CHAPTER 6

Determination of Absolute Molecular Weights Using Sedimentation Equilibrium Analytical Ultracentrifugation

Stephen E. Harding

1. Introduction

One of the most fundamental parameters describing a biological macromolecule is its mol wt, M (unit g/mol), or equivalently the dimensionless "relative molecular mass," M_r. Despite this, it is not always easy to determine or, indeed, for a heterogeneous system, define. With a homogeneous system, the simplest procedure is to calculate it directly from the chemical formula—for example, for a protein whose amino acid composition or sequence is known. In most cases, however, this simple route is not possible; the protein may be glycosylated to an extent that may be difficult to establish precisely, for example, or it may self-associate in solution. Also, the macromolecular system itself may be *polydisperse*, i.e., contain a range of macromolecular species of different mol wt, and in these systems, it is of value to determine the various types of *average* mol wt (usually the number, weight, or z average) or the mol-wt *distribution*. If the system is self-associating, one may also be interested in evaluating the association constant(s).

In molecular biology, the most popular methods for mol-wt determination are the so-called "relative" methods (i.e., requiring calibration using standard macromolecules of known mol wt)—namely, sodium

From: *Methods in Molecular Biology, Vol. 22: Microscopy, Optical Spectroscopy, and Macroscopic Techniques* Edited by: C. Jones, B. Mulloy, and A. H. Thomas
Copyright ©1994 Humana Press Inc., Totowa, NJ

dodecyl gel electrophoresis (SDS-PAGE) and calibrated gel chromatography (GPC, HPLC, and so on)—*see* vol. 1, Chapters 2 and 6 in this series. Both of these methods, although relatively straightforward, have limitations for many systems. Problems may arise from nonuniform binding of SDS in the SDS-PAGE analysis of some proteins (e.g., histones, glycoproteins) or more than just the polypeptide chain mol wt may be desired.

Difficulties may occur in obtaining suitable calibration standards in gel chromatography/high-pressure liquid chromatography analysis. The increasing awareness by the biochemical and molecular biological community of these limitations is a primary reason for the resurgence of interest in absolute techniques for mol-wt measurement, such as light scattering (*see* Chapter 7) and, as we will consider here, sedimentation equilibrium in the analytical ultracentrifuge.

Although the instrumentation is essentially the same as for sedimentation velocity in the analytical ultracentrifuge (*see* Chapter 5), the basic principle of sedimentation equilibrium is somewhat different in that it is not a transport method. In a sedimentation equilibrium experiment, the rotor speed is chosen to be low enough that the forces of sedimentation and diffusion on the macromolecular solute become comparable, so that an equilibrium distribution of solute can be attained; this "equilibrium" can be established after a period of 2–96 h depending on the macromolecule, the solvent, and the run conditions. Since there is no net transport of solute at equilibrium, the recording and analysis of the final equilibrium distribution will give an *absolute* estimate for the mol-wt and associated parameters, since frictional (i.e., shape) effects are not involved. More detailed introductions to the method are given in Price and Dwek *(1)* or van Holde *(2)*, and the experienced user is referred to refs. *3–6*.

2. Summary of Information Available from This Technique

1. Absolute mol-wt values to a precision of up to ±3%.
2. For a multisubunit protein, subunit composition.
3. For heterogeneous systems, an *average* mol wt for the solute distribution: If the absorption or interference optical systems are used, the *weight average*, M^{o}_{w}, is obtained with the highest precision (where the "o" denotes over the whole solute distribution); if the schlieren optical system is used, the most directly obtainable average is the *z-average*, M^{o}_{z}.

For a given optical system, other averages can in principle be obtained, but with lower precision.

