~ CHAPTER 17 ~

ELECTROPHORESIS

Electrophoresis is undoubtedly the most important and useful technique in molecular biology. It is used for most experiments involving either proteins or nucleic acids. **Electrophoresis is simply the movement of charged molecules or particles in an electric field**: a molecule migrates towards the electrode of polarity opposite to its net charge, i.e. positively charged **cations** move towards the **cathode**, negatively charged **anions** towards the **anode**.

The electrophoretic mobility is defined as the ratio of the migration rate (in units such as cm/s) to the electric field strength (in units such as V/cm). The mobility (in units of cm^2/s V) of a macromolecule is a function of its net charge, more specifically the ratio of its surface net charge to its accessible surface area (Section 2.1.A), counteracted by the resistance to migration exerted by the medium in which the separation takes place. More specifically, the electrophoretic mobility of a molecule, u, is proportional to its net charge q and inversely to its frictional coefficient f (Section 15.3.B.2):

$$u = \mathrm{A}\,\frac{q}{f} \tag{17.1}$$

where the proportionality constant A will be unique for each molecule.

This equation is very simple but not very useful with biological macromolecules. One uncertainty is the net charge on a complex macromolecule with numerous charges, such as a protein or nucleic acid. The net charge for electrophoresis must also take into account all the counterions in the solution that are associated to varying extents with the charged groups of the macromolecule. Polyanions like DNA and RNA in particular bind counterions very avidly. During electrophoresis, the macromolecule will tend to migrate in one direction, the loosely held counterions in the other. The relevant net charge on a protein molecule can be measured experimentally by gradually modifying chemically the amino groups on a protein and altering their net charge, until the electrophoretic mobility becomes zero. Even this knowledge, however, is not very useful. In general, the actual value of the electrophoretic mobility of a macromolecule under any one set of conditions is not very informative and is rarely measured. Instead, **electrophoresis is primarily a comparative technique, in which only relative mobilities are significant**.

The pH is the most important parameter for electrophoresis, as it controls the net charge on a molecule (Section 3.4). The electrophoretic medium needs to be buffered at a pH that produces the

appropriate direction and rate of migration. The mobilities of proteins generally increase the further the pH is from their **isoelectric points** (pI), which increases their net charge. On the other hand, the resolution of different species separated on the basis of differences in their net charge improves as the net charge is reduced and is optimal when the net charge of one of them approaches zero, at its pI, as in isoelectric focusing (Section 17.6). Nucleic acids exhibit high mobilities over a wide range of pH values, due to the high surface negative charge densities of their phosphate groups, as do proteins complexed with the anionic detergent sodium dodecyl sulfate (SDS; Section 17.4). In addition to the pH, other crucial parameters are the ionic strength, the composition of the solvent, the presence of detergents and the temperature, all of which influence the electrophoretic mobilities of analytes, their separation and the maintenance of their native conformations; for example, the stabilities of nucleic acid structures depend on the Na^+ and Mg^{2+} concentrations. Proteins tend to aggregate at low ionic strength and at their pI.

The resistance to movement due to the viscosity of the medium affects species of all sizes equally and therefore does not contribute to their separation. Indeed, all DNA molecules ranging in size from 400 to 48,500 base pairs have the same intrinsic electrophoretic mobility in solution, irrespective of their size, because they have the same density of net charge (Figure 17-1). The same is true of proteins coated uniformly with the anionic detergent SDS (Figure 17-2). Such molecules of different sizes and shapes will have different electrophoretic mobilities only if they are subjected to molecular sieving, through a polymer network, such as polyacrylamide (Section 17.1.A).

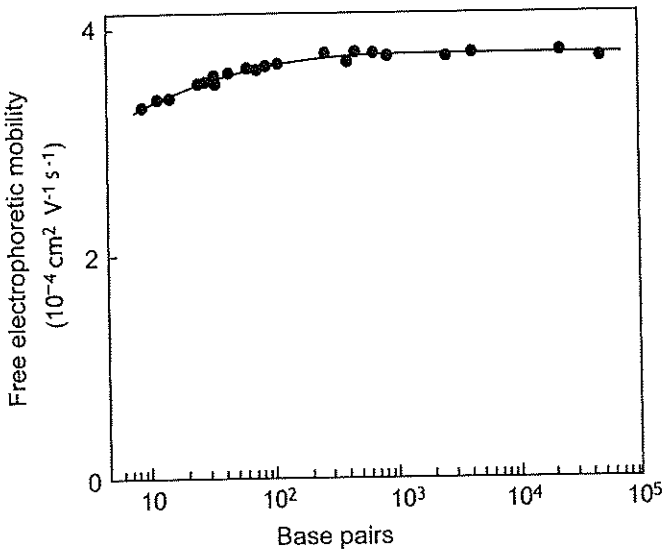

Figure 17-1. Dependence of the free electrophoretic mobility of double-stranded DNA fragments in aqueous solution on the size of the fragment. The electrophoretic mobility is expected to be independent of the size of the fragment, because of the constant ratio of net charge to mass. Why the mobility decreases slightly with fragments containing fewer than 400 base pairs is not known. Data from N. C. Stellwagen *et al.* (1997) *Biopolymers* **42**, 687–703.

Electrophoresis requires only a direct-current power supply to generate the electric field through electrodes immersed in the buffer at opposite ends of the apparatus. The most straightforward electrophoretic method is **moving-boundary** (or **free**) **electrophoresis**, which is analogous to boundary sedimentation in an ultracentrifuge (Section 16.3). A broad zone of the sample in solution is bounded by solvent alone, so that there is a boundary between the two at each end of the sample; when subjected to an electric field, the movement of the boundary is observed. Such electrophoresis is straightforward because only the sample need be present and simple solvents may be used, but large volumes and quantities of sample are required. For practical reasons, samples subjected to electrophoresis need to be protected from convective mixing. This can be accomplished within a

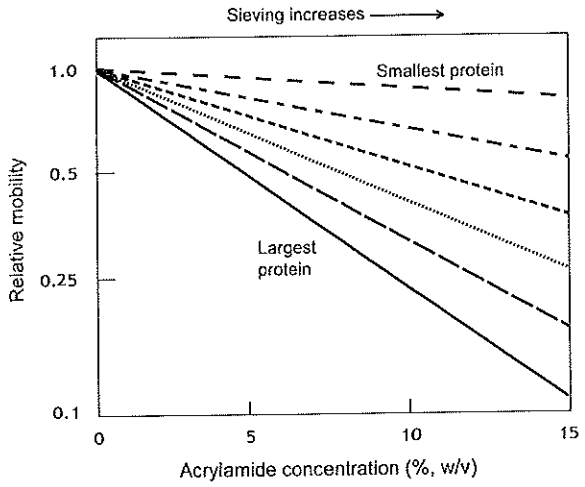

Figure 17-2. Ferguson plot of the SDS complexes of seven different proteins of varying sizes. Their electrophoretic mobilities were measured at several different acrylamide concentrations. Note that all seven extrapolate to the same mobility at zero polyacrylamide concentration, indicating that they all have the same net charge density, due to the SDS molecules bounds. Their mobilities decrease with increasing acrylamide concentration in proportion to their sizes, which is the basis for measuring the molecular weights of proteins by SDS–PAGE (Figure 17-11). The largest protein is that of the bottom curve.

liquid sample by using a density gradient (Section 16.5) or very narrow tubes of less than 0.2 mm in diameter, as in capillary zone electrophoresis (Section 17.9), but **usually electrophoresis is carried out within a gel.** There the sample generally migrates as a narrow zone, becoming separated into different 'bands' of each of the components present. In this case, the electrophoretic rate of migration depends upon the pore size and other properties of the gel. Alternatively, the sample may migrate within a 'stack' of moving boundaries of buffer components, with a mobility regulated by that of the boundary, as in disc electrophoresis or isotachophoresis (Section 17.2).

All the ions in a solution are subjected to movement in an electric field, so **high ionic strengths require large currents to move all the ions and produce substantial heating,** which may 'cook' the sample during electrophoresis. Some cooling is always necessary during electrophoresis, but this requirement can be minimized by using dilute buffers with the minimum of charged species and a low ionic strength (Equation 3.27).

Lattice models of DNA electrophoresis. A. van Heukelum & G. T. Barkema (2002) *Electrophoresis* **23**, 2562–2568.

History and principles of conductive media for standard DNA electrophoresis. J. R. Brody & S. E. Kern (2004) *Anal. Biochem.* **333**, 1–13.

Electrophoresis: the march of pennies, the march of dimes. P. G. Righetti (2005) *J. Chromat. A* **1079**, 24–40.

17.1. GEL ELECTROPHORESIS

The ability to resolve macromolecules that differ in their size, shape or conformation by electrophoresis is dramatically increased when it is carried out in gels rather than simply in a liquid solution. Convection and diffusion of a zone of macromolecules is much reduced in gels relative to solutions, because there are few pores through which the molecules can move. Within a gel, there are no free spaces like those outside molecular sieve resins (Section 18.1) and a macromolecule can move only through spaces in the gel large enough to accommodate it (Figure 18-1). The smaller molecules will find many

more pores accessible to them than will the larger molecules, so on this basis **smaller molecules tend to move faster than larger ones through gels.** The gel retards the movement of the molecules, and the degree to which the electrophoretic mobility changes with varying gel concentrations (the **retardation coefficient**) is directly related to the size and shape of the molecule (Section 17.1.C). A further advantage of gel electrophoresis is that only very small amounts of sample are required, as the molecules can be detected using a variety of staining procedures (Section 17.10).

A crucial advantage of gel electrophoresis is the ability to vary continuously the gel concentration, to achieve any effective pore size. The average pore size in a gel can range from a few Ångstroms, typical of small molecules, up to 500–1000 Å (50–100 nm), corresponding to the diameters of plant viruses; the effective limit is nearly a micron (1000 nm). Moreover, gels can be prepared easily with a gradient of gel concentrations and a corresponding continuum of pore sizes. Gel electrophoresis is able to resolve large numbers of components in a sample: as many as 50 in one-dimensional gels (1-D) and up to several thousand in two-dimensional (2-D) gel electrophoresis (Section 17.7). Numerous samples can be subjected to electrophoresis simultaneously on multiple channels in the same gel slab (Figure 17-3), which is crucial for comparing the mobilities of different bands.

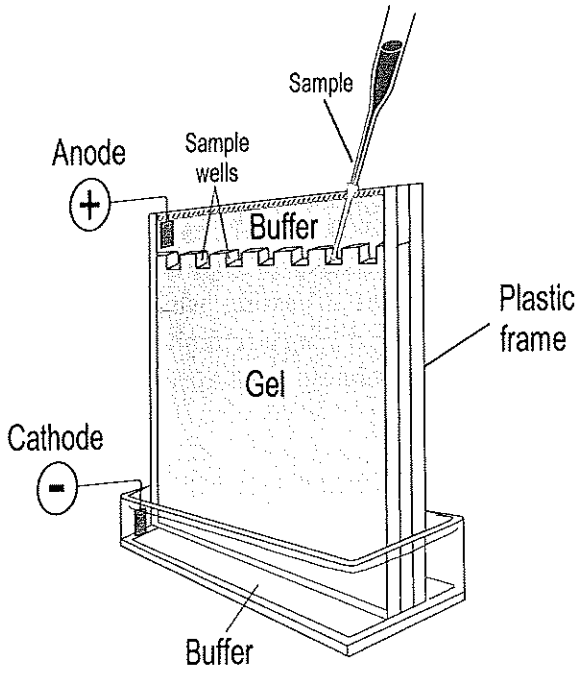

Figure 17-3. Apparatus for polyacrylamide slab gel electrophoresis of basic proteins. Samples are applied in small amounts to the top of the gel, within the various sample wells and under the buffer. Electrophoresis then causes the samples to migrate downwards through the gel. After the electrophoretic separation is complete, the gel is stained or analyzed for the presence of the various components of interest.

When the charge densities of species are equal, as in the case of nucleic acids or with proteins coated with an ionic detergent like SDS (Section 17.4), the electrophoretic mobility in solution is independent of the sizes of the molecules, but **sieving by the gel can cause the rate of migration to depend primarily upon the particle size.** This can provide an estimate of the molecular weight of the molecule. Additional and more reliable information concerning the properties of each species in a sample can be derived from gel electrophoresis at multiple gel concentrations (Section 17.1.C). To be fully effective, gel electrophoresis needs to be conducted with optimized gel concentrations and buffers, which are usually determined empirically.

Two types of gels are most commonly used for electrophoresis: agarose and polyacrylamide.

Agarose and polyacrylamide gel electrophoresis. A. M. Guilliatt (2002) *Methods Mol. Biol.* **187**, 1–11.

Characterization of the proteasome using native gel electrophoresis. S. Elsassr *et al.* (2005) *Methods Enzymol.* **398**, 353–363.

Electrophoresis of a polyelectrolyte through a nanopore. S. Ghosal (2006) *Phys. Rev. E* **74**, 041901.

17.1.A. Polyacrylamide: Free-radical Polymerization

Polyacrylamide is the building material for gel electrophoresis with the widest range of pore sizes. It is most suitable for resolving by electrophoresis, on the basis of their size, shape and conformation, species with molecular weights ranging from a few hundred Daltons to those of large particles the size of cell organelles. Polyacrylamide gel electrophoresis is commonly referred to as PAGE.

Polyacrylamide is produced by polymerizing the monomer acrylamide by a complex reaction that is catalyzed by free radicals (Figure 17-4). Concentrations of acrylamide that will produce practical gels are in the range of 3–50% (w/v), which is often designated as the %T of the gel. **The greater the concentration of acrylamide, the smaller the pores and the greater the sieving effect in decreasing the rate of movement** (Figure 17-5-A). To obtain a reproducible polymer, and a high degree of conversion of monomers to polymer, a considerable number of parameters need to be controlled. The monomers must be pure, without acrylic acid and polymers of acrylamide. The free radical donors that initiate the polymerization must be pure and at the appropriate concentrations. Any substance that reacts with free radicals will inhibit the polymerization process. Most important are O_2 and thiol or other reducing reagents: dissolved oxygen should be kept at a minimum and constant in order to produce reproducible gels; reducing reagents must not be added to the polymerization mixture. Other parameters to be controlled are the temperature, pH, concentrations of divalent metals and reaction time, which roughly determine the degree of polymerization.

Many different species can be used to initiate polymerization, but most common are (1) **riboflavin,** which decomposes in light and produces radicals, (2) tetramethylethylenediamine (**TEMED**), which is a free radical stabilizer, and (3) **persulfate**, which decomposes chemically to free radicals:

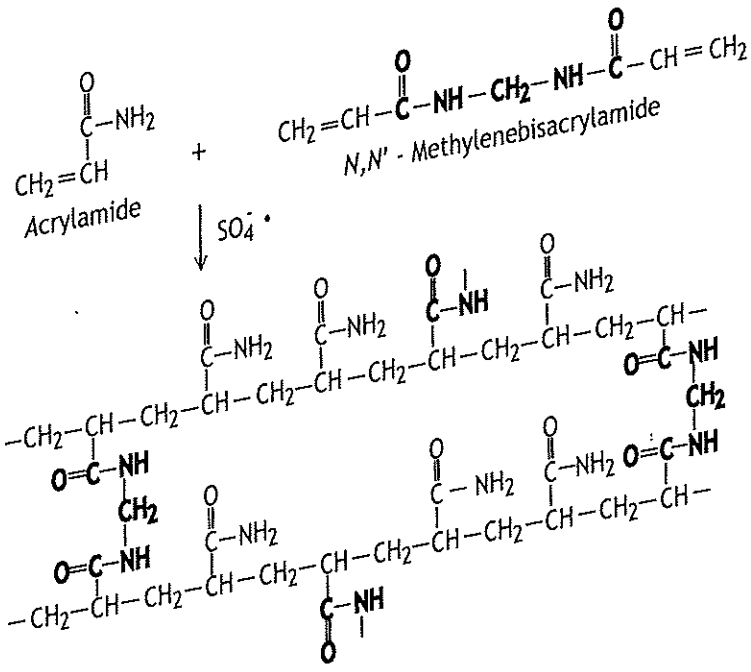

Figure 17-4. Polymerization of acrylamide and methylenebisacrylamide to form a cross-linked polyacrylamide gel. Polymerization is initiated by free radicals, in this case SO_4^{-}• from the chemical decomposition of ammonium persulfate. The cross-links introduced by the methylenebisacrylamide are indicated in bold.

