. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood.

Glycine may be harmful by inhalation, ingestion, or skin absorption. Wear gloves and safety glasses. Avoid breathing the dust.

Guanidine hydrochloride is irritating to the mucous membranes, upper respiratory tract, skin, and eyes. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Avoid breathing the dust.

Guanidine thiocyanate may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Guanidinium hydrochloride, *see* **Guanidine hydrochloride**

Guanidinium isothiocyanate, *see* **Guanidine thiocyanate**

Guanidinium thiocyanate, *see* **Guanidine thiocyanate**

H_3BO_3, *see* **Boric acid**

$H_3CCOONH_4$, *see* **Ammonium acetate**

HCl, *see* **Hydrochloric acid**

HCOH, *see* **Formaldehyde**

H_3COH, *see* **Methanol**

$HCON(CH_3)_2$, *see* **Dimethylformamide**

HCOOH, *see* **Formic acid**

Heptane may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. It is extremely flammable. Keep away from heat, sparks, and open flame.

HNO_3, *see* **Nitric acid**

H_2O_2, *see* **Hydrogen peroxide**

$HOCH_2CH_2NH_2$, *see* **Ethanolamine**

$HOCH_2CH_2SH$, *see* **β-Mercaptoethanol**

H_3PO_2, *see* **Hypophosphorous acid**

H_3PO_4, *see* **Phosphoric acid (concentrated)**

H_2S, *see* **Hydrogen sulfide**

H_2SO_4, *see* **Sulfuric acid**

Hydrazine, N_2H_4, is highly toxic and explosive in the anhydrous state. It may be harmful by inhalation, ingestion, or skin absorption. Avoid breathing the vapors. Wear appropriate gloves, goggles, and protective clothing. Use only in a chemical fume hood. Dispose of solutions containing hydrazine in accordance with MSDS recommendations. Keep away from heat, sparks, and open flame.

Hydrochloric acid, HCl, is volatile and may be fatal if inhaled, ingested, or absorbed through the skin. It is extremely destructive to mucous membranes, upper respiratory tract, eyes, and skin. Wear appropriate gloves and safety glasses. Use with great care in a chemical fume hood. Wear goggles when handling large quantities.

Hydrogen peroxide, H_2O_2, is corrosive, toxic, and extremely damaging to the skin. It may be harmful by inhalation, ingestion, and skin absorption. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood.

Hydrogen sulfide, H_2S, is an extremely toxic gas that causes paralysis of the respiratory center. It is irritating and corrosive to tissues and may cause olfactory fatigue. Do not rely on odor to detect its presence. Take great care when handling it. Keep H_2S tanks in a chemical fume hood or in a room equipped with appropriate ventilation. Wear appropriate gloves and safety glasses. It is also very flammable. Keep away from heat, sparks, and open flame.

***N*-Hydroxysuccinimide** is an irritant and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Hygromycin B is highly toxic and may be fatal if inhaled, ingested, or absorbed through the skin. Wear appropriate gloves and safety goggles. Use only in a chemical fume hood. Do not breathe the dust.

Hypophosphorous acid, H_3PO_2, is usually supplied as a 50% solution, which is corrosive and should be handled with care. It should be freshly diluted immediately before use. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Inositol may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

IPTG, *see* **Isopropyl-β-D-thiogalactopyranoside**

Isoamyl alcohol may be harmful by inhalation, ingestion, or skin absorption and presents a risk of serious damage to the eyes. Wear appropriate gloves and safety goggles. Keep away from heat, sparks, and open flame.

Isobutanol, *see* **Isobutyl alcohol**

Isobutyl alcohol (Isobutanol) is extremely flammable and may be harmful by inhalation or ingestion. Wear appropriate gloves and safety glasses. Keep away from heat, sparks, and open flame.

Isopropanol is irritating and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Do not breathe the vapor. Keep away from heat, sparks, and open flame.

Isopropyl-β-D-thiogalactopyranoside (IPTG) may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Isotope ^{125}I accumulates in the thyroid and is a potential health hazard. Consult the local radiation safety office for further guidance in the appropriate use and disposal of radioactive materials. Wear appropriate gloves when handling radioactive substances. The $^{125}I_2$ formed during oxidation of $Na^{125}I$ is volatile. Work in an approved chemical fume hood with a charcoal filter when exposing the $Na^{125}I$ to oxidizing reagents such as chloramine-T, IODO-GEN, or acids. Because the oxidation proceeds very rapidly and releases large amounts of volatile $^{125}I_2$ when chloramine-T is used, it is important to be well prepared for each step of the reaction, so that the danger of contamination from volatile radiation can be minimized. Shield all forms of the isotope with lead. When handling the isotope, wear one or two pairs of appropriate gloves, depending on the amount of isotope being used and the difficulty of the manipulation required.

KCl, *see* **Potassium chloride**

$K_3Fe(CN)_6$, *see* **Potassium ferricyanide**

$K_4Fe(CN)_6 \cdot 3H_2O$, *see* **Potassium ferrocyanide**

$KH_2PO_4/K_2HPO_4/K_3PO_4$, *see* **Potassium phosphate**

$KMnO_4$, *see* **Potassium permanganate**

KOH, *see* **Potassium hydroxide**

Laser radiation, both visible and invisible, can be seriously harmful to the eyes and skin and may generate airborne contaminants, depending on the class of laser used. High-power lasers produce perma-

nent eye damage, can burn exposed skin, ignite flammable materials, and activate toxic chemicals that release hazardous by-products. Avoid eye or skin exposure to direct or scattered radiation. Do not stare at the laser and do not point the laser at someone else. Wear appropriate eye protection and use suitable shields that are designed to offer protection for the specific type of wavelength, mode of operation (continuous wave or pulsed), and power output (watts) of the laser being used. Avoid wearing jewelry or other objects that may reflect or scatter the beam. Some non-beam hazards include electrocution, fire, and asphyxiation. Entry to the area in which the laser is being used must be controlled and posted with warning signs that indicate when the laser is in use. Always follow suggested safety guidelines that accompany the equipment and contact your local safety office for further information.

LiCl, *see* **Lithium chloride**

Liquid nitrogen can cause severe damage due to extreme temperature. Handle frozen samples with extreme caution. Do not breathe the vapors. Seepage of liquid nitrogen into frozen vials can result in an exploding tube upon removal from liquid nitrogen. Use vials with O-rings when possible. Wear cryo-mitts and a face mask.

Lithium chloride, LiCl, is an irritant to the eyes, skin, mucous membranes, and upper respiratory tract. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety goggles. Use in a chemical fume hood. Do not breathe the dust.

Lysozyme is caustic to mucus membranes. Wear appropriate gloves and safety glasses.

Magnesium chloride, $MgCl_2$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Magnesium sulfate, $MgSO_4$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Manganese chloride, $MnCl_2$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

MeOH or H_3COH, *see* **Methanol**

β-Mercaptoethanol (2-Mercaptoethanol), $HOCH_2CH_2SH$, may be fatal if inhaled or absorbed through the skin and is harmful if ingested. High concentrations are extremely destructive to the mucous membranes, upper respiratory tract, skin, and eyes. β-Mercaptoethanol has a very foul odor. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood.

MES, *see* **2-[*N*-morpholino]ethanesulfonic acid**

Methanol, MeOH or **H_3COH,** is poisonous and can cause blindness. It may be harmful by inhalation, ingestion, or skin absorption. Adequate ventilation is necessary to limit exposure to vapors. Avoid inhaling these vapors. Wear appropriate gloves and goggles. Use only in a chemical fume hood.

Methotrexate (MTX) is a carcinogen and a teratogen. It may be harmful by inhalation, ingestion, or skin absorption. Exposure may cause gastrointestinal effects, bone marrow suppression, liver or kidney damage. It may also cause irritation. Avoid breathing the vapors. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood.

***N,N′*-Methylenebisacrylamide** is a poison and may effect the central nervous system. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Do not breathe the dust.

Methylene blue is irritating to the eyes and skin. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Methylmercuric hydroxide is extremely toxic and may be harmful by inhalation, ingestion, or skin absorption. It is also volatile. Therefore, carry out all manipulations of solutions containing concentrations of methylmercuric hydroxide in excess of 10^{-2} M in a chemical fume hood and wear appropriate gloves when handling such solutions. Treat all solid and liquid wastes as toxic materials and dispose of in accordance with MSDS recommendations.

$MgCl_2$, *see* **Magnesium chloride**

$MgSO_4$, *see* **Magnesium sulfate**

2-[*N*-morpholino]ethanesulfonic acid (MES) may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

3-(*N*-Morpholino)-propanesulfonic acid (MOPS) may be harmful by inhalation, ingestion, or skin absorption. It is irritating to mucous membranes and upper respiratory tract. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

$MnCl_2$, *see* **Manganese chloride**

MOPS, *see* **3-(*N*-Morpholino)-propanesulfonic acid**

MTX, *see* **Methotrexate**

NaF, *see* **Sodium fluoride**

Na_2HPO_4, *see* **Sodium hydrogen phosphate**

NaN_3, *see* **Sodium azide**

$NaNO_3$, *see* **Sodium nitrate**

NaOH, *see* **Sodium hydroxide**

N_2H_4, *see* **Hydrazine**

NH_4Cl, *see* **Ammonium chloride**

$(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$, *see* **Ammonium molybdate**

NH_4OH, *see* **Ammonium hydroxide**

$(NH_4)_2SO_4$, *see* **Ammonium sulfate**

$(NH_4)_2S_2O_8$, *see* **Ammonium persulfate**

Nickel sulfate, $NiSO_4$, is a carcinogen and may cause heritable genetic damage. It is a skin irritant and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood. Do not breathe the dust.

$NiSO_4$, *see* **Nickel sulfate**

Nitric acid, HNO_3, is volatile and must be handled with great care. It is toxic by inhalation, ingestion, and skin absorption. Wear appropriate gloves and safety goggles. Use in a chemical fume hood. Do not breathe the vapors. Keep away from heat, sparks, and open flame.

PEG, *see* **Polyethyleneglycol**

Perchloric acid may be fatal by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood.

Phenol is extremely toxic, highly corrosive, and can cause severe burns. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves, goggles, and protective clothing. Always use in a chemical fume hood. Rinse any areas of skin that come in contact with phenol with a large volume of water and wash with soap and water; do not use ethanol!

Phenylmethylsulfonyl fluoride (PMSF), $C_7H_7FO_2S$ or **$C_6H_5CH_2SO_2F$**, is a highly toxic cholinesterase inhibitor. It is extremely destructive to the mucous membranes of the respiratory tract, eyes, and skin. It may be fatal by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood. In case of contact, immediately flush eyes or skin with copious amounts of water and discard contaminated clothing.

Phosphoric acid, H_3PO_4, is highly corrosive and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Piperidine is highly toxic and is corrosive to the eyes, skin, respiratory tract, and gastrointestinal tract. It reacts violently with acids and oxidizing agents and may be harmful by inhalation, ingestion, or skin absorption. Do not breathe the vapors. Keep away from heat, sparks, and open flame. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

PMSF, *see* **Phenylmethylsulfonyl fluoride**

Polyacrylamide is considered to be nontoxic, but it should be treated with care because it may contain small quantities of unpolymerized material (*see* **Acrylamide**).

Polyethyleneglycol (PEG) may be harmful by inhalation, ingestion, or skin absorption. Avoid inhalation of powder. Wear appropriate gloves and safety glasses.

Polyvinylpyrrolidone may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Potassium cacodylate, *see* **Cacodylate**

Potassium chloride, KCl, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Potassium ferricyanide, $K_3Fe(CN)_6$, may be fatal by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Always use with extreme care in a chemical fume hood. Keep away from strong acids.

Potassium ferrocyanide, $K_4Fe(CN)_6 \cdot 3H_2O$, may be fatal by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Always use with extreme care in a chemical fume hood. Keep away from strong acids.

Potassium hydroxide, KOH and KOH/methanol, can be highly toxic. It may be harmful by inhalation, ingestion, or skin absorption. Solutions are caustic and should be handled with great care. Wear appropriate gloves.

Potassium permanganate, $KMnO_4$, is an irritant and a strong oxidant. It may form explosive mixtures when mixed with organics. Use all solutions in a chemical fume hood. Do not mix with hydrochloric acid.

Potassium phosphate, $KH_2PO_4/K_2HPO_4/K_3PO_4$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Do not breathe the dust. $K_2HPO_4 \cdot 3H_2O$ is *dibasic* and KH_2PO_4 is *monobasic*.

PPO, *see* **Diphenyloxazole**

Probe DNA or RNA, *see* **Radioactive substances**

Proteinase K is an irritant and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Putrescine is flammable and corrosive and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Keep away from heat, sparks, and open flame.

Radioactive substances: When planning an experiment that involves the use of radioactivity, consider the physico-chemical properties of the isotope (half-life, emission type, and energy), the chemical form of the radioactivity, its radioactive concentration (specific activity), total amount, and its chemical concentration. Order and use only as much as needed. Always wear appropriate gloves, lab coat, and safety goggles when handling radioactive material. **X-rays** and **gamma rays** are electromagnetic waves of very short wavelengths either generated by technical devices or emitted by radioactive materials. They might be emitted isotropically from the source or may be focused into a beam. Their potential dangers depend on the time period of exposure, the intensity experienced, and the wavelengths used. Be aware that appropriate shielding is usually made of lead or other similar material. The thickness of the shielding is determined by the energy(s) of the X-rays or gamma rays. Consult the local safety office for further guidance in the appropriate use and disposal of radioactive materials. Always

monitor thoroughly after using radioisotopes. A convenient calculator to perform routine radioactivity calculations can be found at:

http://www.graphpad.com/calculators/radcalc.cfm

S-Adenosylmethionine is toxic and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood. Do not breathe the dust.

SDS, *see* **Sodium dodecyl sulfate**

Silane may be harmful by inhalation, ingestion, or skin absorption. It is extremely flammable. Keep away from heat, sparks, and open flame. The vapor is irritating to the eyes, skin, mucous membranes, and upper respiratory tract. Wear appropriate gloves and safety glasses. Always use in a chemical fume hood.

Silica is an irritant and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Do not breathe the dust.

Silver nitrate, $AgNO_3$, is a strong oxidizing agent and should be handled with care. It may be harmful by inhalation, ingestion, or skin absorption. Avoid contact with skin. Wear appropriate gloves and safety glasses. It can cause explosions upon contact with other materials.

Sodium acetate, *see* **Acetic acid**

Sodium azide, NaN_3, is highly poisonous. It blocks the cytochrome electron transport system. Solutions containing sodium azide should be clearly marked. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety goggles and handle it with great care.

Sodium cacodylate may be carcinogenic and contains arsenic. It is highly toxic and may be fatal by inhalation, ingestion, or skin absorption. It also may cause harm to the unborn child. Effects of contact or inhalation may be delayed. Do not breathe the dust. Wear appropriate gloves and safety goggles. Use only in a chemical fume hood. *See also* **Cacodylate.**

Sodium citrate, *see* **Citric acid**

Sodium dodecyl sulfate (SDS) is toxic, an irritant, and poses a risk of severe damage to the eyes. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety goggles. Do not breathe the dust.

Sodium fluoride, NaF, is highly toxic and causes severe irritation. It may be fatal by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood.

Sodium hydrogen phosphate, Na_2HPO_4, (sodium phosphate, dibasic) may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Sodium hydroxide, NaOH, and solutions containing NaOH are highly toxic and caustic and should be handled with great care. Wear appropriate gloves and a face mask. All concentrated bases should be handled in a similar manner.

Sodium nitrate, $NaNO_2$, is irritating to the eyes, mucous membranes, upper repiratory tract, and skin. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses and always use in a chemical fume hood. Keep away from acids.

Sodium nitrate, $NaNO_3$, may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Sodium pyrophosphate is an irritant and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Do not breathe the dust.

Sodium salicylate is an irritant and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Do not breathe the dust.

Spermidine may be corrosive and harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Streptomycin is toxic and a suspected carcinogen and mutagen. It may cause allergic reactions. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Sulfuric acid, H_2SO_4, is highly toxic and extremely destructive to tissue of the mucous membranes and upper respiratory tract, eyes, and skin. It causes burns and contact with other materials (e.g., paper) may cause fire. Wear appropriate gloves, safety glasses, and lab coat. Use in a chemical fume hood.

SYBR Green I/Gold is supplied by the manufacturer as a 10,000-fold concentrate in DMSO which transports chemicals across the skin and other tissues. Wear appropriate gloves and safety glasses and decontaminate according to Safety Office guidelines. *See* **DMSO**.

TCA, *see* **Trichloroacetic acid**

TEMED, *see* ***N,N,N′,N′*-Tetramethylethylenediamine**

Tetracycline may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses and use in a chemical fume hood. Solutions of tetracycline are sensitive to light.

***N,N,N′,N′*-Tetramethylethylenediamine (TEMED)** is extremely destructive to tissues of the mucous membranes and upper respiratory tract, eyes, and skin. Inhalation may be fatal. Prolonged contact can cause severe irritation or burns. Wear appropriate gloves, safety glasses, and other protective clothing. Use only in a chemical fume hood. Wash thoroughly after handling. Flammable: Vapor may travel a considerable distance to source of ignition and flash back. Keep away from heat, sparks, and open flame.

TFA, *see* **Trifluoroacetic acid**

Thiourea may be carcinogenic and may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use in a chemical fume hood.

Tissues (human), *see* **Blood (human) and blood products**

Trichloroacetic acid (TCA) is highly caustic. Wear appropriate gloves and safety goggles.

Trichlorotrifluoroethane may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety goggles. Use in a chemical fume hood. Keep away from heat, sparks, and open flame.

Trifluoroacetic acid (TFA) (concentrated) may be harmful by inhalation, ingestion, or skin absorption. Concentrated acids must be handled with great care. Decomposition causes toxic fumes. Wear appropriate gloves and a face mask. Use in a chemical fume hood.

Tris may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Triton X-100 causes severe eye irritation and burns. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety goggles.

Trypan blue may be a carcinogen and may be harmful by inhalation, ingestion, or skin absorption. Do not breathe the dust. Wear appropriate gloves and safety glasses.

Tryptophan may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Urea may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

UV light and/or **UV radiation** is dangerous and can damage the retina of the eyes. Never look at an unshielded UV light source with naked eyes. Examples of UV light sources that are common in the laboratory include hand-held lamps and transilluminators. View only through a filter or safety glasses that absorb harmful wavelengths. UV radiation is also mutagenic and carcinogenic. To minimize exposure, make sure that the UV light source is adequately shielded. Wear protective appropriate gloves when holding materials under the UV light source.

Valine may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

X-gal may be toxic to the eyes and skin. Observe general cautions when handling the powder. Note that stock solutions of X-gal are prepared in DMF, an organic solvent. For details, see ***N,N*-dimethylformamide.** See also **5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (BCIG)**.

X-rays, *see* **Radioactive substances**

Xylene is flammable and may be narcotic at high concentrations. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Use only in a chemical fume hood. Keep away from heat, sparks, and open flame.

Xylene cyanol, *see* **Xylene**

Zinc chloride, $ZnCl_2$, is corrosive and poses possible risk to the unborn child. It may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses. Do not breathe the dust.

$ZNCl_2$, *see* **Zinc chloride**

Zymolase may be harmful by inhalation, ingestion, or skin absorption. Wear appropriate gloves and safety glasses.

Appendix 13

Suppliers

Commercial sources and products have been included in the text for the user's convenience and should not necessarily be construed as an endorsement by the authors. With the exception of those suppliers listed in the text with their addresses, all suppliers mentioned in this manual can be found in the *BioSupplyNet Source Book* and on the Web Site at:

http://www.biosupplynet.com

If a copy of *BioSupplyNet Source Book* was not included with this manual, a free copy can be ordered by using any of the following methods:

- Complete the Free *Source Book* Request Form found at the "Get the *Source Book*" link on the Web Site:

  http://www.biosupplynet.com

- E-mail a request to info@biosupplynet.com
- Fax a request to 631-370-8808

Appendix 14

Trademarks

Products mentioned in this book that may be trademarked or registered-trademarked are listed on the www.MolecularCloning.com Web Site for informational purposes only. Any trademarks or names being used are for editorial purposes only, with no intention of infringing upon that trademark. CSHL Press has provided and will continue to provide, to the best of our ability, accurate and up-to-date information on its Web Site. However, all registered trademarks and servicemarks are the property of their respective owners, and are listed on the Web Site solely for identification or description of the applicable goods and services. Therefore, please consult individual manufacturers and other resources for current and specific product information.

APPENDIX REFERENCES

Aaij C. and Borst P. 1972. The gel electrophoresis of DNA. *Biochim. Biophys. Acta* **269:** 192–200.

Adams A., Gottschling D., Kaiser C., and Stearns T. 1998. *Methods in yeast genetics: A laboratory course manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., et al. 2000. The genome sequence of *Drosophila melanogaster. Science* **287:** 2185–2195.

Åkerström B. and Björck L. 1986. A physicochemical study of protein G, a molecule with unique immunoglobulin G-binding properties. *J. Biol. Chem.* **261:** 10240–10247.

———. 1989. Protein L: An immunoglobulin light chain-binding bacterial protein. Characterization of binding and physicochemical properties. *J. Biol. Chem.* **264:** 19740–19746.

Åkerström B., Brodin T., Reis K., and Björck L. 1985. Protein G: A powerful tool for binding and detection of monoclonal and polyclonal antibodies. *J. Immunol.* **135:** 2589–2592.

