Protein Aggregation as Studied by Sedimentation Equilibrium

Recent Developments in Instrumentation and Theory

Harald C. Nielsen* and Alfred C. Beckwith

Advances in instrumentation and theory have extended the usefulness of sedimentation equilibrium in studying reversibly aggregating protein systems. The best starting point in the description of an aggregating protein system is a graph of apparent molecular weight *vs.* concentration in a given solvent. Such a graph can be obtained with two or three sedimentation equilibrium experiments, whereas alternative methods, such as light scattering, require many experiments. The first step in the analysis of a sedimentation equilibrium experiment is the determination of protein concentration as a function of radius. Current methods for calculating concentration as a function of radius are described and procedures for calculating apparent molecular weight *vs.* concentration from such data are discussed and possibilities for automation are mentioned. An apparent molecular weight *vs.* concentration curve is useful in determining the molecular weight of the monomer, detecting paucidispersity and thermodynamic nonideality, and calculating association constants.

Sedimentation equilibrium is an extremely useful method for determining the molecular weight of a protein (or other macromolecule) in solution. The method involves centrifuging a protein solution at low speed, which causes a decrease in concentration at the top of the solution column and an increase at the bottom. The system is at equilibrium when no further concentration redistribution takes place. The concentration distribution so produced is used to determine the molecular weight of the protein.

An aggregating protein system exhibits a spectrum of molecular weights depending on concentration and tendency to aggregate. This review will only deal with reversible concentration dependent aggregation. Irreversible aggregation such as that caused by formation of disulfide bonds will not be covered. Since a system at sedimentation equilibrium exhibits a range of protein concentrations and since the apparent molecular weight of an aggregating protein system is a function of concentration, one can in principle calculate the apparent molecular weight of an aggregating protein over the entire concentration range present in the ultracentrifuge cell at sedimentation equilibrium. Real progress toward this goal has been made during the past few years. Progress has also been made in applying apparent molecular weight *vs.* concentration information to describe the reactions taking place.

The study of protein aggregation by sedimentation equilibrium can be divided into three major parts: (1) determination of concentration as a function of radius for a system at sedimentation equilibrium; (2) calculation of apparent molecular weight as a function of concentration; and (3) interpretation of apparent molecular weight *vs.* concentration data in terms of aggregation.

Northern Regional Research Laboratory, U.S. Department of Agriculture, Peoria, Ill. 61604

DETERMINATION OF CONCENTRATION AS A FUNCTION OF RADIUS

A typical ultracentrifuge cell is shown schematically in Figure 1. The size of the cell is such that only 0.2 ml of protein solution is required for a typical sedimentation equilibrium experiment. Protein solution is placed in one sector and reference solvent in the other. The ends of the cell are fitted with windows to enable the protein distribution to be viewed.

The three ultracentrifuge optical systems currently in use are schlieren, Rayleigh interferometric, and absorbance scanner. Typical outputs of these systems for a protein solution at sedimentation equilibrium are shown in Figure 2.

The schlieren system produces a photographic record of the refractive index gradient of the solution and the solvent *vs.* radius. This pattern is directly convertible into values of concentration gradient *vs.* radius. A reasonable initial concentration for a sedimentation equilibrium run with schlieren optics is 5–10 mg/ml. An additional synthetic boundary run and tedious calculations are necessary to convert values of concentration gradient into concentration values (Schachman, 1957, 1959). The synthetic boundary run can be eliminated by making the equilibrium run at a speed high enough so that the concentration at the meniscus is essentially zero. However, this procedure produces steep concentration gradients at the bottom of the liquid column.

The Rayleigh interferometric optical system produces a fringe pattern which indicates the refractive index difference between solution and solvent at various radii. Moving across the representative pattern in Figure 2, each fringe crossed represents roughly a 0.25 mg/ml concentration increment in a 12-mm cell. With Rayleigh optics a typical initial concentration of 2 mg/ml is used, a level which is an advantage since the behavior of a protein approaches ideality with decreasing protein concentration. To convert concentration increments to actual concentrations across the cell, one must