4. If the system is heterogeneous, regions nearer the base will have relatively larger amounts of higher mol-wt material than regions near the meniscus. Because of this redistribution of solute throughout the cell, the evaluation of local or "point" average mol wts, M_w, M_z, and so on (corresponding to a particular radial displacement in the centrifuge cell) is also useful, providing different information depending on the origin of the heterogeneity. (A) If the heterogeneity is owing to self-association phenomena, association constants may be determined, or if it is owing to complex formation between species of different types, interaction constants may be estimated but this is very difficult. (B) If the heterogeneity is owing to "polydispersity" (i.e., the presence of noninteracting components of different mol wt or density), mol-wt distributions may be obtained (particularly if a simple combined procedure with gel permeation chromatography can be employed *[6]*). (For further details of these, *see* ref. *7*).

3. Limitations

For large macromolecular systems, nonideality may be significant at the concentrations used, and some form of correction for thermodynamic nonideality effects may be necessary. The measured mol wts at a finite concentration will be "apparent" mol wts. Because of instability of rotor systems at very low speeds, the technique may be unsuitable for large macromolecular assemblies ($M \gtrsim 20 \times 10^6$).

4. Availability of Instrumentation

This is as for sedimentation velocity work, and is covered in Chapter 5.

5. Materials

5.1. Choice of Solvent

As for sedimentation velocity studies (*see* Chapter 5), if possible, work with a solvent of a sufficiently high ionic strength, I, to provide adequate suppression of charge effects. Such effects contribute to the thermodynamic nonideality of the system (*see* Note 4 *below*). For sedimentation equilibrium studies, it is advisable to dialyze the solution against the solvent prior to a sedimentation equilibrium run; the dialysate is used as a reference blank (although for proteins under mild conditions near their isoelectric point, only trivial errors will arise if this is not done): The reasons are related to possible redistribution phenomena of the

(aqueous) solvent components (salt ions etc.) themselves (*see*, e.g., ref. *3*). Care has to be expressed with the choice of centerpiece for the centrifuge cell, since some materials are not resistant to extremes of pH, guanidine hydrodrochloride, dithiothreitol, and so forth.

5.2. Concentration and Volume Requirements for the Macromolecular Solute

This depends on the optical system being used, the path length of the cell (which must be of the "double sector" type if Rayleigh interference optics are used), and the optical properties of the macromolecule itself (refractive index or extinction characteristics): Long path length cells (20–30 mm) are generally appropriate for low concentration (0.1–1.0 mg/mL); short path length cells (10–12 mm) are for higher concentrations (≥1.0 mg/mL). To obtain a solution column length of 0.2–0.3 mm, loading volumes are typically 0.1–0.3 mL, again depending on the path length of the cell.

6. Methods

6.1. Choice of Optical System (see *Chapter 5)*

Rayleigh interference optics (*see* Fig. 1A for an example of the optical record) usually provide the best optical records for mol-wt analysis. For proteins and nucleic acids, the absorption optical system is, however, the most convenient (Fig. 1B), and facilitates the multiplexing of up to four samples in a run (in some cases, even more), although problems of anomalous protein absorption on cell windows have to be avoided *(8)*.

6.2. Length of Run

Smaller molecules get to sedimentation equilibrium faster than larger ones. For molecules of $M \leq 10{,}000$, <24 h are required; large macromolecules take 48–72 h, although for the latter, time to equilibrium can be decreased by "overspeeding," i.e., running at a higher speed for a few hours before setting to the final equilibrium speed (*see*, e.g., ref. *9*). It may, in some applications, be desirable to use shorter columns (as low as 0.5 mm); although the accuracy of the mol wts will be lower, this "short column" method offers the advantage of fast equilibrium (<24 h), which may be important if many samples need to be run and/or the macromolecule is relatively unstable.