Alfefs J.J., Salentijn E.M., Krens F.A., and Rouwendal G.J. 1990. Optimization of non-radioactive Southern blot hybridization: Single copy detection and reuse of blots. *Nucleic Acids Res.* **18:** 3099–3100.

Allen R.C. and Loose L.D. 1976. Phagocytic activation of a luminol-dependent chemiluminescence in rabbit alveolar and peritoneal macrophages. *Biochem. Biophys. Res. Commun.* **69:** 245–252.

Alloway J.L. 1932. The transformation *in vitro* of R pneumococci into S forms of different specific types by the use of filtered pneumococcus extracts. *J. Exp. Med.* **55:** 91–99.

———. 1933. Further observations on the use of pneumococcus extracts in effecting transformation of type in vitro. *J. Exp. Med.* **57:** 265–278.

Alm R.A., Ling L.-S., Moir D.T., King B.L., Brown E.D., Doig P.C., Smith D.R., Noonan B., Guild B.C., deJonge B.L., et al. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori* (erratum *Nature* [1999] **397:** 719). *Nature* **397:** 176–180.

Altschul S.F. and Gish W. 1996. Local alignment statistics. *Methods Enzymol.* **266:** 460–480.

Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., and Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25:** 3389–3402.

Anderson S. 1981. Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res.* **9:** 3015–3027.

Andersson S.G., Zomorodipour A., Andersson J.O., Sicheritz-Ponten T., Alsmark U.C., Podowski R.M., Näslund A.K., Eriksson A.-S., Winkler H.H., and Kurland C.G. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396:** 133–140.

Appleyard R.K. 1954. Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K12. *Genetics* **39:** 440–452.

Arakawa H., Maeda M., and Tsuji A. 1982. Chemiluminescence enzyme immunoassay of 17 α-hydroxyprogesterone using glucose oxidase and bis(2,4,6-trichlorophenyl)oxalate: fluorescent dye system. *Chem. Pharm. Bull.* **30:** 3036–3039.

Arber W., Enquist L., Hohn B., Murray N.E., and Murray K. 1983. Experimental methods for use with lambda. In *Lambda II* (ed. R.W. Hendrix et al.), pp. 433–466. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Arnold Jr., L.J., Hammond P.W., Wiese W.A., and Nelson N.C. 1989. Assay formats involving acridinium-ester-labeled DNA probes. *Clin. Chem.* **35:** 1588–1594.

Avery O.T., McLeod C.M., and McCarty M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* **79:** 137–158.

Backman K. 1980. A cautionary note on the use of certain restriction endonucleases with methylated substrates. *Gene* **11:** 169–171.

Bailey T.L. 1995. "Discovering motifs in DNA and protein sequences: The approximate common substring problem." Ph.D. thesis, University of California, San Diego.

Bailey T.L. and Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Ismb* **21:** 28–36.

———. 1995. Unsupervised learning of multiple motifs in biopolymers using EM. *Mach. Learn.* **21:** 51–80.

Bailey T.L. and Gribskov M. 1998. Methods and statistics for combining motif match searches. *J. Comput. Biol.* **5:** 211–221.

Bajorath J., Hinrichs W., and Saenger W. 1988. The enzymatic activity of proteinase K is controlled by calcium. *Eur. J. Biochem.* **176:** 441–447.

Bajorath J., Raghunathan S., Hinrichs W., and Saenger W. 1989. Long-range structural changes in proteinase K triggered by calcium ion removal. *Nature* **337:** 481–484.

Baker W., van den Broek A., Camon E., Hingamp P., Sterk P., Stoesser G., and Tuli M.A. 2000. The EMBL Nucleotide Sequence Database. *Nucleic Acids Res.* **28:** 19–23.

Baldwin T.O., Berends T., Bunch T.A., Holzman T.F., Rausch S.K., Shamansky L., Treat M.L., and Ziegler M.M. 1984. Cloning of the luciferase structural genes from *Vibrio harveyi* and expression of bioluminescence in *Escherichia coli. Biochemistry* **23:** 3663–3667.

Barany F. and Gelfand D.H. 1991. Cloning, overexpression and nucleotide sequence of a thermostable DNA ligase-encoding gene. *Gene* **109:** 1–11.

Baret A. and Fert V. 1990. T_4 and ultrasensitive TSH immunoassays using luminescent enhanced xanthine oxidase assay. *J. Biolumin. Chemilumin.* **4:** 149–153.

Baret A., Fert V., and Aumaille J. 1990. Application of a long-term enhanced xanthine oxidase-induced luminescence in solid-phase immunoassays. *Anal. Biochem.* **187:** 20–26.

Bassam B.J., Caetano-Anollés G., and Gresshoff P.M. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal. Biochem.* **196:** 80–83.

Bateman A., Birney E., Eddy S.R., Finn R.D., and E.L. Sonnhammer. 1999. Pfam 3.1: 1313 multiple alignments match the majority of proteins. *Nucleic Acids Res.* **27:** 260–262.

Bateman A., Birney E., Durbin R., Eddy S.R., Howe K.L., and Sonnhammer E.L. 2000. The Pfam protein families database. *Nucleic Acids Res.* **28:** 263–266.

Baughman G. and Howell S.H. 1988. Cauliflower mosaic virus 35S RNA leader region inhibits translation of downstream genes. *Virology* **167:** 125–135.

Beaven G.H., Holiday E.R., and Johnson E.A. 1955. Optical properties of nucleic acids and their components. In *The nucleic acids* (ed. E. Chargaff and J.N. Davidson), vol. 1, pp. 493–553. Academic Press, New York.

Beck S. and Köster H. 1990. Applications of dioxetane chemiluminescent probes to molecular biology (erratum *Anal. Chem.* [1991] **63:** 848). *Anal. Chem.* **62:** 2258–2270.

Beck S., O'Keeffe T., Coull J.M., and Köster H. 1989. Chemiluminescent detection of DNA: Application for DNA sequencing and hybridization. *Nucleic Acids Res.* **17:** 5115–5123.

Bej A.K. and Mahbubani M.H. 1994. Thermostable DNA polymerases for *in vitro* DNA amplifications. In *PCR technology: Current innovations* (ed. H.G. Griffin and A.M. Griffin), pp. 219–238. CRC Press, Boca Raton, Florida.

Belas R., Mileham A., Cohn D., Hilmen M., Simon M., and Silverman M. 1982. Bacterial luminescence: Isolation and expression of the luciferase genes from *Vibrio harveyi*. *Science* **218:** 791–793.

Bennetzen J.L. and Hall B.D. 1982. Codon selection in yeast. *J. Biol. Chem.* **257:** 3026–3031.

Bensaude O. 1988. Ethidium bromide and safety — Readers suggest alternative solutions (letter). *Trends Genet.* **4:** 89–90.

Berger S.L., Wallace D.M., Puskas R.S., and Eschenfeldt W.H. 1983. Reverse transcriptase and its associated ribonuclease H: Interplay of two enzyme activities controls the yield of single-stranded complementary deoxyribonucleic acid. *Biochemistry* **22:** 2365–2372.

Berk A.J. and Sharp P.A. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* **12:** 721–732.

Berkner K.L. and Folk W.R. 1977. Polynucleotide kinase exchange reaction. Quantitative assay for restriction endonuclease-generated 5′-phosphoryl termini in DNAs. *J. Biol. Chem.* **252:** 3176–3184.

Berman H.M., Zardecki C., and Westbrook J. 1998. The nucleic acid database: A resource for nucleic acid science. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54:** 1095–1104.

Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., and Bourne P.E. 2000. The Protein Data Bank. *Nucleic Acids Res.* **28:** 235–242.

Bernardi G. 1961. Chromatography of native deoxyribonucleic acid on calcium phosphate. *Biochem. Biophys. Res. Commun.* **6:** 54–57.

. 1965. Chromatography of nucleic acids on hydroxyapatite. *Nature* **206:** 779–783.

Bernardi G. and Bernardi G. 1986. Compositional constraints and genome evolution. *J. Mol. Evol.* **24:** 1–11.

Bernardi G. and Timasheff S.N. 1961. Chromatography of Ehrlich ascites tumor cell high molecular weight ribonucleic acid on calcium phosphate. *Biochem. Biophys. Res. Commun.* **6:** 58–61.

Bernardi A., Gaillard C., and Bernardi G. 1975. The specificity of five DNAses as studied by the analysis of 5′ terminal doublets. *Eur. J. Biochem.* **52:** 451–457.

Bertani G. and Weigle J.J. 1953. Host controlled variation in bacterial viruses. *J. Bacteriol.* **65:** 113–121.

Betzel C., Pal G.P., and Saenger W. 1988. Three-dimensional structure of proteinase K at 0.15nm resolution. *Eur. J. Biochem.* **178:** 155–171.

Bickle T.A. 1987. DNA restriction and modification systems. In Escherichia coli *and* Salmonella typhimurium: *Cellular and molecular biology* (ed. F.C. Neidhardt et al.), vol. 1, pp. 692–697. American Society for Microbiology, Washington, D.C.

Bierwerth S., Kahl G., Weigand F., and Weising K. 1992. Oligonucleotide fingerprinting of plant and fungal genomes: A comparison of radioactive, colorigenic and chemiluminescent detection methods. *Electrophoresis* **13:** 115–122.

Bird A.P. and Southern E.M. 1978. Use of restriction enzymes to study eukaryotic DNA methylation. I. The methylation pattern in ribosomal DNA from *Xenopus laevis*. *J. Mol. Biol.* **118:** 27–47.

Bird A.P., Taggart M., Frommer M., Miller O.J., and Macleod D. 1985. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* **40:** 91–99.

Birren B., Green E.D., Klapholz S., Myers R.M., and Roskams J. 1997. *Genome analysis: A laboratory manual*, vol. 1. *Analyzing DNA* (ed E.D. Green et al.). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Bitler B. and McElroy W.D. 1957. The preparation and purification of crystalline firefly luciferase. *Arch. Biochem. Biophys.* **72:** 358–368.

Bjerrum O.J. and Schafer-Nielsen C. 1986. Buffer systems and transfer parameters for semidry electroblotting with a horizontal apparatus. In *Electrophoresis '86: Proceedings of the 5th Meeting of the International Electrophoresis Society* (ed. M.J. Dunn), pp. 315–327. VCH Publishers, Deerfield Beach, Florida.

Bjerrum O.J., Larsen K.P., and Heegaard N.H. 1988. Non-specific binding and artifacts-specificity problems and troubleshooting with an atlas of immunoblotting artifacts. In *CRC handbook of immunoblotting of proteins: Technical descriptions* (ed. O.J. Bjerrum and N.H. Heegaard), vol. 1, pp. 227–254. CRC Press, Boca Raton, Florida.

Björck L. 1988. Protein L: A novel bacterial cell wall protein with affinity for IgL chains. *J. Immunol.* **140:** 1194–1197.

Björck L. and Kronvall G. 1984. Purification and some properties of streptococcal protein G, a novel IgG-binding reagent. *J. Immunol.* **133:** 969–974.

Björck, L., Kastern W., Lindahl G., and Wideback K. 1987. Streptococcal protein G, expressed by streptococci or by *Escherichia coli*, has separate binding sites for human albumin and IgG. *Mol. Immunol.* **24:** 1113–1122.

Blackburn P., Wilson G., and Moore S. 1977. Ribonuclease inhibitor from human placenta. Purification and properties. *J. Biol. Chem.* **252:** 5904–5910.

Blake C.C.F., Johnson L.N., Mair G.A., North A.C.T., Phillips D.C., and Sarma V.R. 1967. Crystallographic studies of the activity of hen egg-white lysozyme. *Proc. R. Soc. Lond. B Biol. Sci.* **167:** 378–388.

Blake M.S., Johnston K.H., Russell-Jones G.J., and Gotschlich E.C. 1984. A rapid, sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on Western blots. *Anal. Biochem.* **136:** 175–179.

Blattner F.R., Plunkett III, G., Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., et al. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **277:** 1453–1474.

Bodenteich A., Chissoe S., Wang Y.-F., and Roe B.A. 1994. Shotgun cloning as the strategy of choice to generate templates for high-throughput dideoxynucleotide sequencing. In *Automated DNA sequencing and analysis* (ed. M.D. Adams et al.), pp. 42–50. Academic Press, San Diego, California.

Bolivar F. and Backman K. 1979. Plasmids of *Escherichia coli* as cloning vectors. *Methods Enzymol.* **68:** 245–267.

Bolivar F., Rodriguez R.L., Green P.J., Betlach M.C., Heyneker H.L., Boyer H.W., Crosa J.H., and Falkow S. 1977. Construction and characterization of new cloning vehicles.

II. A multipurpose cloning system. *Gene* **2:** 95–113.

Bollum F.J. 1974. Terminal deoxynucleotidyl transferase. In *The enzymes*, 3rd edition (ed. P.D. Boyer), vol. 10, p. 145. Academic Press, New York.

Bolton A.E. and W.M. Hunter. 1973. The labelling of proteins to high specific radioactivities by conjugation to a ^{125}I-containing acylating agent. *Biochem. J.* **133:** 529–539.

Bonhoeffer S., McCaskill J.S., Stadter P.F., and Schuster P. 1993. RNA multistructure landscapes. *Biophys. J.* **22:** 13–24.

Bonner W.M. and Laskey R.A. 1974. A film detection method for tritium-labelled proteins and nucleic acids in polyacrylamide gels. *Eur. J. Biochem.* **46:** 83–88.

Borck K., Beggs J.D., Brammar W.J., Hopkins A.S., and Murray N.E. 1976. The construction in vitro of transducing derivatives of phage lambda. *Mol. Gen. Genet.* **146:** 199–207.

Borodovsky M. and McInich J. 1993. GeneMark: Parallel gene recognition for both DNA strands. *Comput. Chem.* **17:** 123–133.

Bouvet J.P. 1994. Immunoglobin Fab fragment-binding proteins. *Int. J. Immunopharmacol.* **16:** 419–424.

Bowman R.A., O'Neil G.L., and Riley T.V. 1991. Non-radioactive restriction fragment length polymorphism (RFLP) typing of *Clostridium dificile*. *FEMS Microbiol. Lett.* **63:** 269–272.

Boyer H.W. and Roulland-Dussoix D. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* **41:** 459–472.

Boylan M., Pelletier J., and Meighen E.A. 1989. Fused bacterial luciferase subunits catalyze light emission in eukaryotes and prokaryotes. *J. Biol. Chem.* **264:** 1915–1918.

Boyle M.D.P. 1990. *Bacterial immunoglobin-binding proteins*. Academic Press, San Diego, California.

Boyle M.D.P. and Reis K.J. 1987. Bacterial Fc-receptors. *Bio/Technology* **5:** 697–703.

Brawerman G., Mendecki J., and Lee S.Y. 1972. A procedure for the isolation of mammalian messenger ribonucleic acid. *Biochemistry* **11:** 637–641.

Brent R. and Ptashne M. 1981. Mechanism of action of the *lexA* gene product. *Proc. Natl. Acad. Sci.* **78:** 4204–4208.

Brickman E. and Beckwith J. 1975. Analysis of the regulation of *Escherichia coli* alkaline phosphatase synthesis using deletions and phi80 transducing phages. *J. Mol. Biol.* **96:** 307–316.

Britten R.J. and Kohne D.E. 1966. Nucleotide sequence repetition. *Carnegie Inst. Wash. Year Book* **65:** 78–106.

———. 1967. Repeated nucleotide sequences. *Carnegie Inst. Wash. Year Book* **66:** 73–88.

———. 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science* **161:** 529–540.

Britten R.J., Graham D.E., and Neufeld B.R. 1974. Analysis of repeating DNA sequences by reassociation. *Methods Enzymol.* **29**: 363–418.

Bronstein I. 1990. Chemiluminescent 1,2-dioxetane-based enzyme substrates and their applications. In *Luminescence immunoassays and molecular applications* (ed. K. Van Dyke and R. Van Dyke), pp. 255–274. CRC Press, Boca Raton, Florida.

Bronstein I. and Kricka L.J. 1989. Clinical applications of luminescent assays for enzymes and enzyme labels. *J. Clin. Lab. Anal.* **3:** 316–322.

Bronstein I. and McGrath P. 1989. Chemiluminescence lights up. *Nature* **338:** 599–600.

Bronstein I. and Voyta J.C. 1989. Chemiluminescent detection of herpes simplex virus I DNA in blot and in situ hybridization assays. *Clin. Chem.* **35:** 1856–1860.

Bronstein I., Edwards B., and Voyta J.C. 1989. 1,2-dioxetanes: Novel chemiluminescent enzyme substrates. Applications to immunoassays. *J. Biolumin. Chemilumin.* **4:** 99–111.

Bronstein I., Juo R.R., Voyta J.C., and Edwards B. 1991. Novel chemiluminescent adamantyl 1,2-dioxetane enzyme substrates. In *Bioluminescence and chemiluminescence: Current status* (ed. P.E. Stanley and L.J. Kricka), pp. 73–82. Wiley, Chichester, United Kingdom.

Bronstein I., Voyta J.C., Murphy O.J., Bresnick L., and Kricka L.J. 1992. Improved chemiluminescent Western blotting procedure. *BioTechniques* **12:** 748–753.

Bronstein I., Voyta J.C., Lazzari K.G., Murphy O., Edwards B., and Kricka L.J. 1990. Rapid and sensitive detection of DNA in Southern blots with chemiluminescence. *BioTechniques* **8:** 310–314.

Bronstein I., Voyta J.C., Murphy O.J., Tizard R., Ehrenfels C.W., and Cate R.L. 1993. Detection of DNA in Southern blots with chemiluminescence. *Methods Enzymol.* **217:** 398–414.

Brunak S., Engelbrecht J., and Knudsen J. 1991. Prediction of human mRNA donor and acceptor sites from DNA sequence. *J. Mol. Biol.* **220:** 49–65.

Brunk C.F. and Simpson L. 1977. Comparison of various ultraviolet sources for fluorescent detection of ethidium bromide-DNA complexes in polyacrylamide gels. *Anal. Biochem.* **82:** 455–462.

Bucher P. 1990. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* **212:** 563–578.

Bullock W.O., Fernandez J.M., and Short J.M. 1987. XI1-blue: A high efficiency plasmid transforming recA *Escherichia coli* strain with β-galactosidase selection. *BioTechniques* **5:** 376–379.

Bulmer M. 1990. The effects of context on synonymous codon usage in genes with low codon bias. *Nucleic Acids Res.* **18:** 2869–2873.

Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D., et al. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273:** 1058–1073.

Burge C.B. and Karlin S. 1997. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268:** 78–94.

Burke B., Griffiths G., Reggio H., Louvard D., and Warren G. 1982. A monoclonal antibody against a 135-K Golgi membrane protein. *EMBO J.* **1:** 1621–1628.

Burnette W.N. 1981. Western blotting: Electrophoretic transfer of proteins from sodium dodecyl sulphate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* **112:** 195–203.

Butler E.T. and Chamberlin M.J. 1982. Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. *J. Biol. Chem.* **257:** 5772–5778.

C. elegans Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans:* A platform for investigating biology (errata *Science* [1999] **283:** 35; **283:** 2103; **285:** 1493). *Science* **282:** 2012–2018.

Campbell A. 1965. The steric effect in lysogenization by bacteriophage lambda. I. Lysogenization of a partially diploid strain of *Escherichia coli* K12. *Virology* **27:** 329–339.

Campbell A.K. 1988. *Chemiluminescence*. Ellis Horwood, Chichester, United Kingdom.

Campbell J.L., Richardson C.C., and Studier F.W. 1978. Genetic recombination and complementation between bacterio-

phage T7 and cloned fragments of T7 DNA. *Proc. Natl. Acad. Sci.* **75:** 2276–2280.

Campbell V.W. and Jackson D.A. 1980. The effect of divalent cations on the mode of action of DNase I. The initial reaction products produced from covalently closed circular DNA. *J. Biol. Chem.* **255:** 3726–3735.

Cardwell H.M.E. and Smith S. 1953. The structure of digoxygenin. *Experientia* **9:** 267–368.

———. 1954. Digitalis gluosides. VII. The structure of the digitalis anhydrogenins and the orientation of the hydroxy-groups in digoxygenin and gitoxigenin. *J. Chem. Soc.* **1954:** 2012–2023.

Carlson D.P., Superko C., Mackey J., Gaskill M.E., and Hansen P. 1990. Chemiluminescent detection of nucleic acid hybridization. *Focus* (Life Technologies) **12:** 9–12.

Cate R.L., Ehrenfels C.W., Wysk M., Tizard R., Voyta J.C., Murphy O.J., and Bronstein I. 1991. Genomic Southern analysis with alkaline phosphatase-conjugated oligonucleotide probes and the chemiluminescent substrate AMPPD. *Genet. Anal. Tech. Appl.* **8:** 102–106.

Ceci L.R., Volpicella M., Liuni S., Volpetti V., Licciulli F., and Gallerani R. 1999. PLMItRNA, a database for higher plant mitochondrial tRNAs and tRNA genes. *Nucleic Acids Res.* **27:** 156–157.

Cesarone C.F., Bolognesi C., and Santi L. 1979. Improved microfluorimetric DNA determination in biological material using 33258 Hoechst. *Anal. Biochem.* **100:** 188–197.

———. 1980. Fluorometric DNA assay at nanogram levels in biological material. *Boll. Soc. Ital. Biol. Sper.* **56:** 1666–1672.

Chalfie M., Tu Y., Euskirchen G., Ward W.W., and Prasher D.C. 1994. Green fluorescent protein as a marker for gene expression. *Science* **263:** 802–805.

Chalon, M.P., R.W. Milne, and J.P. Vaerman. 1979. Interactions between mouse immunoglobulins and staphylococcal protein A. *Scand. J. Immunol.* **9:** 359–364.