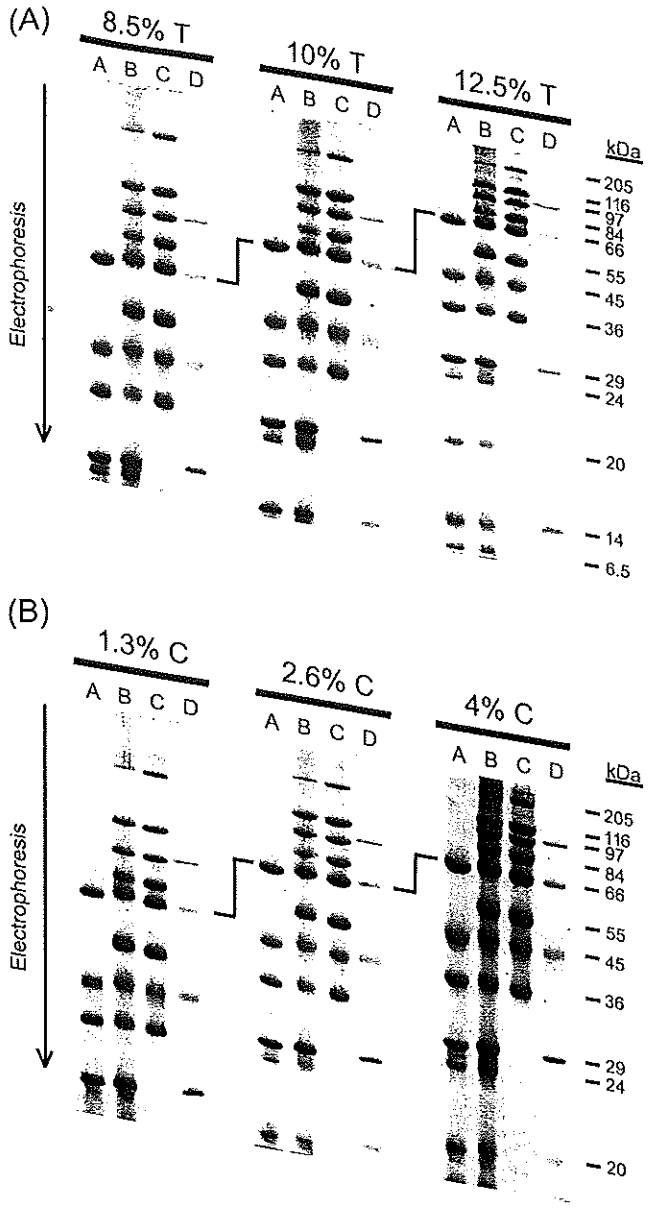

Figure 17-5. SDS–PAGE patterns of four mixtures of standard proteins (A–D), at varying concentrations of (A) polyacrylamide (%T) and (B) the cross-linker bisacrylamide (%C). The mobilities decrease with increasing gel concentrations, and the albumin bands in each are connected by the lines between gels. The molecular weights of the proteins in the gel on the *right* are indicated. The buffer system was that described in Figure 17-6. The bisacrylamide concentration in (A) was 2.6%C. The polyacrylamide gel concentration in (B) was 10%T. Data from G. S. Makowski & M. L. Ramsby (1997) in *Protein Structure: a practical approach* (T. E. Creighton, ed.), IRL Press, Oxford, pp. 15 and 18.

$$S_2O_8^{2-} \rightarrow 2\,SO_4^{-}\bullet \qquad (17.2)$$

Persulfate and TEMED together are effective in the alkaline pH range, but acidic gels are more effectively polymerized in the presence of all three catalysts combined, in appropriate ratios that should be determined experimentally. Photopolymerization by riboflavin in the presence of a small amount of persulfate is advantageous even in the alkaline pH range: it will produce gelation within about 10 min for analytical-scale gels and permits precise control of the time at which polymerization is initiated by when the illumination is initiated.

Polyacrylamide that is freshly prepared contains oxidizing compounds, including some types of peroxides. Some, but not all, can be removed by pre-electrophoresis of the gel before adding the

sample, especially if charged reducing agents, such as thioglycolate ($HS–CH_2–CO_2^-$), are loaded prior to or with the sample. Any such pre-electrophoresis is possible only with homogeneous, continuous buffer systems.

Polyacrylamide adheres well to glass walls, and the gels can be oriented vertically (Figure 17-3). The sample is applied in a small volume on top of the gel and beneath the buffer of the electrode compartment. In horizontally oriented apparatus, polyacrylamide can be firmly bonded to thin plastic sheets; such thin gels can be used at high voltages without heating, which enhances the resolution.

Normally, polyacrylamide is cross-linked by including a small amount of N,N′-bisacrylamide (commonly abbreviated to Bis) in the polymerization mixture (Figure 17-4). **The amount of cross-linker added is commonly expressed as a percentage of the acrylamide monomer, %C.** Small extents of cross-linking decrease the sizes of the pores, and the rate of migration of macromolecules is retarded progressively as the fraction of Bis is increased from 0 to 5% of the total acrylamide monomers (Figure 17-5-B); beyond that, increased cross-linking diminishes the retardation and the effect becomes negligible when the concentration of Bis approaches that of the acrylamide.

The pore structures of cross-linked vinyl polymers like polyacrylamide can be varied greatly by using a large number of alternative cross-linking agents or initiator conditions. Inclusion in the polymerization mixture of hydrophilic polymers, such as **polyethylene glycol** (Section 9.1.A.3), causes the polyacrylamide gel fibers to aggregate laterally and produce large pores. Furthermore, **the retardation of molecules in a polyacrylamide gel varies with the average length of the polyacrylamide chains.** The chain length and effective pore size are inversely related, and the average chain length can be controlled through the polymerization conditions; this is one reason why casting reproducible gels requires careful control of the polymerization conditions. The effective **pore size** of polyacrylamide is determined by the ratio of the diameter of one covalent polyacrylamide particle to the average distance in three dimensions between polymer chains. If this ratio is too small, or if it exceeds unity, molecular sieving decreases, even ceasing. Decreasing the average chain length by using excessive initiator concentrations or the presence of inhibitors (especially oxygen) during polymerization, causes the effectiveness of sieving to decrease. The distribution of cross-links is also important for molecular sieving. Cross-linking agents that are incorporated into the growing polymer chain more slowly than the monomer tend to generate an uneven distribution of cross-links ('nodules'), which increases the effective pore size.

Polyacrylamide chains without cross-links form an insoluble gel at monomer concentrations greater than 10% (w/v), whereas those with cross-links gel at monomer concentrations as low as 3%. Polyacrylamide without cross-links produces very homogeneous networks but it is less effective than cross-linked polyacrylamide in the molecular sieving of proteins. It is able, however, to separate some DNA conformations that cross-linked polyacrylamide cannot.

Protein diffusion in charged polyacrylamide gels. Visualization and analysis. R. K. Lewus & G. Carta (1999) *J. Chromat. A* **865**, 155–168.

On the characteristics of migration of oligomeric DNA in polyacrylamide gels and in free solution. U. Mohanty & L. McLaughlin (2001) *Ann. Rev. Phys. Chem.* **52**, 93–106.

Using *in situ* rheology to characterize the microstructure in photopolymerized polyacrylamide gels for DNA electrophoresis. J. Wang & V. M. Ugaz (2006) *Electrophoresis* **27**, 3349–3358.

produces the bulk flow of liquid in that direction. Although EEO can enhance electrophoretic resolution, it usually has the opposite effect and is suppressed by removing or blocking the stationary charged groups.

Electroendosmosis correction for electrophoretic mobility determined in gels. S. Ghosh & D. B. Moss (1974) *Anal. Biochem.* **62**, 365–370.

The effects of electroendosmosis in agarose electrophoresis. Y. Guo *et al.* (1998) *Electrophoresis* **19**, 1311–1313.

Fluid mechanics of electroosmotic flow and its effect on band broadening in capillary electrophoresis. S. Ghosal (2004) *Electrophoresis* **25**, 214–228.

17.2. DISC ELECTROPHORESIS: BUFFER DISCONTINUITIES

The preceding discussion of electrophoresis has assumed that the medium in which it is taking place is homogeneous. In contrast, gel electrophoresis in **discontinuous** buffer systems (known as **disc electrophoresis**) provides much better resolution than is usually possible with systems in which the buffer is uniform throughout, because **the macromolecules can be concentrated into very thin zones at the start of the electrophoresis.** At least two different electrophoresis buffers are used, separated by a discontinuity. They differ in either a cation or anion, while the counterion is the same; one of the differing ions has an electrophoretic mobility less than the species of interest to be separated, and is known as the **trailing ion**; the other, the **leading ion**, has a greater mobility. The discontinuous buffer system most commonly used is appropriate for acidic molecules and has glycine as the trailing anion and Cl^- as the leading ion; Tris is the common counterion (Figure 17-6). The leading ion would tend to move more rapidly than the trailing ion, but this is impossible because of the need to maintain electroneutrality and the conservation of mass, which produces an increased field strength behind the fastest migrating (leading) ion in the electric field (Figure 17-7). The gradient in field strength that develops behind the leading ion produces similarly oriented gradients of heat and pH, plus an opposite gradient of specific conductance (Figure 17-7). These gradients accelerate the trailing ions until they have attained the net mobility of the leading ion.

At steady-state, a macromolecule within such a system will migrate within a moving boundary between the leading and trailing ions, at a field strength (conductance) and concentration that are regulated by the mobilities and concentrations of the leading, trailing and common buffer ions that make up the boundary. All trailing species with net mobilities equal to, or greater than, that of the trailing ion will have the same regulated mobility; consequently, the system is said to be 'equal-mobility-electrophoretic' or **isotachophoretic** (Section 17.2.B). Multiple species migrating isotachophoretically in order of their intrinsic mobilities is known as a 'stack'.

Regulation of the trailing buffer phase by the leading ion causes a macromolecular ion present in that phase to migrate within the stack at a very high concentration, in the order of 100 mg/ml for a 100-kDa protein; this is independent of how diluted the macromolecule was in the original sample. Therefore, **very dilute samples can produce very condensed zones of sample within the gel, and high resolution.** This is the greatest advantage of disc electrophoresis.

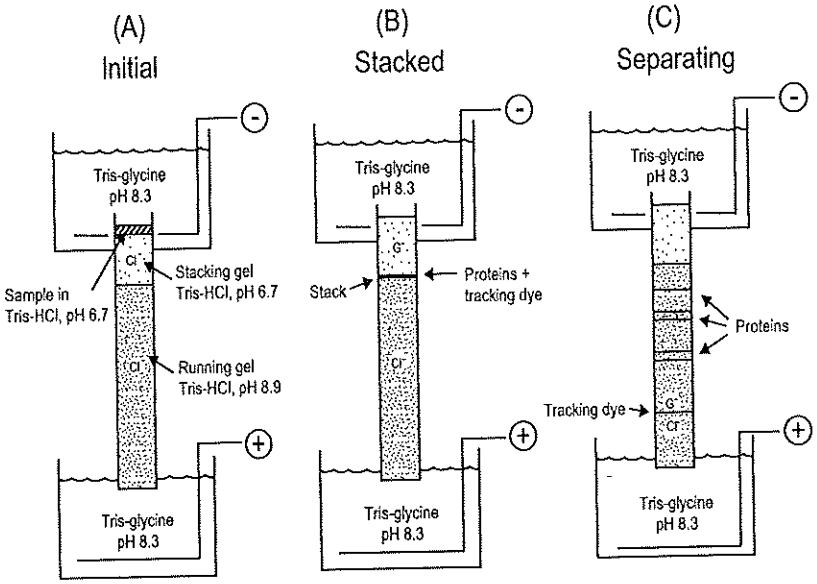

Figure 17-6. Gel electrophoresis for anions in a discontinuous buffer system, illustrated with the system used most commonly. (A) The initial gel electrophoresis system of four phases. (1) A cathode buffer containing the trailing ion, T (glycine, G^-), as a salt with any counterion, X (Tris in this case), plus a tracking dye; (2) a stacking gel of large pore size that does not retard migration of macromolecular particles and contains the leading ion, L (Cl^-), at a pH lower than that of the resolving gel; (3) a resolving gel at a pH higher than that of the stacking gel and with a gel concentration that decreases the mobilities of macromolecules because of molecular sieving; and (4) an anode buffer. The sample is layered on the stacking gel. (B) Upon initiating electrophoresis, the macromolecules in the sample are concentrated into a 'stack' at the boundary between the leading and trailing ions in the stacking gel. (C) Upon arrival at the resolving gel, the tracking dye and the boundary between the leading and trailing ions proceed as a stack at a velocity increased by the higher pH and not diminished by sieving. In contrast, the velocities of the macromolecules in the sample are decreased due to molecular sieving and they migrate in order of their net mobilities.

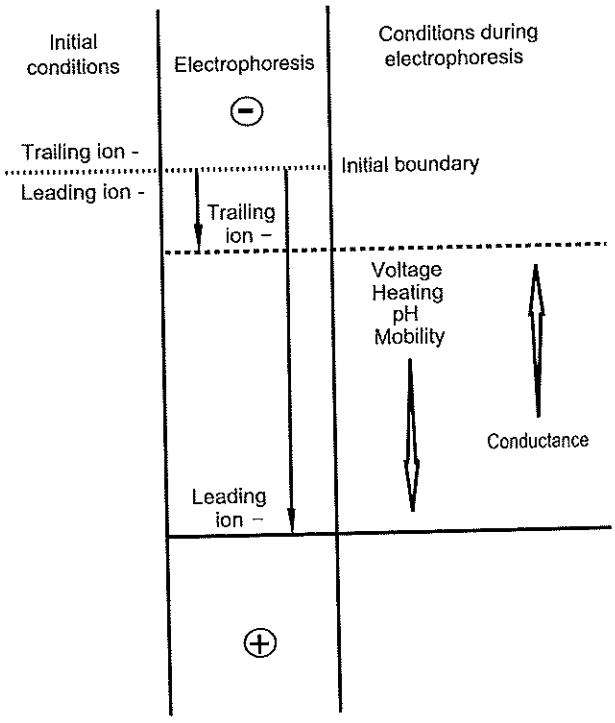

Figure 17-7. Discontinuous buffer systems. Prior to electrophoresis, a boundary (*dotted line*) separates the trailing ion (T) in the cathode buffer from the leading ion (L) in the gel buffer. As T and L electrophorese out of the starting zone, the greater instantaneous velocity of L would tend to create behind it an 'ion vacuum', which would be in violation of the laws of conservation of mass and maintenance of electroneutrality. These two laws produce a 'regulation' in the zone behind L, which leads to an increase in field strength, heating and pH, and a decrease in conductance. These, in turn, produce an increased mobility of all species migrating between T and L.

The concentrating (stacking) buffer phase is contained within a **stacking gel** (Figure 17-8) with large pores that does not sieve but serves only to prevent convection. Contiguous with this gel is a more concentrated gel, known as the **resolving gel**, that retards, or sieves, the migrating macromolecules. As the stack of moving boundaries containing the concentrated sample enters the resolving gel, the buffer boundaries are unaffected by the elevated gel concentration and altered pH, while the migrations of the macromolecules are altered by both. The increased gel concentration decreases the electrophoretic mobilities of the macromolecules to less than that of the trailing ion. Consequently, they 'unstack' and separate on the basis of their new mobilities. The stack of buffer ions remains (Figure 17-6) and produces a moving boundary 'front' that can be marked by an appropriate **tracking dye** (Section 17.2.A). The electrophoretic mobility of a macromolecule is usually measured relative to that of the tracking dye, which is known as the **relative mobility**, R_f.

