EQUILIBRIUM ULTRACENTRIFUGATION OF POLYMERS: A NEW METHOD OF DATA REDUCTION WITH AN APPLICATION TO TMV PROTEIN

Charles L. STEVENS

Department of Biophysics and Microbiology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

Received 2 June 1975

1. Introduction

Equilibrium centrifugation is a useful method for determining stoichiometry and stability constants of associating macromolecules, and a number of such systems are of considerable biological interest. Very often, however, the most important information on the association must come from the data for which solute concentration is very low and thus relative error is high. Also, to reduce data from such experiments, absolute concentration across the cell must be known. But with the Rayleigh interference system, only concentration differences are directly obtainable.

For the association of identical molecules (polymerization) where certain conditions are met, there is a function of concentration *differences* which appears to overcome these limitations in a straightforward way. The function can be expanded in a power series in a dimensionless positional parameter, X. The expansion has a mathematical form characteristic of the polymerization and takes particularly simple form for monomer–dimer, monomer–trimer, monomer–n-mer and indefinite (condensation) polymerization. Absolute concentration is then obtained as a parameter of the fit. If the polymerization is reversible, stability constants are directly obtainable. Using a simple graphical procedure, weight- and number-average degree of polymerization can also be obtained, and for the latter, there is no general requirement of negligible solute concentration at the meniscus [1,2]. For simple polymerizations, parameters can be evaluated without complicated multivariate computer fits to sums of exponentials. The most conspicuous limitation of the method is that molecular weight of the polymerizing unit must be known beforehand. However, this information usually can be obtained from established procedures. Indeed, in many cases, including the example used here, the subunit molecular weight is known from chemical procedures to an accuracy exceeding that ordinarily obtainable from physical measurements.

Using the procedure outlined below, the earlier conclusion that trimer is the first intermediate in the reassembly of tobacco mosaic virus (TMV) protein [3,4], is shown to be correct.

2. Results

In a centrifugal field, the equilibrium distribution of a solute species is given by equation (1). The species may be in association–

$$C(R) = \sum_{i=1}^{n} C_i(R_0) \exp\left[M_i(1-\bar{\nu}_i \rho)\omega^2(R^2-R_0^2)/2RT\right] \quad (1)$$

dissociation equilibrium with each other [5] but should be otherwise ideal. Here, $C(R)$ is the total weight concentration of solute at the radial position R. Where C is subscripted, C refers to the concentration of the species identified by the subscript. R_0 refers to an arbitrary reference position in the cell. The other symbols have the usual significance*.

* M_i is molecular weight of the ith species, $\bar{\nu}_i$ is partial specific volume of the ith species, ρ is solvent density, ω is angular speed of the rotor, R is the gas constant and T is temperature.

If the species are polymers, M_i equals iM_1; if the solvent is incompressible and $\bar{v}_i$ is the same for all species and independent of pressure, $\bar{v}$ and ρ are constants independent of R and the subscript on $\bar{v}$ can be dropped. The quantity X, a reduced positional variable, is defined by equation (2).

$$X = \exp\,[M_1(1-\bar{v}\,\rho)\omega^2(R^2-R_0^2)/2RT] \qquad (2)$$

We are interested in an expression with concentration differences; let ΔC be defined as $C(R_0) - C(R)$. Using this and equation (1) and (2), one can write $\Delta C = C_1(R_0)(1-X) + C_2(R_0)(1-X^2) + \ldots + C_n(R_0)(1-X^n)$. Finally, using the well-known formula for the sum of a finite geometric series and rearranging, equation (3) is obtained. The coefficients

$$\frac{\Delta C}{1-X} = A_1 + A_2X + \ldots + A_nX^{n-1} \qquad (3)$$

$$A_1 = C(R_0);\ A_j = C(R_0) - \sum_{i=1}^{j-1} C_i(R_0);\ j = 2,3,\ldots,n \qquad (4)$$

A are defined in equation (4) which, together with (3) are the results upon which the present method of analysis is based. Many useful formulae, all of which stem directly from equation (3) and the various definitions, have been collected in table 1.