Chamberlain J.P. 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. *Anal. Biochem.* **98:** 132–135.

Chang L.M.S. and Bollum F.J. 1986. Molecular biology of terminal transferase. *Crit. Rev. Biochem.* **21:** 27–52.

Chao K.-M., Pearson W.R., and Miller W. 1992. Aligning two sequences within a specified diagonal band. *Comput. Appl. Biosci.* **8:** 481–487.

Chao S., Chao L., and Chao J. 1989. Enhanced specificity in immunoscreening of expression cDNA clones using radiolabelled antigen technology. *BioTechniques* **7:** 68–72.

Chen W., Tabor S., and Struhl K. 1987. Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 RNA polymerase. *Cell* **50:** 1047–1055.

Cheng H. and Rudick M.J. 1991. A membrane blotting method for following the time course of protein radioiodination using Iodobeads. *Anal. Biochem.* **198:** 191–193.

Chien A., Edgar D.B., and Trela J.M. 1976. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.* **127:** 1550–1557.

Chiu W.L., Niwa Y., Zeng W., Hirano T., Kobayashi H., and Sheen J. 1996. Engineered GFP as a vital marker in plants. *Curr. Biol.* **6:** 325–330.

Chothia C. 1976. The nature of the accessible and buried surfaces in proteins. *J. Mol. Biol.* **105:** 1–12.

Chrambach A. and Rodbard D. 1972. Polymerization of polyacrylamide gels: Efficiency and reproducibility as a function of catalyst concentrations. *Sep. Sci.* **7:** 663–703.

Church G.M. and Gilbert W. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci.* **81:** 1991–1995.

Church G.M. and Kieffer-Higgins S. 1988. Multiplex DNA sequencing. *Science* **240:** 185–188.

Cohn D.H., Mileham A.J., Simon M.I., Nealson K.H., Rausch S.K., Bonam D., and Baldwin T.O. 1985. Nucleotide sequence of the *luxA* gene of *Vibrio harveyi* and the complete amino acid sequence of the α subunit of bacterial luciferase. *J. Biol. Chem.* **260:** 6139–6146.

Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E., et al. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence (erratum *Nature* [1998] **396:** 190). *Nature* **393:** 537–544.

Cook R.F., Cook S.J., and Issel C.J. 1992. A nonradioactive microassay for released reverse transcriptase activity of a lentivirus. *BioTechniques* **13:** 380–386.

Corpet F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16:** 10881–10890.

Corpet F., Gouzy J., and Kahn D. 1998. The ProDom database of protein domain families. *Nucleic Acids Res.* **26:** 323–326.

Cozzarelli N.R., Kelly R.B., and Kornberg A. 1968. A minute circular DNA from *Escherichia coli* 15. *Proc. Natl. Acad. Sci.* **60:** 992–999.

———. 1969. Enzymic synthesis of DNA. XXXIII. Hydrolysis of a 5′-triphosphate-terminated polynucleotide in the active center of DNA polymerase. *J. Mol. Biol.* **45:** 513–531.

Creasey A., D'Angio L.J., Dunne T.S., Kissinger C., O'Keefe T., Perry-O'Keefe H., Moran L., Roskey M., Schildkraut I., Sears L.E., and Slatko B. 1991. Application of a novel chemiluminescent-based DNA detection method to a single vector and multiplex DNA sequencing. *BioTechniques* **11:** 102–109.

Crooke S.T. 1998. Molecular mechanisms of antisense drugs: RNase H. *Antisense Nucleic Acid Drug Dev.* **8:** 133–134.

Crouch R. 1990. Ribonuclease H: From discovery to 3D structure. *New Biol.* **2:** 771–777.

Daniels D.L., Schroeder J.L., Szybalski W., Sanger F., and Blattner F.R. 1983. Appendix I: A molecular map of coliphage lambda. In *Lambda II* (ed. R.W. Hendrix et al.), pp. 469–517. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Davanloo P., Rosenberg A.H., Dunn J.J., and Studier F.W. 1984. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci.* **81:** 2035–2039.

Davidson J.N. 1972. *The biochemistry of the nucleic acids*, 7th edition. Academic Press, New York.

Davidson P.F. 1959. The effect of hydrodynamic shear on the deoxyribonucleic acid from T2 and T4 bacteriophages. *Proc. Natl. Acad. Sci.* **45:** 1560–1568.

———. 1960. Biophysics: Sedimentation of deoxyribonucleic acid isolated under low hydrodynamic shear. *Nature* **185:** 918–920.

Davis B.J. 1964. Disc electrophoresis. II. Method and application to human serum proteins. *Ann. N.Y. Acad. Sci.* **121:** 404–427.

Davis M.M. 1986. Subtractive cDNA hybridization and the T-cell receptor genes. In *Handbook of experimental immunology* (ed. D.M. Weir et al.), pp. 76.1–76.13. Blackwell Scientific, Oxford, United Kingdom.

Daxhelet G.A., Coene M.M., Hoet P.P., and Cocito C.G. 1989. Spectrofluorometry of dyes with DNAs of different base composition and conformation. *Anal. Biochem.* **179:** 401–403.

Dayhoff M.O., ed. 1972. *Atlas of protein sequence and structure*, vol. 5. National Biomedical Research Foundation,

Georgetown University Medical Center, Washington, D.C.

Dayhoff M.O., Schwartz R.M., and Orcutt B.C. 1978. A model of evolutionary change in proteins. In *Atlas of protein sequence and structure* (ed. M.O. Dayhoff et al.), vol. 5, suppl. 3, pp. 345–352. National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C.

de Blas A.L. and Cherwinski H.M. 1983. Detection of antigens on nitrocellulose: In situ staining of alkaline phosphatase conjugated antibody. *Anal. Biochem.* **133:** 214–219.

Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R., Feldman R.A., Short J.M., Olsen G.J., and Swanson R.V. 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392:** 353–358.

Decorte R. and Cassiman J.J. 1991. Detection of amplified VNTR alleles by direct chemiluminescence: Application to the genetic identification of biological samples in forensic cases. *Exper. Suppl.* **58:** 371–390.

Deininger P.L. 1983. Approaches to rapid DNA sequence analysis. *Anal. Biochem.* **135:** 247–263.

Deisenhofer J. 1981. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8-Å resolution. *Biochemistry* **20:** 2361–2370.

DeLuca M. and McElroy W.D. 1978. Purification and properties of firefly luciferase. *Methods Enzymol.* **57:** 3–14.

. 1984. Two kinetically distinguishable ATP sites in firefly luciferase. *Biochem. Biophys. Res. Commun.* **123:** 764–770.

Deng G. and Wu R. 1983. Terminal transferase: Use in the tailing of DNA and for in vitro mutagenesis. *Methods Enzymol.* **100:** 96–116.

Denhardt D.T. 1966. A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* **23:** 641–646.

DenHollander N. and Befus D. 1989. Loss of antigens from immunoblotting membranes. *J. Immunol. Methods* **122:** 129–135.

Dente L., Cesareni G., and Cortese R. 1983. pEMBL: A new family of single-stranded plasmids. *Nucleic Acids Res.* **11:** 1645–1655.

De Rijk P., Neefs J.M., Van de Peer Y., and De Wachter R. 1992. Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res.* **20:** 2075–2089.

De Rijk P., Robbrecht E., de Hoog S., Caers A., Van de Peer Y., and De Wachter R. 1999. Database on the structure of large subunit ribosomal RNA. *Nucleic Acids Res.* **27:** 174–178.

Derrick J.P. and Wigley D.B. 1992. Crystal structure of a streptococcal protein G domain bound to a Fab fragment. *Nature* **359:** 752–754.

de Wet J.R., Wood K.V., Helinski D.R., and DeLuca M. 1985. Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **82:** 7870–7873.

. 1986. Cloning firefly luciferase. *Methods Enzymol.* **133:** 3–14.

de Wet J.R., Wood K.V., DeLuca M., Helinski D.R., and Subramani S. 1987. Firefly luciferase gene: Structure and expression in mammalian cells. *Mol. Cell. Biol.* **7:** 725–737.

DiLella A.G., Hope D.A., Chen H., Trumbauer M., Schwartz R.J., and Smith R.G. 1988. Utility of firefly luciferase as a reporter gene for promoter activity in transgenic mice. *Nucleic Acids Res.* **16:** 4159.

Dooley S., Welter C., and Blin N. 1992. Nonradioactive southwestern analysis using chemiluminescent detection. *BioTechniques* **13:** 540–542.

Dresel H.A. and Schettler G. 1984. Characterization and visualisation of the low density lipoprotein receptor by ligand blotting using anti-low density lipoprotein enzyme-linked immunoabsorbent assay (ELISA). *Electrophoresis* **5:** 372–373.

Dubitsky A., Brown J., and Brandwein H. 1992. Chemiluminescent detection of DNA on nylon membranes. *BioTechniques* **13:** 392–400.

Durrant I. 1990. Light-based detection of biomolecules. *Nature* **346:** 297–298.

Earnshaw W.C. and Rothfield N. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* **91:** 313–321.

Ebeling W., Hennrich N., Klockow M., Metz H., Orth H.D., and Lang H. 1974. Proteinase K from *Tritirachium album* Limber. *Eur. J. Biochem.* **47:** 91–97.

Economou J.S., Rhoades K., Essner R., McBride W.H., Gasson J.C., and Morton D.L. 1989. Genetic analysis of human tumor necrosis factor α/cachectin promoter region in a macrophage cell line. *J. Exp. Med.* **170:** 321–326.

Eddy S.R. and Durbin R. 1994. RNA sequence analysis using covariance models. *Nucleic Acids Res.* **22:** 2079–2088.

Edgar R.S. 1966. Conditional lethals. In *Phage and the origins of molecular biology* (ed. J. Cairns et al.), pp. 166–170. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Edgar R.S. and Lielausis I. 1964. Temperature-sensitive mutants of bacteriophage T4D: Their isolation and genetic characterization. *Genetics* **49:** 649–662.

Efstratiadis A., Kafatos F.C., Maxam A.M., and Maniatis T. 1976. Enzymatic in vitro synthesis of globin genes. *Cell* **7:** 279–288.

Ehrlich S.D., Devillers-Thiery A., and Bernardi G. 1973. Further investigations on the specificity of an acid deoxyribonuclease from *Helix aspersa* (Mull.). *Eur. J. Biochem.* **40:** 139–141.

Eliasson M., Andersson R., Olsson A., Wigzell H., and Uhlén M. 1989. Differential IgG-binding characteristics of staphylococcal protein A, streptococcal protein G, and a chimeric protein AG. *J. Immunol.* **142:** 575–581.

Eliasson M., Olsson A., Palmcrantz E., Wiberg K., Inganas M., Guss B., Lindberg M., and Uhlén M. 1988. Chimeric IgG-binding receptors engineered from staphylococcal protein A and streptococcal protein G. *J. Biol. Chem.* **263:** 4323–4327.

Engler M.J. and Richardson C.C. 1982. DNA ligases. In *The enzymes*, 3rd edition (ed. P.D. Boyer), vol. 15, pp. 3–29. Academic Press, New York.

Engler-Blum G., Meier M., Frank J., and Muller G.A. 1993. Reduction of background problems in nonradioactive northern and southern blot analyses enables higher sensitivity than ^{32}P-based hybridizations. *Anal. Biochem.* **210:** 235–244.

Engvall E. 1980. Enzyme immunoassay ELISA and EMIT. *Methods Enzymol.* **70:** 419–439.

Epstein R.H., Bolle A., Steinberg C.M., Kellenberger E., Boy de la Tour E., Chevalley R., Edgar R.S., Susman M., Denhardt G.H., and Lielausis A. 1964. Physiological studies of conditional lethal mutants of bacteriophage T4D. *Cold Spring Harbor Symp. Quant. Biol.* **28:** 375–394.

Erntell M., Myhre E.B., and Kronvall G. 1985. Non-immune IgG F(ab′)2 binding to group C and G streptococci is mediated by structures on gamma chains. *Scand. J. Immunol.* **21:** 151–157.

Eschenfeldt W.H., Puskas R.S., and Berger S.L. 1987. Homopolymeric tailing. *Methods Enzymol.* **152:** 337–342.

Ey P.L., Prowse S.J., and Jenkin C.R. 1978. Isolation of pure IgG_1, IgG_{2a} and IgG_{2b} immunoglobulins from mouse serum using protein A-Sepharose. *Immunochemistry* **15:** 429–436.

Eyre-Walker A.C. 1991. An analysis of codon usage in mammals: Selection or mutation-bias? *J. Mol. Evol.* **33:** 442–449.

Favaloro J., Treisman R., and Kamen R. 1980. Transcription maps of polyoma virus-specific RNA: Analysis by two-dimensional nuclease S1 gel mapping. *Methods Enzymol.* **65:** 718–749.

Fazekas de St. Groth S., Webster R.G. and Datyner A. 1963. Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. *Biochim. Biophys. Acta* **71:** 377–391.

Felton J. 1983. M13 host strain JM103 contains two restriction systems. *BioTechniques* **1:** 42–43.

Ferguson W.J., Braunschweiger K.I., Braunschweiger W.R., Smith J.R., McCormick J.J., Wasmann C.C., Jarvis N.P., Bell D.H., and Good N.E. 1980. Hydrogen ion buffers for biological research. *Anal. Biochem.* **104:** 300–310.

Fichant G.A. and C. Burks. 1991. Identifying potential tRNA genes in genomic DNA sequences. *J. Mol. Biol.* **220:** 659–671.

Fields-Berry S.C., Halliday A.L., and Cepko C.L. 1992. A recombinant retrovirus encoding alkaline phosphatase confirms clonal boundary assignment in lineage analysis of murine retina. *Proc. Natl. Acad. Sci.* **89:** 693–697.

Fierz-David H.E. and Blangey L. 1949. *Fundamental processes of dye chemistry.* Interscience, New York.

Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M., et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269:** 496–512.

Fleming A. 1922. On a remarkable bacteriolytic element found in tissues and secretions. *Proc. R. Soc. Lond. B Biol. Sci.* **93:** 306–317.

Fontana W., Konings D.A.M., Stadler P.F., and Schuster P. 1993. Statistics of RNA secondary structures. *Biopolymers* **13:** 1389–1404.

Forsgren A. and Sjöquist J. 1966. "Protein A" from *S. aureus.* I. Pseudo-immune reaction with human gamma-globulin. *J. Immunol.* **97:** 822–827.

———. 1967. "Protein A" from *Staphylococcus aureus.* III. Reaction with rabbit gamma-globulin. *J. Immunol.* **99:** 19–24.

Fraker P.J. and Speck Jr., J.C. 1978. Protein and cell membrane iodinations with a sparingly soluble chloroamide, 1,3,4,6-tetrachloro-3a,6a-diphrenylglycoluril. *Biochem. Biophys. Res. Commun.* **80:** 849–857.

Franci C. and Vidal J. 1988. Coupling redox and enzymic reactions improves the sensitivity of the ELISA-spot assay. *J. Immunol. Methods* **107:** 239–244.

Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., et al. 1995. The minimal gene complement of *Mycoplasma genitalium. Science* **270:** 397–403.

Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R., Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., et al. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi. Nature* **390:** 580–586.

Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., et al. 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* **281:** 375–388.

Freifelder D. 1971. Electron microscopic study of the ethidium bromide-DNA complex. *J. Mol. Biol.* **60:** 401–403.

French B.T., Maul H.M., and Maul G.G. 1986. Screening cDNA expression libraries with monoclonal and polyclonal antibodies using an amplified biotin-avidin-peroxidase technique. *Anal. Biochem.* **156:** 417–423.

Frischauf A.M., Garoff H., and Lehrach H. 1980. A subcloning strategy for DNA sequence analysis. *Nucleic Acids Res.* **8:** 5541–5549.

Frischauf A.M., Lehrach H., Poustka A., and Murray N. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* **170:** 827–842.

Fursht A. 1985. *Enzyme structure and mechanism*, 2nd edition. W.H. Freeman, New York.

Gaillard C. and Strauss F. 1990. Ethanol precipitation of DNA with linear acrylamide as a carrier. *Nucleic Acids Res.* **18:** 378.

Galas D.J. and Schmitz A. 1978. DNAase footprinting: A simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **5:** 3157–3170.

Gallimore P.H., Sharp P.A., and Sambrook J. 1974. Viral DNA in transformed cells. II. A study of the sequences of adenovirus 2 DNA in nine lines of transformed rat cells using specific fragments of the genome. *J. Mol. Biol.* **89:** 49–72.

Garner R.C. 1975. Testing of some benzidine analogues for microsomal activation to bacterial mutagens. *Cancer Lett.* **1:** 39–42.

Gaugain B., Barbet J., Oberlin R., Roques B.P., and LePecq J.B. 1978. DNA bifunctional intercalators. I. Synthesis and conformational proerties of an ethidium homodimer and of an acridine ethidium heterodimer. *Biochemistry* **17:** 5071–5078.

Gelb L.D., Kohne D.E., and Martin M.A. 1971. Quantitation of simian virus 40 sequences in African green monkey, mouse and virus-transformed cell genomes. *J. Mol. Biol.* **57:** 129–145.

Gerard G.F. 1983. Reverse transcriptase. In *Enzymes of nucleic acid synthesis and modification: DNA enzymes* (ed. S.T. Jacob), vol. 1, p. 1. CRC Press, Boca Raton, Florida.

Gershoni J.M. 1988. Protein blotting: A manual. *Methods Biochem. Anal.* **33:** 1–58.

Gershoni J.M. and Palade G.E. 1982. Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to a positively charged membrane filter. *Anal. Biochem.* **124:** 396–405.

Ghosh D. 1998. OOTFD (Object-Oriented Transcription Factors Database): An object-oriented successor to TFD. *Nucleic Acids Res.* **26:** 360–362.

Gibson T.J. 1984. "Studies on the Epstein-Barr virus genome." Ph.D. thesis, Cambridge University, United Kingdom.

Gillespie P.G. and Hudspeth A.J. 1991. Chemiluminescence detection of proteins from single cells. *Proc. Natl. Acad. Sci.* **88:** 2563–2567.

Gish W. and States D.J. 1993. Identification of protein coding regions by database similarity search. *Nat. Genet.* **3:** 266–272.

Glasel J. 1995. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *BioTechniques* **18:** 62–63.

Glazer A.N. and Rye H.S. 1992. Stable dye-DNA intercalation complexes as reagents for high sensitivity fluorescence detection. *Nature* **359:** 859–861.

Glazer A.N., Peck K., and Mathies R.A. 1990. A stable soubble-stranded DNA-ethidium homodimer complex: Application to picogram fluorescence detection of DNA in agarose gels. *Proc. Natl. Acad. Sci.* **87:** 3851–3855.

Goding J.W. 1978. Use of staphylococcal protein A as an immunological reagent. *J. Immunol. Methods* **20:** 241–253.

Godson G.N. and Vapnek D. 1973. A simple method of preparing large amounts of ϕX174 RFI supercoiled DNA. *Biochim.*

Biophys. Acta **299:** 516–520.

Goffeau A., Barrell B.G., Bussey H., Davis R.W., Dujon B., Feldmann H., Galibert F., Hoheisel J.D., Jacq C., Johnston M., Louis E.J., Mewes H.W., Murakami Y., Philippsen P., Tettelin H., and Oliver S.G. 1996. Life with 6000 genes. *Science* **274:** 546, 563–567.

Goldberg D.A. 1980. Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene. *Proc. Natl. Acad. Sci.* **77:** 5794–5798.

Gomori G. 1946. Buffers in the range of pH 6.5 to 9.6. *Proc. Soc. Exp. Biol. Med.* **6:** 233–234.

———. 1955. Preparation of buffers for use in enzyme studies. *Methods Enzymol.* **1:** 138–146.

Gonnett G.H., Cohen M.A., and Benner S.A. 1992. Exhaustive matching of the entire sequence database. *Science* **256:** 1443–1445.

Good N.E. and Izawa S. 1972. Hydrogen ion buffers. *Methods Enzymol.* **24:** 53–68.

Good N.E., Winget G.D., Winter W., Connolly T.N., Izawa S., and Singh R.M.M. 1966. Hydrogen ion buffers for biological research. *Biochemistry* **5:** 467–477.

Gotoh O. 1995. A weight system and algorithm for aligning many phylogenetically related sequences. *Comput. Appl. Biosci.* **11:** 543–551.

Goudswaard J., Van der Donk J.A., Noordzij A., van Dam R.H., and Vaerman J.-P. 1978. Protein A reactivity of various mammalian immunoglobulins. *Scand. J. Immunol.* **8:** 21–28.

Gough J.A. and Murray N.E. 1983. Sequence diversity among related genes for recognition of specific targets in DNA molecules. *J. Mol. Biol.* **166:** 1–19.

Gould S.J. and Subramani S. 1988. Firefly luciferase as a tool in molecular and cell biology. *Anal. Biochem.* **175:** 5–13.

Gouy M. and Gautier C. 1982. Codon usage in bacteria: Correlation with gene expressivity. *Nucleic Acids Res.* **10:** 7055–7074.

Graham Jr., R.C. and Karnovsky M.J. 1966. The early stages of absorption of injected horseradish peroxidase in the proximal tubules of mouse kidney: Ultrastructural cytochemistry by a new technique. *J. Histochem. Cytochem.* **14:** 291–302.

Grantham R., Gautier C., Gouy M., Jacobzone M., and Mercier R. 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucleic Acids Res.* **9:** r43–r74.