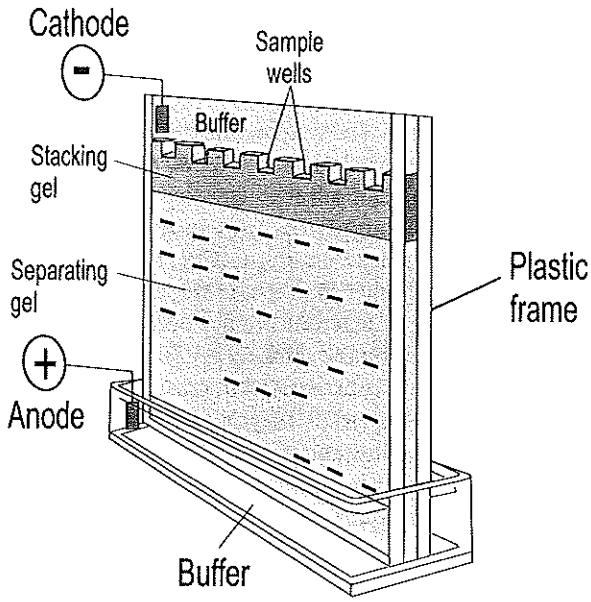

Figure 17-8. A disc electrophoresis apparatus. The three zones of buffer, stacking gel and separating gel have different buffers present, which produce the stacking and separation phenomena that are characteristic of disc electrophoresis (Figures 17-7 and 17-6). The samples are applied to the sample wells, under the buffer; they need not have small volumes, because they are concentrated as they 'stack' in the stacking gel. Otherwise, the operation is comparable to that in a normal polyacrylamide gel (Figure 17-3). Bands of some of the solutes being separated are indicated in the separating gel.

The greatest advantage of disc electrophoresis is that it is largely independent of the volume of the original sample, as the stacking process concentrates the sample into very thin bands. With a continuous gel system, the width of an electrophoretic band cannot be less than that of the sample applied to the gel, and it subsequently increases due to diffusion. The stacking process with disc electrophoresis avoids the necessity to concentrate a sample prior to its electrophoresis, which is important for most applications in which only dilute samples are available. The stacking gel can also assist separation by excluding species with very high or low mobilities from the stack.

Discontinuous buffer systems are available from a computer program at http://www.buffers.nichd.nih.gov for electrophoresis at widely varying pH values and for both negatively and positively charged macromolecules. Most discontinuous electrophoresis experiments, however, use the buffer system described in Figure 17-6.

Moving boundaries in sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Z. Buzas & A. Crambach (2004) *Electrophoresis* **25**, 970–972.

Discontinuous buffer systems operative at pH 2.5–11.0, 0°C and 25°C, available on the Internet. D. Wheeler *et al.* (2004) *Electrophoresis* **25**, 973–974.

Simul 5: free dynamic simulator of electrophoresis. V. Hruska *et al.* (2006) *Electrophoresis* **27**, 984–991.

17.2.A. Keeping Track: Tracking Dyes

Tracking dyes migrate in disc electrophoresis with a mobility that is intermediate between those of the leading and trailing ions; consequently, they accumulate at, and mark, the moving boundary. Being colored and visible, they are used to observe the progress of electrophoresis. The mobilities of electrophoretic bands are often expressed relative to that of the tracking dye. Tracking dyes cannot bind to any of the materials present in a sample, in contrast to those dyes that bind to and stain macromolecules (Section 17.10).

The most commonly used tracking dye is **bromphenol blue**; it is used generally in SDS–PAGE (Section 17.4) to mark the chloride/glycine moving boundary. It may not do so accurately, however, if it is displaced by the micelles of SDS that are present in a typical moderate- or low-concentration polyacrylamide gel and migrate more rapidly. Bromphenol blue can occupy and correctly mark the ion boundary only after molecular sieving effects start to retard micelles of SDS, which generally requires at least a 12% T (w/v) polyacrylamide gel.

The dye **pyroninY** complexed with SDS has a mobility greater than that of micellar SDS and marks the boundary correctly at low polyacrylamide concentrations. It has a large size when bound to SDS, however, and is retarded by molecular sieving at gel concentrations greater than about 12% T, when it migrates behind the moving boundary front. Still greater gel concentrations retard even a small dye like bromphenol blue by sieving, and it migrates behind the moving boundary. Alternatively, the position of the boundary can be determined by precipitating the leading or trailing ions or by using one in a radioactive form. A moving boundary front indicated by a tracking dye is frequently used to measure the relative mobility (R_f) of a band, so that R_f is strictly valid only in the gel concentration range in which the tracking dye migrates with the moving boundary front. Otherwise, it is still a useful reference for that one set of conditions.

Gel electrophoresis in a continuous buffer has no buffer discontinuities and no stacking, so any dye can be used as a reference. In the absence of stacking, however, the band of a relatively small dye molecule spreads rapidly by diffusion, so its usefulness decreases with the time of electrophoresis.

17.2.B. Isotachophoresis: Stacking

Isotachophoresis involves the 'stacking' phenomenon that takes place in discontinuous buffer systems. Sequential moving boundaries of all the components present are produced, generating a stack, and they all migrate electrophoretically in the electric field at the same velocity (Figure 17-9). The name comes from the Greek: *iso* = equal, *tachos* = speed, *phoresis* = migration. Such a system is the same in principle as the stacking phase of disc electrophoresis described above, but it can be used in isolation, especially with capillary electrophoresis (Section 17.9). Isotachophoresis is essentially the same as isoelectric focusing (Section 17.6) when the concentration of the common counterion approaches zero.

Detection of conformational transformation of antithrombin in blood with crossed immunoelectrophoresis: new application for a classical method. J. Corral *et al.* (2003) *J. Lab. Clin. Med.* **142**, 298–305.

17.4. SODIUM DODECYL SULFATE–POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS–PAGE)

Electrophoresis of fully denatured and dissociated polypeptide chains derived from proteins by binding of the negatively charged detergent sodium dodecyl sulfate (SDS):

$$CH_3 - (CH_2)_{11} - SO_4^- Na^+ \qquad (17.4)$$

is usually designated as **SDS–PAGE**. It is by far the most popular mode of polyacrylamide gel electrophoresis, in part because it provides exceptional resolving capacity when conducted as disc electrophoresis (Figure 17-5) but also because the mobilities of complexes of SDS and linear polypeptide chains are observed to be inversely proportional to the lengths of the polypeptide chains (Figure 17-2); **under ideal conditions, SDS–PAGE can give an accurate estimate of the molecular weight of an unknown polypeptide chain by comparing its mobility with those of standard proteins** (Figure 17-11).

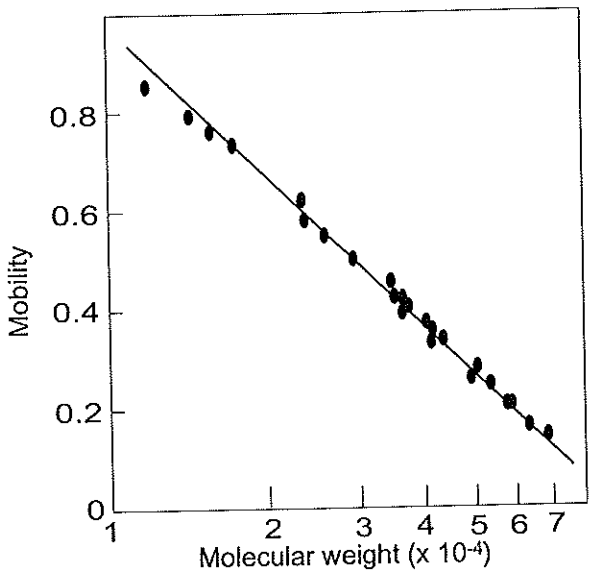

Figure 17-11. Typical correlation between the mobility in SDS–PAGE and molecular weight of the polypeptide chain. Proteins of known molecular weight within the range 14,000–70,000 were subjected to electrophoresis under conditions where the linear relationship applied. Data from K. Weber & M. Osborne.

SDS unfolds proteins and binds to them at a mass ratio of 1.4:1, which corresponds to roughly one detergent molecule per two amino acid residues. The negative charges of the SDS molecules usually overwhelm the intrinsic charge on the protein and cause the charge of the SDS–protein complex to be determined primarily by the bound SDS. In this case, the free electrophoretic mobilities of all protein–SDS complexes should be identical, and consequently their Ferguson plots intersect at zero polyacrylamide concentration (Figure 17-2). **The dependence of mobility on the molecular weight of the polypeptide chain is due entirely to the sieving effect of the polyacrylamide.** This requires that all the protein–SDS complexes have the same shape so that their sizes are determined solely by the lengths of their polypeptide chains; this requires that all disulfide bonds be broken (this

is usually accomplished by boiling the sample in SDS and a reducing agent like mercaptoethanol or dithiothreitol) and there be no other covalent cross-links. Molecular weights measured for proteins in this way are valid only if all the above conditions apply.

Users of SDS–PAGE should be aware of the possible complications. The relative mobilities are usually not constant at pH values lower than 7. Unfolding proteins with SDS produces polypeptides of uniform random-coil conformations only if the reaction conditions are sufficiently severe, which may require prolonged boiling with SDS, and if all cross-links are broken. Covalently attached groups, such as carbohydrates, also affect the size, shape and mobility. The surface charge densities of SDS-polypeptides are not always uniform, especially if a protein has an extreme intrinsic net charge, or if it contains other moieties, such as the carbohydrate side-chains of glycoproteins, attached lipid groups, etc. In this case, the Ferguson plots of many SDS–polypeptides do not intersect at or near the ordinate (Figure 17-2) and accurate molecular weights of SDS–polypeptides can be derived only from the slope of a Ferguson plot. Finally, the plot of log(molecular weight) versus migration distance over a wide range is sigmoidal, not linear (Figure 17-12). The approximate linearity in the central segment of the sigmoidal curve only applies to a relatively narrow range of migration distances.

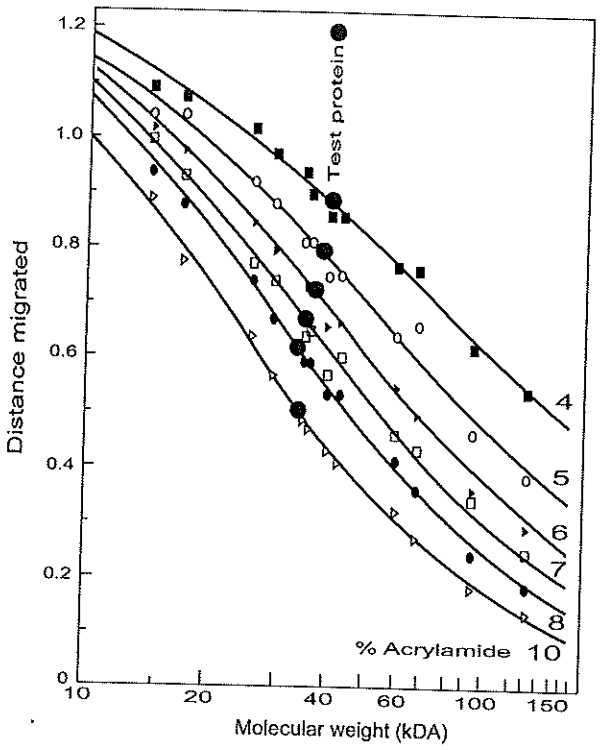

Figure 17-12. Typical standard curve of log (molecular weight) versus relative distance migrated in SDS–PAGE. Twelve standard proteins of known size were separated by SDS–PAGE in polyacrylamide gels of six different concentrations. The protein whose molecular weight was to be determined was rhodopsin (circled dot); slightly different molecular weights were indicated by the various experiments. Data from R. N. Frank & D. Rodbard.

In spite of its importance and widespread use, **the denaturing action of SDS is not understood.** This detergent has anomalous properties when compared with homologous detergents with longer or shorter nonpolar chains. Nor is the structure of the complex of proteins with SDS known. There appear to be at least two independent electrophoretic migration mechanisms for SDS–protein complexes: (1) for proteins in the 14–65-kDa range at <15% polyacrylamide gels, linear Ferguson plots suggest that they migrate ideally and that their effective radii can be estimated in this manner; (2) concave Ferguson plots at higher gel concentrations, and for larger proteins, indicate that migration of the protein–SDS complexes in these cases should be described by the reptation of flexible chains. Migration of

large proteins at lower gel concentrations and small proteins at higher gel concentrations exhibit intermediate behavior. All but the smallest SDS–protein complexes appear to adopt a necklace-like structure in which spherical micelles are distributed along the unfolded polypeptide chain.

In view of the limitations and uncertainties of SDS–PAGE, it is always advisable to establish the accuracy of any important molecular weight value by other means. The most straightforward method is to compare its Ferguson plot with that of the standard proteins (Figure 17-2). Ultimately, the protein should be purified and its mass measured extremely accurately by mass spectrometry (Chapter 6).

Reevaluation of the electrophoretic migration behavior of soluble globular proteins in the native and detergent-denatured states in polyacrylamide gels. W. H. Westerhuis *et al.* (2000) *Anal. Biochem.* **284**, 143–152.

SDS polyacrylamide gel electrophoresis. J. V. Maizel (2000) *Trends Biochem. Sci.* **25**, 588–590.

New method for analyzing the molecular weights of proteins by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. S. Ma *et al.* (2001) *Biochem. Biophys. Res. Commun.* **281**, 976–978.

17.5. HYDROPHOBIC ELECTROPHORESIS

Aqueous buffers are the appropriate solvents for gel electrophoresis of globular proteins and nucleic acids, but **electrophoresis of hydrophobic membrane proteins requires the presence of detergents or miscible organic solvents.** Detergents that are uncharged or amphoteric with no net charge (Section 3.3.A) can be simply incorporated into a polyacrylamide electrophoresis gel and buffer and will have little effect on the electrophoresis. Charged detergents, such as SDS, can be included at low concentrations, although they will migrate in the electrical field. Charged detergent micelles can be subjected to molecular sieving at high gel concentrations, and moving boundaries can be generated in discontinuous buffer systems (Section 17.2) in which the detergent is the leading or trailing ion. Miscible organic solvents are especially suitable for relatively hydrophobic gels, but the relevant pH of the solution is then uncertain (Figure 3-14).

To preserve the native or relatively undissociated structures of membrane proteins and their biological activities, nondenaturing detergents must be used under conditions that do not disrupt or inactivate the native structures and complexes. A number of amphoteric or uncharged detergents are particularly useful in that regard. Their effectiveness in solubilizing membrane proteins depends upon their concentration, critical micelle concentration (cmc) and size of their micelles, as well as the ratio of detergent to protein. Blue native PAGE has been developed for this purpose (Section 17.7.A).

Capillary zone electrophoresis in non-aqueous solutions: pH of the background electrolyte. S. P. Porras & E. Kenndler (2004) *J. Chromat. A* **1037**, 455–465.

Are the asserted advantages of organic solvents in capillary electrophoresis real? A critical discussion. S. P. Porras & E. Kenndler (2005) *Electrophoresis* **26**, 3203–3220.

17.6. ISOELECTRIC FOCUSING (IEF)

When macromolecules are subjected to electrophoresis in a continuous gradient of pH values, they will migrate electrophoretically towards the pH of their isoelectric point (pI), where they have zero mobility. When the system comes to equilibrium, each species of macromolecule will be tightly focused at the pH corresponding to its pI; **any diffusion away from this equilibrium position will be reversed by electrophoretic migration back to the pI.** Unlike conventional chromatographic and electrophoretic separation techniques, where zones of concentrated material are constantly dissipated by diffusion, IEF and the similar techniques of isopycnic centrifugation (Section 16.5.D) and isotachophoresis (Section 17.2.B) have mechanisms that oppose this. When the analyte reaches an environment in which its physicochemical parameters are equal (*iso*) to those of its surroundings, it focuses in a very thin zone that is kept stable and sharp by the external fields (voltage gradients in IEF or centrifugal fields in isopycnic centrifugation) that force any deviating analyte back into its zone of focusing.