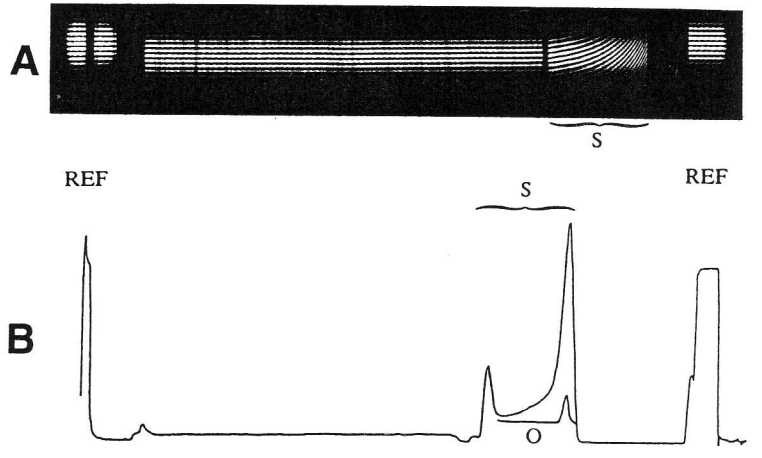

Fig. 1. Optical records of solute distributions at sedimentation equilibrium. The direction of the centrifugal field in both cases is from left to right. (A) Rayleigh interference profiles for human immunoglobulin G. (B) Absorption optical profile (280 nm) for porcine titin. REF: reference marks allowing calibration of the optical records in terms of actual distances from the rotor center. S: solution record. O: (absorption) optical baseline.

6.3. Data Capture and Analysis

If scanning absorption optics are used, the equilibrium patterns can be digitized directly on-line into a microcomputer or off-line via a graphics digitizing pattern. The average slope of a plot of Ln (absorbance) vs r^2, the square of the radial distance from the center of the rotor, will yield the weight average mol wt:

$$M = (d\mathrm{Ln}A/dr^2) \times 2RT/(1 - \bar{\nu}\rho)\omega^2 \quad (1)$$

where $\bar{\nu}$ is the partial specific volume* (typically ~0.73 mL/g for proteins, ~0.61 mL/g for saccharides), ω is the angular velocity (rad/s), and ρ the solution density (in general the solvent density ρ_o can be used instead without giving rise to serious error in M).

With Rayleigh interference optics, the corresponding records of fringe displacement, J, vs radial displacement, r, can be obtained either

*An accurate estimate for $\bar{\nu}$ is normally required since, for proteins, an error of ±1% in $\bar{\nu}$ results in an error of ~ ±3% in M. $\bar{\nu}$ can be found by precision densimetry (*see* ref. *8*) or by direct calculation from the composition data (e.g., for proteins, from the amino acid sequence/composition *[10]*).

manually using "microcomparators," off-line using a laser densitometer *(11)*, or now directly on-line into a microcomputer *(12)*. An average slope of a plot of Ln[J] vs r^2 can be used in much the same way as Ln[A] vs r^2, yielding the weight average mol wt. Various other manipulations can be used to give the number and z-average mol wt *(3)*. If schlieren optics are used, the average slope of a plot of Ln[$(1/r)\cdot dn/dr$] vs r^2, where dn/dr is the refractive index gradient at a given radial position, r, yields the z-average mol wt.

7. Notes

1. Provided that an adequate baseline is available, absorption optical traces (when applicable) give a parameter (the absorbance) directly proportional to solute concentration (within the limits of the Lambert-Beer law). If the Rayleigh interference system is used, each "solution" fringe profile (cf. region "S" of Fig. 1A) corresponds to a plot of solute concentration *relative to the solute concentration at the air/solution meniscus* vs radial distance. To obtain the "absolute" fringe concentration, J, the meniscus concentration needs to be obtained. The various procedures can be found in refs. *3* and *13*. The easiest procedure—where applicable—is to run at sufficiently high speed so as to deplete the meniscus of solute. This "meniscus depletion" method *(14)* is by far the most popular for the analysis of reasonably monodisperse protein systems, although considerable caution has to be expressed, particularly when dealing with heterogeneous systems (*see*, e.g., ref. *3*)
2. For a simple monodisperse system—a situation approached by, for example, a dilute solution of a small enzyme or a polypeptide hormone, plots of Ln[A] or Ln[J] vs r^2 will be linear (Fig. 2A,B): The slope yields the mol wt as indicated earlier.
3. For a heterogeneous system, plots of Ln[A] or Ln[J] vs r^2 will be curved upward (Fig. 2C). A simple average can be taken over the whole distribution (corresponding to the "weight" average, M_w^o), or local slopes can be taken to give point average mol-wt information, such as the point weight average, M_w (Fig. 3). Obtaining the average over the whole distribution can be difficult, particularly if the optical pattern near the cell base is poorly defined: For these systems, and for the nonideal systems considered below, more advanced methods of analyzing the data are recommended (*see*, e.g., ref. *13)* rather than just trying to measure a simple slope.
4. For larger macromolecules ($M \gtrsim 100{,}000$) and/or for more concentrated solutions, nonideality may become significant, and this will tend to cause downward curvature in the Ln[A] or Ln[J] vs r^2 plots. If the