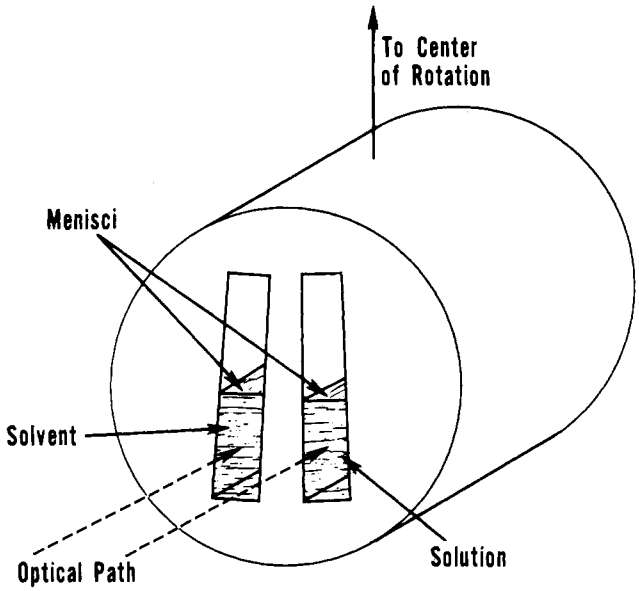

Figure 1. Schematic of an analytical ultracentrifuge cell

either locate the hinge point (Richards *et al.*, 1968) or make an additional synthetic boundary run plus some involved calculations (Van Holde, 1967; Chervenka, 1969).

The ultracentrifuge absorbance scanner has now developed to the point where it offers several important advantages for recording results of sedimentation equilibrium experiments. The output of the scanner is a graph of absorbance of solution *vs.* solvent as a function of radius at a selected wavelength. Values of concentration *vs.* radius are directly obtained from this graph. This method is much easier than the calculations associated with schlieren or interference optical systems. Since almost all proteins strongly absorb around 2800 Å, very low initial protein concentrations can be used (typically 0.5 mg/ml). Data at very low protein concentration are useful in extrapolating the behavior of the system to zero protein concentration. A very important advantage of the absorbance scanner is that its output is an electrical signal. Thus standard electronic equipment can be used to convert the output of the absorbance scanner into a digital form that can be accepted by a computer (Beckwith *et al.*, 1971). This automation bypasses the time-consuming steps of manually measuring and recording chart values and increases accuracy.

Whichever method is used to obtain concentration values through the liquid column, it is important to derive a smooth function of concentration *vs.* radius before making further calculations. For nonaggregating systems, it is common to fit a straight line to values of natural logarithm of concentration (ln c) *vs.* radius squared (r^2) values, *i.e.*,

$$\ln c = A_0 + A_1 (r^2) \quad (1)$$

For aggregating systems, we feel it necessary to use a nonlinear least squares procedure to fit the data. The relative merits of commonly used nonlinear least squares procedures is discussed by Wampler (1969). It is also important to use an equation with some basis in sedimentation theory as a basis for fitting. We have found several aggregating protein systems for which concentration *vs.* radius data can be fit by the equation

$$\ln c + A_2 c = A_0 + A_1 (r^2) \quad (2)$$

(eq 2 describes a system in which the reciprocal of weight-average molecular weight is a linear function of concentration.) A least squares procedure for determining the coefficients (A_0, A_1, A_2) in eq 2 has been described (Beckwith, *et al.*, 1971). More complex systems can be handled by adding power series terms in even powers of radius (r) to eq 2. Figure 3 shows concentration as a function of radius for a typical aggregating protein system. The dots represent every other experimental point and the line is the fitted curve.

CALCULATION OF APPARENT MOLECULAR WEIGHT AS A FUNCTION OF CONCENTRATION

Concentration *vs.* radius data, such as that shown in Figure 2, is used to calculate apparent molecular weight values. An apparent molecular weight value can be calculated at each protein concentration in the solution column at sedimentation equilibrium. Several methods of calculation can be used (Van Holde, 1967). The method most frequently chosen is

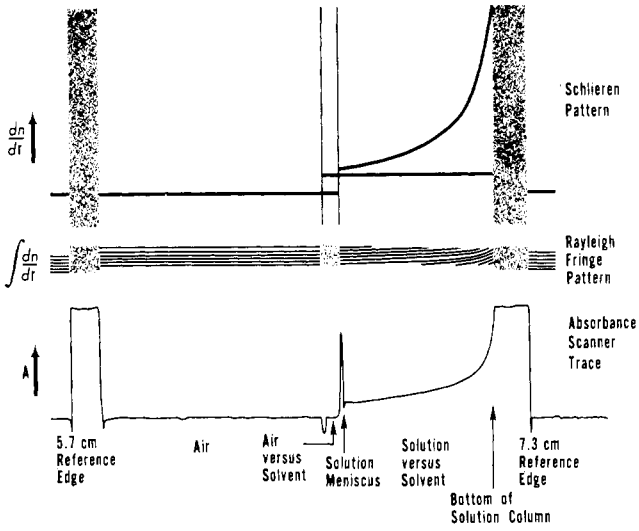

Figure 2. Outputs of the three common ultracentrifuge optical systems for a protein solution at sedimentation equilibrium