The resolving power of IEF is expressed as ΔpI, the minimum difference in isoelectric points between two macromolecules that may be resolved:

$$\Delta pI = 3.17\ (D_t[\partial(pH)/\partial x]/E[\partial u/\partial(pH)])^{1/2} \qquad (17.5)$$

where D_t is the protein translational diffusion coefficient (Section 15.3), E is the voltage gradient (in units of V/cm) applied, $\partial(pH)/\partial x$ is the slope of the pH gradient along the separation axis x, and $\partial u/\partial(pH)$ is the titration curve of the protein expressed in terms of its net charge, which determines its electrophoretic mobility u, as a function of pH. It is obvious that the experimental conditions that minimize ΔpI, and offer the greatest separation of molecules with different pIs, are shallow pH gradients and high voltage gradients. **The best ΔpI attainable in practice is in the order of 0.001 pH unit**, and mixtures of macromolecules can usually be resolved into very many species. What appears to be a single species by many separation techniques is frequently resolved into several by IEF (Figure 17-13).

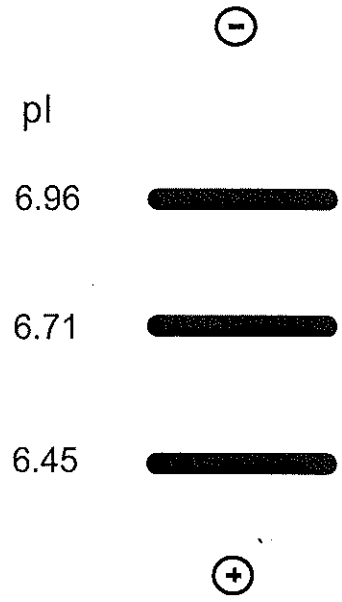

Figure 17-13. Example of the separations possible with IEF. A soluble form of the receptor for epidermal growth factor (EGF) gave a single band on SDS–PAGE, but it was resolved by an immobilized pH gradient from pH 5 to 8 into three major isoforms plus several minor bands. The positions of the major components are indicated schematically, and their pI values are given. Data from W. Weber *et al.* (1994) *J. Chromat.* **679**, 181–189.

There are two primary methods of generating the pH gradients required for IEF, using either soluble or immobilized amphoteric reagents.

The isoelectric focusing problem analytic solution. L. L. Frumin *et al.* (2000) *J. Biochem. Biophys. Methods* 45, 205–209.

Recent applications of capillary isoelectric focusing. F. Kilar (2003) *Electrophoresis* 24, 3908–3916.

Isoelectric focusing. R. Westermeier (2004) *Methods Mol. Biol.* 244, 225–232.

17.6.A. Carrier Ampholytes: Soluble Amphoteric Buffers

Carrier ampholytes are mixtures of soluble amphoteric buffers (Figure 17-14) **that set up the pH gradient for IEF by distributing themselves at their isoelectric points when in an electric field.** To generate a gradient over a range of pH values, an appropriate mixture of ampholytes is necessary. Besides being amphoteric, having both acidic and basic ionized groups, they must also have suitable buffering power and conductivity at their pI. **The most important aspect of a carrier ampholyte is the absolute value of the difference between its pI and the pK_a of its buffering groups;** the smaller this difference, the greater its conductivity and buffering capacity at its pI.

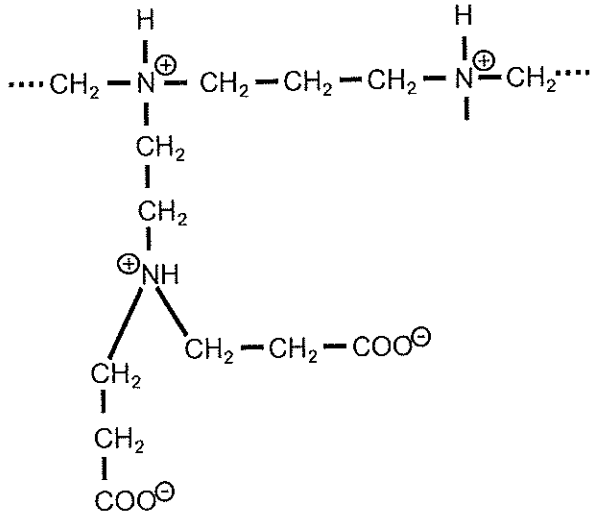

Figure 17-14. Representative chemical formula of the aliphatic oligo-amino, oligo-carboxylic acid ampholines.

Suitable synthetic carrier ampholytes are available commercially. They are amino carboxylic acids that each contain at least four weakly ionizing groups; at least one is a carboxyl group and at least one a basic N atom, but they should contain no peptide bonds (Figure 17-14). The synthesis of ampholytes is a random, chaotic process, in that very heterogeneous mixtures of different oligoamines (typically pentaethylene hexamine, tetraethylene pentamine and triethylene tetramine, including their branched isomers) are reacted with an α-β unsaturated acid (typically acrylic acid) at an appropriate ratio (usually two N atoms/carboxyl group). The result is a great mixture of molecules with β-propionic acid residues:

$$-NH_2 + CH_2 = CH - CO_2^- \rightarrow -NH - CH_2 - CH_2 - CO_2^- \quad (17.6)$$

By this synthetic approach, more than 5000 chemically distinct amphoteres can be produced with pIs varying between 2.5 and 11. Very narrow pH ranges are generated by subfractionating mixtures with wider pH ranges.

The oligoamines have pK_a values well distributed along the pH scale, so they can provide the buffering power and conductivity required for IEF. Additionally, the extra ionizable group of the acrylic acid grafted onto the oligoamino backbone provides acidic carrier ampholytes with extra buffering power. As a result, **focusing with carrier ampholyte buffers is usually more successful in acidic pH ranges.** Other reasons are that fewer alkaline ampholytes are generated, and that IEF gel slabs at pH >8 open to the atmosphere absorb CO_2 and become more acidic; this can be minimized by submerging them under a thin layer of light paraffin oil.

Carrier ampholyte IEF is a technique with one of the greatest resolving powers. Two particular applications are its use in 2-D electrophoresis (Section 17.7) and for visualizing the titration curves of macromolecules. The latter technique uses an IEF slab gel that has been focused to generate the desired pH gradient. The slab is then turned 90° and the sample loaded onto one side of the gel, from the original anode to the original cathode. It is then subjected to electrophoresis perpendicular to the stationary pH gradient. So long as this separation does not perturb the original pH gradient, **each protein will exhibit its own pH/mobility curve in the gel slab that gives a graphic representation of its titration curve.**

Carrier ampholyte IEF gels have, however, a number of problems: their chemical environments of low ionic strength and uneven buffering capacity and conductivity are not well-defined; proteins of the sample can be lost at the cathode during prolonged runs; and only limited pH gradients are possible with the carrier ampholytes that are available. Even in analytical runs with small samples, proteins tend to precipitate when they reach their isolectric pH, where they are at their minimum solubility, which is exacerbated by the low ionic strength within the gel. Immobilized pH gradients tend to overcome most of these problems.

The transitional isoelectric focusing process. T. Huang *et al.* (2005) *Anal. Bioanal. Chem.* **382**, 783–788.

Isoelectric buffers. III. Determination of pK_a and pI values of diamino sulfate carrier ampholytes by indirect UV-detection capillary electrophoresis. S. Lalwani *et al.* (2005) *Electrophoresis* **26**, 2503–2510.

Loading capacity of carrier ampholyte-based buffers in capillary electrophoresis. J. M. Busnel *et al.* (2006) *Electrophoresis* **27**, 563–571.

Mass distribution and focusing properties of carrier ampholytes for isoelectric focusing. I. Novel and unexpected results. R. Sebastiano *et al.* (2006) *Electrophoresis* **27**, 3919–3934.

17.6.B. Immobilized pH Gradients (IPG)

In IPG-IEF, the pH gradient is copolymerized, and thus immobilized, within the polyacrylamide matrix. The pH is determined using a set of nonamphoteric weak acids and bases, with the general chemical composition CH_2=CH–CO–NH–R, where R denotes either weak carboxyl groups or

tertiary amino groups, each with unique pK_a values. The buffers are not amphoteric but bifunctional: at one end of the molecule is located the buffering (or titrating) group, R, and at the other is the acrylic double bond, which will participate in the acrylamide polymerization reaction (Figure 17-4). During polymerization of a polyacrylamide gel, these buffering species are efficiently incorporated into the gel: a conversion efficiency of 84–86% is generally obtained after 1 h at 50°C. The pH gradients can be cast in the same way as conventional gradient polyacrylamide gels, using a density gradient to stabilize the buffer concentration gradient (Figure 20-2). Contrary to the situation where many soluble, amphoteric buffers must create and sustain the pH gradient during the electrophoretic run, IPGs are based on only a few well-defined, nonamphoteric buffers and titrants, and the pH gradients are highly reproducible because they are insoluble and immobilized.

The acidic and alkaline buffering groups have different temperature coefficients for their ionization ($\partial pK_a/\partial T$), so temperature affects the pH gradients (Section 3.4), as do the ionic strength and any additives that alter the water structure (e.g. chaotropic agents such as urea; Section 3.8) or lower its dielectric constant. The inclusion of urea produces the largest changes: the pK_a values of acidic groups are increased in 8 M urea by as much as 0.9 pH units, basic groups by 0.45 pH unit. Detergents at up to 2% (w/v) in the gel do not alter the pK_a values, suggesting that acidic and basic groups attached to the gel are not incorporated into surfactant micelles. Extended pH gradients can be generated using additional strong titrants having pK_a values well outside the desired pH range, such as quaternary amino ethyl (QAE)-acrylamide, with a pK_a greater than12, and 2-acrylamido-2-methyl propane sulfonic acid (AMPS), with a pK_a of approximately 1.

The IPG technique has advantages over IEF in that it is highly reproducible, the resolution is increased by at least one order of magnitude, the pH gradient is stable indefinitely, many different pH intervals and buffers are possible, the sample can be larger, salts in the sample do not distort the pH gradient, and the sample can be recovered without contamination by the IEF reagents. The pH gradients can cover intervals of 0.1–8.0 pH units, and recipes are available for a wide variety of pH gradients. IPG gels are effective at very alkaline pH values, up to pH 12, where carrier ampholytes do not function. Also, the gradients need not be linear but can be concave, convex or sigmoidal.

After the IPG gels have been formed, they are washed extensively to remove nonpolymerized material, salts and polymerization catalysts, which could react with thiol and amino groups of proteins. IPG gels are cast onto plastic supports and are thin (0.5 mm) and porous, so they can be stored dry and then reswollen with any desired additives just prior to use.

Isoelectric focusing in immobilized pH gradients: recent analytical and preparative developments. P. G. Righetti & A. Bossi (1997) *Anal. Biochem.* **247**, 1–10.

Isoelectric focusing in immobilized pH gradients: an update. P. G. Righetti & A. Bossi (1997) *J. Chromat. B* **699**, 77–89.

17.7. TWO-DIMENSIONAL (2-D) GEL ELECTROPHORESIS

2-D gel electrophoresis is one of the most powerful techniques for analyzing samples that contain many components. One of the reasons is simple geometry. One lane of a 1-D electrophoresis gel covered with evenly spaced ideal bands might be able to resolve 50 such bands in a space of 10 cm.

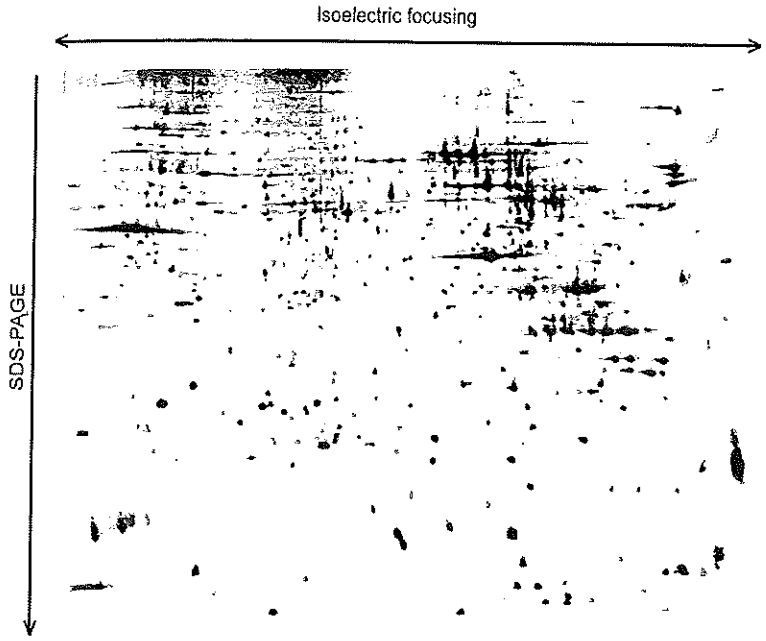

Figure 17-15. Typical example of a two-dimensional gel, separating total proteins from mammalian cells. The first dimension, horizontal, was isoelectric focusing using wide-range ampholytes between pH 3 and 10. The second dimension was SDS–PAGE. The proteins were detected by silver staining.

Correspondingly, a square of the same dimensions can theoretically resolve 50^2, or 2500, analytes in individual spots. 2-D gels can resolve thousands of proteins, including those from an entire cell (Figure 17-15).

Proteins differ from each other primarily in their size and net charge, so IEF is normally used in the first dimension to provide a separation based on net charge (Section 17.6). The second dimension is normally SDS–PAGE (Section 17.4), to separate on the basis of size. **The position of each protein on the gel is defined by its isoelectric point (pI) and its apparent molecular weight (M_r).**

With so many spots, comparing different gels is limited by the reproducibility of their positions. Immobilized pH gradients for the IEF (Section 17.6.B) are very reproducible and help to produce constant positions for each protein at its pI in the first dimension. Comparing different gels generally involves densitometry and computerized pattern displacement techniques. Alternatively, two samples may be compared on the same gel if the proteins in them are labeled in two different ways, either with dyes or radioactive labels. When specific antibodies are available, immunological detection of spots after blotting (Section 21.1) has made it unnecessary to identify each spot from just its position. The identities of proteins in spots can also be determined by their molecular weights measured using mass spectrometry (Chapter 6). The identities of the proteins observed on 2-D gels are compiled in databases.

Other combinations of separation techniques can be appropriate in other instances. For example, agarose gel electrophoresis of viruses uses a low-concentration gel for the first dimension, separating on the basis of charge, while a more concentrated second-dimension gel separates on the basis of size.

Fluorescent two-dimensional difference gel electrophoresis unveils the potential of gel-based proteomics. G. van den Bergh & L. Arckens (2004) *Curr. Opinion Biotechnol.* 15, 38–43.

Current two-dimensional electrophoresis technology for proteomics. A. Gorg *et al.* (2004) *Proteomics* 4, 3665–3685.

Recent advances in 2D electrophoresis: an array of possibilities. G. van den Bergh & L. Arckens (2005) *Expert Rev. Proteomics* **2**, 243–252.

Two-dimensional gel electrophoresis as tool for proteomics studies in combination with protein identification by mass spectrometry. B. Wittmann-Liebold *et al.* (2006) *Proteomics* **6**, 4688–4703.

17.7.A. Blue Native PAGE

The application of 2-D electrophoresis to hydrophobic membrane proteins is hampered by precipitation of many of these proteins during IEF in the first dimension. Therefore, new strategies for identifying and characterizing membrane proteins have been developed. **Blue native polyacryamide gel electrophoresis (often abbreviated as BN–PAGE) is a special type of native electrophoresis for high-resolution separation of enzymatically active protein complexes from membranes and tissue homogenates and cell fractions.** Membrane protein complexes are solubilized using mild neutral detergents that do not disrupt the complexes. The electrophoretic separation relies on the gentle binding of the dye Coomassie blue G250 to all membrane and many soluble protein complexes, which maintains their structures and functions and adds negative charges to the surface of the proteins. During migration to the anode at pH 7.5, protein complexes are separated according to their molecular mass and/or size, and high resolution is obtained using the decreasing pore size of a polyacrylamide-gradient gel. This method is useful for complexes of molecular weights between 10,000 and 10,000,000 using PAGE, and even larger weights using agarose electrophoresis.

The blue native electrophoresis is usually followed by a second electrophoresis in the second dimension. This is usually SDS–PAGE, which dissociates the complexes into their constituent polypeptide chains and aids their identification (Figure 17-16). If the function is to be characterized, the second dimension can be nondenaturing and similar to the first but omitting the blue dye and charge change. This procedure is known as **colorless** or **clear native PAGE (CN–PAGE)**. It uses the normal charge properties of the protein complex and has proved to be an especially mild separation capable of preserving weak protein–protein interactions.