Grantham R., Gautier C., Gouy M., Mercier R., and Pave A. 1980. Codon catalog usage and the genome hypothesis. *Nucleic Acids Res.* **8:** r49–r62.

Grassmann W. and Deffner G. 1953. Verteilungschromatographisches Verhalten von Proteinen und Peptiden in phenolhaltigen Lösungsmitteln. *Hoppe-Seylor's Z. Physiol. Chem.* **293:** 89–98.

Gray Jr., H.B., Winston T.P., Hodnett J.L., Legerski R.J., Nees D.W., Wei C.-F., and Robberson D.L. 1981. The extracellular nuclease from *Alteromonas espejiana:* An enzyme highly specific for nonduplex structure in nominally duplex DNAs. In *Gene amplification and analysis: Structural analysis of nucleic acids* (ed. J.G. Chirikjian and T.S. Papas), vol. 2, pp. 169–203. Elsevier, New York.

Green A.A. 1933. The preparation of acetate and phosphate buffer solutions of known pH and ionic strength. *J. Am. Chem. Soc.* **55:** 2331–2336.

Green N.M. 1975. Avidin. *Adv. Protein Chem.* **29:** 85–133.

Green N.M. and Toms E.J. 1972. The dissociation of avidin-biotin complexes by guanidinium chloride. *Biochem. J.* **130:** 707–711.

———. 1973. The properties of subunits of avidin coupled to sepharose. *Biochem. J.* **133:** 687–700.

Greene P.H., Poonian M.S., Nussbaum A.L., Tobias L., Garfin D.E., Boyer H.W., and Goodman H.M. 1975. Restriction and modification of a self-complementary octanucleotide containing the EcoRI substrate. *J. Mol. Biol.* **99:** 237–261.

Greenfield L., Simpson L., and Kaplan D. 1975. Conversion of closed circular DNA molecules to single-nicked molecules by digestion with DNAase I in the presence of ethidium bromide. *Biochim. Biophys. Acta* **407:** 365–375.

Greenwood F.C., Hunter W.M., and Glover J.S. 1963. The preparation of ^{131}I-labelled human growth hormone of high specific radioactivity. *Biochem. J.* **89:** 114–123.

Gronenborn A.M., Filpula D.R., Essig N.Z., Achari A., Whitlow M., Wingfield P.T., and Clore G.M. 1991. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* **253:** 657–661.

Grundy W.N., Bailey T.L. , Elkan C.P., and Baker M.E. 1997. Meta-MEME: Motif-based hidden Markov models of proteins families. *Comput. Appl. Biosci.* **13:** 397–406.

Guarente L. and Ptashne M. 1981. Fusion of *Escherichia coli lacZ* to the cytochrome *c* gene of *Saccharomyces cerevisiae. Proc. Natl. Acad. Sci.* **78:** 2199–2203.

Gundermann K.-D. and McCapra F. 1987. *Chemiluminescence in organic chemistry.* Springer-Verlag, Berlin.

Gunkel F.A. and Gassen H.G. 1989. Proteinase K from *Tritirachium album* Limber. Characterization of the chromosomal gene and expression of the cDNA in *Escherichia coli. Eur. J. Biochem.* **179:** 185–194.

Gutell R.R. 1994. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. *Nucleic Acids Res.* **22:** 3502–3507.

Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166:** 557–580.

———. 1985. Techniques for transformation of *E. coli.* In *DNA cloning: A practical approach* (ed. D.M. Glover), vol. 1, p. 109–135. IRL Press, Oxford, United Kingdom.

Hancock K. and Tsang V.C. 1983. India ink staining of proteins on nitrocellulose paper. *Anal. Biochem.* **133:** 157–162.

Harlow E. and Lane D. 1988. *Antibodies: A laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

———. 1999. *Using antibodies: A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Harper D.R. and Murphy G. 1991. Nonuniform variation in band pattern with luminol/horse radish peroxidase Western blotting. *Anal. Biochem.* **192:** 59–63.

Hattman S., Brooks J.E. and Masurekar M. 1978. Sequence-specificity of the P1 modification methylase (M.*Eco* P1) and the DNA methylase (M.*Eco* dam) controlled by the *Escherichia coli dam* gene. *J. Mol. Biol.* **126:** 367–380.

Hausen P. and Stein H. 1970. Ribonuclease H. An enzyme degrading the RNA moiety of DNA-RNA hybrids. *Eur. J. Biochem.* **14:** 278–283.

Haussler D., Krogh A., Mian S., and Sjölander K. 1992. Protein modeling using hidden markov models. In *C.I.S. Technical Report UCSC-CRL-92-93.* University of California, Santa Cruz.

Hawkes R. 1986. The dot immunobinding assay. *Methods Enzymol.* **121:** 484–491.

Hawkes R., Niday E., and Gordon J. 1982. A dot-immunobinding assay for monoclonal and other antibodies. *Anal. Biochem.* **119:** 142–147.

Hayashi K., Nakazawa M., Ishizaki Y., Hiraoka N., and Obayashi

A. 1986. Regulation of inter- and intramolecular ligation with T4 DNA ligase in the presence of polyethylene glycol. *Nucleic Acids Res.* **14:** 7617–7631.

Hegenauer R. 1971. Pflanstoffe und Pflanzensystematik. *Naturwissenschaften* **58:** 585–598.

Heinemeyer T., Chen X., Karas H., Krel A.E., Kel O.V., Liebach I., Meinhardt T., Reuter I., Schacherer E., and Wingender E. 1999. Expanding the TRANSFAC database towards an expert system of regulatory molecular mechanisms. *Nucleic Acids Res.* **27:** 318–322.

Heitman J. and Model P. 1987. Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. *J. Bacteriol.* **169:** 3243–3250.

Hemmati-Brivanlou A., Frank D., Bolce M.E., Brown B.D., Sive H.L., and Harland R.M. 1990. Localization of specific mRNAs in *Xenopus* embryos by whole-mount *in situ* hybridization. *Development* **110:** 325–330.

Hengen P.N. 1997. Shearing DNA for genomic library construction. *Trends Biochem. Sci.* **22:** 273–274.

Henikoff J.G., Henikoff S., and Pietrokovski S. 1999a. New features of the Blocks Database servers. *Nucleic Acids Res.* **27:** 226–228.

———. 1999b. Blocks+: A non-redundant database of protein alignment blocks derived from multiple compilations. *Bioinformatics* **15:** 471–479.

Henikoff S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:** 351–359.

Henikoff S. and Henikoff J.G. 1991. Automated assembly of protein blocks for database searching. *Nucleic Acids Res.* **19:** 6565–6572.

———. 1992. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci.* **89:** 10915–10919.

———. 1993. Performance evaluation of amino acid substitution matrices. *Protein* **17:** 49–61.

———. 1994. Protein family classification based on searching a database of blocks. *Genomics* **19:** 97–107.

Henikoff S., Henikoff J.G., Alford W.J., and Pietrokovksi S. 1995. Automated construction and graphical presentation of protein blocks from unaligned sequences. *Gene* **163:** GC17–26.

Herrin D.L. and Schmidt G.W. 1988. Rapid, reversible staining of northern blots prior to hybridization. *BioTechniques* **6:** 196–200.

Hertz G.Z. and Stormo G.D. 1995. Identification of consensus patterns in unaligned DNA and protein sequences: A large deviation statistical basis for penalizing gaps. In *Proceedings of the 3rd International Conference on Bioinformatics and Genome Research* (ed. H.A. Lim and C.R. Cantor), pp. 201–216. World Scientific, Singapore.

———. 1999. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15:** 563–577.

Higgins D.G. and Sharp P.M. 1988. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* **73:** 237–244.

Higgins D.G., Bleasby A.J., and Fuchs R. 1992. CLUSTAL V: Improved software for multiple sequence alignment. *Comput. Appl. Biosci.* **8:** 189–191.

Higo K., Ugawa Y., Iwamoto M., and Korenaga T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* **27:** 297–300.

Hilwig I. and Gropp A. 1975. pH-dependent fluorescence of DNA and RNA in cytologic staining with "33258 Hoechst". *Exp. Cell Res.* **91:** 457–460.

Hilz H., Wiegers U., and Adamietz P. 1975. Stimulation of proteinase K action by denaturing agents: Application to the isolation of nucleic acids and the degradation of 'masked' proteins. *Eur. J. Biochem.* **56:** 103–108.

Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.C., and Herrmann R. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res.* **24:** 4420–4449.

Hjelm H., Hjelm K., and Sjöquist J. 1972. Protein A from *Staphylococcus aureus*. Its isolation by affinity chromatography and its use as an immunosorbent for isolation of immunoglobulins. *FEBS Lett.* **28:** 73–76.

Hodgson M. and Jones P. 1989. Enhanced chemiluminescence in the peroxidase-luminol-H_2O_2 system: Anomalous reactivity of enhancer phenols with enzyme intermediates. *J. Biolumin. Chemilumin.* **3:** 21–25.

Hofacker I.L., Fontana W., Stadler P.F., Bonhoeffer S., Tacker M., and Schuster P. 1994. Fast folding and comparison of RNA secondary structures. *Monatsh. Chem.* **125:** 167–188.

Hohn B. 1979. *In vitro* packaging of λ and cosmid DNA. *Methods Enzymol.* **68:** 299–309.

Hohn B. and Murray K. 1977. Packaging recombinant DNA molecules into bacteriophage particles *in vitro*. *Proc. Natl. Acad. Sci.* **74:** 3259–3263.

Holm L. and Sander C. 1996. Mapping the protein universe. *Science* **273:** 595–603.

Holmstrom K., Rossen L., and Rasmussen O.F. 1993. A highly sensitive and fast nonradioactive method for detection of polymerase chain reaction products. *Anal. Biochem.* **209:** 278–283.

Höltke H.-J. and Kessler C. 1990. Non-radioactive labeling of RNA transcripts in vitro with the hapten digoxigenin (DIG): Hybridization and ELISA-based detection. *Nucleic Acids Res.* **18:** 5843–5851.

Höltke H.-J., Ettl J., Obermaier J., and Schmitz G. 1991. Sensitive chemiluminescent detection of digoxigenin (DIG) labeled nucleic acids. A fast and simple protocol and its applications. In *Bioluminescence and chemiluminescence: Current status* (ed. P.E. Stanley and L.J. Kricka), pp. 179–182. Wiley, Chichester, United Kingdom.

Höltke H.-J., Seibl R., Burg J., Mühlegger K., and Kessler C. 1990. Non-radioactive labeling and detection of nucleic acids. II. Optimization of the digoxigenin system. *Biol. Chem. Hoppe-Seyler* **371:** 929–938.

Hong G.F. 1982. A systemic DNA sequencing strategy. *J. Mol. Biol.* **158:** 539–549.

Hostomsky Z., Hostomska Z., and Matthews D.A. 1993. Ribonucleases H. In *Nucleases*, 2nd edition (ed. S.M. Linn et al.), pp. 341–376. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Huang X. and Miller W. 1991. A time-efficient linear-space local similarity algorithm. *Adv. Appl. Math.* **12:** 337-357.

Huang X., Hardison R.C., and Miller W. 1990. A space-efficient algorithm for local similarities. *Comput. Appl. Biosci.* **6:** 373–381.

Hudson L.G., Santon J.B., and Gill G.N. 1989. Regulation of epidermal growth factor receptor gene expression. *Mol. Endocrinol.* **3:** 400–408.

Hughey R. and Krogh A. 1995. SAM: Sequence alignment and modeling software system. In *C.I.S. Technical Report UCSC-CRL-95-7*. University of California, Santa Cruz.

———. 1996. Hidden Markov models for sequence analysis:

Extension and analysis of the basic method. *Comput. Appl. Biosci.* **12:** 95–107.

Hunkapiller M.W., Lujan E., Ostrander F., and Hood L.E. 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. *Methods Enzymol.* **91:** 227–236.

Hunter W.M. and Greenwood F.C. 1962. Preparation of iodine-131 labelled human growth hormone of high specific activity. *Nature* **194:** 495–496.

Hurn B.A. and Chantler S.M. 1980. Production of reagent antibodies. *Methods Enzymol.* **70:** 104–142.

Ikeda S. and Oda T. 1993. Nonisotopic gel-mobility shift assay using chemiluminescent detection. *BioTechniques* **14:** 878–880.

Ikemura T. 1982. Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. *J. Mol. Biol.* **158:** 573–597.

. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* **2:** 13–34.

Isaacson U. and Wettermark G. 1974. Chemiluminescence in analytical chemistry. *Anal. Chim. Acta* **68:** 339–362.

ISCO. 1982. *ISCOTABLES: A handbook of data for biological and physical scientists,* 8th edition. ISCO, Inc., Lincoln, Nebraska.

Jacobsen H., Klenow H., and Overgaard-Hansen K. 1974. The N-terminal amino-acid sequences of DNA polymerase I from *Escherichia coli* and of the large and the small fragments obtained by a limited proteolysis. *Eur. J. Biochem.* **45:** 623–627.

Jaeger J.A., Turner D.H., and Zuker M. 1989. Improved predictions of secondary structures for RNA. *Proc. Natl. Acad. Sci.* **86:** 7706–7710.

Jain V.K. and Magrath I.T. 1991. A chemiluminescent assay for quantitation of β-galactosidase in the femtogram range: Application to quantitiation of β-galactosidase in *lac* Z-transfected cells. *Anal. Biochem.* **199:** 639–650.

Janin J. 1979. Surface and inside volumes in globular proteins. *Nature* **277:** 491–492.

Jany K.-D., Lederer G., and Mayer B. 1986. Amino acid sequence of proteinase K from the mold *Tritirachium album* Limber. *FEBS Lett.* **199:** 139–144.

Jendrisak J., Young R.A., and Engel J.D. 1987. Cloning cDNA into λgt10 and λgt11. *Methods Enzymol.* **152:** 359–371.

Jerpseth B., Greener A., Short J.M., Viola J., and Kretz P.L. 1992. XL1-Blue MRF *E. coli* cells. In *Strategies Newsletter,* vol. 5, pp. 81–83. Stratagene, La Jolla, California.

Johansson S. and Skoog B. 1987. Rapid silver staining of polyacrylamide gels. *J. Biochem. Biophys. Methods* (suppl.) **14:** 33.

Johnson D.A., Gautsch J.W., Sportsman J.R., and Elder J.H. 1984. Improved technique utilizing nonfat dry milk for analysis of proteins and nucleic acids transferred to nitrocellulose. *Gene Anal. Tech.* **1:** 3–8.

Johnson F.H., Shimomura O., Saiga Y., Gershman L.C., Reynolds G.T., and Waters J.R. 1962. Quantum efficiency of *Cypridina* luminescence, with a note on that of *Aequorea. J. Cell. Comp. Physiol.* **60:** 85–103.

Johnston T.C., Thompson R.B., and Baldwin T.O. 1986. Nucleotide sequence of the *luxB* gene of *Vibrio harveyi* and the complete amino acid sequence of the β subunit of bacterial luciferase. *J. Biol. Chem.* **261:** 4805–4811.

Jollès P. 1960. Lysozyme. In *The Enzymes,* 2nd edition (ed. P.D. Boyer et al.), pp. 431–445. Academic Press, New York.

Jones D.T., Taylor W.R., and Thornton J.M. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8:** 275–282.

Jónsson Z.O., Thorbjarnardóttir S.H., Eggertsson G., and Palsdottir A. 1994. Sequence of the DNA ligase-encoding gene from *Thermus scotoductus* and conserved motifs in DNA ligases. *Gene* **151:** 177–180.

Joyce C.M. and Grindley N.D.F. 1983. Construction of a plasmid that overproduces the large proteolytic fragment (Klenow fragment) of DNA polymerase I of *Escherichia coli. Proc. Natl. Acad. Sci.* **80:** 1830–1834.

Jung Y.H. and Lee Y. 1995. RNases in ColE1 DNA metabolism. *Mol. Biol. Rep.* **22:** 195–200.

Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W., Olinger L., Grimwood J., Davis R.W., and Stephens R.S. 1999. Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis. Nat. Genet.* **21:** 385–389.

Kan N.C., Lautenberger J.A., Edgell M.H., and Hutchison C.A. 1979. The nucleotide sequence recognized by the *Escherichia coli* K12 restriction and modification enzymes. *J. Mol. Biol.* **130:** 191–209.

Kanaya S. and Ikehara M. 1995. Functions and structures of ribonuclease H enzymes. *Subcell. Biochem.* **24:** 377–422.

Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., et al. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3:** 109–136.

Karger A.E., Weiss R., and Gesteland R.F. 1993. Line scanning system for direct digital chemiluminescence imaging of DNA sequencing blots. *Anal. Chem.* **65:** 1785–1793.

Karlin S. and Altschul S.F. 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci.* **87:** 2264–2268.

———. 1993. Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc. Natl. Acad. Sci.* **90:** 5873–5877.

Karlin S. and Bucher P. 1992. Correlation analysis of amino acid usage in protein classes. *Proc. Natl. Acad. Sci.* **89:** 12165–12169.

Karn J., Brenner S., Barnett L., and Cesareni G. 1980. Novel bacteriophage λ cloning vector. *Proc. Natl. Acad. Sci.* **77:** 5172–5176.

Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y., Yamamoto S., Sekine M., Baba S., Kosugi H., Hosoyama A., et al. 1998. Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. *DNA Res.* **5:** 55–76.

Keller G.-A., Gould S., DeLuca M., and Subramani S. 1987. Firefly luciferase is targeted to peroxisomes in mammalian cells. *Proc. Natl. Acad. Sci.* **84:** 3264–3268.

Kelley W.S. and Stump K.H. 1979. A rapid procedure for isolation of large quantities of *Escherichia coli* DNA polymerase I utilizing a λpolA transducing phage. *J. Biol. Chem.* **254:** 3206–3210.

Kerkhof L. 1992. A comparison of substrates for quantifying the signal from a nonradiolabeled DNA probe. *Anal. Biochem.* **205:** 359–364.

Kessler C. 1991. The digoxigenin:anti-digoxigenin (DIG) technology — A survey on the concept and realization of a novel bioanalytical indicator system. *Mol. Cell. Probes* **5:** 161–205.

Kessler C. and Manta V. 1990. Specificity of restriction endonuclease and DNA modification methyltransferases — A review.

Gene **92:** 1–248.

Kessler C., Höltke H.-J., Seibl R., Burg J., and Mühlegger K. 1989. A novel DNA labeling and detection system based on digoxigenin: Anti-digoxigenin ELISA principle. *J. Clin. Chem. Clin. Biochem.* **27:** 130–131.

——. 1990. Non-radioactive labeling and detection of nucleic acids. I. A novel DNA labeling and detection system based on digoxigenin: Anti-digoxigenin ELISA principle (digoxigenin system). *Biol. Chem. Hoppe-Seyler* **371:** 917–927.

Kim Y., Grable J.C., Love R.J., Greene P.J., and Rosenberg J.M. 1990. Refinement of EcoRI endonuclease crystal structure: A revised chain tracing. *Science* **249:** 1307–1309.

Kinkel T., Lubbers H., Schmidt E., Molz P., and Skrzipczyk H.J. 1989. Synthesis and properties of new luminescent acridinium-9-carboxylic acid derivatives and their application in luminescence immunoassays (LIA). *J. Biolumin. Chemilumin.* **4:** 136–139.

Kirby K.S. 1956. A new method for the isolation of ribonucleic acids from mammalian tissues. *Biochem. J.* **64:** 405.

——. 1957. A new method for the isolation of deoxyribonucleic acids: Evidence on the nature of bonds between deoxyribonucleic acid and protein. *Biochem. J.* **66:** 495–504.

——. 1964. Isolation and fractionation of nucleic acids. *Prog. Nucleic Acid Res. Mol. Biol.* **3:** 1–31.

Kirchner G., Roberts J.L., Gustafson G.D., and Ingolia T.D. 1989. Active bacterial luciferase from a fused gene: Expression of a *Vibrio harveyi luxAB* translational fusion in bacteria, yeast and plant cells. *Gene* **81:** 349–354.

Klenk H.P., Clayton R.A., Tomb J.F., White O., Nelson K.E., Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D., et al. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus* (erratum *Nature* [1998] **394:** 101). *Nature* **390:** 364–370.

Klenow H. and Henningsen I. 1970. Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from *Escherichia coli* B by limited proteolysic digestion. *Proc. Natl. Acad. Sci.* **65:** 168–175.

Klimasauskas S., Kumar S., Roberts R.J., and Cheng X. 1994. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76:** 357–369.

Klug A., Jack A., Viswamitra M.A., Kennard O., Shakked Z., and Steitz T.A. 1979. A hypothesis on a specific sequence-dependent conformation of DNA and its relation to the binding of the *lac*-repressor protein. *J. Mol. Biol.* **131:** 669–680.

Knecht D.A. and Dimond R.L. 1984. Visualization of antigenic proteins on Western blots. *Anal. Biochem.* **136:** 180–184.

Kobatake E., Ikariyama Y., and Aizawa M. 1995. Production of the chimeric-binding protein, maltose-binding protein-protein A, by gene fusion. *J. Biotechnol.* **38:** 263–268.

Kohen F., Pazzagli M., Kim J.B., and Lindner H.R. 1979. An assay procedure for plasma progesterone based on antibody enhanced chemiluminescence. *FEBS Lett.* **104:** 140–143.

Kolchanov N.A., Ananko E.A., Podkolodnaya O.A., Ignatieva E.V., Stepanenko I.L., Kel-Margoulis O.V., Kel A.E., Merkulova T.I., Goryachkovskaya T.N., Busygina T.V., Kolpakov F.A., Podkolodny N.L., Naumochkin A.N., and Romashchenko A.G. 1999. Transcription Regulatory Regions Database (TRRD): Its status in 1999. *Nucleic Acids Res.* **27:** 303–306.