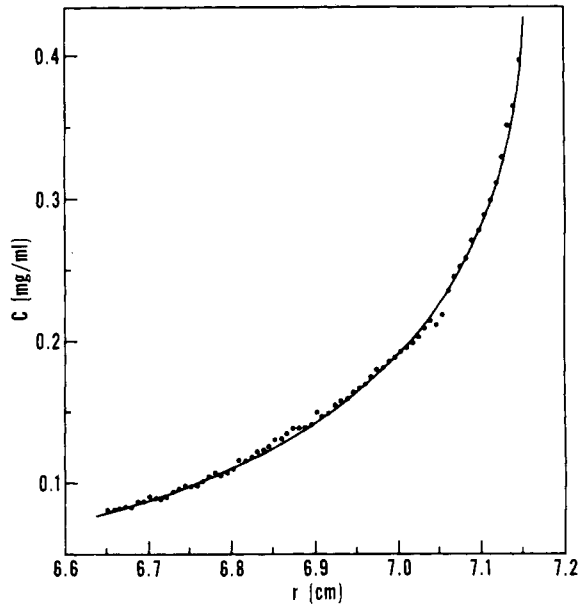

Figure 3. Concentration *vs.* radius plot for an aggregating protein system at sedimentation equilibrium. The dots represent every other experimental point, and the line is the calculated curve

based on the equation

$$\bar{M}_{w\,\mathrm{app}\,(c)} = \frac{2RT}{\omega^2(1 - \bar{v}\rho)} \frac{d \ln c}{d(r^2)} \quad (3)$$

in which $\bar{M}_{w\,\mathrm{app}\,(c)}$ is the apparent weight-average molecular weight of the protein system at concentration c; R is the gas constant (8.314×10^7 ergs/deg/mol); T is the absolute temperature of the system; ω is the angular velocity of the rotor (radians/sec); $\bar{v}$ is the partial specific volume of the protein; ρ is the density of the solution; and r is the radius. The quantity $d \ln c/d(r^2)$ is the slope of a ln c *vs.* r^2 plot at a given concentration. Data for such a plot are obtained by differentiating the equation (*e.g.*, eq 2) which was fitted to the concentration *vs.* radius data. A typical ln c *vs.* r^2 plot for an aggregating protein system appears in Figure 4. (Data for the plot were derived from Figure 3 as described above.) Note that the plot has only a slight curvature. Determination of this curvature places severe demands on the accuracy of the experimental data and method of calculation.

A typical plot of apparent molecular weight *vs.* concentration obtained from a single sedimentation equilibrium experiment is shown in the solid portion (a–b) of Figure 5. This was obtained by applying equation 3 to data similar to that of Figure 4. It should be pointed out that such a curve can be obtained from a heterogeneous nonassociating mixture as well as from an associating system. In order to distinguish between the two possibilities, a second equilibrium experiment is run using a different initial concentration. If one has a system in which only association is taking place, both sets of apparent molecular weight *vs.* concentration values will lie on one curve, an example of which is the solid (a–b) and broken (c–d) portions of Figure 5. If the system is heterogeneous, equilibrium experiments at different initial concentrations will yield separate curves (Squire and Li, 1961). Since heterogeneity is common in associating systems, it is important to run sedimentation equilibrium experiments at two or more different initial concentrations. With presently available ultracentrifuges several cells containing different initial concentrations can be run at the same time.

Plots of apparent molecular weight *vs.* concentration can also be made from either light scattering data or osmotic pressure measurements. These methods, however, only yield one point per experiment.

INTERPRETATION OF APPARENT MOLECULAR WEIGHT *vs.* CONCENTRATION DATA

A useful starting point in describing an aggregating protein system is a heuristic examination of an apparent weight-average molecular weight *vs.* concentration curve combined with other information about the protein, such as conditions of solubility and monomer size. The curve shown in Figure 5 is characteristic of a system containing a 28,000 mol wt monomer that seems to associate without limit. Further information may be obtained from experiments in different solvent systems; *e.g.*, if a protein aggregates in dilute salt solutions but not in strong urea solutions, the aggregation may be due to hydrogen bonding.