Advantages and limitations of clear-native PAGE. I. Wittig & H. Schagger (2005) *Proteomics* **5**, 4338–4346.

Analysis of membrane protein complexes by blue native PAGE. V. Reisinger & L. A. Eichacker (2006) *Proteomics* **6**, 6–15.

Detection and analysis of protein–protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: (membrane) protein complexes and supercomplexes. F. Krause (2006) *Electrophoresis* **27**, 2759–2781.

17.8. TRANSVERSE-GRADIENT GEL ELECTROPHORESIS (TGGE)

Electrophoresis is primarily a comparative technique, and the size and shape of a molecule are two of the main determinants of its electrophoretic mobility through sieving gels. Consequently, **electrophoresis through polyacrylamide gels is especially useful to monitor changes in the conformations of proteins and nucleic acids.** Phenomena where the conformation can be induced to change under the influence of some agent, such as a denaturant or a change in temperature, can be readily studied by electrophoresis in transverse-gradient gels.

~ CHAPTER 18 ~

MOLECULAR SIEVES: GEL FILTRATION/SIZE EXCLUSION CHROMATOGRAPHY

The simplest technique for obtaining information about the sizes and hydrodynamic shape of macromolecules is known as **gel filtration** or **size exclusion chromatography** (SEC). Historically, SEC has been referred to as **gel filtration chromatography** when applied to the separation of proteins in aqueous (native) conditions; the two terms are used interchangeably in molecular biology. The method is suitable for both analytical and preparative-scale separations of biological macromolecules.

SEC measures the tendency of a solute to enter the pores of a **molecular sieve** by monitoring how long it takes the solute to pass through a chromatography column containing the molecular sieve (Figure 18-1). **Solutes are separated on the basis of their size, because small molecules tend to enter the pores of the sieving material, whereas large molecules are excluded and remain in the solvent outside the sieves; consequently, the largest molecules emerge first.** SEC also differs from migration through a continuous sieving material, such as electrophoresis through a polyacrylamide gel, where the larger molecules migrate more slowly (Figure 18-1). **Obtaining quantitative information about the size, however, requires comparison with an appropriate set of standard molecules of known size and the same shape;** there is no way to calculate the size from first principles. A less appreciated, but very powerful, application of the method is for the analysis of the interactions of macromolecules, to determine their stoichiometries and energetics (Chapter 19). A great advantage of SEC is that the macromolecule being studied need not be pure, if it can be identified specifically in some way, such as by some biological activity, after it has emerged from the molecular sieve column.

SEC differs from other types of chromatography, such as ion-exchange and affinity chromatography (Chapter 20), in that the analytes are not separated because they interact differently with the column resin. On the contrary, SEC is most powerful when there are no interactions between the macromolecular solute and the stationary phase.

Some potentialities and drawbacks of contemporary size-exclusion chromatography. K. Stuklik *et al.* (2003) *J. Biochem. Biophys. Methods* **56**, 1–13.

Analytical exclusion chromatography. D. J. Winzor (2003) *J. Biochem. Biophys. Methods* **56**, 15–52.

Size-exclusion chromatography. P. Cutler (2004) *Methods Mol. Biol.* **244**, 239–252.

Gel filtration chromatography. P. Staunton (2004) *Methods Mol. Biol.* **251**, 55–74.

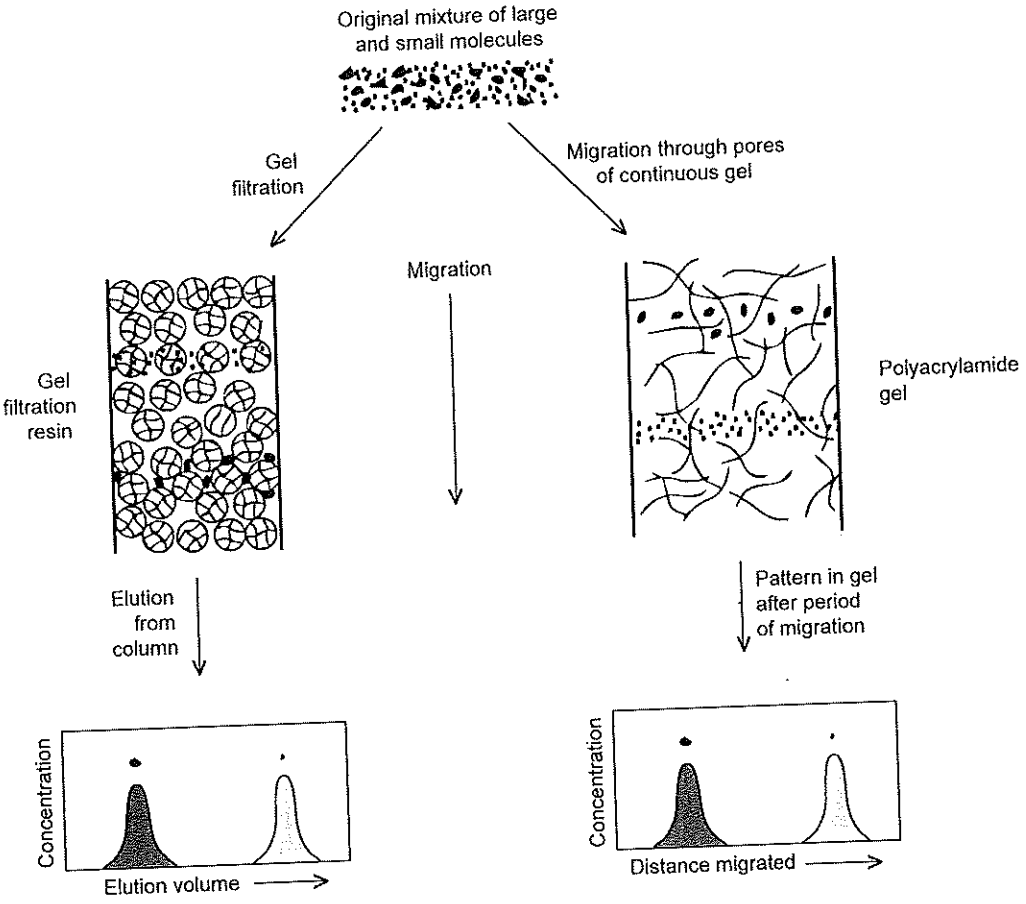

Figure 18-1. Comparison of gel filtration through a column of particles and molecular sieving through a continuous medium, such as a polyacrylamide gel. At the top is a mixture of two types of molecules, one large and one small, that is subjected to gel filtration (left) or sieving through a continuous polyacrylamide gel (right). During gel filtration, the larger molecule is excluded from the resin and consequently migrates more rapidly than the smaller molecule and elutes first. In contrast, the larger molecule moves more slowly through the pores of the continuous polyacrylamide gel. From T. E. Creighton (1993) *Proteins: structures and molecular properties*, 2nd edn, W. H. Freeman, NY, p. 25.

18.1. MOLECULAR SIEVE MATERIALS

Molecular sieve resins are porous beads that are synthesized from noninteracting, inert polymeric materials in order to have pores of varying sizes (Figure 18-2). When a very small solute is transported through a column packed with this resin, it will have access to all the pores. A very large solute, in contrast, would not be accommodated in the pores at all and would be totally excluded from the interiors of the beads. Solutes of intermediate size would partition (Section 3.2.A) into the pores of the molecular sieve. The extent of partitioning of the solute into the pores is inversely proportional to its size, with larger solutes partitioning to a lesser extent than smaller solutes. The sizes of the pores in the beads are generally not uniform, but are best described as a distribution.

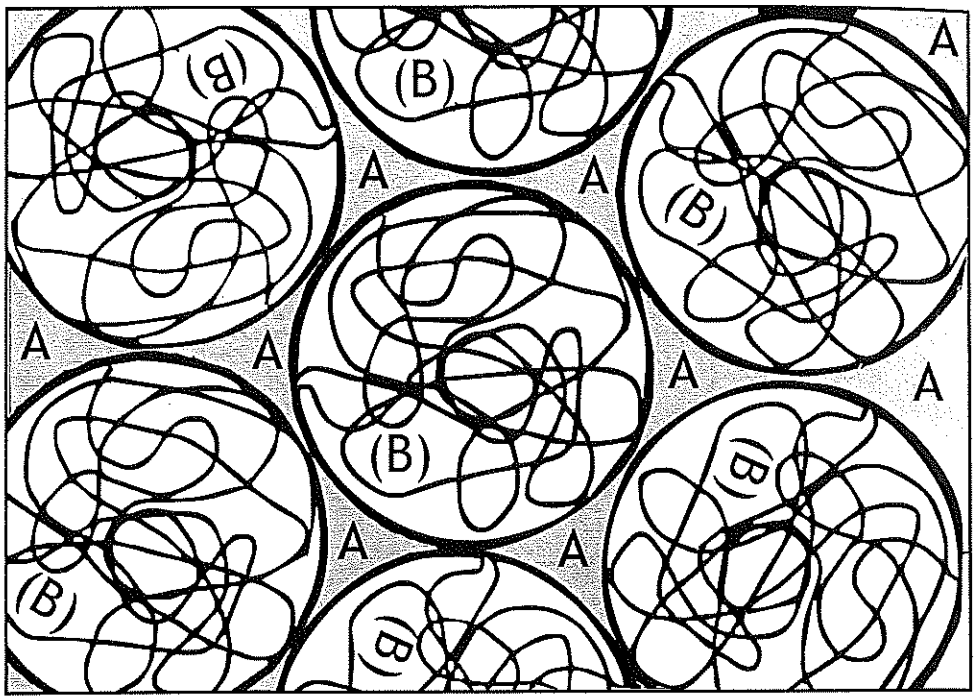

Figure 18-2. Schematic representation of a molecular sieve partitioning system. The *shaded* area represents the mobile phase exterior to the resin particle. The porous gel particles enclosed by the *circles* represent the stationary phase; each consists of two regions, the internal solvent (B) and the polymeric gel matrix (*curved lines*). The liquid phase outside the resin particles is A. Solute molecules introduced into the system partition between A and B, with the extent of partitioning into B depending on the size of the solute.

Molecular sieves can be constructed from a range of materials, but the two most common types are cross-linked gels and derivatized glass or silica. The cross-linked gels can be natural, based on polysaccharides, or synthetic, made by polymerization of hydrophilic monomers. The most common are cross-linked dextrans and agarose, polyacrylamide and combinations of polyacrylamide and agarose. The glass materials have the advantages of being rigid and having uniform sizes of pores and particles; their major disadvantages are their tendencies to adsorb molecules and degrade with use. Adsorption of macromolecules to glass results from its SiOH groups; incomplete blockage of these groups results in a negatively charged surface to which proteins can bind. This is a disadvantage, as SEC is most straightforward if the solute being chromatographed does not bind to the resin. Degradation of the silica limits the types of liquids that may be used. Polymeric resins are not totally inert, however, and the liquid phase used in SEC experiments should have minimal effects on the resin.

The most crucial factor for molecular sieves is the size of the pores, which governs what sizes of molecules will be able to enter them. Consequently, a range of molecular sieves with varying pore sizes are available commercially, and it is important to choose one appropriate for the molecules to be analyzed (Table 18-1). Only the most popular cross-linked gels will be described here. Sephadex, Sepharose, and Sephacryl are trade names of Pharmacia Fine Chemicals Company for three types of gel filtration matrices that are widely used.

Table 18-1. Properties of some common gel filtration resins

Resin	Molecular weight range for globular proteins (kDa)
Sephadex (dextran)	
G-10	<0.7
G-25	1–5
G-50	1.5–30
G-75	3–80
G-100	4–150
G-150	5–300
G-200	5–600
Sepharose (agarose)	
6B	10–4000
4B	60–20,000
2B	70–40,000
Sephacryl (cross-linked dextran)	
S-100 HR	1–100
S-200 HR	5–250
S-300 HR	10–1500
S-400 HR	20–8000
S-500 HR	80–80,000
Biogel (polyacrylamide)	
P-2	0.1–1.8
P-4	0.8–4
P-6	1–6
P-10	1.5–20
P-30	2.5–40
P-60	3–60
P-100	5–100
P-150	15–150
P-200	30–200
P-300	60–400

Table 18-1. Properties of some common gel filtration resins – *Continued*

Resin	Molecular weight range for globular proteins (kDa)
Superdex (dextran bonded to highly-cross-linked agarose)[a]	
75	3–70
200	10–600
Superose (cross-linked agarose)[a]	
12	1–300
6	5–5000
Fractogel (polyvinyl chloride)[b]	
TSK HW-40	0.1–10
TSK HW-55	1–700
TSK G2000SW	5–100
TSK G3000SW	10–500
TSK HW-65	50–5000
TSK HW-75	500–50,000

[a] For use with FPLC instrumentation.

[b] For use with HPLC instrumentation.

Determination of pore size distributions of porous chromatographic adsorbents by inverse size-exclusion chromatography. Y. Lao & A. M. Lenhoff (2004) *J. Chromatogr. A* **1037**, 273–282.

Three-dimensional DNA crystals as molecular sieves. P. J. Paukstelis (2006) *J. Am. Chem. Soc.* **128**, 6794–6795.

^{17}O solid-state NMR spectroscopic studies of the involvement of water vapor in molecular sieve formation by dry-gel conversion. B. Chen & Y. Huang (2006) *J. Am. Chem. Soc.* **128**, 6437–6446.

18.1.A. Sephadex

Sephadex is a gel prepared by cross-linking dextran using epichlorohydrin and forming it into beads. The large number of hydroxyl groups of dextran means that the gel is extremely hydrophilic, so it swells readily in aqueous solutions, as well as in formamide and dimethylsulfoxide. It is insoluble in other nonaqueous solvents, unless chemically degraded. It is stable in organic solvents and in alkaline and weakly acidic aqueous solutions. Sterilizing it in the dry or wet state at neutral pH by autoclaving for 30 min at 120°C should not affect its gel filtration properties. Sephadex does not melt.

18.1.B. Sepharose

Sepharose is prepared from agarose by removing the charged polysaccharides to produce a gel with only a very small number of residual charged groups; it is then formed into beads. Sepharose is stable in aqueous solution over the pH range 4–9 but is degraded by oxidizing agents. It melts on heating above 40°C, and the bead structure may be irreversibly damaged on freezing. Sepharose can be cross-linked, by reaction with 2,3-dibromopropanol under strongly alkaline conditions, to produce the product known as Sepharose CL. Its thermal and chemical stabilities are greatly enhanced; for example Sepharose CL can be used in aqueous media between pH 3 and 14.

Sepharose is frequently linked covalently to other molecules, to produce resins for affinity chromatography (Section 20.2.G) and other types of chromatography in which the molecules being separated interact with the covalently attached groups. The Sepharose in this case is not involved in molecular sieving but is simply the structural backbone of the resin.

18.1.C. Sephacryl

Sephacryl is allyl dextran cross-linked covalently with N,N′-methylene bisacrylamide, the cross-linker used to prepare polyacrylamide gels (Section 17.1.A). The result is a rigid gel with a well-defined range of pore sizes. Aqueous solutions are generally used, although organic solvents have a much smaller effect on the pore size than in the case of Sephadex. Chemically intact Sephacryl is insoluble in all common solvents. Aqueous solutions can be used in the pH range 2–11 and can contain detergents such as SDS (Section 17.4) and denaturants such as 6 M guanidinium chloride and 8 M urea. Sephacryl does not melt, and its gel filtration properties are not affected by repeated autoclaving at 120°C and neutral pH.

18.2. MOLECULAR SIEVING

During molecular sieving, **solute molecules partition between the solvent spaces in the pores of the resin particles and those outside the resin particles** (Figure 18-2). The resin particles represent the stationary phase in the chromatographic column, while the liquid outside the particles is the mobile phase (Figure 20-1). The stationary phase consists of the sieving resin equilibrated with the solvent used for the mobile phase. The three compartments within the experimental system are: (1) the volume exterior to the resin particles, the **void volume**, V_0; (2) the solvent region in the interior of the gel particles, the **internal volume**, V_i; and (3) the volume occupied by the hydrated gel matrix that excludes even the smallest solute molecule, V_g.