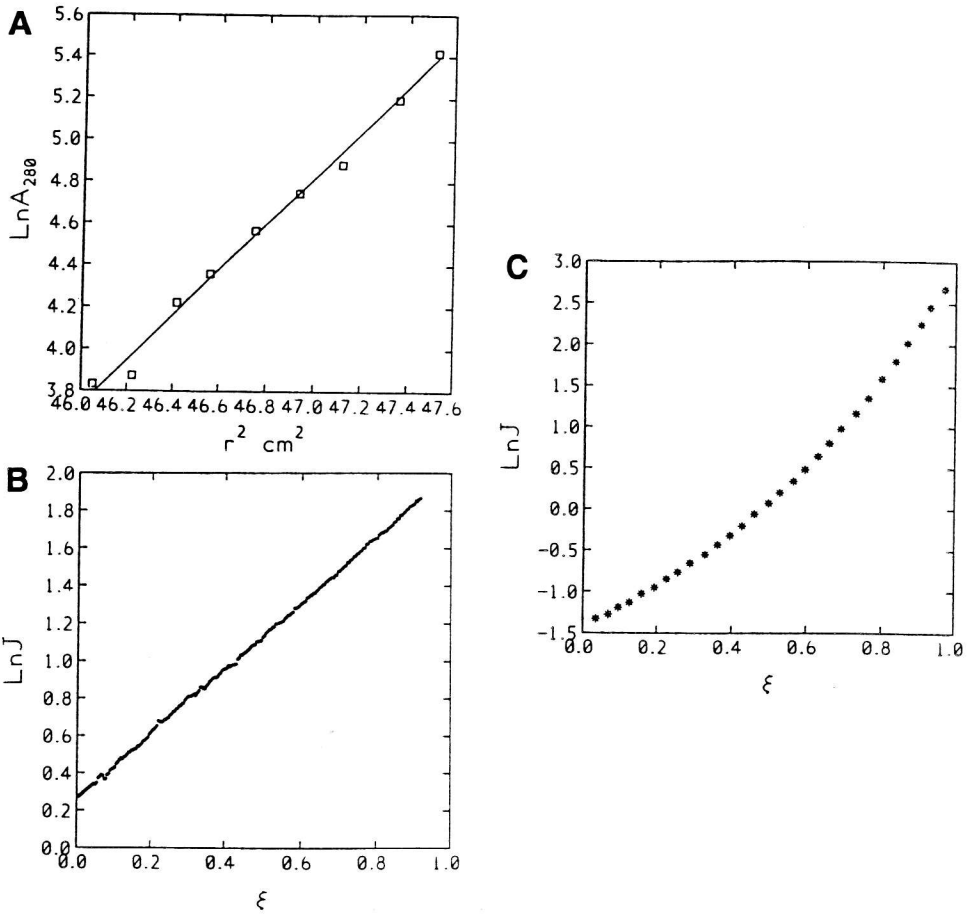

Fig. 2. (A) Log absorbance at 280 nm vs the square of the radial displacement from the center of the rotor, r^2 [for porcine titin, $M^{\circ}_w = (2.5 \pm 0.1) \times 10^6$]. (B) Log fringe concentration, J, vs a function, ξ, of r^2, normalized so it has a value of 0 at the meniscus and 1 at the cell base [$\xi = (r^2 - a^2)/(b^2 - a^2)$] for a small polypeptide, recombinant hirudin. [M_w° (from sedimentation equilibrium) = (7000 ± 200); M (from the amino acid sequence) = 6964.] (C) Log fringe concentration vs ξ for a glycoprotein (BM GRE) from the bronchial mucus of a chronic bronchitis patient. $M_w^{\circ} = (6.2 \pm 0.4) \times 10^6$ (Reproduced from ref. *15*, with permission.)