One may also try to describe an apparent molecular weight *vs.* concentration curve in numerical terms; that is, write a scheme of association reactions and then derive numerical values for the association constants under defined conditions of solvent and temperature. For the example shown in Figure 5, equations can be written for two monomers going to dimer, dimer plus monomer to trimer, trimer plus monomer to tetramer, etc. At a concentration of 0.2 mg/ml the apparent molecular weight is 50,000 (Figure 5). Assuming that

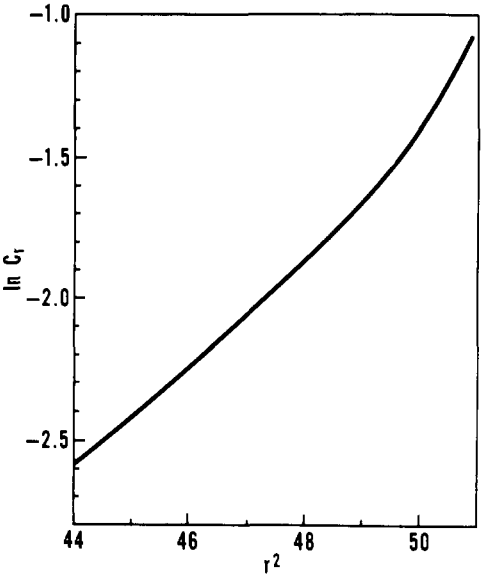

Figure 4. Log concentration *vs.* radius squared plot for an aggregating protein system at sedimentation equilibrium derived from the data of Figure 3

only monomer and dimer are present, the concentration of each can be calculated and then the monomer–dimer association constant can be calculated. At a concentration of 0.3 mg/ml the apparent molecular weight of the system is 82,000. If the presence of monomer, dimer, and trimer is assumed, a dimer–trimer association constant can be calculated. This line of reasoning is used to calculate a set of association constants that fits the entire molecular weight *vs.* concentration curve. In some special cases all the association constants have the same value. It is also possible to write other reaction schemes such as monomer going to dimer going to tetramer, or perhaps monomer going directly to tetramer. Procedures for obtaining numerical values for association constants and for choosing among possible association schemes are complex, and often the available data do not

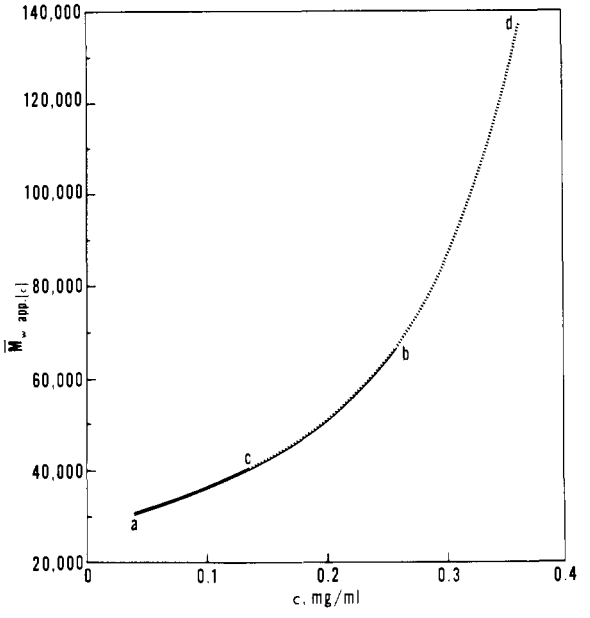

Figure 5. Apparent weight-average molecular weight *vs.* concentration curve for an aggregating protein system. Solid curve (a–b) represents data from one sedimentation equilibrium experiment. Overlapping broken curve (c–d) represents data from another sedimentation equilibrium experiment at a higher initial concentration

allow an unambiguous choice among possible schemes. Methods for calculating association constants are discussed in some detail by Adams (1965, 1967a,b, 1969).

In addition to heterogeneity which was discussed above, thermodynamic nonideality often complicates the analysis of aggregating protein systems. In a nonaggregating system thermodynamic nonideality is revealed by a decrease in $\bar{M}_{w\,app}$ with increasing concentration. There are examples in which thermodynamic nonideality in an aggregating system is the cause of a maximum in a plot of $\bar{M}_{w\,app}$ *vs.* c (Van Holde and Rossetti, 1967; Van Holde *et al.*, 1969; Payens *et al.*, 1969).

Another possible complication in the study of associating systems is the effect of pressure. There is a large increase in pressure from meniscus to bottom of the sedimentating liquid column. Since there is little if any volume change when a protein aggregates, pressure would not be expected to have any effect; however, there are at least two examples in the literature (Kegeles *et al.*, 1969; Josephs and Harrington, 1967) where pressure does affect an aggregating protein system.

SELECTED EXAMPLES OF PROTEIN AGGREGATION

The pioneering work in the mathematical analysis of reversibly aggregating protein systems was that of Steiner (1952). He studied insulin by light scattering at acid pH. His important contribution came in the form of equations that could be used to calculate association constants from apparent molecular weight *vs.* concentration data if the molecular weight of the monomer was known and if thermodynamic nonideality could be ignored. Insulin at acid pH was also studied by Jeffrey and Coates (1966) but by sedimentation equilibrium. Their results were not very different from those of Steiner in that the aggregation involved monomers (mol wt = 5800), dimers, tetramers, and hexamers but no trimers nor pentamers.