The total volume of the column, V_t, is simply the sum of these three volumes:

$$V_t = V_0 + V_i + V_g \tag{18.1}$$

The value of V_0 can be determined experimentally as the elution volume of a solute that is so large as to be excluded completely from the interior of the resin. The value of V_i is given by the increased elution volume of a solute so small as to have access to all the pores of the resin (Figure 18-3). The difference between the sum of these two volumes and the total volume of the column is V_g. Often the total volume of the column is taken to be that due to the solvent alone, $V_0 + V_i$, ignoring the volume occupied by the gel, V_g.

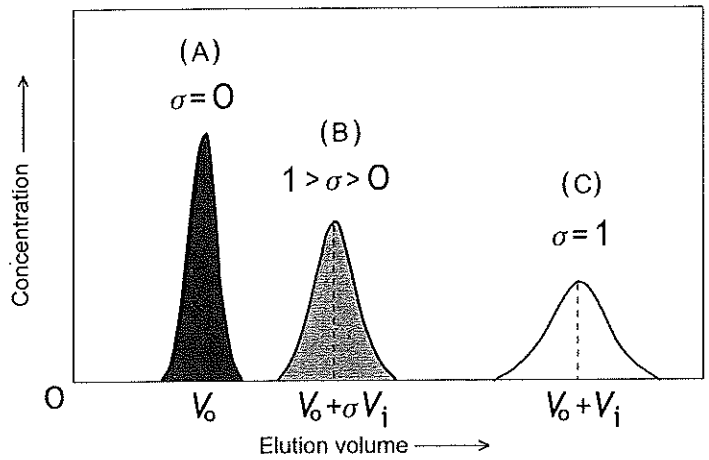

Figure 18-3. Schematic diagram of elution from a gel filtration column, illustrating the three types of solute behavior: (A) total exclusion, which defines the void volume, V_0; (B) a molecule that partially penetrates the molecular sieve; (C) a small molecule that is not excluded, and defines the internal volume, V_i.

The gel filtration behavior of a solute is usually expressed in terms of the extent of its partitioning into the pores of the resin by the **partition coefficient σ:**

$$\sigma = \frac{V_S}{V_i} \tag{18.2}$$

where V_s is the penetrable volume of the gel that can be occupied by the solute molecules at equilibrium. **The partition coefficient σ for a molecule is the fraction of the internal volume of the resin that is penetrable by that molecule. The values of s range between 0 and 1.** Very large solutes that are totally excluded from the pores have σ = 0, whereas σ = 1 for very small solutes that can enter all the pores freely.

The volume at which a solute molecule elutes from a column, V_e, is given by (Figure 18-3):

$$V_e = V_0 + \sigma V_i \tag{18.3}$$

and the value of σ is given by

$$\sigma = \frac{V_e - V_0}{V_i} \tag{18.4}$$

The partition coefficient σ is also known as the **distribution coefficient K_d**. When the volume occupied by the resin is ignored ($V_t = V_0 + V_i$), the distribution coefficient is often expressed as

$$K_d = \frac{V_e - V_0}{V_t - V_0} = \frac{V_e - V_0}{V_i} \tag{18.5}$$

In analysis of size exclusion chromatography data, the partition coefficient used is often K_{av}, which is expressed in terms of the total volume of the gel phase, $V_i + V_g (=V_t - V_0)$. K_{av} is defined as the volume fraction of the stationary phase that can be occupied by the solute:

$$K_{av} = \frac{V_S}{V_t - V_0} \tag{18.6}$$

In this case, K_{av} must be less than one, as no molecule can occupy the total volume of the column:

$$K_{av} = \frac{V_e - V_0}{V_t - V_0} \tag{18.7}$$

K_{av} appears to be the same as K_d, but the latter ignores V_g. K_{av} is related to σ by:

$$\sigma = K_{av}\left(1+\frac{\bar{v}_g}{S_r}\right) \tag{18.8}$$

where $\bar{v}_g$ is the partial specific volume of the gel matrix and S_r is the volume of the solvent taken up by the gel material. Consequently, for any type of sieving resin, the values of K_{av} and σ have a constant ratio. If the values of $\bar{v}_g$ and S_r are known for two different gel materials, a value for a partition coefficient measured in one system can be converted to the value in the other.

In all such quantitative treatments of SEC, corrections must be introduced with high solute concentrations to take into account the effects of nonideality.

Thermodynamic equilibrium of the solute distribution in size-exclusion chromatography. A. M. Striegel (2004) *J. Chromatogr. A* **1033**, 241–245.

Stochastic theory of size exclusion chromatography: peak shape analysis on single columns. A. Felinger *et al.* (2005) *Anal. Chem.* **77**, 3138–3148.

Application of the general height equivalent to a theoretical plate equation to size exclusion chromatography. Study of the mass transfer of high-molecular-mass compounds in liquid chromatography. F. Gritti & G. Guiochon (2007) *Anal. Chem.* **79**, 3188–3198.

18.3. SIZE EXCLUSION CHROMATOGRAPHY (SEC)

To perform gel filtration, the molecular sieving resin is equilibrated with the appropriate buffer and packed into a chromatography column. The sample is loaded onto the top of the resin, as either a small or large zone, and allowed to enter the resin at a flow rate determined by gravity or a pump connected at the column outlet. Once the sample has entered the column, it is followed by the column buffer and transport of the macromolecules proceeds (Figure 18-4). Standard simple low-pressure chromatographic apparatus has been used traditionally for SEC, but the more complex fast-protein liquid chromatography (FPLC) and high-performance liquid chromatography (HPLC) systems have many advantages, especially in giving reproducible results.

When the analytes reach the bottom of the column, they **are eluted in decreasing order of size,** so long as they are within the fractionation range of the molecular sieve (Table 18-1). The eluted molecules can be detected as they emerge from the column, most conveniently by spectroscopic methods. The material eluted can be collected in numerous fractions and characterized later, using any

technique appropriate for the sample. Radioactive molecules can be detected by scintillation counting (Section 5.4.A.2), and all molecules can be detected and identified by electrospray ionization mass spectrometry (Section 6.1).

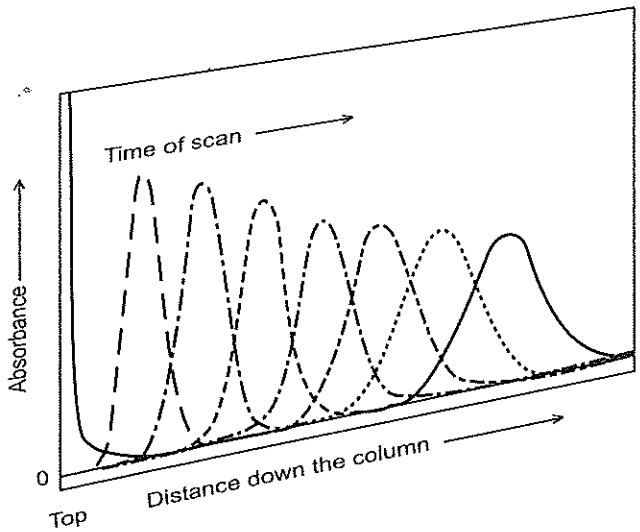

Figure 18-4. Typical migration of a small zone of a small solute through a SEC column. The initial sample applied to the column is the very sharp peak on the *left* represented by scan 1; as the zone moves down the column, it broadens and the maximum concentration decreases.

Some potentialities and drawbacks of contemporary size-exclusion chromatography. K. Stulik *et al.* (2003) *J. Biochem. Biophys. Methods* **56**, 1–13.

High-performance size-exclusion chromatography of peptides. G. B. Irvine (2003) *J. Biochem. Biophys. Methods* **56**, 233–242.

Fast size-exclusion chromatography: theoretical and practical considerations. S. T. Popovici & P. J. Shoenmakers (2005) *J. Chromatogr. A* **1099**, 92–112.

18.3.A. Small-zone Filtration

SEC usually involves loading a small volume of solute onto the column (Figure 18-5-A), referred to as a **small-zone experiment** or **zonal analysis.** It may be used for either analytical or preparative separations. The volume of the sample should be no more than 1–4% of the total column bed volume. This initially small zone spreads as it flows through the column, due to diffusion (Figure 18-4). A single homogeneous peak should elute from the column with a Gaussian shape, and its apex marks the elution volume of the solute. A mixture of solutes with different elution properties should produce multiple peaks with Gaussian shapes (Figure 18-3).

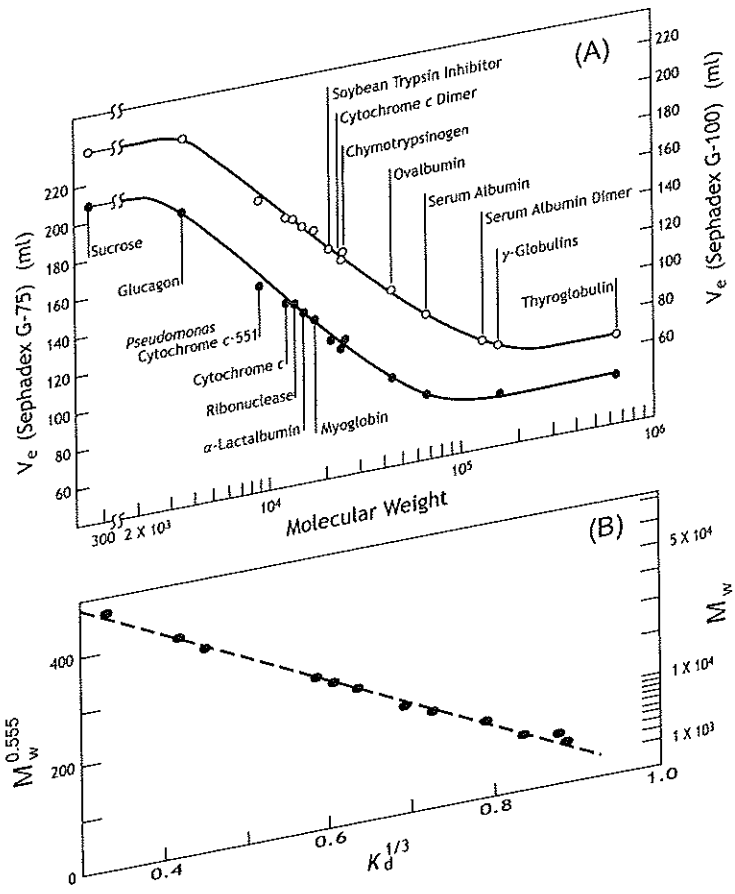

Figure 18-6. Determining the molecular weights of proteins on the basis of their gel filtration properties. The relationship between the molecular weights of native globular proteins and their elution volumes is illustrated in (A), using either Sephadex G-75 (*solid circles*) or Sephadex G-100 (*open circles*); Sephadex G-75 has the smaller pores. The plateaus on the *left* and *right* give the total internal (V_i) and total void (V_0) volumes of the column, respectively. This relationship between molecular weight and elution volume is observed in this case only because all of these proteins are roughly spherical in shape. Data from P. Andrews (1964) *Biochem. J.* **91**, 222–233. In (B) the proteins were all unfolded in 6 M guanidinium chloride. The elution volumes of the unfolded polypeptide chains are expressed as distribution coefficients, K_d (Equation 18.5) and the value of $K_d^{1/3}$ is linearly related to the 0.555 power of the molecular weight, M_w. Data from N. Ui (1979) *Anal. Biochem.* **97**, 65–71.

Experimental data are consistent with these expectations (Figure 18-7). Values of the inverse error function complement are compiled in tables provided by the National Institute of Standards and Technology (NIST) or they may be readily calculated.

18.4.A. Estimating the Molecular Weight

A linear relationship between the elution volumes and molecular weights of standard molecules permits the molecular weight of a comparable molecule to be estimated from its elution volume. Such a linear relationship will be observed only if all the molecules have the same physical properties and shapes, and only over a limited range of sizes appropriate for the sieving material used. This relationship is apparent in the two plots in Figure 18-6 only because each is for molecules with the same characteristics. In the top plot, all the proteins are roughly spherical folded globular proteins. In the bottom plot, all the proteins are unfolded random coils in a denaturing solvent. The correct molecular weight will not be obtained for an unfolded protein analyzed on a molecular sieving column that was calibrated with folded proteins.

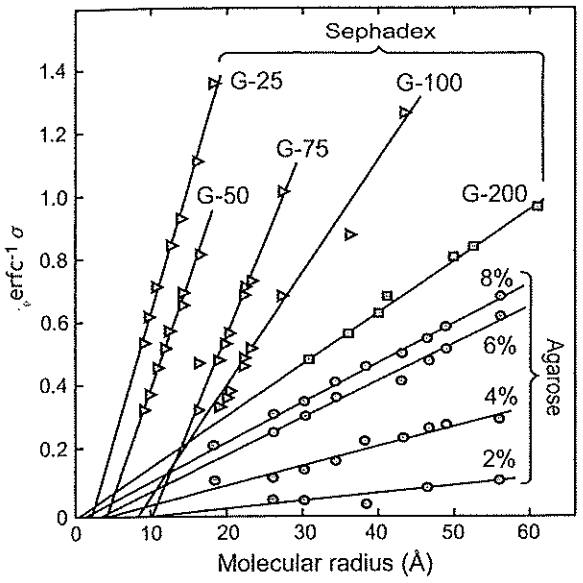

Figure 18-7. Correlation of the partition coefficient σ with the molecular radius of the molecules, using different molecular sieve resins. The elution position of the molecules is expressed as the reciprocal error function of the parameter σ. Data from G. K. Ackers.

Another approach to estimating the molecular weight is to **combine the apparent Stokes radius obtained from gel filtration, r_H, which provides the frictional coefficient, f** (Equation 15.21), **with the sedimentation coefficient derived from centrifugation, s** (Equation 16.13), **to calculate the mass of a protein independently of its shape or hydration:**

$$M_{s,r} = \frac{6\pi N_A\, s r_H \eta}{(1 - \bar{\nu}\rho)} \tag{18.12}$$

where N_A is Avogadro's number, η the viscosity of the solvent and ρ its density, and $\bar{\nu}$ the partial specific volume of the molecule. Both s and r_H are sensitive to the shape and hydration of the macromolecule, but in opposite ways: an asymmetric shape decreases the sedimentation coefficient, making the molecule appear smaller, while an asymmetric shape causes it to enter the pores less readily, and thus to appear larger. To the extent that the two phenomena cancel out, the two measurements can provide a reasonable measure of the molecule's molecular weight.

Size-exclusion chromatography: a review of calibration methodologies. L. K. Kostanski *et al.* (2004) *J. Biochem. Biophys. Methods* **58**, 159–186.

Determining the subunit structure of phosphodiesterases using gel filtration and sucrose density gradient centrifugation. W. Richter (2005) *Methods Mol. Biol.* **307**, 167–180.

18.5. INTERACTING MOLECULES

Gel filtration is ideally suited for monitoring the binding of one molecule to another, as at least one of the molecules will have its apparent molecular weight and hydrodynamic volume increased. Zonal analysis will, however, tend to separate two such reactants, and there will be no indication that they interact unless their complex is so very stable that it survives the gel filtration. Relatively weak and dynamic interactions must be studied using analysis of the boundaries of a large-

~ CHAPTER 6 ~

MASS SPECTROMETRY

A major concern in molecular biology is determining the size of a macromolecule. Proteins and nucleic acids have a very wide range of sizes, and it is important to know whether one is working with a large molecule or a relatively small one. Consequently, a great many techniques have been devised to provide some estimate of the sizes of macromolecules, including light scattering (Chapter 7), microscopy (Chapter 8), diffusion and viscosity (Chapter 15), sedimentation (Chapter 16), polyacrylamide gel electrophoresis (Chapter 17) and size exclusion chromatography (Chapter 18). The most accurate technique by far, however, is mass spectrometry (MS).

Mass spectrometry requires that the sample be in the gas phase and ionized, by either the loss or the gain of at least one charge, due to electron ejection, protonation or deprotonation. The mass spectrometer then separates the ions on the basis of their ratio of mass (m) to charge (z) (Figure 6-1). The molecules being separated must not collide with other molecules or atoms, so the separation must take place in a high vacuum. The result is a mass spectrum that can provide very accurate measures of the molecule's mass. Even structural information can be inferred from the charge distribution on a protein. In some cases, the ionization process breaks the molecule into pieces, which can yield information about its covalent structure.