The TMV protein polymerizing system has been reviewed recently by Lauffer [4]. Polymerization occurs reversibly [3,6] and thermodynamic nonideality is negligible for ordinary conditions of centrifugation [3]. There is a change in $\bar{v}$ of less than

Table 1
Formulae generated from the general centrifuge equation for polymerizing systems

1t. $\lim_{X\to 0} \frac{\Delta C}{1-X} = C(R_0)$;	2t. $\lim_{X\to 1} \frac{\Delta C}{1-X} = C(R_0)\,\bar{N}_w(R_0)$
3t. $C(R) = C(R_0) - \Delta C$;	4t. $\lim_{X\to 1} \frac{dC(R)}{dX} = C(R_0)\,\bar{N}_w(R_0)$
5t. $\bar{N}_n(R_0) = \frac{C(R_0)}{\int_0^1 \frac{C(R)}{X}dX}$;	6t. $\lim_{X\to 0} \frac{C(R)}{X} = C_1(R_0)$
7t. For a monomer–dimer association:	$\frac{\Delta C}{1-X} = C(R_0) + C_2(R_0)X$
8t. For a monomer–trimer association:	$\frac{\Delta C}{1-X} = C(R_0) + C_3(R_0)X(1+X)$
9t. For a monomer–n-mer association:	$\frac{\Delta C}{1-X} = C(R_0) + \frac{C_n(R_0)X(1-X^{n-1})}{(1-X)}$
9ta. For large n and X below 1:	$\frac{\Delta C}{1-X} \cong C(R_0) + C_n(R_0)\,\frac{X}{1-X}$
10t. For condensation polymerization:	$\frac{\Delta C}{1-X} = C(R_0)\left[\frac{1-P^2X}{(1-PX)^2}\right]$ where $P = KC_1(R_0)$

$\Delta C/(1-X) = A_1 + A_2X + A_3X^2 + \ldots + A_nX^{n-1}$ for which $A_1 = C(R_0)$, $A_j = C(R_0) - \sum_{i=1}^{j-1} C_i(R_0)$;

where $j = 2,3,\ldots n$. Also $A_1 \geqslant A_2 \geqslant \ldots \geqslant A_n$ $\bar{N}_w(R_0)$ and $\bar{N}_n(R_0)$ are weight- and number-average degree of polymerization respectively, each prevailing at the reference position R_0. K is the association constant for condensation polymerization. The other parameters are defined in the texy.

1% upon polymerization [7] but this is not large enough to affect the calculations. A sample of TMV protein was sedimented to equilibrium such that all polymers except the smallest were at the bottom of the cell and not observed. The parameters X, ΔC and $\Delta C/(1-X)$ were calculated as functions of radial position R. In fig.1, $\Delta C/(1-X)$ is plotted against both X (open symbols) and $X(1+X)$ (filled symbols). For the upper set of curves, R_0 was taken near the cell bottom, but in the lower set, R_0 was moved centripetally to where concentration dropped to about half. It is clear that the plots are linear in $X(1+X)$ at all values of X except 0.8 and above. With reference to equations 7t and 8t of table 1, this shows that the polymerization is largely monomer–trimer at the concentration appropriate to both reference positions. Actually, data simulation has shown that up to about 15% dimer relative to trimer could be present without noticeable departure of the curves from linearity. From the least squares lines of the figure, the calculated values of the apparent association constant, $K_{1,3}$, for the upper and lower graphs are the same; the value, 7.1×10^9 l^2 $mole^{-2}$, compares well with that obtained by Westover under similar conditions but using a different analysis procedure [3,4].

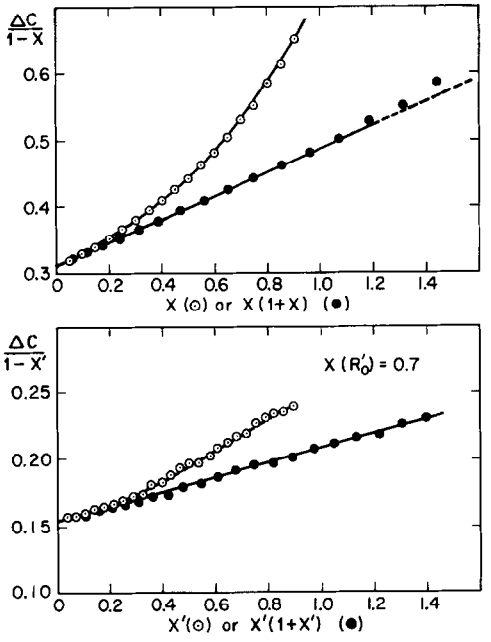

Fig.1. TMV protein, run XIX. Subunit mol. wt is 17 500. The function $\Delta X/(1-X)$ is plotted against both X, (open symbols) and against $X(1+X)$, (filled symbols). The reference position, R_0, was taken both at 7.118 cm, (upper graph) and at 7.098 cm, (lower graph). (Only points centripetal to R_0 are used). The lines through the filled symbols are least square lines. Upper graph: intercept = 0.311 ± 0.001; slope = 0.178 ± 0.001. Lower graph: intercept = 0.153 ± 0.001; slope = 0.056 ± 0.001. All units are mg/ml. This is a run with a long solution column (7.2 mm) and a long optical path, (30 mm centerpiece). Centrifuge speed was 34 000 rev/min; equilibrium at 4°C was achieved in 5 days. Solute distribution was obtained from Rayliegh interference patterns after subtracting the solvent blank. Fringe displacement per unit solute concentration was determined using a synthetic boundary cell. TMV protein was dialyzed against 0.005 M phosphate buffer, pH 6.0 containing 0.1 M KCl, and loaded at a concentration of 0.5 mg/ml.