Koncz C., Olsson O., Langridge W.H.R., Shell J., and Szalay A.A. 1987. Expression and assembly of functional bacterial luciferase in plants. *Proc. Natl. Acad. Sci.* **84:** 131–135.

Korab-Laskowska M., Rioux P., Brossard N., Littlejohn T.G., Gray M.W., Lang B.F., and Burger G. 1998. The Organelle Genome Database Project (GOBASE). *Nucleic Acids Res.* **26:** 138–144.

Koren J.F., Melo T.B., and Prydz S. 1970. Enhanced photoemulsion sensitivity at low temperatures used in radiochromatography. *J. Chromatogr.* **46:** 129–131.

Kornberg A. and Baker T.A. 1992. Topoisomerases. In *DNA replication,* 2nd edition, pp. 379–401. W.H. Freeman, New York.

Kornberg T. and Kornberg A. 1974. Bacterial DNA polymerases. In *The enzymes,* 3rd edition (ed. P.D. Boyer), vol. 10, p. 119. Academic Press, New York.

Kotewicz M.L., Sampson C.M., D'Alessio J.M., and Gerard G.F. 1988. Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.* **16:** 265–277.

Kramer J.A., Singh G.B., and Krawetz S.A. 1996. Computer-assisted search for sites of nuclear matrix attachment. *Genomics* **33:** 305–308.

Krause A., Nicodeme P., Bornberg-Bauer E., Rehmsmeier M., and Vingron M. 1999. WWW access to the SYSTERS protein sequence cluster set. *Bioinformatics* **15:** 262–263.

Kricka L.J. 1991. Chemiluminescent and bioluminescent techniques. *Clin. Chem.* **37:** 1472–1481.

——. 1992. *Nonisotopic DNA probe techniques.* Academic Press, San Diego, California.

Kricka L.J. and DeLuca M. 1982. Effects of solvents on the catalytic activity of firefly luciferase. *Arch. Biochem. Biophys.* **217:** 674–681.

Kricka L.J. and Stanley P.E. 1992. Bioluminescence and chemiluminescence literature–1991: Part I. *J. Biolumin. Chemilumin.* **7:** 47–73.

Kricka L.J. and Thorpe G.H. 1986. Photographic detection of chemiluminescent and bioluminescent reactions. *Methods Enzymol.* **133:** 404–420.

Kricka L.J., Nozaki O., and Stanley P.E. 1993. Bioluminescence and chemiluminescence literature–1992: Part I. *J. Biolumin. Chemilumin.* **8:** 169–182..

Kroeker W.D. and Kowalski D. 1978. Gene-sized pieces produced by digestion of linear duplex DNA with mung bean nuclease. *Biochemistry* **17:** 3236–3243.

Kronvall G. 1973. A surface component in group A, C, and G streptococci with non-immune reactivity for immunoglobulin G. *J. Immunol.* **111:** 1401–1406.

Kronvall G., Grey H.M., and Williams Jr., R.C. 1970a. Protein A reactivity with mouse immunoglobulins. Structural relationship between some mouse and human immunoglobulins. *J. Immunol.* **105:** 1116–1123.

Kronvall G., Seal U.S., Finstad J., and Williams Jr., R.C. 1970b. Phylogenetic insight into evolution of mammalian Fc fragment of λG globulin using staphylococcal protein A. *J. Immunol.* **104:** 140–147.

Kulp D., Haussler D., Reese M.G., and Eeckman F.H. 1996. A generalized hidden Markov model for the recognition of human genes in DNA. *Ismb* **4:** 134–142.

Kumar S., Cheng X., Klimasauskas S., Mi S., Posfai J., Roberts R.J., and Wilson G.G. 1994. The DNA (cytosine-5) methyltransferases. *Nucleic Acids Res.* **22:** 1–10.

Kunitz M. 1950. Crystalline desoxyribonuclease. II. Digestion of thymus nucleic acid (desoxyribonucleic acid). The kinetics of the reaction. *J. Gen. Physiol.* **33:** 363–377.

Kunkel T.A., Roberts J.D., and Zakour R.A. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154:** 367–382.

Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessières P., Bolotin A., Borchert S., et al. 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390:** 249–256.

Kyte J. and Doolittle R.F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157:** 105–132.

Labarca C. and Paigen K. 1980. A simple, rapid and sensitive DNA assay procedure. *Anal. Biochem.* **102:** 344–352.

Laemmli U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680–685.

Lafontaine D.A., Deschenes P., Bussiere F., Poisson V., and Perreault J.P. 1999. The viroid and viroid-like RNA database. *Nucleic Acids Res.* **27:** 186–187.

Langone J.J. 1980. Radioiodination by use of the Bolton-Hunter and related reagents. *Methods Enzymol.* **70:** 221–247.

———. 1982. Protein A of *Staphylococcus aureus* and related immunoglobulin receptors produced by *streptococci* and *pneumonococci*. *Adv. Immunol.* **32:** 157–252.

Laskey R.A. 1980. The use of intensifying screens or organic scintillators for visualizing radioactive molecules resolved by gel electrophoresis. *Methods Enzymol.* **65:** 363–371.

Laskey R.A. and Mills A.D. 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. *Eur. J. Biochem.* **56:** 335–341.

———. 1977. Enhanced autoradiographic detection of ^{32}P and ^{125}I using intensifying screens and hypersensitized film. *FEBS Lett.* **82:** 314–316.

Laskowski Sr., M. 1980. Purification and properties of the mung bean nuclease. *Methods Enzymol.* **65:** 263–276.

Lathe R. 1985. Synthetic oligonucleotide probes deduced from amino acid sequence data: Theoretical and practical considerations. *J. Mol. Biol.* **183:** 1–12.

Latt S.A. and Wohlleb J.C. 1975. Optical studies of the interaction of 33258 Hoechst with DNA, chromatin and metaphase chromosomes. *Chromosoma* **52:** 297–316.

Lau P.P. and Gray Jr., H.B. 1979. Extracellular nucleases of *Alteromonas espejiana* BAL 31. IV. The single strand-specific deoxyriboendonuclease activity as a probe for regions of altered secondary structure in negatively and positively supercoiled closed circular DNA. *Nucleic Acids Res.* **6:** 331–357.

Lauer G., Rudd E.A., McKay D.L., Ally A., Ally D., and Backman K.C. 1991. Cloning, nucleotide sequence, and engineered expression of *Thermus thermophilus* DNA ligase, a homolog of *Escherichia coli* DNA ligase. *J. Bacteriol.* **173:** 5047–5053.

Law S.-J., Miller T., Piran U., Klukas C., Chang S., and Unger J. 1989. Novel poly-substituted aryl acridinium esters and their use in immunoassay. *J. Biolumin. Chemilumin.* **4:** 88–98.

Lawrence C.E., Altschul S.F., Boguski M.S., Liu J.S., Neuwald A.F., and Wootton J.C. 1993. Detecting subtle sequence signals: A Gibbs sampling strategy for multiple alignment. *Science* **262:** 208–214.

Leary J.J., Brigati D.J., and Ward D.C. 1983. Rapid and sensitive colorimetric method for visualizing biotin-labeled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose: Bio-blots. *Proc. Natl. Acad. Sci.* **80:** 4045–4049.

Leder P., Tiemeier D., and Enquist L. 1977. EK2 derivatives of bacteriophage lambda useful in the cloning of DNA from higher organisms: The λgt*WES* system. *Science* **196:** 175–177.

Legerski R.J., Hodnett J.L., and Gray Jr., H.B. 1978. Extracellular nucleases of *Pseudomonas* BAL 31. III. Use of the double-strand deoxyriboexonuclease activity as the basis of a convenient method for the mapping of fragments of DNA produced by cleavage with restriction enzymes. *Nucleic Acids Res.* **5:** 1445–1464.

Legocki R.P., Legocki M., Baldwin T.O., and Szalay A.A. 1986. Bioluminescence in soybean root nodules: Demonstration of a general approach to assay gene expression *in vivo* by using bacterial luciferase. *Proc. Natl. Acad. Sci.* **83:** 9080–9084.

Lehman I.R. 1981. DNA polymerase I of *Escherichia coli*. In *The enzymes*, 3rd edition (ed. P.D. Boyer), vol. 14, pp. 15–37. Academic Press, New York.

Lehrach H., Diamond D., Wozney J.M., and Boedtker H. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry* **16:** 4743–4751.

LePecq J.B. and Paoletti C. 1967. A fluorescent complex between ethidium bromide and nucleic acids. Physical-chemical characterization. *J. Mol. Biol.* **27:** 87–106.

Lerner R.A. 1984. Antibodies of predetermined specificity in biology and medicine. *Adv. Immunol.* **36:** 1–44.

Levinson A., Silver D., and Seed B. 1984. Minimal size plasmids containing an M13 origin for production of single-stranded transducing particles. *J. Mol. Appl. Genet.* **2:** 507-517.

Limbach P.A., Crain P.F., and McCloskey J.A. 1994. Summary: The modified nucleosides of RNA. *Nucleic Acids Res.* **22:** 2183–2196.

Lin J.J., Smith M., Jessee J., and Bloom F. 1992. DH11S: An *Escherichia coli* strain for preparation of single-stranded DNA from phagemid vectors. *BioTechniques* **12:** 718–721.

Lin W. and Kasamatsu H. 1983. On the electrotransfer of polypeptides from gels to nitrocellulose membranes. *Anal. Biochem.* **128:** 302–311.

Lind I., Live I., and Mansa B. 1970. Variation in staphylococcal protein A reactivity with gamma G-globulins of different species. *Acta Pathol. Microbiol. Scand. B Microbiol. Immunol.* **78:** 673–682.

Lindbladh C., Persson M., Bulow L., Stahl S., and Mosbach K. 1987. The design of a simple competitive ELISA using human proinsulin-alkaline phosphatase conjugates prepared by gene fusion. *Biochem. Biophys. Res. Commun.* **149:** 607–614.

Lindsay S. and Bird A.P. 1987. Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature* **327:** 336–338.

Lipman D.J., Altschul S.F., and Kececioglu J. 1989. A tool for multiple sequence alignment. *Proc. Natl. Acad. Sci.* **86:** 4412–4415.

Lis J.T. 1980. Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Methods Enzymol.* **65:** 347–353.

Lis J.T. and Schleif R. 1975a. Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Res.* **2:** 383–389.

———. 1975b. The regulatory region of the L-arabinose operon: Its isolation on a 1000 base-pair fragment from DNA heteroduplexes. *J. Mol. Biol.* **95:** 409–416.

Little J.W. 1981. Lambda exonuclease. *Gene Amplif. Anal.* **2:** 135–145.

Little J.W., Lehman I.R., and Kaiser A.D. 1967. An exonuclease induced by bacteriophage lambda. I. Preparation of the crystalline enzyme. *J. Biol. Chem.* **242:** 672–678.

Lomonossoff G.P., Butler P.J., and Klug A. 1981. Sequence-dependent variation in the conformation of DNA. *J. Mol. Biol.* **149:** 745–760.

Loontiens F.G., Regenfuss P., Zechel A., Dumortier L., and Clegg

R.M. 1990. Binding characteristics of Hoechst 33258 with calf thymus DNA, poly[d(A-T)], and d(CCGGAATTCCGG): Multiple stoichiometries and determination of tight binding with a wide spectrum of site affinities. *Biochemistry* **29:** 9029–9039.

Low B. 1968. Formation of merodiploids in matings with a class of Rec⁻ recipient strains of *Escherichia coli* K12. *Proc. Natl. Acad. Sci.* **60:** 160–167.

Lowe T.M. and Eddy S.R. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **22:** 955–964.

———. 1999. A computational screen for methylation guide snoRNAs in yeast. *Science* **283:** 1168–1171.

Luckow B., Renkawitz R., and Schütz G. 1987. A new method for constructing linker scanning mutants. *Nucleic Acids Res.* **15:** 417–429.

Lukashin A.V. and Borodovsky M. 1998. GeneMark.hmm: New solutions for gene finding. *Nucleic Acids Res.* **26:** 1107–1115.

Lunn G. and Sansone E.B. 1987. Ethidium bromide: Destruction and decontamination of solutions. *Anal. Biochem.* **162:** 453–458.

Maanich C. and Schneider W. 1941. Uber die Glykoside con *Digitalis orientalis* L. *Arch. Pharm.* **279:** 223–248.

MacGregor J.T. and Johnson I.J. 1977. In vitro metabolic activation of ethidium bromide and other phenanthridinium compounds: Mutagenic activity in *Salmonella typhimurium*. *Mutat. Res.* **48:** 103–107.

Mahler H.R. and Bastos R.N. 1974. Coupling between mitochondrial mutation and energy transduction. *Proc. Natl. Acad. Sci.* **71:** 2241–2245.

Main R.K. and Cole L.J. 1957. Chromatography of deoxyribonucleic acids on calcium phosphate columns. *Arch. Biochem. Biophys.* **68:** 186–195.

Main R.K., Wilkins M.J., and Cole L.J. 1959. A modified calcium phosphate for column chromatography of polynucleotides and proteins. *J. Am. Chem. Soc.* **81:** 6490–6495.

Malone R.W., Felgner P.L., and Verma I. M. 1989. Cationic liposome-mediated RNA transfection. *Proc. Natl. Acad. Sci.* **86:** 6077–6081.

Manchester K.L. 1995. Value of $A_{260/280}$ ratios for measurement of purity of nucelic acids. *BioTechniques* **19:** 208–219.

———. 1996. Use of UV methods for the measurement of protein and nucleic acid concentrations. *BioTechniques* **20:** 968–970.

Maniatis T., Jeffrey A., and Kleid D.G. 1975. Nucleotide sequence of the rightward operator of phage lambda. *Proc. Natl. Acad. Sci.* **72:** 1184–1188.

Maniatis T., Hardison R.C., Lacy E., Lauer J., O'Connell C., Quon D., Sim G.K., and Efstratiadis A. 1978. The isolation of structural genes from libraries of eucaryotic DNA. *Cell* **15:** 687–701.

Manoil C., Mekalanos J.J., and Beckwith J. 1990. Alkaline phosphatase fusions: Sensors of subcellular location. *J. Bacteriol.* **172:** 515–518.

Marinus M.G. 1973. Location of DNA methylation genes on the *Escherichia coli* K-12 genetic map. *Mol. Gen. Genet.* **127:** 47–55.

———. 1987. Methylation of DNA. In Escherichia coli *and* Salmonella typhimurium: *Cellular and molecular biology* (ed. F.C. Neidhardt et al.), vol. 1, pp. 697–702. American Society for Microbiology, Washington, D.C.

Marinus M.G. and Morris N.R. 1973. Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12. *J. Bacteriol.* **114:** 1143–1150.

Markwell M.A. 1982. A new solid-state reagent to iodinate proteins. I. Conditions for the efficient labeling of antiserum. *Anal. Biochem.* **125:** 427–432.

Marshall J., Molloy R., Moss G.W., Howe J.R., and Hughes T.E. 1995. The jellyfish green fluorescent protein: A tool for studying ion channel expression and function. *Neuron* **14:** 211–215.

Martin C., Bresnick L., Juo R.R., Voyta J.C., and Bronstein I. 1991. Improved chemiluminescent DNA sequencing. *BioTechniques* **11:** 110–113.

Martinson H.G. and Wagenaar E.B. 1974. Thermal elution chromatography and the resolution of nucleic acids on hydroxylapatite. *Anal. Biochem.* **61:** 144–154.

Masuya H., Kondo K., Aramaki Y., and Ishimori Y. 1992. Pyridopyridazine compounds and their use. European Patent No. EP0491477.

Mattingly P.G. 1991. Chemiluminescent 10-methyl-acridinium-9-(N-sulphonylcarboxamide) salts. Synthesis and kinetics of light emission. *J. Biolumin. Chemilumin.* **6:** 107–114.

Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G., Markelov M.K., and Lukyanov S.A. 1999. Fluorescent proteins from nonbioluminescent *Anthozoa* species. *Nat. Biotechnol.* **17:** 969–973.

Maxam A.M. and Gilbert W. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci.* **74:** 560–564.

May M.S. and Hattman S. 1975. Analysis of bacteriophage deoxyribonucleic acid sequences methylated by host-and R-factor-controlled enzymes. *J. Bacteriol.* **123:** 768–770.

McCann J., Choi E., Yamasaki E., and Ames B.N. 1975. Detection of carcinogens as mutagens in the *Salmonella*/microsome test: Assay of 300 chemicals. *Proc. Natl. Acad. Sci.* **72:** 2241–2245.

McCarty M. 1985. *The transforming principle: Discovering that genes are made of DNA*, 1st edition, pp. 83–85. W.W. Norton, New York.

McCaskill J.J. 1990. The equilibrium partition function and base pair binding probabilities for RNA secondary structures. *Biopolymers* **29:** 1105–1119.

McClelland M. 1981. Purification and characterization of two new modification methylases: MClaI from *Caryophanon latum* L and MTaqI from *Thermus aquaticus* YTI. *Nucleic Acids Res.* **9:** 6795–6804.

———. 1983. The frequency and distribution of methylatable DNA sequences in leguminous plant protein coding genes. *J. Mol. Evol.* **19:** 346–354.

McClelland M. and Nelson M. 1987. Enhancement of the apparent cleavage specificities of restriction endonucleases: Applications to megabase mapping of chromosomes. In *Gene amplification and analysis: Restriction endonucleases and methylases* (ed. J.G. Chirikjian), vol. 5, pp. 257–282. Elsevier, New York.

———. 1988. The effect of site-specific DNA methylation on restriction endonucleases and DNA modification methyltransferases — A review. *Gene* **74:** 291–304.

———. 1992. Effect of site-specific methylation on DNA modification methyltransferases and restriction endonucleases. *Nucleic Acids Res.* (suppl.) **20:** 2145–2157.

McClelland M., Kessler L.G., and Bittner M. 1984. Site-specific cleavage of DNA at 8- and 10-base-pair sequences. *Proc. Natl. Acad. Sci.* **81:** 983–987.

McComb R.B. and Bowers Jr., G.N. 1972. Study of optimum buffer conditions for measuring alkaline phosphatase activity in human serum. *Clin. Chem.* **18:** 97–104.

McConahey P.J. and Dixon F.J. 1966. A method of trace iodination of proteins for immunologic studies. *Int. Arch. Allergy Appl. Immunol.* **29:** 185–189.

. 1980. Radioiodination of proteins by the use of the chloramine-T method. *Methods Enzymol.* **70:** 210–213.

McCracken S. 1989. Preparation of RNA transcripts using SP6 RNA polymerase. In *DNA probes* (ed. G.H. Keller and M.M. Manak), pp. 119–120. Stockton Press, New York.

McFarland W.N. and Norris K.S. 1958. Control of pH and CO_2 by buffers in fish transport. *Calif. Fish Game* **44:** 291–310.

McGadey J. 1970. A tetrazolium method for non-specific alkaline phosphatase. *Histochemie* **23:** 180–184.

McKee J.R., Christman C.L., O'Brien Jr., W.D., and Wang S.Y. 1988. Effects of ultrasound on nucleic acid bases. *Biochemistry* **16:** 4651–4654.

McKimm-Breschkin J.L. 1990. The use of tetramethylbenzidine for solid phase immunoassays. *J. Immunol. Methods* **67:** 1–11.

Medgyesi G.A., Fust G., Gergely J., and Bazin H. 1978. Classes and subclasses of rat immunoglobulins: Interaction with the complement system and with staphylococcal protein A. *Immunochemistry* **15:** 125–129.

Meier T., Arni S., Malarkannan S., Poincelet M., and Hoessli D. 1992. Immunodetection of biotinylated lymphocyte-surface proteins by enhanced chemiluminescence: A nonradioactive method for cell-surface protein analysis. *Anal. Biochem.* **204:** 220–226.

Meissner P.S., Sisk W.P., and Berman M.L. 1987. Bacteriophage λ cloning system for the construction of directional cDNA libraries. *Proc. Natl. Acad. Sci.* **84:** 4171–4175.

Melgar E. and Goldthwait D.A. 1968. Deoxyribonucleic acid nucleases. II. The effects of metals on the mechanism of action of deoxyribonuclease I. *J. Biol. Chem.* **243:** 4409–4416.

Mellon P.L., Clegg C.H., Correll L.A., and McKnight G.S. 1989. Regulation of transcription by cyclic AMP-dependent protein kinase. *Proc. Natl. Acad. Sci.* **86:** 4487–4891.

Melton D.A., Krieg P.A., Rebagliati M.R., Maniatis T., Zinn K., and Green M.R. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12:** 7035–7056.

Menozzi F.D., Michel A., Pora H., and Miller A.O.A. 1990. Absorption method for rapid decontamination of solutions of ethidium bromide and propidium iodide. *Chromatographia* **29:** 167–169.

Merril C.R. 1987. Detection of proteins separated by electrophoresis. *Adv. Electrophor.* **1:** 111–139.

. 1990. Silver staining of proteins and DNA. *Nature* **343:** 779–780.

Merril C.R., Goldman D., and Van Keuren M.L. 1984. Gel protein stains: Silver stain. *Methods Enzymol.* **104:** 441–447.

Meselson M. and Yuan R. 1968. DNA restriction enzyme from *E. coli*. *Nature* **217:** 1110–1114.

Messing J. 1979. A multipurpose cloning system based on the single-stranded DNA bacteriophage M13. *Recomb. DNA Tech. Bull.* **2:** 43–48.

Messing J., Crea R., and Seeburg P.H. 1981. A system for shotgun sequencing. *Nucleic Acids Res.* **9:** 309–321.