The ions are detected by either **electron multipliers** or **photomultipliers**. The ion strikes either a dynode that emits secondary electrons or a phosphorous screen that releases photons. The secondary electrons of the dynode are accelerated by a voltage and attracted to a second dynode that emits more electrons. Passage through additional dynodes can amplify the original signal about 10^6-fold. The photons released by a phosphorous screen are detected by a photomultiplier.

Molecules are distinguished in mass spectrometry solely on the basis of their masses, which are expressed in Daltons (Da, 1 Da = 1 g/mol).*

---

* *The absolute masses of molecules are measured in Daltons (Da). Mass spectrometers, however, measure relative masses, because the instrument has been calibrated with molecules of known mass. Such masses are relative and should have no units, but they should have the same magnitude as the absolute mass, so they are commonly referred to as absolute masses. The term 'molecular weight' is widely used, even though strictly speaking a weight is a force (mass × gravity), not a mass, but it has the same magnitude as the mass and is used very commonly to refer to the mass.*

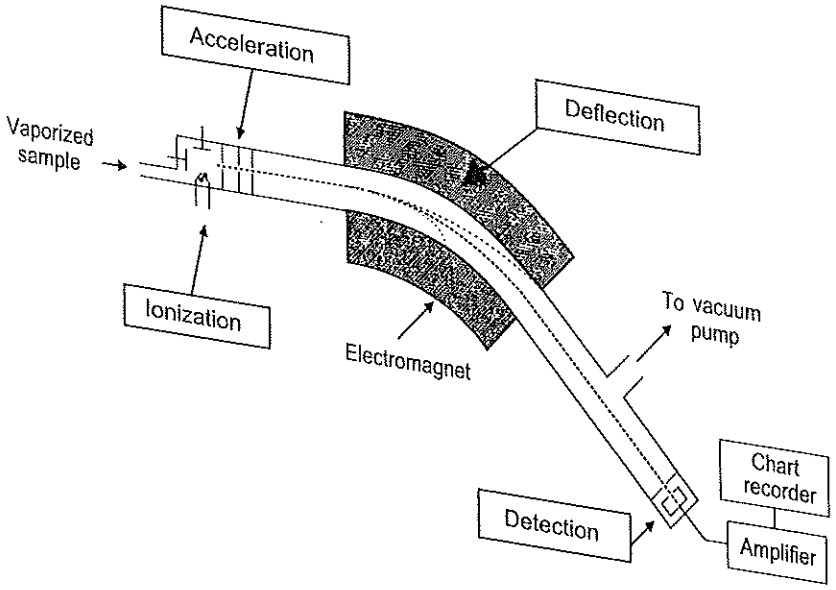

Figure 6-1. Schematic diagram of a mass spectrometer. The molecules of the sample are ionized by knocking off one or more electrons to give a positive ion. The ions are accelerated in an electric field to have the same kinetic energy, then deflected by a magnetic field. The degree of deflection depends upon the ratio of their mass to their net charge: lighter ions are deflected more than heavier ones with the same net charge. Which ions reach the detector depends upon the strength of the magnetic field, which is varied to detect a range of molecules. When positive ions reach the detector, they acquire electrons from it, which is detected by an electric current to the detector. So that the molecules are not affected by collisions with gas molecules, all of this must take place in a vacuum.

Masses of the natural amino acid residues and nucleotides are unique, except for the amino acid residues Leu and Ile, which are isomers and have the same molecular weight. They can be distinguished by mass spectrometry only if they differ in the way that the polypeptide chain is fragmented.

Mass spectrometry has been used routinely for many years with small molecules, but its use with biological macromolecules was prevented for a long time by the need to generate intact ions in the gas phase; biological macromolecules are not detectably volatile. The developments of **electrospray ionization (ESI)** and **matrix-assisted laser desorption/ionization (MALDI)** overcame this problem, and mass spectrometry has become an integral and important part of biological research. ESI and MALDI are fundamentally different ionization techniques, but they achieve essentially the same end result, namely the generation of gas-phase ions of the molecules of the sample by their vaporization and ionization without destroying them. Ionization occurs in both techniques by the addition or abstraction of protons from the molecule M, to produce either $[M+H]^+$ or $[M-H]^-$ ions. Peptides and proteins are generally studied as $[M+H]^+$ ions, nucleic acids as $[M-H]^-$ ions.

Proteins, peptides, carbohydrates and oligonucleotides can now be analyzed routinely by mass spectrometry and examined structurally in very small (picomole to femtomole) amounts. Mass spectrometry can be used to measure the masses of very large biomolecules, up to 10^6 Da in the case of proteins, but about 90 kDa for DNA and 150 kDa for RNA. It can also provide sequence information on unknown peptides and proteins and detect noncovalent complexes, with a molecular weight accuracy of the order of ±0·01% or better. With very high accuracy measurements, the various

minor isotopes present naturally, such as ^{13}C and ^{15}N (Section 5.1), become apparent (Figure 5-3). Proteins and nucleic acids with masses greater than about 8000 Da have at least one such isotope in virtually every molecule.

Molecular weight determination of peptides and proteins by ESI and MALDI. K. Strupat (2005) *Methods Enzymol.* **405**, 1–36.

Mass spectrometry of peptides and proteins. V. H. Wysocki *et al.* (2005) *Methods* **35**, 211–222.

Mass spectrometry and protein analysis. B. Domon & R. Aebersold (2006) *Science* **312**, 212–217.

Mass spectrometry of RNA. B. Thomas & A. V. Akoulitchev (2006) *Trends Biochem. Sci.* **31**, 173–181.

Mass spectrometry of RNA: linking the genome to the proteome. Z. Meng & P. A. Limbach (2006) *Brief Funct. Genomic Proteomic* **5**, 87–95.

6.1. ELECTROSPRAY IONIZATION (ESI)

ESI generates ions of macromolecules directly from an aqueous or aqueous/organic solvent that does not contain much salt. A fine spray of highly charged droplets is created in the presence of a strong electric field (Figure 6-2). These charged droplets vaporize as they move into the vacuum of the mass spectrometer, which concentrates the ionized molecules. When the electrostatic repulsion between them becomes sufficiently great, they leave the droplet and enter the gas phase individually as charged ions. The number of charges on a molecule depends on factors such as the composition and pH of the solvent and the chemical nature of the sample. Proteins with multiple positive charges are usually obtained by ESI from acidic solutions of pH 2–4, while negatively charged protein molecules are produced from alkaline solutions of pH 8–10. ESI usually produces a series of multiply charged species of large molecules. Proteins from acidic solution often contain one positive charge on each accessible basic group, i.e. those of the lysine and arginine side-chains, plus the terminal amino group.

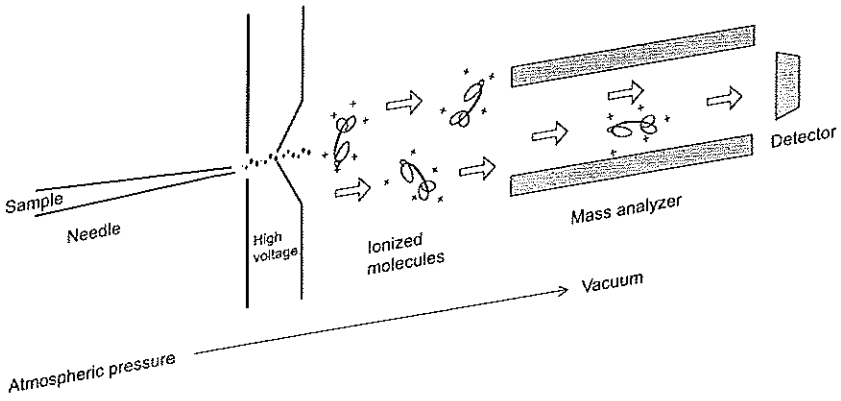

Figure 6-2. The electrospray ionization process. An aqueous solution of the analyte (at a concentration of about 5 μM) is placed in the needle. A voltage of several kilovolts is applied to the gold-plated needle, producing an electrospray of fine droplets from 1 to 2 μl of the sample. The positively charged droplets are desolvated by the vacuum, and the individual molecules are separated and detected by the mass spectrometer.

The **ESI mass spectrum of a homogeneous macromolecule contains multiple peaks corresponding to the different charged states** and different mass-to-charge (m/z) ratios (Figure 6-3). Adjacent peaks differ by one integral charge, plus one proton, so the spectrum can be deconvoluted to determine the net charge of each peak and the molecular mass of the original molecule.

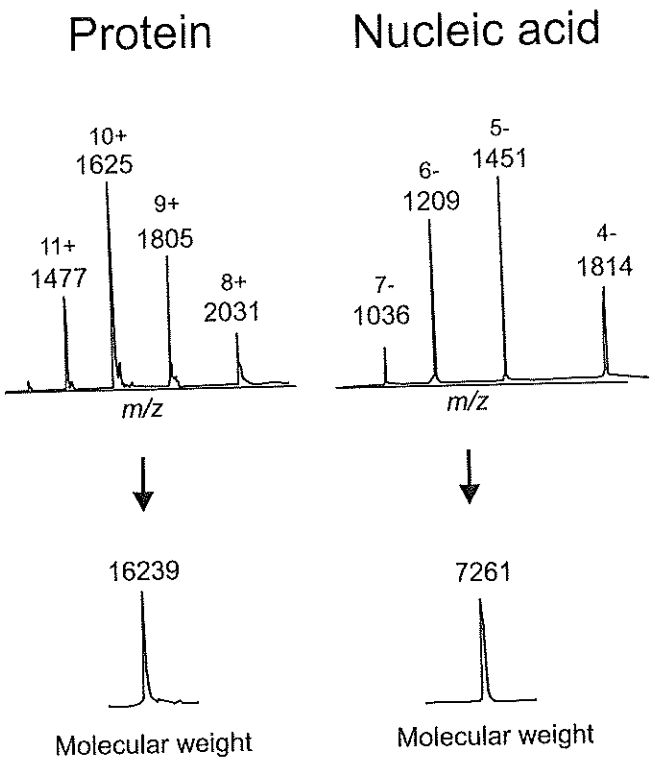

Figure 6-3. Examples of data generated with an ESI mass spectrometer. Proteins usually produce ions with multiple positive charges (*left*), while oligonucleotides generate ions with multiple negative charges (*right*). The number of charges and the m/z-value measured for each peak are given. Below each spectrum is an example of the type of molecular weight information generated by deconvoluting the data above. The molecular weight of the original molecule is given.

The multiple charging that occurs in ESI is a unique and useful characteristic of the technique. It enables a molecule's mass to be determined with great precision, because masses can be calculated independently from several different charged states. The multiple charging also permits the analysis of large molecules, even using conventional mass analyzers that are normally limited to the detection of relatively small molecules. For example, a 70-kDa protein subjected to ESI will typically contain 40–70 charges. This will produce multiply charged species with m/z ratios of between only 1000 and 2000, which can be detected readily with quadrupole mass analyzers (Section 6.3.B).

ESI is the gentlest ionization technique, and **the ionization of a macromolecule by ESI can reflect its conformational properties**. For example, folded proteins usually have smaller numbers of charges than unfolded proteins, and the two forms can be distinguished (Figure 6-4). Complexes in which one molecule binds to another noncovalently (Chapter 19) can survive the ionization, persist in the gas phase, and be detected. In some cases, however, complexes of the macromolecule with salt ions are detected.

Another advantage of ESI-MS is that it is compatible with liquid chromatography (Section 20.2), so molecules emerging from a chromatography column can be injected directly into the mass spectrometer.

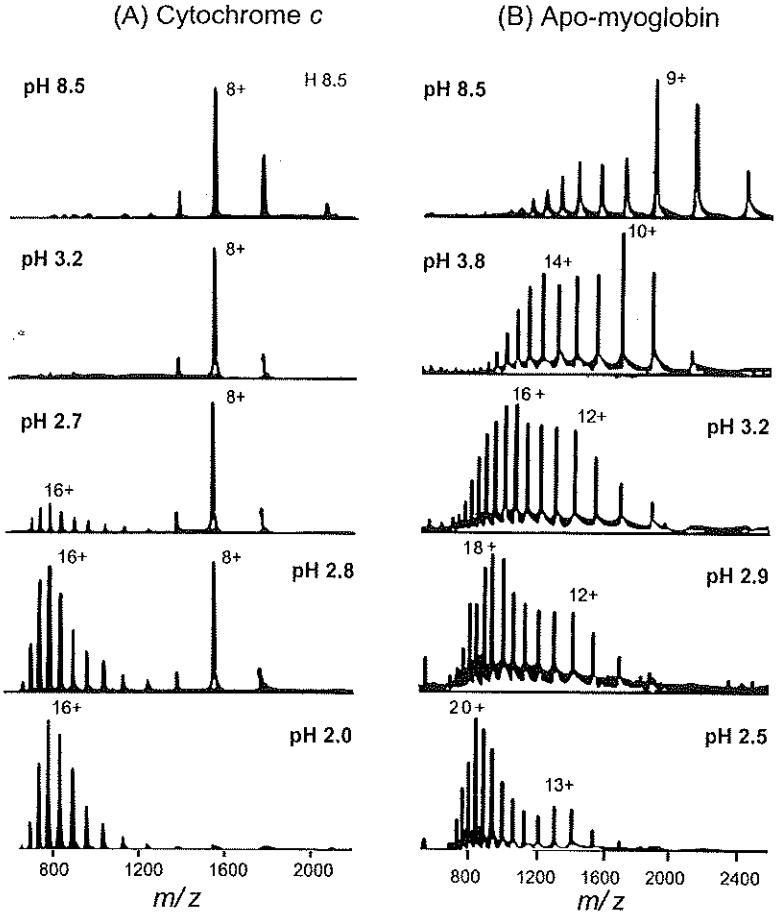

Figure 6-4. ESI spectra of cytochrome *c* (*left*) and apo-myoglobin (*right*) when initially folded and unfolded. The samples were initially at the indicated pH values. At pH 8.5, both proteins are folded in their native conformations. At acidic pH values, however, both unfold. In both cases, the unfolded form has a much greater net charge than the folded form. Native cytochrome *c* molecules generally have a net charge of +8, whereas the unfolded molecules average +16. In the region of pH 2.7, only two populations of molecules appear to be present, either folded or unfolded. Apo-myoglobin, in contrast, appears to have an intermediate form populated from pH 3.8 to 2.5, as the charge spectrum of the molecules gradually shifts. Data from L. Konermann & D. J. Douglas.

Recent developments in electrospray ionisation mass spectrometry: noncovalently bound protein complexes. A. E. Ashcroft (2005) *Nat. Prod. Rep.* **22**, 452–464.

Future directions for electrospray ionization for biological analysis using mass spectrometry. R. D. Smith (2006) *Biotechniques* **41**, 147–148.

Protein structures under electrospray conditions. A. Patriksson *et al.* (2007) *Biochemistry* **46**, 933–945.

6.2. MATRIX-ASSISTED LASER DESORPTION/IONIZATION (MALDI)

Gas-phase ions are generated using MALDI by the vaporization of a mixture of the molecule of interest in a solid matrix upon radiation with intense light from a laser (Figure 6-5). The macromolecule is mixed and embedded in the solid matrix, which often consists of an organic material that absorbs light, such as *trans*-3-indoleacrylic acid, and inorganic salts, such as sodium chloride and silver trifluoroacetate. This solid sample is then irradiated with a laser producing light with a wavelength that is absorbed by the matrix. Short laser pulses of 10–20 ns duration and a power of about 10^6 W/cm^2 eject electronically excited matrix ions, cations and neutral macromolecules into the gaseous phase. The macromolecules become ionized by collisions with small cations, such as H^+, Na^+ and Ag^+. They usually acquire relatively few charges.