material is not significantly heterogeneous, then a simple extrapolation from a single experiment of point (apparent) mol wt to zero concentration can be made, to give the infinite dilution "ideal" value (in general, reciprocals are usually plotted—Fig. 4). Whole distribution "number" and "z"-averages, and point number and z-averages can also *in principle* be

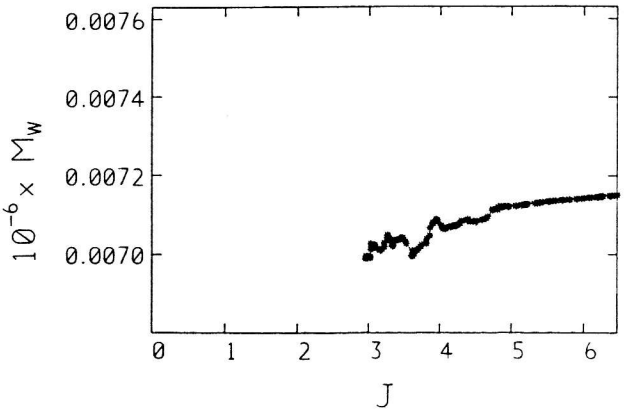

Fig. 3. Point weight average mol wt vs (fringe) concentration plot for recombinant hirudin. Note there is no evidence for dimerization behavior, i.e., M_w is ~ constant across the solute distribution.

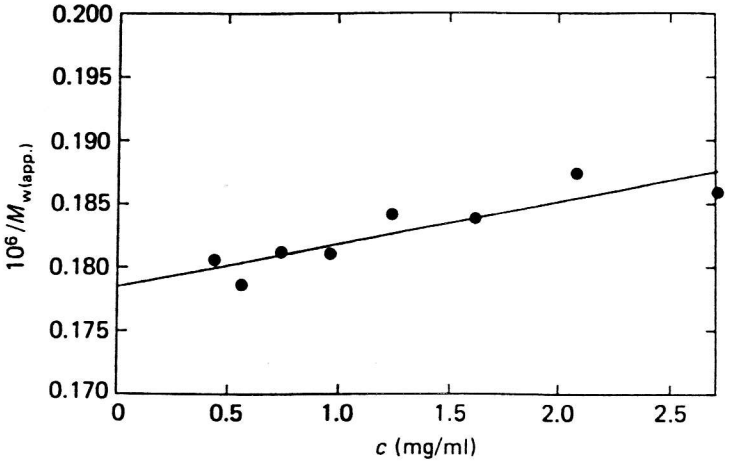

Fig. 4. Plot of the reciprocal of the (apparent) mol wt vs concentration for turnip yellow mosaic virus (TYMV). M_w [from extrapolation of point average values to zero concentration] = $(5.8 \pm 0.2) \times 10^6$; $M°_w$ (the weight average over the whole distribution, uncorrected for nonideality) = $(5.5 \pm 0.2) \times 10^6$. (Reproduced from ref. *16*, with permission).

obtained, although this usually requires data of the highest precision (from, e.g., on- or off-line multiple data acquisition). For this, the reader is referred to more advanced texts (e.g., ref. *7*).

5. For those heterogeneous systems where nonideality is severe, several sedimentation equilibrium experiments performed at different loading concentrations and extrapolation of "whole-cell" mol wt to zero concen-

tration are necessary. For these systems, it is worth adding that the effects of nonideality and heterogeneity can partly cancel each other out and, in some cases, yield a "pseudo-ideal" linear plot of Ln[A] or Ln[J] vs r^2, which can be misleading.