Messing J.B., Gronenborn B., Müller-Hill B., and Hofschneider P.H. 1977. Filamentous coliphage M13 as a cloning vehicle: Insertion of a *Hin*dII fragment of the *lac* regulatory region in M13 replicative form *in vitro*. *Proc. Natl. Acad. Sci.* **74:** 3642–3246.

Meyer T.S. and Lamberts B.L. 1965. Use of Coomassie brilliant blue R250 for the electrophoresis of microgram quantities of parotid saliva proteins on acrylamide gel strips. *Biochim. Biophys. Acta* **107:** 144–145.

Michaelis S., Guarente L., and Beckwith J. 1983. *In vitro* construction and characterization of *phoA-lacZ* gene fusions in *Escherichia coli*. *J. Bacteriol.* **154:** 356–365.

Michelson A.M. and Orkin S.H. 1982. Characterization of the homopolymer tailing reaction catalyzed by terminal deoxynucleotidyl transferase. Implications for the cloning of cDNA. *J. Biol. Chem.* **257:** 14773–14782.

Miele E.A., Mills D.R., and Kramer F.R. 1983. Autocatalytic replication of a recombinant RNA. *J. Mol. Biol.* **171:** 281–295.

Mierendorf R.C., Percy C., and Young R.A. 1987. Gene isolation by screening λgt11 libraries with antibodies. *Methods Enzymol.* **152:** 458–469.

Miller J.H. 1972. *Experiments in molecular genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Mirouchoud D. and Gautier C. 1988. High codon-usage changes in mammalian genes. *Mol. Biol. Evol.* **5:** 192–194.

Miska W. and Geiger R. 1987. Synthesis and characterization of luciferin derivatives for use in bioluminescence enhanced enzyme immunoassays. New ultrasensitive detection systems for enzyme immunoassays, I. *J. Clin. Chem. Clin. Biochem.* **25:** 23–30.

Misteli T., Caceres J.F., and Spector D.L. 1997. The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387:** 523–527.

Mitchell L.G., Bodenteich A., and Merril C.R. 1994. Use of silver-staining to detect nucleic acids. *Methods Mol. Biol.* **31:** 197–203.

Modrich P. and Lehman I.R. 1970. Enzymatic joining of polynucleotides. IX. A simple and rapid assay of polynucleotide joining (ligase) activity by measurement of circle formation from linear deoxyadenylate-deoxythymidylate copolymer. *J. Biol. Chem.* **245:** 3626–3631.

Moeremans M., Daneels G., and De Mey J. 1985. Sensitive colloidal metal (gold or silver) staining of protein blots on nitrocellulose membranes. *Anal. Biochem.* **145:** 315–321.

Monstein H.-J., Geijer T., and Bakalkin G.Y. 1992. BLOTTO-MF, an inexpensive and reliable hybridization solution in northern blot analysis using complementary RNA probes. *BioTechniques* **13:** 842–843.

Moos M. 1992. Isolation of proteins for microsequence analysis. In *Current protocols in immunology* (ed. J.E. Coligan et al.), pp. 8.7.1–8.7.12. Wiley, New York.

Morgenstern B., Dress A., and Werner T. 1996. Multiple DNA and protein sequence alignment based on segment-to-segment comparison. *Proc. Natl. Acad. Sci.* **93:** 12098–12103.

Morise H., Shimomura O., Johnson F.H., and Winant J. 1974. Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* **13:** 2656–2662.

Morris C.E., Klement J.F., and McAllister W.T. 1986. Cloning and expression of the bacteriophage T3 RNA polymerase gene. *Gene* **41:** 193–200.

Mount D. 2001. *Bioinformatics: Sequence and genome analysis.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. (In press.)

Muilerman H.G., ter Hart H.G., and Van Dijk W. 1982. Specific detection of inactive enzyme protein after polyacrylamide gel electrophoresis by a new enzyme-immunoassay method using unspecific antiserum and partially purified active enzyme: Application to rat liver phosphodiesterase I. *Anal. Biochem.* **120:** 46–51.

Müller A., Hinrichs W., Wolf W.M., and Saenger W. 1994. Crystal structure of calcium-free proteinase K at 1.5Å resolution. *J. Biol. Chem.* **269:** 23108–23111.

Müller W. and Gautier F. 1975. Interactions of heteroaromatic compounds with nucleic acids. A-T-specific non-intercalating DNA ligands. *Eur. J. Biochem.* **54:** 385–394.

Murray N.E., Brammar W.J., and Murray K. 1977. Lambdoid phages that simplify the recovery of *in vitro* recombinants. *Mol. Gen. Genet.* **150:** 53–61.

Murzin A.G., Brenner S.E., Hubbard T., and Chothia C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247:** 536–540.

Myers R.M., Larin Z., and Maniatis T. 1985. Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science* **230:** 1242–1246.

Myhre E.B. and Kronvall G. 1977. Heterogeneity of nonimmune immunoglobulin Fc reactivity among gram-positive cocci: Description of three major types of receptors for human immunoglobulin G. *Infect. Immun.* **17:** 475–482.

———. 1980a. Immunochemical aspects of Fc-medicated binding of human IgG subclasses to group A, C and G streptococci. *Mol. Immunol.* **17:** 1563–1573.

———. 1980b. Binding of murine myeloma proteins of different Ig classes and subclasses to Fc-reactive surface structures in gram-positive cocci. *Scand. J. Immunol.* **11:** 37–46.

———. 1981. Specific binding of bovine, ovine, caprine and equine IgG subclasses to defined types of immunoglobulin receptors in gram-positive cocci. *Comp. Immunol. Microbiol. Infect. Dis.* **4:** 317–328.

Nader W.F., Edlind T.D., Huettermann A., and Sauer H.W. 1985. Cloning of *Physarum* actin sequences in an exonuclease-deficient bacterial host. *Proc. Natl. Acad. Sci.* **82:** 2698–2702.

Nagai K. and Thøgerson H.C. 1984. Generation of β-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. *Nature* **309:** 810–812.

Nakamura Y., Gojobori T., and Ikemura T. 1999. Codon usage tabulated from the international DNA sequence databases: Its status 1999. *Nucleic Acids Res.* **27:** 292.

Nakane P.K. and Pierce Jr., G.B. 1967. Enzyme-labeled antibodies for the light and electron microscopic localization of tissue antigens. *J. Cell Biol.* **33:** 307–318.

Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., et al. 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399:** 323–329.

Nelson M. and McClelland M. 1987. The effect of site-specific methylation on restriction-modification enzymes. *Nucleic Acids Res.* (suppl.) **15:** 219–230.

Nelson M. and Schildkraut I. 1987. The use of DNA methylases to alter the apparent recognition specificities of restriction endonucleases. *Methods Enzymol.* **155:** 41–48.

Nelson M., Christ C., and Schildkraut I. 1984. Alteration of apparent restriction endonuclease recognition specificities by DNA methylases. *Nucleic Acids Res.* **12:** 5165–5173.

Nelson T. and Brutlag D. 1979. Addition of homopolymers to the 3′-ends of duplex DNA with terminal transferase. *Methods Enzymol.* **68:** 41–50.

Neuwald A.F. and Green P. 1994. Detecting patterns in protein sequences. *J. Mol. Biol.* **239:** 698–712.

Nevill-Manning C.G., Wu T.D., and Brutlag D.L. 1998. Highly specific protein sequence motifs for genome analysis. *Proc. Natl. Acad. Sci.* **95:** 5865–5871.

Nguyen N., Hansen C., and Braman J. 1992. Detection of *bcl-2* t(14;18) chromosomal translocation using PCR and chemiluminescent probes. *Clin. Chem.* **38:** 469.

Nguyen V.T., Morange M., and Bensaude O. 1988. Firefly luciferase luminescence assays using scintillation counters for quantitation in transfected mammalian cells. *Anal. Biochem.* **171:** 404–408.

———. 1989. Protein denaturation during heat shock and related stress. *Escherichia coli* β-galactosidase and *Photinus pyralis* luciferase inactivation in mouse cells. *J. Biol. Chem.* **264:** 10487–10492.

Niedenthal R.K., Riles L., Johnston M., and Hegemann J.H. 1996. Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. *Yeast* **12:** 773–786.

Nilsson B. and Abrahmsén L. 1990. Fusions to staphylococcal protein A. *Methods Enzymol.* **185:** 144–161.

Nilsson B., Abrahmsén L., and Uhlén M. 1985. Immobilization and purification of enzymes with staphylococcal protein A gene fusion vectors. *EMBO J.* **4:** 1075–1080.

Nilsson R., Myhre E., Kronvall G., and Sjogren H.O. 1982. Fractionation of rat IgG subclasses and screening for IgG Fc-binding to bacteria. *Mol. Immunol.* **19:** 119–126.

Nordeen S.K. 1988. Luciferase reporter gene vectors for analysis of promoters and enhancers. *BioTechniques* **6:** 454–458.

Normore W.M. 1976. Guanine-plus-cytosine (GC) composition of the DNA of bacteria, fungi, algae and protozoa. In *Handbook of biochemistry and molecular biology: Nucleic acids*, 3rd edition (ed. G.D. Fasman), vol. 2, p. 65. CRC Press, Cleveland, Ohio.

Nossal N.G. 1974. DNA synthesis on a double-stranded DNA template by the T4 bacteriophage DNA polymerase and the T4 gene 32 DNA unwinding protein. *J. Biol. Chem.* **249:** 5668–5676.

Notredame C. and Higgins D. 1996. SAGA: Sequence alignment by genetic algorithm. *Nucleic Acids Res.* **24:** 1515–1524.

Notredame C., Holm L., and Higgins D.F. 1998. COFFEE: An objective function for multiple sequence alignments. *Bioinformatics* **14:** 407–422.

Nozaki O., Kricka L.J., and Stanley P.E. 1992. Nucleic acid hybridization assays. *J. Biolumin. Chemilumin.* **7:** 223–228.

Nygren P.A., Eliasson M., Abrahmsén L., Uhlén M., and Palmcrantz E. 1988. Analysis and use of the serum albumin binding domains of streptococcal protein G. *J. Mol. Recognit.* **1:** 69–74.

Nygren P.A., Flodby P., Andersson R., Wigzell H., and Uhlén M. 1991. *In vivo* stabilization of a human recombinant CD4 derivative by fusion to a serum-albumin-binding receptor. In *Vaccines 91: Modern approaches to new vaccines including prevention of aids* (ed. R.M. Channock et al.), pp. 363–368. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Oakley B.R., Kirsch D.R., and Morris. N.R. 1980. A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels. *Anal. Biochem.* **105:** 361–363.

Ochs D.C., McConkey E.H., and Sammons D.W. 1981. Silver stains for proteins in polyacrylamide gels: A comparison of six methods. *Electrophoresis* **2:** 304–307.

Oefner P.J., Hunicke-Smith S.P., Chiang L., Dietrich F., Mulligan J., and Davis R.W. 1996. Efficient random subcloning of DNA sheared in a recirculating point-sink flow system. *Nucleic Acids Res.* **24:** 3879–3886.

O'Farrell P.H., Kutter E., and Nakanishi M. 1980. A restriction map of the bacteriophage T4 genome. *Mol. Gen. Genet.* **179:** 421–435.

Ogata R. and Gilbert W. 1977. Contacts between the *lac* repressor and the thymines in the *lac* operator. *Proc. Natl. Acad. Sci.* **74:** 4973–4976.

Oi V.T., Glazer A.N., and Stryer L. 1982. Fluorescent phycobiliprotein conjugates for analyses of cells and molecules. *J. Cell Biol.* **93:** 981–986.

Oka T., Sakamoto S., Miyoshi K., Fuwa T., Yoda K., Yamasaki M., Tamura G., and Miyake T. 1985. Synthesis and secretion of human epidermal growth factor by *Escherichia coli*. *Proc. Natl. Acad. Sci.* **82:** 7212–7216.

Okayama H. and Berg P. 1982. High efficiency cloning of full-length cDNA. *Mol. Cell. Biol.* **2:** 161–170.

Olmsted J.B. 1981. Affinity purification of antibodies from diazotized paper blots of heterogeneous protein samples. *J. Biol. Chem.* **256:** 11955–11957.

Olsson A., Eliasson M., Guss B., Nilsson B., Hellman U., Lindberg M., and Uhlén M. 1987. Structure and evolution of the repetitive gene encoding streptococcal protein G. *Eur. J. Biochem.* **168:** 319–324.

Ornstein L. 1964. Disc electrophoresis. I. Background and theory. *Ann. N.Y. Acad. Sci.* **121:** 321–349.

Palmer B.R. and Marinus M.G. 1994. The *dam* and *dcm* strains of *Escherichia coli* — A review. *Gene* **143:** 1–12.

Panasenko S.M., Alazard R.J., and Lehman I.R. 1978. A simple, three-step procedure for the large scale purification of DNA ligase from a hybrid λ lysogen constructed in vitro. *J. Biol. Chem.* **253:** 4590–4592.

Panasenko S.M., Cameron J.R., Davis R.W., and Lehman I.R. 1977. Five hundredfold overproduction of DNA ligase after induction of a hybrid λ lysogen constructed in vitro. *Science* **196:** 188–189.

Pascarella S. and Argos P. 1992. A data bank merging related protein structures and sequences. *Protein Eng.* **5:** 121–137.

Pataki S., Meyer K., and Reichstein T. 1953. Die Konfiguration des Digoxygenins. *Experientia* **9:** 253–254.

Patterson T.A. and Dean M. 1987. Preparation of high titer lambda phage lysates. *Nucleic Acids Res.* **15:** 6298.

Pavesi A., Contero F., Bolchi A., Dieci G., and Ottonello S. 1994. Identification of new eucaryotic tRNA genes in genomic databases by a multistep weight matrix analysis of transcriptional control regions. *Nucleic Acids Res.* **22:** 1247–1256.

Pazzagli M., Bolelli G.F., Messeri G., Martinazzo G., Tomassi A., Salerno R., and Serio M. 1982. Homogeneous luminescent immunoassay for progesterone: A study on the antibody-enhanced chemiluminescence. In *Luminescent assays: Perspectives in endocrinology and clinical chemistry* (ed. M. Serio and M. Pazzagli), pp. 191–200. Raven Press, New York.

Peacock S.L., McIver C.M., and Monahan J.J. 1981. Transformation of *E. coli* using homopolymer-linked plasmid chimeras. *Biochim. Biophys. Acta* **655:** 243–250.

Pearson W.R. and Lipman D.J. 1988. Improved tools for biological sequence analysis. *Proc. Natl. Acad. Sci.* **85:** 2444–2448.

Perier R.C., Junier T., Bonnard C., and Bucher P. 1999. The Eukaryotic Promoter Database (EPD): Recent developments. *Nucleic Acids Res.* **27:** 307–309.

Perler F.B., Kumar S., and Kong H. 1996. Thermostable DNA polymerases. *Adv. Protein Chem.* **48:** 377–435.

Peterhans A., Mecklenburg M., Meussdoerffer F., and Mosbach K. 1987. A simple competitive enzyme-linked immunosorbent assay using antigen-beta-galactosidase fusions. *Anal. Biochem.* **163:** 470–475.

Pheiffer B.H. and Zimmerman S.B. 1983. Polymer-stimulated ligation: Enhanced blunt- or cohesive-end ligation of DNA or deoxyribooligonucleotides by T4 DNA ligase in polymer solutions. *Nucleic Acids Res.* **11:** 7853–7871.

Phillips H.J. 1973. Dye exclusion tests for cell viability. In *Tissue culture: Methods and applications* (ed. P.F. Kruse, Jr. and M.K. Paterson, Jr.), pp. 406–408. Academic Press, New York.

Piloty O. and Ruff O. 1897. Über die Reduktion des tertiären Nitroisobutylglycerins und das Oxim des Dioxyacetons. *Ber. Dtsch. Chem. Ges.* **30:**1656–1665.

Pluskal M.G., Przekop M.B., Kavonian M.R., Vecoli C., and Hicks D.A. 1986. Immobilon PVDF transfer membrane — A new membrane substrate for western blotting of proteins. *BioTechniques* **4:** 272–283.

Pollard-Knight D., Simmonds A.C., Schaap A.P., Akhavan H., and Brady M.A. 1990. Nonradioactive DNA detection on Southern blots by enzymatically triggered chemiluminescence. *Anal. Biochem.* **185:** 353–358.

Pouliot J.J., Yao K.C., Robertson C.A., and Nash H.A. 1999. Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. *Science* **286:** 552–555.

Poxton I.R. 1990. Immunoblotting techniques. *Curr. Opin. Immunol.* **2:** 905–909.

Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G., and Cormier M.J. 1992. Primary structure of the *Aequorea victoria* fluorescent green protein. *Gene* **111:** 229–233.

Prell B. and Vosberg H.P. 1980. Analysis of covalent complexes formed between calf thymus DNA topoisomerase and single-stranded DNA. *Eur. J. Biochem.* **108:** 389–398.

Premaratne S., Helms M., and Mower H.F. 1993. Enhancement of ethidium bromide fluorescence in double-stranded DNA reacted with chloroacetaldehyde. *BioTechniques* **15:** 394–395.

Price P.A. 1972. Characterization of Ca^{++} and Mg^{++} binding to bovine pancreatic deoxyribonuclease A. *J. Biol. Chem.* **247:** 2895–2899.

Puskas R.S., Manley N.R., Wallace D.M., and Berger S.L. 1982. Effect of ribonucleoside-vanadyl complexes on enzyme-catalyzed reactions central to recombinant deoxyribonucleic acid technology. *Biochemistry* **21:** 4602–4608.

Quandt K., Grote K., and Werner T. 1996a. GenomeInspector: Basic software tools for analysis of spatial correlations between genomic structures within megabase sequences. *Genomics* **15:** 301–304.

———. 1996b. GenomeInspector: A new approach to detect correlation patterns of elements on genomic sequences. *Comput. Appl. Biosci.* **12:** 405–413.

Quillardet P. and Hofnung M. 1988. Ethidium bromide and safety — Readers suggest alternative solutions (letter). *Trends Genet.* **4:** 89–90.

Radloff R., Bauer W., and Vinograd J. 1967. A dye-buoyant density method for the detection and isolation of closed circular duplex DNA: The closed circular DNA in HeLa cells. *Proc. Natl. Acad. Sci.* **57:** 1514–1521.

Raleigh E.A. and Wilson G. 1986. *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. *Proc. Natl. Acad. Sci.* **83:** 9070–9074.

Raleigh E.A., Murray N.E., Revel H., Blumenthal R.M., Westaway D., Reith A.D., Rigby P.W., Elhai J., and Hanahan D. 1988. McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning (erratum *Nucleic Acids Res.* [1995] **23:** 3612). *Nucleic Acids Res.* **16:** 1563–1575.

Rassoulzadegan M., Binetruy B., and Cuzin F. 1982. High fre-

quency of gene transfer after fusion between bacteria and eukaryotic cells. *Nature* **295:** 257–259.

Reese M.G. 1994. "Erkennung von Promotoren in pro- und eukaryontischen DNA-Sequenzen durch künstliche Neuronale Netze." Ph.D. thesis, German Cancer Research Center, Heidelberg, Germany.

Reese M.G. and F.H. Eeckman. 1995. New neural network algorithms for improved eukaryotic promoter site recognition. *Genome Sci. Technol.* **1:** 45.

Reese M.G., N.L. Harris, and F.H. Eeckman. 1996. Large scale sequencing specific neural networks for promoter-splice site recognition. In *Biocomputing: Proceedings of the 1996 Pacific Symposium* (ed. L. Hunter and T.E. Klein). World Scientific, Singapore.

Reese M.G., F.H. Eeckman, D. Kulp, and D. Haussler. 1997. Improved splice site detection in Genie. *J. Comput. Biol.* **4:** 311–323.

Reichstein T. 1962. Besonderheiten der Zucker von herzaktiven Glkosiden. *Angew Chem.* **74:** 887–918.

Remaut E., Stanssens P., and Fiers W. 1981. Plasmid vectors for high-efficiency expression controlled by the p_L promoter of coliphage lambda. *Gene* **15:** 81–93.

Renz M. and Kurz C. 1984. A colorimetric method for DNA hybridization. *Nucleic Acids Res.* **12:** 3435–3444.

Rice P., Craigie R., and Davies D.R. 1996. Retroviral integrases and their cousins. *Curr. Opin. Struct. Biol.* **6:** 76–83.

Richardson C.C. 1971. Polynucleotide kinase from *Escherichia coli* infected with bacteriophage T4. In *Procedures in nucleic acid research* (ed. G.L. Cantoni and D.R. Davies), vol. 2, p. 815–828. Harper and Row, New York.

. 1981. Bacteriophage T4 polynucleotide kinase. In *The enzymes*, 3rd edition (ed. P.D. Boyer), vol. 14, pp. 299–314. Academic Press, New York.

Richardson J.S. and Richardson D.C. 1989. Principles and patterns of protein conformation. In *Prediction of protein structure and the principles of protein conformation* (ed. G.D. Fasman), pp. 1–97. Plenum Press, New York.

Richman D.D., Cleveland P.H., Oxman M.N., and Johnson K.M. 1982. The binding of staphylococcal protein A by the sera of different animal species. *J. Immunol.* **128:** 2300–2305.

Richterich P. and Church G.M. 1993. DNA sequencing with direct transfer electrophoresis and nonradioactive detection. *Methods Enzymol.* **218:** 187–222.

Rigby P.W., Dieckmann M., Rhodes C., and Berg P. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* **113:** 237–251.