Bovine adrenocorticotropin was studied by Squire and Li (1961). They found little association at pH 1.3. At pH 3.5 the data were best explained by an indefinite aggregation; *i.e.*, aggregation continuing without limit with increasing concentration. They also found one of their samples to be paucidisperse and the procedure for detecting this by sedimentation equilibrium is described.

The aggregation of milk proteins is quite complex. Roark and Yphantis (1969) found that β-lactoglobulin A could be separated by gel filtration into two fractions having different aggregation behavior. Payens *et al.* (1969) studied the self association of β-casein at pH 7. They found increasing aggregation with increasing ionic strength, and they also found a high degree of thermodynamic nonideality.

Hemoglobin is an interesting system because it forms a stable tetramer (M = 64,000) over a wide concentration range at neutral pH. In most aggregating protein systems the degree of aggregation increases with increasing protein concentration. Schachman and Edelstein (1966) made the interesting observation that at neutral pH and at the very low protein concentrations which can be studied by the ultracentrifuge absorbance scanner (0.002 mg/ml), the tetramer will dissociate all the way to monomers.

CONCLUSION

Many proteins exhibit reversible aggregation reactions and the quantitative study of these reactions is in a fairly early stage of development. This review has tried to outline some of the important developments in sedimentation equilibrium practice and theory that have taken place over the past few years—developments which have made sedimentation equilibrium one of the important tools in studying protein aggregation.

LITERATURE CITED

Adams, E. T., *Biochemistry* **4,** 1646 (1965).
Adams, E. T., *Biochemistry* **6,** 1864 (1967a).
Adams, E. T., "Fractions," No. 3, Spinco Division of Beckman Instruments, Inc., Palo Alto, Calif., 1967b.
Adams, E. T., *Ann. N.Y. Acad. Sci.* **164,** 226 (1969).
Beckwith, A. C., Nielsen, H. C., Butterfield, R. O., *Anal. Chem.* accepted for publication (1971).
Chervenka, C. H., "A Manual of Methods for the Analytical Ultracentrifuge," Spinco Division of Beckman Instruments, Inc., Palo Alto, Calif., 1969, pp 69–72.
Jeffrey, P. D., Coates, J. H., *Biochemistry* **5,** 489 (1966).
Josephs, R., Harrington, W., *Proc. Nat. Acad. Sci. U.S.A.* **58,** 1587 (1967).
Kegeles, G., Kaplan, S., Rhodes, L., *Ann. N.Y. Acad. Sci. U.S.A.* **164,** 183 (1969).
Payens, T. A. J., Brinkhuis, J. A., Van Markwijk, B. W., *Biochim. Biophys. Acta* **175,** 434 (1969).
Richards, E. G., Teller, D. C., Schachman, H. K., *Biochemistry* **7,** 1054 (1968).
Roark, D. E., Yphantis, D. A., *Ann. N.Y. Acad. Sci.* **164,** 245 (1969).
Schachman, H. K., *Methods Enzymol.* **4,** 32 (1957).
Schachman, H. K., in "Ultracentrifugation in Biochemistry," Academic Press, New York, N.Y., 1959, pp 186–188.
Schachman, H. K., Edelstein, S. J., *Biochemistry* **5,** 2681 (1966).
Squire, P. G., Li, C. H., *J. Amer. Chem. Soc.* **83,** 3521 (1961).
Steiner, R. F., *Arch. Biochem. Biophys.* **39,** 333 (1952).
Van Holde, K. E., "Fractions," No. 1, Spinco Division of Beckman Instruments, Inc., Palo Alto, Calif., 1967.
Van Holde, K. E., Rossetti, G. P., *Biochemistry* **6,** 2189 (1967).
Van Holde, K. E., Rossetti, G. P., Dyson, R. D., *Ann. N.Y. Acad. Sci. U.S.A.* **164,** 279 (1969).
Wampler, R. H., *J. Res. Nat. Bur. Stand. (U.S.)* **73B,** 59 (1969).

Received for review December 1, 1970. Accepted May 17, 1971. Presented at Division of Agricultural and Food Chemistry Symposium on Characterization of Proteins, 160th Meeting, ACS, Chicago, Ill., September 1970. The Northern Regional Research Laboratory is headquarters for the Northern Marketing and Nutrition Research Division, Agricultural Research Service, U.S. Department of Agriculture.