Glossary of Symbols

M, Mol wt (g/mol); M_r, Relative molecular mass; M°_w, Weight average mol wt for the distribution of solute within a centrifuge cell (g/mol); M°_n, Number average mol wt for the distribution of solute within a centrifuge cell (g/mol); M°_z, z-Average mol wt for the distribution of solute within a centrifuge cell (g/mol); M_w, "Point" (i.e., at a given radial position in the centrifuge cell) weight average mol wt (g/mol); M_n, Point number average mol wt (g/mol); M_z, Point z-average mol wt (g/mol); r, Radial displacement from the center of the rotor (cm); A, Absorbance; J, Displacement of the Rayleigh interference fringes (corrected for any finite solute concentration at the air/solution meniscus) normal to the direction of the centrifugal field. At any given radial position in the solute distribution, J is directly proportional to the weight concentration, c (g/mL), of the solute at that radial position; ω, Angular velocity (rad/s); I, Ionic strength (mol/L or mol/mL); $\bar{v}$, Partial specific volume (mL/g); T, Temperature (K); R, Universal gas constant (8.314×10^7 erg/mol/K).

References

1. Price, N. C. and Dwek, R. A. (1978) *Principles and Problems in Physical Chemistry for Biochemists*. Clarendon, Oxford.
2. Van Holde, K. E. (1985) *Physical Biochemistry*. Prentice-Hall, Englewood Cliffs, NJ.
3. Creeth, J. M. and Pain, R. H. (1967) The determination of molecular weights of biological macromolecules by ultracentrifuge methods. *Progr. Biophys. Mol. Biol.* **17,** 217–287.
4. Teller, D. C. (1973) Characterization of proteins by sedimentation equilibrium in the analytical ultracentrifuge. *Methods Enzymol.* **27,** 346–441.
5. Fujita, H. (1975) *Foundations of Ultracentrifuge Analysis*. Wiley, New York.
6. Harding, S. E., Ball, A., and Mitchell, J. R. (1988) Combined low speed sedimentation equilibrium/gel permeation chromatography approach to molecular weight distribution analysis. *Int. J. Biol. Macromol.* **10,** 259–264.
7. Harding, S. E., Rowe, A. J., and Horton, J. C. (eds.) (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Royal Society of Chemistry, Cambridge, UK.
8. Rowe, A. J. (1984) Techniques for determining molecular weight. *Protein Enzyme Biochem.* **BS106,** 1–37.

9. Chervenka, C. H. (1969) *A Manual of Methods for the Analytical Ultracentrifuge*, Beckman Instruments, Palo Alto, CA, p. 43.
10. Perkins, S. J. (1986) Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur. J. Biochem.* **157,** 169–180.
11. Rowe, A. J. and Harding, S. E. (1992) Off-line schlieren and Rayleigh data capture for sedimentation velocity and equilibrium analysis, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., eds.), Royal Society of Chemistry, Cambridge, UK, pp. 49–62.
12. Laue, T. (1992) On-line data acquisition and analysis from the Rayleigh interferometer, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, (Harding, S. E., Rowe, A. J., and Horton, J. C., eds.) Royal Society of Chemistry, Cambridge, UK, pp. 63–89.
13. Creeth, J. M. and Harding, S. E. (1982) Some observations on a new type of point average molecular weight. *J. Biochem. Biophys. Meth.* **7,** 25–34.
14. Yphantis, D. A. (1964) Equilibrium ultracentrifugation of dilute solutions. *Biochemistry* **3,** 297–317.
15. Harding, S. E. (1984) An analysis of the heterogeneity of mucins. No evidence for self-association phenomena. *Biochem. J.* **220,** 117–123.
16. Harding, S. E. and Johnson, P. (1985) Physicochemical studies on turnip yellow mosaic virus: Homogeneity, molecular weights, hydrodynamic radii and concentration dependence parameters. *Biochem. J.* **231,** 549–555.