Robbins D., Fahimi H.D., and Cotran R.S. 1971. Fine structural cytochemical localization of peroxidase activity in rat peritoneal cells: Mononuclear cells, eosinophils and mast cells. *J. Histochem. Cytochem.* **19:** 571–575.

Roberts I.M., Jones S.L., Premier R.R., and Cox J.C. 1991. A comparison of the sensitivity and specificity of enzyme immunoassays and time-resolved fluoroimmunoassay. *J. Immunol. Methods* **143:** 49–56.

Roberts T.M., Swanberg S.L., Poteete A., Riedel G., and Backman K. 1980. A plasmid cloning vehicle allowing a positive selection for inserted fragments. *Gene* **12:** 123–127.

Robinson M.O., McCarrey J.R., and Simon M.I. 1989. Transcriptional regulatory regions of testis-specific *PGK2* defined in transgenic mice. *Proc. Natl. Acad. Sci.* **86:** 8437–8441.

Rogers S.G. and Weiss B. 1980. Exonuclease III of *Escherichia coli* K-12, an AP endonuclease. *Methods Enzymol.* **65:** 201–211.

Rohringer R. and Holden D.W. 1985. Protein blotting: Detection of proteins with colloidal gold, and of glycoproteins and lectins with biotin-conjugated and enzyme probes. *Anal. Biochem.* **144:** 118–127.

Rombauts S., Dehais P., Van Montagu M., and Rouze P. 1999. PlantCARE, a plant *cis*-acting regulatory element database. *Nucleic Acids Res.* **27:** 295–296.

Rose G.D., Geselowitz A.R., Lesser G.J., Lee R.H., and Zehfus M.N. 1985. Hydrophobicity of amino acid residues in globular proteins. *Science* **229:** 834–838.

Rosenberg S.M. 1985. *EcoK* restriction during in vitro packaging of coliphage lambda DNA. *Gene* **39:** 313–315.

Roswell D.F. and White E.H. 1978. The chemiluminescence of luminol and related hydrazides. *Methods Enzymol.* **57:** 409–423.

Roychoudhury R. and Wu R. 1980. Terminal transferase-catalyzed addition of nucleotides to the 3′-termini of DNA. *Methods Enzymol.* **65:** 43–62.

Roychoudhury R., Jay E., and Wu R. 1976. Terminal labeling and addition of homopolymer tracts to duplex DNA fragments by terminal deoxynucleotidyl transferase. *Nucleic Acids Res.* **3:** 101–116.

Rozenski J., Crain P.F., and McCloskey J.A. 1999. The RNA Modification Database: 1999 update. *Nucleic Acids Res.* **27:** 196–197.

Rüther U. and Müller-Hill B. 1983. Easy identification of cDNA clones. *EMBO J.* **2:** 1791–1794.

Salinovitch O. and Montelaro R.C. 1986. Reversible staining and peptide mapping of proteins transferred to nitrocellulose after separation by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Anal. Biochem.* **156:** 341–347.

Sambrook J., Fritsch E., and Maniatis T. 1989. *Molecular cloning: A laboratory manual*, 2nd edition, p. 18.15. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sambrook, J., Sharp P.A., and Keller W. 1972. Transcription of simian virus 40. I. Separation of the strands of SV40 DNA and hybridization of the separated strands to RNA extracted from lytically infected and transformed cells. *J. Mol. Biol.* **70:** 57–71.

Sammons D.W., Adams L.D., and Nishizawa E.E. 1981. Ultrasensitive silver-based color staining of polypeptides in polyacrylamide gels. *Electrophoresis* **2:** 135–141.

Sander C. and Schneider R. 1991. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9:** 56–68.

Sanger F., Nicklen S., and Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74:** 5463–5467.

Sano T. and Cantor C.R. 1991. A streptavidin-protein A chimera that allows one-step production of a variety of specific antibody conjugates. *Bio/Technology* **9:** 1378–1381.

Schaap A.P., Akhavan H., and Romano L.J. 1989. Chemiluminescent substrates for alkaline phosphatase: Application to ultrasensitive enzyme-linked immunoassays and DNA probes. *Clin. Chem.* **35:** 1863–1864.

Schaap A.P., Sandison M.D., and Handley R.S. 1987. Chemical and enzymatic triggering of 1,2-dioxetanes. 3. Alkaline phosphatase-catalyzed chemiluminescence from an aryl phosphate substituted dioxetane. *Tetrahedron Lett.* **28:** 1159–1162.

Schaffner W. 1980. Direct transfer of cloned genes from bacteria to mammalian cells. *Proc. Natl. Acad. Sci.* **77:** 2163–2167.

Scheffler I.E., Elson E.L., and Baldwin R.L. 1968. Helix formation by dAT oligomers. I. Hairpin and straight-chain helices. *J. Mol. Biol.* **36:** 291–304.

Schmitz A. and Galas D.J. 1979. The interaction of RNA polymerase and *lac* repressor with the *lac* control region. *Nucleic Acids Res.* **6:** 111–137.

Schmitz G.G., Walter T., Seibl R., and Kessler C. 1991. Nonradioactive labeling of oligonucleotides in vitro with the hapten digoxigenin by tailing with terminal transferase. *Anal. Biochem.* **192:** 222–231.

Schnare M.N., Damberger S.H., Gray M.W., and Gutell R.R. 1996. Comprehensive comparison of structural characteristics in eukaryotic cytoplasmic large subunit (23 S-like) ribosomal RNA. *J. Mol. Biol.* **256:** 701–719.

Schneppenheim R. and Rautenberg P. 1987. A luminescence Western blot with enhanced sensitivity for antibodies to human immunodeficiency virus. *Eur. J. Clin. Microbiol.* **6:** 49–51.

Schneppenheim R., Budde U., Dahlmann N., and Rautenberg P. 1991. Luminography — A new highly sensitive visualization method for electrophoresis. *Electrophoresis* **12:** 367–372.

Schoenle E.J., Adams L.D., and Sammons D.W. 1984. Insulin-induced rapid decrease of a major protein in fat cell plasma membranes. *J. Biol. Chem.* **259:** 12112–12116.

Schriefer L.A., Gebauer B.K., Qui L.Q., Waterston R.H., and Wilson R.K. 1990. Low pressure DNA shearing: A method for random DNA sequence analysis. *Nucleic Acids Res.* **18:** 7455–7456.

Schroeder H.R., Boguslaski R.C., Carrico R.J., and Buckler R.T. 1978. Monitoring specific binding reactions with chemiluminescence. *Methods Enzymol.* **57:** 424–445.

Schug J. and Overton G.C. 1997. TESS: Transcription element search software on the www. Computational Biology and Informatics Laboratory, University of Pennsylvania School of Medicine (Technical Report CBIL-TR-1997-1001-v0.0).

Schuler G.D., Altschul S.F., and Lipman D.J. 1991. A workbench for multiple alignment construction and analysis. *Proteins* **9:** 180–191.

Schwartz O., Virelizier J.-L., Montagnier L., and Hazan U. 1990. A microtransfection method using the luciferase-encoding reporter gene for the assay of human immunodeficiency virus LTR promoter activity. *Gene* **88:** 197–205.

Schwartz R.M. and Dayhoff M.O. 1978. Matrices for detecting distant relationships. In *Atlas of protein sequence and structure* (ed. M.O. Dayhoff et al.), vol. 5, suppl. 3, pp. 353–358. National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C.

Seevers R.R.H. and Counsell R.E. 1982. Radioiodination techniques for small organic molecules. *Chem. Rev.* **82:** 575–590.

Seibl R., Höltke H.-J., Ruger R., Meindl A., Zachau H.G., Rasshofer R., Roggendorf M., Wolf H., Arnold N., Wienberg J., et al. 1990. Non-radioactive labeling and detection of nucleic acids. III. Applications of the digoxigenin system. *Biol. Chem. Hoppe-Seyler* **371:** 939–951.

Sekiya K., Saito Y., Ikegami T., Yamamoto M., Sato Y., Maeda M., and Tsuji A. 1991. Fully-automated analyzer for chemiluminescent enzyme immunoassay. In *Bioluminescence and chemiluminescence: Current status* (eds. P.E. Stanley and L.J. Kricka), pp. 123–126. Wiley, Chichester, United Kingdom.

Semenza G. 1957. Chromatography of deoxyribonucleic acid in calcium phosphate. *Ark. Kemi* **11:** 89–96.

Setlow P. 1976. Nearest neighbor frequencies in deoxyribonucleic acids. In *Handbook of biochemistry and molecular biology: Nucleic acids*, 3rd edition (ed. G.D. Fasman), vol. 2, p. 312. CRC Press, Cleveland, Ohio.

Sevag M.G. 1934. Eine neue physikalische Enteiweißungs-methode zur Darstellung biologisch wirksamer Substanzen. *Biochem. Z.* **213:** 419–429.

Sevag M.G., Lackman D.B., and Smolens J. 1938. Isolation of components of streptococcal nucleoproteins in serologically active form. *J. Biol. Chem.* **124:** 425.

Shannon L.M., Kay E., and Lew J.Y. 1966. Peroxidase isozymes from horseradish roots. I. Isolation and physical properties. *J. Biol. Chem.* **241:** 2166–2172.

Shapiro B.A. 1988. An algorithm for comparing multiple RNA secondary structures. *Comput. Appl. Biosci.* **4:** 381–393.

Shapiro B.A. and Zhang K. 1990. Comparing multiple RNA secondary structures using tree comparison. *Comput. Appl. Biosci.* **6:** 309–318.

Shapiro H.S. 1976. Distribution of purines and pyrimidines in deoxyribonucleic acids. In *Handbook of biochemistry and molecular biology: Nucleic acids*, 3rd edition (ed. G.D. Fasman), vol. 2, p. 241. CRC Press, Cleveland, Ohio.

Sharp P.A., Sugden B., and Sambrook J. 1973. Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose–ethidium bromide electrophoresis. *Biochemistry* **12:** 3055–3063.

Sharp P.M. and Cowe E. 1991. Synonymous codon usage in *Saccharomyces cerevisiae*. *Yeast* **7:** 657–678.

Sharp P.M. and Devine K.M. 1989. Codon usage and gene expression level in *Dictyostelium discoideum*: Highly expressed genes do "prefer" optimal codons. *Nucleic Acids Res.* **17:** 5029–5039.

Shields D.C. and Sharp P.M. 1987. Synonymous codon usage in *B. subtilis* refelects both translational selection and mutational bias. *Nucleic Acids Res.* **15:** 3023–3040.

Shimomura O. 1979. Structure of the chromophore of *Aequorea* green fluorescent protein. *FEBS Lett.* **104:** 220–222.

Shine J., Dalgarno L., and Hunt J.A. 1974. Fingerprinting of eukaryotic ribosomal RNA labelled with tritiated nucleosides. *Anal. Biochem.* **59:** 360–365.

Siegel L.I. and Bresnick E. 1986. Northern hybridization analysis of RNA using diethylpyrocarbonate-treated nonfat milk. *Anal. Biochem.* **159:** 82–87.

Siezen R.J. and Leunissen J.A. 1997. Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Sci.* **6:** 501–523.

Siezen R.J., de Vos W.M., Leunissen J.A., and Dijkstra B.W. 1991. Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases. *Protein Eng.* **4:** 719–731.

Simatake H. and Rosenberg M. 1981. Purified λ regulatory protein cII positively activates promoters for lysogenic development. *Nature* **292:** 128–132.

Singer V.L., Lawler T.E., and Yue S. 1999. Comparison of SYBR® Green I nucleic acid gel stain mutagenicity and ethidium bromide mutagenicity in the *Salmonella* mammalian microsome reverse mutation assay (Ames test). *Mutat. Res.* **439:** 37–47.

Singh G.B., Kramer J.A., and Krawetz S.A. 1997. Mathematical model to predict regions of chromatin attachment to the nuclear matrix. *Nucleic Acids Res.* **25:** 1419–1425.

Singh L. and Jones K.W. 1984. The use of heparin as a simple cost-effective means of controlling background in nucleic acid hybridization procedures. *Nucleic Acids Res.* **12:** 5627–5638.

Sjöbring U., Björck L., and Kastern W. 1989. Protein G genes:

Structure and distribution of IgG-binding and albumin-binding domains. *Mol. Microbiol.* **3:** 319–327.

———. 1991. Streptococcal protein G. Gene structure and protein binding properties. *J. Biol. Chem.* **266:** 399–405.

Sjödahl J. 1977. Structural studies on the four repetitive Fc-binding regions in protein A from *Staphylooocccus aureus*. *Eur. J. Biochem.* **78:** 471–479.

Sjölander A., Stahl S., Lovgren K., Hansson M., Cavelier L., Walles A., Helmby H., Wahlin B., Morein B., Uhlén M., et al. 1993. *Plasmodium falciparum*: The immune response in rabbits to the clustered asparagine-rich protein (CARP) after immunization in Freund's adjuvant or immunostimulating complexes (ISCOMs). *Exp. Parasitol.* **76:** 134–145.

Sjöquist J., Meloun B., and Hjelm H. 1972. Protein A isolated from *Staphylococcus aureus* after digestion with lysostaphin. *Eur. J. Biochem.* **29:** 572–578.

Skinner M.K. and Griswold M.D. 1983. Fluorographic detection of radioactivity in polyacrylamide gels with 2,5-diphenyloxazole in acetic acid and its comparison with existing procedures. *Biochem. J.* **209:** 281–284.

Smith D.F. and Fisher P.A. 1984. Identification, developmental regulation, and response to heat shock of two antigenically related forms of a major nuclear envelope protein in *Drosophila* embryos: Application of an improved method for affinity purification of antibodies using polypeptides immobilized on nitrocellulose blots. *J. Cell Biol.* **99:** 20–28.

Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H., Dubois J., Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K., et al. 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: Functional analysis and comparative genomics. *J. Bacteriol.* **179:** 7135–7155.

Smith R.F. and Smith T.F. 1990. Automatic generation of primary sequence patterns from sets of related protein sequences. *Proc. Natl. Acad. Sci.* **87:** 118–122.

———. 1992. Pattern-induced multi-sequence alignment (PIMA) algorithm employing secondary structure-dependent gap penalties for comparative protein modelling. *Protein Eng.* **5:** 35–41.

Smith H.O. 1980. Recovery of DNA from gels. *Methods Enzymol.* **65:** 371–380.

Smith H.O., Annau T.M., and Chandrasegaran S. 1990. Finding sequence motifs in groups of functionally related proteins. *Proc. Natl. Acad. Sci.* **87:** 826–830.

Snyder E.E. and Stormo G. 1993. Identification of coding regions in genomic DNA sequences: An application of dynamic programming and neural networks. *Nucleic Acids Res.* **21:** 607–613.

———. 1995. Identification of coding regions in genomic DNA. *J. Mol. Biol.* **248:** 1–18.

Sonnhammer E.L., Eddy S.R., and Durbin R. 1997. Pfam: A comprehensive database of protein families based on seed alignments. *Proteins* **28:** 405–420.

Sonnhammer E.L., Eddy S.R., Birney E., Bateman A., and Durbin R. 1998. Pfam: Multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res.* **26:** 320–322.

Sørensen M.A., Kurland C.G., and Pedersen S. 1989. Codon usage determines translation rate in *Escherichia coli*. *J. Mol. Biol.* **207:** 365–377.

Southern E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:** 503–517.

Souza A.E. and Göringer H.U. 1998. The guide RNA database. *Nucleic Acids Res.* **26:** 168–169.

Spector D.L., Goldman R.D., and Leinwand L.A. 1998. *Cells: A laboratory manual*, vol. 1. *Culture and biochemical analysis of cells*, pp. 2.8–2.10. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Spingola M., Grate L., Haussler D., and Ares Jr., M. 1999. Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. *RNA* **5:** 221–234.

Stahl S., Nygren P.A., Sjölander A., and Uhlén M. 1993. Engineered bacterial receptors in immunology. *Curr. Opin. Immunol.* **5:** 272–277.

Stanley P.E. 1992. A survey of more than 90 commercially available luminometers and imaging devices for low-light measurements of chemiluminescence and bioluminescence including instruments for manual, automatic and specialized operation, for HPLC, LC, GLC, and microtiter plates. I. Descriptions. *J. Biolumin. Chemilumin.* **7:** 77–108.

———. 1993. A survey of some commercially available kits and reagents which include bioluminescence or chemiluminescence for their operation: Including immunoassays, hybridization, labels, probes, blots and ATP-based rapid microbiology. *J. Biolumin. Chemilumin.* **8:** 51–63.

Stein H. and Hausen P. 1969. Enzyme from calf thymus degrading the RNA moiety of DNA-RNA hybrids: Effect on DNA-dependent RNA polymerase. *Science* **166:** 393–395.

Stephens R.S., Kalman S., Lammel C., Fan J., Marathe R., Aravind L., Mitchell W., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., and Davis R.W. 1998. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282:** 754–759.

Sternberg N. 1985. Evidence that adenine methylation influences DNA-protein interactions in *Escherichia coli*. *J. Bacteriol.* **164:** 490–493.

Sternberg N., Smoller D., and Braden T. 1994. Three new developments in P1 cloning. Increased cloning efficiency, improved clone recovery, and a new P1 mouse library. *Genet. Anal. Tech. Appl.* **11:** 171–180.

Stokke T. and Steen H.B. 1985. Multiple binding modes for Hoechst 33258 to DNA. *J. Histochem. Cytochem.* **33:** 333–338.

Stone G.C., Sjöbring U., Björck L., Sjöquist J., Barber C.V., and Nardella F.A. 1989. The Fc binding site for streptococcal protein G is in the Cγ2-Cγ3 interface region of IgG and is related to the sites that bind staphylococcal protein A and human rheumatoid factors. *J. Immunol.* **143:** 565–570.

Stone T. and Durrant I. 1991. Enhanced chemiluminescence for the detection of membrane-bound nucleic acid sequences: Advantages of the Amersham system. *Genet. Anal. Tech. Appl.* **8:** 230-237.

Stott D.I. 1989. Immunoblotting and dot blotting. *J. Immunol. Methods* **119:** 153–187.

Strauss F. and Varshavsky A. 1984. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell* **37:** 889–901.

Studier F.W. 1973. Analysis of bacteriophage T7 early RNAs and proteins on slab gels. *J. Mol. Biol.* **79:** 237–248.

Studier F.W. and Moffatt B.A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189:** 113–130.

Stulnig T.M. and Amberger A. 1994. Exposing contaminating phenol in nucleic acid preparations. *BioTechniques* **16:** 403–404.

Subramani S. and DeLuca M. 1988. Applications of the firefly luciferase as a reporter gene. *Genet. Eng.* **10:** 75–89.

Suck D. and Oefner C. 1986. Structure of DNase I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature* **321:** 620–625.

Suck D., Lahm A., and Oefner C. 1988. Structure refined to 2Å of a nicked DNA octanucleotide complex with DNase I. *Nature* **332:** 464–468.

Suck D., Oefner C., and Kabsch W. 1984. Three-dimensional structure of bovine pancreatic DNase I at 2.5 Å resolution. *EMBO J.* **3:** 2423–2430.

Sugino A., Goodman H.M., Heyneker H.L., Shine J., Boyer H.W., and Cozzarelli N.R. 1977. Interaction of bacteriophage T4 RNA and DNA ligases in joining of duplex DNA at base-paired ends. *J. Biol. Chem.* **252:** 3987–3994.

Sühnel J. 1997. Views of RNA on the World Wide Web. *Trends Genet.* **13:** 206–207.

Sutcliffe J.G. 1978. Nucleotide sequence of the ampicillin resistance gene of *Escherichia coli* plasmid pBR322. *Proc. Natl. Acad. Sci.* **75:** 3737–3741.

———. 1979. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. *Cold Spring Harbor Symp. Quant. Biol.* **43:** 77–90.

Sutherland M.W. and Skerritt J.H. 1986. Alkali enhancement of protein staining on nitrocellulose membranes. *Electrophoresis* **7:** 401–406.

Suzuki S., Craddock B.P., Kana T., and Steigbigel R.T. 1993. Chemiluminescent enzyme-linked immunoassay for reverse transcriptase illustrated by detection of HIV reverse transcriptase. *Anal. Biochem.* **210:** 277–281.

Svaren J., Inagami S., Lovegren E., and Chalkley. 1987. DNA denatures upon drying after ethanol precipitation. *Nucleic Acids Res.* **15:** 8739–8754.

Swanstrom R. and Shank P.R. 1978. X-ray intensifying screens greatly enhance the detection by autoradiography of the radioactive isotopes ^{32}P and ^{125}I. *Anal. Biochem.* **86:** 184–192.

Switzer III, R.C., Merril C.R., and Shifrin S. 1979. A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal. Biochem.* **98:** 231–237.

Szymanski M., Barciszewska M.Z., Barciszewski J., and Erdmann V.A. 1999. 5S Ribosomal RNA Data Bank. *Nucleic Acids Res.* **27:** 158–160.

Tabor S. and Richardson C.C. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci.* **82:** 1074–1078.

Tabor S., Huber H.E., and Richardson C.C. 1987. *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem.* **262:** 16212–16223.

Tai M.S., Mudgett-Hunter M., Levinson D., Wu G.M., Haber E., Oppermann H., and Huston J.S. 1990. A bifunctional fusion protein containing Fc-binding fragment B of staphlyococcal protein A amino terminal to antidigoxygenin single-chain Fv. *Biochemistry* **30:** 8024–8030.

Takahashi M. and Uchida T. 1986. Thermophilic HB8 DNA ligase: Effects of polyethylene glycols and polyamines on blunt-end ligation of DNA. *J. Biochem.* **100:** 123-131.