These data can be used also as an example of the calculation of weight- and number-average degree of polymerization from equations 1t, 2t, 3t, and 5t of table 1. From the upper graph of fig.2, one obtains the value of about 0.77 for the limit of $\Delta C/(1-X)$

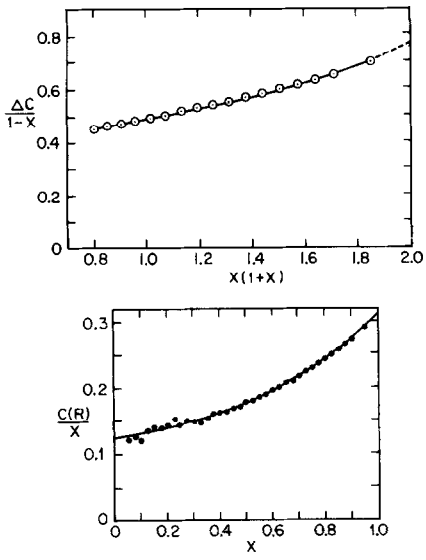

Fig.2. TMV protein, run XIX. Upper graph: data are the same as for the upper graph of fig.1. They are plotted here, however, in a form more convenient for evaluating $\Delta C/(1-X)$ as $X \to 1$. Lower graph: from the same data, $C(R)$ was calculated according to equation 3t of table 1 and $C(R)/X$ plotted against X for the evaluation of the integral in equation 5t.

as $X \to 1$. (Here, the function is plotted against $X(1 + X)$ rather than X merely to make the extrapolation easier.) From fig.1, $C(R_0)$ is found to be 0.31, giving a value of 2.5 for $\bar{N}_w(R_0)$. The value of the integral in equation 5t was found from the lower graph of fig.2, 0.190 ± 0.006 mg/ml, which gives a value of 1.6 for $\bar{N}_n(R_0)$.

3. Discussion

When one is assured that the rather rigid prerequisites are met by the system under investigation, the method outlined here can be a valuable supplement to more general techniques already available. The utility of the method is illustrated by the TMV protein system; the nature of the polymerization is immediately apparent from the form of the curves.

Although a complete development of the method including applications will be reported elsewhere, a few points are especially worthy of note. It is generally desirable to choose conditions of centrifugation that yield values of X near 0 for points near the meniscus. This allows for the evaluation of $C(R_0)$ by a short extrapolation. From equation (2) one sees that this occurs with a long solution column and high speed; the optimum must be chosen with regard to time available for equilibration, subunit molecular weight, sample dispersity and strength of the association.

From definition, $A_1 \geqslant A_2 \geqslant \ldots \geqslant A_n$; if an experiment does not yield data with this property, no meaningful interpretation is possible and the data must be re-examined or discarded.

Although the error in $C(R_0)$ tends to 0 as X tends to 0, it should be emphasized that a ± 10% error in the value of X is sufficient to make the curve for monomer–dimer polymerization noticeably nonlinear, and will result in a ∓ 10% error in both weight- and number-average degree of polymerization. For most experiments, random errors do not get unmanageable until X becomes greater than about 0.8 or more. This can be a problem in evaluating these averages, but as fig.2 shows, it can be done without complication for some data. Usually, the analysis can be improved by some form of data smoothing.

Acknowledgements

This work was made possible by a research grant, GM 10403, from the National Institutes of Health. The author gratefully acknowledges the technical assistance of Sanda Loga. Thanks are expressed also to Max A. Lauffer for his encouragement and for many helpful discussions on the TMV protein system.

References

[1] Yphantis, D. A. (1964) Biochemistry 3, 297–317.
[2] Roark, D. E. and Yphantis, D. A. (1969) Ann. N.Y. Acad. Sci. 164, 245–278.
[3] Westover, C. J., Jr. (1971) Ph.D. Dissertation, University of Pittsburgh.
[4] Lauffer, M. A. (1971) in: Biological Macromolecules IV, (Timasheff, S. N. and Fasman, G. F., eds.) pp. 149–199, Marcel Dekker.
[5] Adams, E. T., Jr. (1969) Ann. N.Y. Acad. Sci. 164, 226–244.
[6] Smith, C. E. and Lauffer, M. A. (1967) Biochemistry 6, 2457–2465.
[7] Stevens, C. L. and Lauffer, M. A. (1965) Biochemistry 4, 31–37.