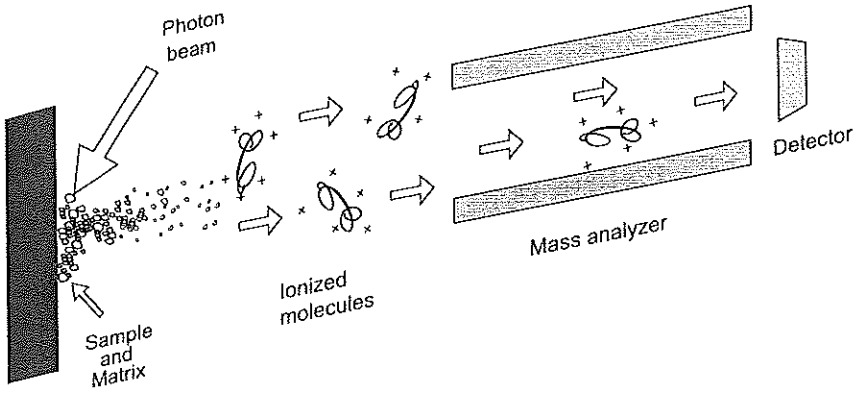

Figure 6-5. The MALDI process. The photon beam is absorbed by the matrix and ejects ionized macromolecules that were embedded in it; they are analyzed in the mass analyzer.

Low molecular-weight molecules of less than 20 kDa typically acquire only one or two charges, while larger molecules can have as many as 3–5, depending on the specific conditions used, the type of matrix material and the power of the laser. The relatively low number of charge states observed in MALDI makes the technique especially well-suited for the analysis of multi-component mixtures, because individual components can be identified easily by the signal generated by their +1 charge state (Figure 6-6). Dimers and trimers of molecules, plus complexes with materials of the matrix, can also be observed in the spectrum. Molecules with masses of up to 300,000 Da can be analyzed.

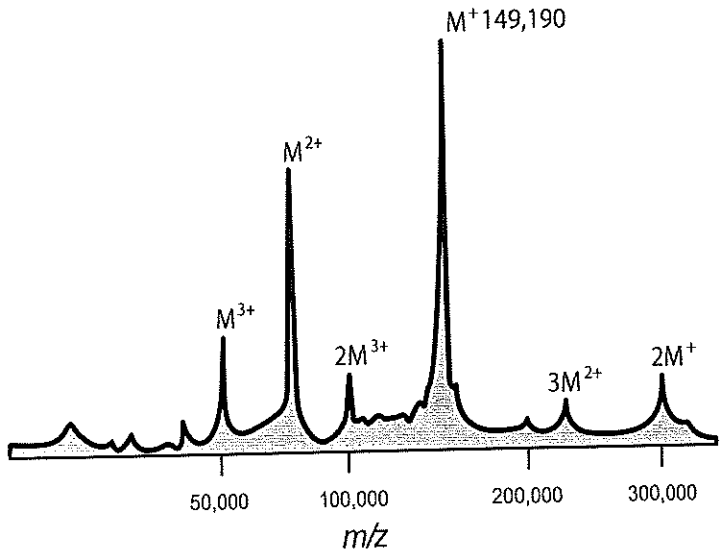

Figure 6-6. The MALDI mass spectrum of a monoclonal antibody. Monomers (M), dimers (2M) and trimers (3M) are apparent, with up to 3+ charges. Data from F. Hillenkamp & M. Karas.

MALDI: more than peptide mass fingerprints. K. Stuhler & H. E. Meyer (2004) *Curr. Opinion Mol. Ther.* **6**, 239–248.

DNA analysis by MALDI-TOF mass spectrometry. I. G. Gut (2004) *Human Mutat.* **23**, 437–441.

MALDI-TOF mass spectrometry: a versatile tool for high-performance DNA analysis. C. Jurinke *et al.* (2004) *Mol. Biotechnol.* **26**, 147–164.

Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry in genomics research. J. Ragoussis *et al.* (2006) *PLoS Genet.* **2**, e100.

6.3. MASS ANALYZERS

Mass analyzers separate ions according to their mass to charge ratio (m/z); they are the most cruc part of a mass spectrometer.

ESI-MS commonly uses **quadrupole** mass analyzers, whereas MALDI-MS uses **time-of-flight** (TO mass analyzers. The resolution offered by TOF mass analyzers is less than that of quadrupole ma analyzers. This, combined with the complication of adduct formation, results in MALDI-MS having lower accuracy, of the order of 0.1%. In contrast, ESI-MS typically has an accuracy of roughly 0.01 Higher resolution mass analyzers, such as the ultra-high-resolution **Fourier-transform ion cyclotro resonance mass analyzer**, to produce FTMS (Section 6.3.D), give accuracies of better than 0.001%.

6.3.A. Magnetic Focusing

The mass spectrometer shown in Figure 6-1 separates ions with varying ratios of m/z by their differen trajectories within a magnetic field. The radius r followed by ions with a particular m/z depends upo the accelerating voltage (V) and the strength of the magnetic field (H) according to:

$$r = \frac{1}{\mathrm{H}} (2\mathrm{V}\frac{m}{z})^{1/2} \tag{6.1}$$

Consequently, varying the strength of the magnetic field permits ions with varying m/z-values to pass through the detector and generate the mass spectrum. Good accuracy requires that all the ions have the same energy after acceleration, which is accomplished in **double-focusing magnetic sector spectrometers** by subjecting the ions to a constant voltage either prior to or subsequent to the magnetic field. This produces high resolution and the ability to analyze molecules with m/z-values of up to 15,000 Da, but these instruments are very expensive.

6.3.B. Quadrupole Mass Analyzers

A quadrupole mass analyzer is depicted in Figure 6-7. Only electric fields, not magnetic, are used to separate the ions. The quadrupole consists of four parallel metal rods, and the ions pass down the middle between them. The four rods are linked in two pairs, A and B, and each pair is connected in order to have the same electrical properties. The two pairs have opposite constant direct-current (dc) voltages. Superimposed on each dc voltage is an oscillating voltage, with opposite phases in the A and B rods. For any particular electrical field, only ions with a specified value of m/z pass between the rods and through the slit to the detector. The other ions are deflected and collide with the rods. All the ions in the sample are detected sequentially by varying the applied voltages, permitting ions with varying values of m/z up to 4000 to be detected.

A quadrupole mass spectrometer can be coupled with an **ion trap**, in which the ions are held within another type of quadrupole and manipulated before being permitted to reach the detector. The ions are physically trapped between the electrodes and subjected to both constant and oscillating electric fields, so that ions of specific m/z precess within the trap. As the magnitudes of the electric fields are varied, ions of various m/z are ejected from the trap and allowed to reach the detector. The ion trap increases both the resolution and the sensitivity, and molecules with masses of up to 100,000 Da can be analyzed with an accuracy of up to 0.003%.

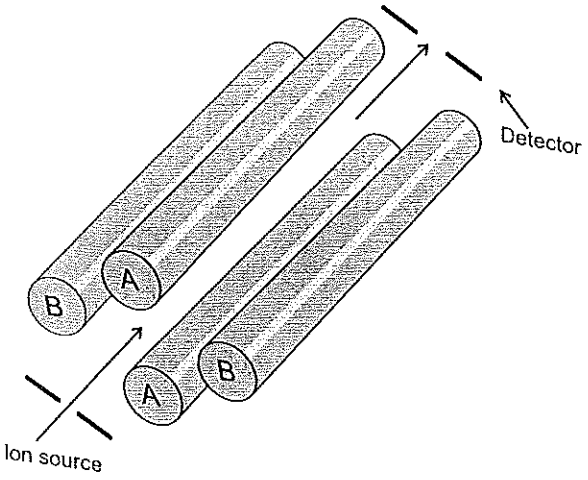

Figure 6-7. A quadrupole mass analyzer. The ions pass between the four parallel rods, which have a fixed voltage, plus one varying with radio frequency. Rods labeled A are connected and have the same voltages, as are those labeled B. The B rods have the opposite fixed voltage to the A rods, and the radio frequency phase is shifted 180° from the A rods. Depending upon the voltages, most ions are deflected; only those with a certain ratio of *m/z* continue through the middle of the rods and pass through the slit to the detector.

Quadrupole and quadrupole ion trap mass analyzers are ideally suited to be coupled with ESI; they have three primary advantages. First, they are tolerant of relatively high pressures of up to 10^{-7} atmospheres, which is well-suited to electrospray ionization, where the ions are produced under atmospheric pressure conditions. Secondly, quadrupoles and ion traps can analyze molecules with an *m/z* ratio of up to 4000, within the range produced by ESI of proteins and other biomolecules. Finally, these mass analyzers are relatively inexpensive, so most successful commercial electrospray instruments use quadrupole mass analyzers.

C-terminal peptide sequencing using acetylated peptides with MS^n in a quadrupole ion trap. A. H. Payne *et al.* (2000) *Analyst* **125**, 635–640.

Tandem mass spectrometry in quadrupole ion trap and ion cyclotron resonance mass spectrometers. A. H. Payne & G. L. Glish (2005) *Methods Enzymol.* **402**, 109–148.

6.3.C. Time-of-Flight (TOF) Analyzers

One of the simplest mass analyzers is the TOF analyzer. A set of ions is accelerated towards a detector with the same amount of energy. If the ions have the same energy, but different masses, they will have different velocities and reach the detector at different times. The smaller ions will reach the detector first because of their greater velocities; the larger ions will take longer. Their mass is determined by the ions' flight time through the analyzer. TOF analyzers are commonly used with MALDI ionization.

The ions are accelerated over a short distance *d* by an electrical field of strength E and then travel through a drift region of length *l*. The time, *t*, required for an ion of mass *m* and charge number *z* to reach the detector at the end of the drift region is given by:

$$t = l\,(m/2z\,\mathrm{e}\mathrm{E}d)^{1/2} \tag{6.2}$$

where e is the unit of fundamental electrical charge. Because all the other parameters are fixed and known, the ratio *m*/z is determined by the time of flight, *t*. The difference in the time of arrival of ions is not great, usually in the microsecond time range, so a complete spectrum can be measured in a very

short time. For example, an accelerating voltage of 20 kV will cause a singly charged ion with a mass of 1000 Da to have a velocity of about 6×10^4 m s^{-1}, and the time to travel 1 m will be 1.7×10^{-5} s.

The ions must enter the flight tube at exactly the same time, which is generally accomplished by generating ions in short bursts, using a pulsed laser with MALDI. The sensitivity and resolution can be increased by slowing the ions with a series of electric field 'lenses' until they stop, then accelerating them in the opposite direction. This 'reflection' increases the path length the ions travel, thereby increasing their separation. The lenses also focus the ions with a specific m/z by reducing the spread in their kinetic energies.

TOF analyzers have the advantages of being very sensitive and able to analyze molecules of essentially any mass.

Tandem time-of-flight mass spectrometry. M. L. Vestal & J.M. Campbell (2005) *Methods Enzymol.* **402**, 79–108.

6.3.D. Fourier-Transform Ion Cyclotron Resonance (FTMS)

FTMS offers high resolution and the ability to perform experiments involving multiple collisions, MS^n, where n can be as high as 4 (Section 6.4), but it requires a cyclotron and a super-conducting magnet. The ions are injected into a small volume in the cyclotron, and a strong magnetic field is applied so that the ions precess in circular orbits that depend upon the magnetic field and their m/z ratio. The ions are kept within the cell by an electric field that is applied to front and rear plates of the sample cell. They are not detected directly, but by their absorption of energy when subjected to an electric field with a frequency that matches their precession frequency, analogous to what happens in magnetic resonance experiments (Section 13.1). The ions transmit a radio frequency current at the detector plates that contains the frequency components of each of the ions. This is converted to a free ion decay signal, which can be transformed into the mass spectrum.

FTMS has the ability to analyze very large molecules, with masses of $> 10^6$ Da. Combined with ESI and MALDI, FTMS offers high accuracy, with errors less than ±0.001%.

Examples of Fourier transform ion cyclotron resonance mass spectrometry developments: from ion physics to remote access biochemical mass spectrometry. A. Rompp *et al.* (2005) *Eur. J. Mass Spectrom.* **11**, 443–456.

Nucleic acid analysis by Fourier transform ion cyclotron resonance mass spectrometry at the beginning of the twenty-first century. J. L. Frahm & D. C. Muddiman (2005) *Curr. Pharm. Des.* **11**, 2593–2613.

Protein primary structure using orthogonal fragmentation techniques in Fourier transform mass spectrometry. R. Zubarev (2006) *Expert Rev. Proteomics* **3**, 251–261.

6.4. TANDEM MASS SPECTROMETRY (MS^N)

The development of tandem mass spectrometry combined with ESI has made determining the complete sequences of biopolymers such as proteins and oligonucleotides routine. ESI generates the intact molecular ion in the gas phase, which is then collided with neutral atoms such as argon or helium.

Collision-induced dissociation (CID) results, producing fragment ions that can be analyzed by their masses. This approach of inducing fragmentation and performing successive mass spectrometry experiments on the fragment ions is known as **tandem mass spectrometry**. It is usually abbreviated as MS^n, where n refers to the number of generations of fragment ions being analyzed (Figure 6-8). The sequence of the peptide can be assembled from the masses of the fragments produced because the collision-induced fragmentation of peptides is well-characterized. Tandem mass spectrometry is used routinely to acquire partial or total sequences of small peptides with fewer than 30 amino acid residues, and of short oligonucleotides.

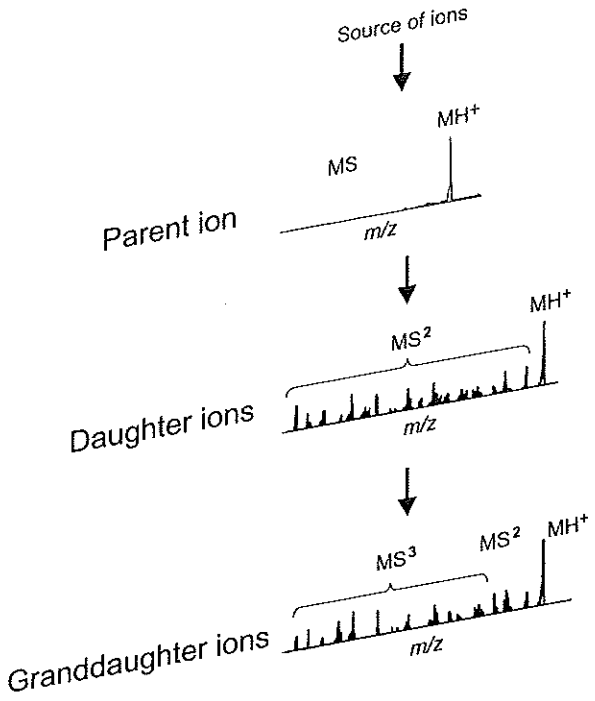

Figure 6-8. Tandem mass spectrometry: generation of fragment ions via collision-induced dissociation and mass analysis (MS^n) of the progeny fragment ions. The parent ion is selected on the basis of its mass. For the MS^2 experiment, the molecular ion MH^+ can be selected by the analyzer and caused to undergo collision-induced dissociation that results in its fragmentation; the products are then analyzed. In an MS^3 experiment, a daughter fragment ion is selected and exposed to collision-induced dissociation, generating granddaughter fragment ions. The terms 'parent', 'daughter' and 'granddaughter' ions are used here, but 'precursor', 'product' and 'second-generation product' ions are also commonly used terms.

Electron capture dissociation (ECD) is a new fragmentation technique that is used in Fourier transform ion cyclotron resonance mass spectrometry (Section 6.3.D) that is complementary to traditional tandem mass spectrometry techniques. Fragmentation is fast and specific, and labile post-translational modifications and noncovalent bonds often remain intact after backbone bond breakage. Disulfide bonds are normally stable to vibrational excitation but are cleaved preferentially in ECD. ECD provides extensive sequence information with polypeptides, and at high electron energies even Ile and Leu residues are distinguishable.

Detection and localization of protein modifications by high resolution tandem mass spectrometry. F. Meng *et al.* (2005) *Mass Spectrom. Rev.* **24**, 126-134.

Tandem mass spectrometry for peptide and protein sequence analysis. J. J. Coon *et al.* (2005) *Biotechniques* **38**, 519, 521, 523.

Analysis of posttranslational modifications of proteins by tandem mass spectrometry. M. R. Larsen *et al.* (2006) *Biotechniques* **40**, 790–798.