Takahashi M., Yamaguchi E., and Uchida T. 1984. Thermophilic DNA ligase. Purification and properties of the enzyme from *Thermus thermophilus* HB8. *J. Biol. Chem.* **259:** 10041–10047.

Tartof K.D. and Hobbs C.A. 1987. Improved media for growing plasmid and cosmid clones. *Focus* (Life Technologies) **9:** 12.

Tatsumi H., Masuda T., and Nakao E. 1988. Synthesis of enzymatically active firefly luciferase in yeast. *Agric. Biol. Chem.* **52:** 1123–1127.

Tatusova T.A. and Madden T.L. 1999. BLAST 2 sequences, a new tool for comparing protein and nucleotide sequences (erratum *FEMS Microbiol. Lett.* [1999] **177:** 187–188). *FEMS Microbiol. Lett.* **174:** 247–250.

Tautz D. and Pfeifle C. 1989. A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98:** 81–85.

Taylor J.M., Illmensee R., and Summers J. 1976. Efficient transcription of RNA into DNA by avian sarcoma virus polymerase. *Biochim. Biophys. Acta* **442:** 324–330.

Taylor W.R. 1986. The classification of amino acid conservation. *J. Theor. Biol.* **119:** 205–218.

Thompson D. and Larson G. 1992. Western blots using stained protein gels. *BioTechniques* **12:** 656–658.

Thompson J.D., Higgins D.G., and Gibson T.J. 1994. CLUSTAL W: Improving the sensitivity of multiple sequence alignment through sequence weighting, position-specfic gap penalties, and weight matrix choice. *Nucleic Acids Res.* **22:** 4673–4680.

Thorpe G.H. and Kricka L.J. 1986. Enhanced chemiluminescent reactions catalyzed by horseradish peroxidase. *Methods Enzymol.* **133:** 331–353.

Thorpe G.H.G., Kricka L.J., Moseley S.B., and Whitehead T.P. 1985. Phenols as enhancers of the chemiluminescent horseradish peroxidase-luminol-hydrogen peroxide reaction: Application in luminescence monitored enzyme immunoassays. *Clin. Chem.* **31:** 1335–1341.

Thorstenson Y.R., Hunicke-Smith S.P., Oefner P.J., and David R. 1998. An automated, hydrodynamic process for controlled, unbiased DNA shearing. *Genome Res.* **8:** 848–855.

Tiselius, A., Hjertén S., and Levin Ö. 1956. Protein chromatography on calcium phosphate columns. *Arch. Biochem. Biophys.* **65:** 132–155.

Tizard R., Cate R.L., Ramachandran K.L., Wysk M., Voyta J.C., Murphy O.J., and Bronstein I. 1990. Imaging of DNA sequences with chemiluminescence. *Proc. Natl. Acad. Sci.* **87:** 4514–4518.

Tomb J.F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G., Fleischmann R.D., Ketchum K.A., Klenk H.P., Gill S., Dougherty B.A., et al. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori* (erratum *Nature* [1997] **389:** 412). *Nature* **388:** 539–547.

Towbin H. and Gordon J. 1984. Immunoblotting and dot immunobinding — Current status and outlook. *J. Immunol. Methods* **72:** 313–340.

Towbin H., Staehelin T., and Gordon J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci.* **76:** 4350–4354.

Triman K.L. and Adams B.J. 1997. Expansion of the 16S and 23S ribosomal RNA mutation databases (16SMDB and 23SMDB). *Nucleic Acids Res.* **25:** 188–191.

Tsunoda T. and Takagi T. 1999. Estimating transcription factor bindability on DNA. *Bioinformatics* **15:** 622–630.

Tu C.-P. and Cohen S.N. 1980. 3′-End labeling of DNA with [α-^{32}P]cordycepin-5′-triphosphate. *Gene* **10:** 177–183.

Tuma R.S., Beaudet M.P., Jin X., Jones L.J., Cheung C.Y., Yue S., and Singer V.L. 1999. Characterization of SYBR Gold nucleic acid gel stain: A dye optimized for use with 300-nm ultraviolet transilluminators. *Anal. Biochem.* **268:** 278–288.

Tung W.L. and Chow K.C. 1995. A modified medium for efficient electrotransformation of *E. coli*. *Trends Genet.* **1:** 128–129.

Uberbacher E.C. and Mural R.J. 1991. Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc. Natl. Acad. Sci.* **88:** 11261–11265.

Uhlén M. 1989. Magnetic separation of DNA. *Nature* **340:** 733–734.

Uhlén M. and Moks T. 1990. Gene fusions for purpose of expression: An introduction. *Methods Enzymol.* **185:** 129–143.

Uhlén, M., Forsberg G., Moks T., Hartmanis M., and Nilsson B. 1992. Fusion proteins in biotechnology. *Curr. Opin. Biotechnol.* **3:** 363–369.

Uhlén M., Guss B., Nilsson B., Gatenbeck S., Philipson L., and Lindberg M. 1984. Complete sequence of the staphylococcal gene encoding protein A. A gene evolved through multiple duplications. *J. Biol. Chem.* **259:** 1695–1702.

Uhlén M., Nilsson B., Guss B., Lindberg M., Gatenbeck S., and Philipson L. 1983. Gene fusion vectors based on the gene for staphylococcal protein A. *Gene* **23:** 369–378.

Uhlenbeck O. and Gumport R.I. 1982. T4 RNA ligase. In *The enzymes*, 3rd edition (ed. P.D. Boyer), vol. 15, p. 31–58. Academic Press, New York.

Urdea M.S., Warner B.D., Running J.A., Stempien M., Clyne J., and Horn T. 1988. A comparison of non-radioisotopic hybridization assay methods using fluorescent, chemiluminescent and enzyme labeled synthetic oligodeoxyribonucleotide probes. *Nucleic Acids Res* **16:** 4937–4956.

Van Dyke K. and Van Dyke R., eds. 1990. *Luminescence immunoassay and molecular applications.* CRC Press, Boca Raton, Florida.

Van Lancker M. and Gheyssens L.C. 1986. A comparison of four frequently-used assays for quantitative determination of DNA. *Anal. Lett.* **19:** 615–623.

van Oss C.J., Good R.J., and Chaudhury M.K. 1987. Mechanisms of DNA (Southern) and protein (Western) blotting on cellulose nitrate and other membranes. *J. Chromatogr.* **391:** 53–65.

van Zonneveld A.J., Curriden S.A., and Loskutoff D.J. 1988. Type I plasminogen activator inhibitor gene: Functional anlaysis and glucocorticoid regulation of its promoter. *Proc. Natl. Acad. Sci.* **85:** 5525–5529.

Vari F. and Bell K. 1996. A simplified silver diamine method for the staining of nucleic acids in polyacrylamide gels. *Electrophoresis* **17:** 20–25.

Verdine G.L. 1994. The flip side of DNA methylation. *Cell* **76:** 197–200.

Verma I.M. 1981. Reverse transcriptase. In *The enzymes*, 3rd edition (ed. P.D. Boyer), vol. 14, pp. 87–103. Academic Press, New York.

Vieira J. and Messing J. 1987. Production of single-stranded plasmid DNA. *Methods Enzymol.* **153:** 3–11.

Vogt V.M. 1973. Purification and further properties of single-strand-specific nuclease from *Aspergillus oryzae*. *Eur. J. Biochem.* **33:** 192–200.

Waalwijk C. and Flavell R.A. 1978. MspI, an isoschizomer of HpaII which cleaves both unmethylated and methylated HpaII sites. *Nucleic Acids Res.* **5:** 3231–3236.

Wada K., Wada Y., Ishibashi F., Gojobori T., and Ikemura T. 1992. Codon usage tabulated form the GenBank genetic sequence data. *Nucleic Acids Res.* **20:** 2111–2118.

Wallace J.C. and Henikoff S. 1992. PATMAT: A searching and extraction program, sequence, and block queries and databases. *Comput. Appl. Biosci.* **8:** 249–254.

Walter A.E., Turner D.H., Kim J., Lyttle M.H., Muller P., Mathews D.H., and Zuker M. 1994. Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Natl. Acad. Sci.* **91:** 9218–9222.

Wang J.C. 1969. Variation of the average rotation angle of the DNA helix and the superhelical turns of covalently closed cyclic λ DNA. *J. Mol. Biol.* **43:** 25–39.

———. 1979. Helical repeat of DNA in solution. *Proc. Natl. Acad. Sci.* **76:** 200–203.

———. 1996. DNA topoisomerases. *Annu. Rev. Biochem.* **65:** 635–692.

Wang R.Y.-H., Zhang X.-Y., Khan R., Zhou Y., Huang L.-H., and Ehrlich M. 1986. Methylated DNA-binding protein from human placenta recognizes specific methylated sites on several prokaryotic DNAs. *Nucleic Acids Res.* **14:** 9843–9860.

Wang S. and Hazelrigg T. 1994. Implications for *bcd* mRNA localization from spatial distribution of *exu* protein in *Drosophila* oogenesis. *Nature* **369:** 400–403.

Warburg O. and Christian W. 1942. Isolierung und kristallisation des gärungsferments enolase. *Biochem. Z.* **310:** 384–421.

Waring M.J. 1965. Complex formation between ethidium bromide and nucleic acids. *J. Mol. Biol.* **13:** 269–282.

———. 1966. Structural requirements for the binding of ethidium bromide to nucleic acids. *Biochim. Biophys. Acta* **114:** 234–244.

Waterman M.L., Adler S., Nelson C., Greene G.L., Evans R.M., and Rosenfeld M.G. 1988. A single domian of the estrogen receptor confers deoxyribonucleic acid binding and transcriptional activation of the rat prolactin gene. *Mol. Endocrinol.* **2:** 14–21.

Watkins T.I. and Wolfe G. 1952. Effect of changing the quaternizing group on the trypanocidal activity of dimidum bromide. *Nature* **169:** 506.

Weeks I., Beheshti I., McCapra F., Campbell A.K., and Woodhead J.S. 1983. Acridinium esters as high-specific activity labels in immunoassay. *Clin. Chem.* **29:** 1474–1479.

Wei C.-F., Alianell G.A., Bencen G.H., and Gray Jr., H.B. 1983. Isolation and comparison of two molecular species of the BAL 31 nuclease from *Alteromonas espejiana* with distinct kinetic properties. *J. Biol. Chem.* **258:** 13506–13512.

Weil M.D. and McClelland M. 1989. Enzymatic cleavage of a bacterial genome at a 10-base-pair recognition site. *Proc. Natl. Acad. Sci.* **86:** 51–55.

Weisblum B. and Haenssler E. 1974. Fluorometric properties of the bibenzimidazole derivative Hoechst 33258, a fluorescent probe specific for AT concentration in chromosomal DNA. *Chromosoma* **46:** 255–260.

Weisburger E.K., Russfield A.B., Homburger F., Weisburger J.H., Boger E., Van Dongen C.G., and Chu K.C. 1978. Testing of twenty-one environmental aromatic amines or derivatives for long-term toxicity or carcinogenicity. *J. Environ. Pathol. Toxicol.* **2:** 325–356.

Weiss B. 1976. Endonuclease II of *Escherichia coli* is exonuclease III. *J. Biol. Chem.* **251:** 1896–1901.

Weiss B., Jacquemin-Sablon A., Live T.R., Fareed G.C., and Richardson C.C. 1968. Enzymatic breakage and joining of deoxyribonucleic acid. IV. Further purification and properties of polynucleotide ligase from *Escherichia coli* infected with bacteriophage T4. *J. Biol. Chem.* **243:** 4543–4555.

Werstuck G. and Capone J.P. 1989. Identification of a domain of the herpes simplex virus *trans*-activator Vmw65 required for protein-DNA complex formation through the use of protein A fusion proteins. *J. Virol.* **63:** 5509–5513.

White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., et al. 1999. Genome sequence of the

radioresistant bacterium *Deinococcus radiodurans* R1. *Science* **286:** 1571–1577.

Whitehead T.P., Thorpe G.H.G., Carter T.J.N., Groucutt C., and Kricka L.J. 1983. Enhanced luminescence procedure for sensitive determination of peroxidase-labelled conjugates in immunoassay. *Nature* **305:** 158–159.

Whittaker P.A., Campbell A.J.B., Southern E.M., and Murray N.E. 1988. Enhanced recovery and restriction mapping of DNA fragments cloned in a new λ vector. *Nucleic Acids Res.* **16:** 6725–6736.

Wick R.A. 1989. Photon counting imaging: Applications in biomedical research. *BioTechniques* **7:** 262–268.

Wiegers U. and Hilz H. 1971. A new method using "proteinase K" to prevent mRNA degradation during isolation from HeLa cells. *Biochem. Biophys. Res. Commun.* **44:** 513–519.

. 1972. Rapid isolation of undegraded polysomal RNA without phenol. *FEBS Lett.* **23:** 77–84.

Wigle D.A., Radakovic N.N., Venance S.L., and Pang S.C. 1993. Enhanced chemiluminescence with catalyzed reporter deposition for increasing the sensitivity of Western blotting. *BioTechniques* **14:** 562–563.

Wilchek M. and Bayer E.A. 1990. Biotin-containing reagents. *Methods Enzymol.* **184:** 123–138.

Wilfinger W.W., Mackey K., and Chomczynski P. 1997. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *BioTechniques* **22:** 474–481.

Wilkie N.M. and R. Cortini. 1976. Sequence arrangement in herpes simplex virus type 1 DNA: Identification of terminal fragments in restriction endonuclease digests and evidence for inversions in redundant and unique sequences. *J. Virol.* **20:** 211–221.

Wilkinson D. 1998. Products for the chemiluminescent detection of DNA. *Scientist* **12:** 20–22.

Williams B.G. and Blattner F.R. 1979. Construction and characterization of the hybrid bacteriophage lambda Charon vectors for DNA cloning. *J. Virol.* **29:** 555–575.

Wilson A.T. 1958. Tritium and paper chromatography. *Nature* **182:** 524.

. 1960. Detection of tritium on paper chromatograms. *Biochim. Biophys. Acta* **40:** 522–526.

Wilson C.M. 1979. Studies and critique of Amido Black 10B, Coomassie Blue R, and Fast Green FCF as stains for proteins after polyacrylamide gel electrophoresis. *Anal. Biochem.* **96:** 263–278.

. 1983. Staining of proteins on gels: Comparison of dyes and procedures. *Methods Enzymol.* **91:** 236–247.

Winkler F.K. 1994. DNA totally flipped out by methylase. *Structure* **2:** 79–83.

Winkler F.K., Banner D.W., Oefner C., Tsernoglou D., Brown R.S., Heathman S.P., Bryan R.K., Martin P.D., Petratos K., and Wilson K.S. 1993. The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.* **12:** 1781–1795.

Winter E., Yamamoto F., Almoguera C., and Perucho M. 1985. A method to detect and characterize point mutations in transcribed genes: Amplification and overexpression of the mutant c-Ki-*ras* allele in human tumor cells. *Proc. Natl. Acad. Sci.* **82:** 7575–7579.

Wintersberger U. 1990. Ribonucleases H of retroviral and cellular origin. *Pharmacol. Ther.* **48:** 259–280.

Wolfenden R., Andersson L., Cullis P.M., and Southgate C.C. 1981. Affinities of amino acid side chains for solvent water. *Biochemistry* **20:** 849–855.

Woltanski K.P., Besch W., Keilacker H., Ziegler M., and Kohnert K.D. 1990. Radioiodination of peptide hormones and immunoglobulin preparations: Comparison of the chloramine T and iodogen method. *Exp. Clin. Endocrinol.* **95:** 39–46.

Wolters G., Kuijpers L., Kacaki J., and Schuurs A. 1976. Solid-phase enzyme-immunoassay for detection of hepatitis B surface antigen. *J. Clin. Pathol.* **29:** 873–879.

Wood K.V. 1991. The origin of beetle luciferases. In *Bioluminescence and chemiluminescence: Current status* (ed. P. Stanley and L. Kricka), pp. 11–14. Wiley, Chichester, United Kingdom.

Wood K.V. and DeLuca M. 1987. Photographic detection of luminescence in *Escherichia coli* containing the gene for firefly luciferase. *Anal. Biochem.* **161:** 501–507.

Wood W.B. 1966. Host specificity of DNA produced by *Escherichia coli:* Bacterial mutations affecting the restriction and modification of DNA. *J. Mol. Biol.* **16:** 118–133.

Worley K.C., Wiese B.A., and Smith R.F. 1995. BEAUTY: An enhanced BLAST-based search tool that integrates multiple biological information resources into sequence similarity search results. *Genome Res.* **5:** 173–184.

Worley K.C., Culpepper P., Wiese B.A., and Smith R.F. 1998. BEAUTY-X: Enhanced BLAST searches for DNA queries. *Bioinformatics* **14:** 890–891.

Wower J. and Zwieb C. 1999. The tmRNA database (tmRDB). *Nucleic Acids Res.* **27:** 167.

Wu J.C. and Santi D.V. 1987. Kinetic and catalytic mechanism of HhaI methyltransferase. *J. Biol. Chem.* **262:** 4778–4786.

Wu R., Jay E., and Roychoudhury R. 1976. Nucleotide sequence analysis of DNA. *Methods Cancer Res.* **12:** 87–176.

Wuchty S., Fontana W., Hofacker I.L., and Schuster P. 1999. Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers* **49:** 145–165.

Wyman A.R. and Wertman K.F. 1987. Host strains that alleviate underrepresentation of specific sequences: Overview. *Methods Enzymol.* **152:** 173–180.

Xue G.P., Denman S.E., Glassop D., Johnson J.S., Dierens L.M., Gobius K.S., and Aylward J.H. 1995. Modification of a xylanase cDNA isolated from an anaerobic fungus *Neocallimastix patriciarum* for high-level expression in *Escherichia coli*. *J. Biotechnol.* **38:** 269–277.

Yamamoto K.R., Alberts B.M., Benzinger R., Lawhorne L., and Treiber G. 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology* **40:** 734–744.

Yanisch-Perron C., Vieira J., and Messing J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33:** 103–119.

Yeh E., Skibbens R.V., Cheng J.W., Salmon E.D., and Bloom K. 1995. Spindle dynamics and cell cycle regulation of dynein in the budding yeast *Saccharomyces cerevisiae*. *J. Cell Biol.* **130:** 687–700.

Yona G., Linial N., Tishby N., and Linial M. 1998. A map of the protein space — An automatic hierarchical classification of all protein sequences. *Ismb* **6:** 212–221.

Young R.A. and Davis R. 1983a. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci.* **80:** 1194–1198.

———. 1983b. Yeast RNA polymerase II genes: Isolation with antibody probes. *Science* **222:** 778–782.

Young R.A., Bloom B.R., Grosskinsky C.M., Ivanyi J., Thomas D., and Davis R.W. 1985. Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. *Proc. Natl. Acad. Sci.* **82:** 2583–2587.

Yuen K.C.L., Johnson T.K., Denell R.E. and Consigli R.A. 1982. A silver-staining technique for detecting minute quantities of proteins on nitrocellulose paper: Retention of antigenicity of stained proteins. *Anal. Biochem.* **126:** 398–402.

Yuksel K.U. and Gracy R.W. 1985. The quantitation of proteins in silver stained polyacrylamide gels. *Electrophoresis* **6:** 361–366.

Zagursky R.J. and Berman R.L. 1984. Cloning vectors that yield high levels of single-stranded DNA for rapid DNA sequencing. *Gene* **27:** 183–191.

Zhang M.Q. 1997. Identification of protein coding regions in the human genome based on quadratic discriminant analysis. *Proc. Natl. Acad. Sci.* **94:** 565–568.

. 1998. Identification of human gene core-promoters in silico. *Genome Res.* **8:** 319–326.

Zhang S., Zubay G., and Goldman E. 1991. Low-usage codons in *Escherichia coli*, yeast, fruitfly and primates. *Gene* **105:** 61–72.

Zhang Z., Schaffer A.A., Miller W., Madden T.L, Lipman D.J., Koonin E.V., and Altschul S. 1998. Protein sequence similarity searches using patterns as seeds. *Nucleic Acids Res.* **26:** 3986–3990.

Zhu J., Liu J.S., and Lawrence C. 1997. Bayesian and adaptive alignment and inference. *Ismb* **5:** 358–368.

———. 1998. Bayesian adaptive sequence alignment algorithms. *Bioinformatics* **14:** 25–39.

Ziegler M.M. and Baldwin T.O. 1981. Biochemistry of bacterial bioluminescence. *Curr. Top. Bioenerg.* **12:** 65–113.

Zimmer R. and Verrinder Gibbins A.M. 1997. Construction and characterization of a large-fragment chicken bacterial artificial chromosome library. *Genomics* **42:** 217–226.

Zimmerman S.B. and Minton A.P. 1993. Macromolecular crowding: Biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* **22:** 27-65.

Zimmerman S.B. and Pheiffer B.H. 1983. Macromolecular crowding allows blunt-end ligation by DNA ligases from rat liver or *Escherichia coli*. *Proc. Natl. Acad. Sci.* **80:** 5852– 5856.

Zoller M.J. and Smith M. 1984. Oligonucleotide-directed mutagenesis: A simple method using two oligonucleotide primers and a single-stranded DNA template. *DNA* **3:** 479–488.

———. 1987. Oligonucleotide-directed mutagenesis: A simple method using two oligonucleotide primers and a single-stranded DNA template. *Methods Enzymol.* **154:** 329–350.

Zuker M. 1989. On finding all suboptimal foldings of an RNA molecule. *Science* **244:** 48–52.

Zwieb C. 1997. The uRNA database. *Nucleic Acids Res.* **25:** 102–103.