# Equilibrium Ultracentrifugation of Dilute Solutions*

DAVID A. YPHANTIS

*From the Rockefeller Institute, New York, N. Y.*

*Received October 24, 1963*

This paper describes a technique of determining molecular weight by equilibrium ultracentrifugation that is especially applicable to the study of paucidisperse solutions. The ultracentrifuge is operated at speeds about three times higher than usual so that the meniscus concentration may be neglected. Concentrations in the cell are then determined as being proportional to the refractive index differences between the meniscus region and the points of interest. The results of such high speed operation are (a) the ability to study lower initial concentration than usual, 0.003–0.2%, (b) a centrifugal fractionation of disperse systems which, in many cases, permits estimation of the size of the smallest macromolecular component present, and (c) moderate equilibrium times. The experimental and theoretical aspects of the technique are discussed in detail: Cells are described for the simultaneous observation of three solution-solvent pairs; procedures are presented for calculations of molecular weight averages and their errors; and the pitfalls of the method, especially the possibility of convection, are discussed. Illustration of the technique is provided by experiments on "good" β-lactoglobulin and on samples of bovine plasma albumin, ribonuclease-A, and southern bean mosaic virus that are contaminated with components larger than the species of interest. The technique appears most useful in studies of limited distributions, such as monomer-polymer systems. The low concentrations required and the possibility of extracting information about the size of the smallest species present suggest application to studies of subunits, and the moderately large concentration range spanned with a single initial concentration suggests application to studies of nonideality.

Equilibrium ultracentrifugation is one of the more attractive methods of determining molecular weight. The method is firmly based on thermodynamics, it covers a broad spectrum of sizes, and requires only small quantities of material. Recent advances in optics (Trautman and Burns, 1954; Richards and Schachman, 1959) and the demonstration by Van Holde and Baldwin (1958) that the time required to attain equilibrium need not be a serious limitation have made equilibrium ultracentrifugation a practical and common biochemical tool.

The concentration range usually employed with refractometric optics, from 0.1% up, is satisfactory for most routine characterizations. However, many interesting systems exhibit strong concentration dependence, some because of association, others as a result of repulsive interactions. In such cases it is desirable to be able to explore as low concentrations as possible with the view in mind not only of determining sizes of molecules, but also of measuring the strengths of their interactions. There are also many systems where it is difficult to obtain even the few milligrams required for the equilibrium techniques or for which it is awkward to prepare sufficiently concentrated solutions. Here again it would be useful to be able to examine dilute solutions. The automatic scanning absorption system recently reported by Schachman *et al.* (1962) and Schachman (1963) provides an almost ideal method of examining such low concentrations. However, until this system is commonly available, equilibrium studies at low concentrations will probably rely on the refractometric optics.

The apparent weight-average molecular weight, $\bar{M}_{w,app}$ of a system is given by:

$$\bar{M}_{w,app} = \frac{c(b) - c(a)}{c_o} \frac{2}{b^2 - a^2} \frac{RT}{\omega^2(1 - \bar{V}\rho)} \quad (1)$$

In this equation $c_o$ is the initial concentration, and $c(a)$ and $c(b)$ are the concentrations at the radii $a$ and $b$ corresponding, respectively, to the meniscus and base of the solution column, $R$ is the gas constant, $T$ is the absolute temperature, $\omega$ is the angular velocity, $\bar{V}$ is the partial specific volume of the solute, and $\rho$ is the density of the solution. Customarily the use of refractometric optics depends on the observation that for dilute solutions the refractive index of the solution is proportional to concentration. The value of $c(b) - c(a)$ is determined in terms of arbitrary refractive index units, making allowance for the effects of the centrifugal field on the solvent and the cell, and the value of $c_o$ is determined as the refractive increment of the initial solution in the same units as $c(b) - c(a)$. An implicit assumption is made in this procedure: the assumption of conservation of solute. Should material be removed from solution in the interval between the determination of the initial refractive increment and the equilibrium observation, then the value of the apparent weight-average molecular weight determined will be too low. There are a number of ways that material may be removed from solution in this interval; a common one is the aggregation of unstable components with subsequent centrifugation out of solution. Such precipitation usually is evident by a thickening of the base meniscus during a run. Another way of removing solute from solution is by adsorption onto the cell walls and interfaces. Ordinarily only an insignificant fraction is adsorbed, but as more dilute solutions are examined the fraction *may* become significant. Experiments with multichannel cells (Yphantis, 1960) which have high surface area–to–volume ratios (~110 $cm^{-1}$) have indicated that significant adsorption may occur with some solutes at concentrations of ~0.05% and below. Low-concentration experiments in the usual double-sector cells appear subject to the same uncertainty of solute depletion by adsorption. These cells also have a comparatively high surface area–to–volume ratio (~50 $cm^{-1}$) when operated with 3-mm columns of solution. Further, the presence of a few per cent of aggregated material may be difficult to detect as a thickening of the base

* This work has been supported by a grant (A-2493) from the National Institute of Arthritis and Metabolic Diseases, U. S. Public Health Service. Part of this work was presented at the 140th meeting, Chicago, Ill., 1961, and at the 141st meeting, Washington, D. C., 1962, of the American Chemical Society.

simply because of the small quantities involved at low concentrations.

The absorption optics give concentrations directly. On the other hand the refractometric optics, as usually employed, give only measures of concentration differences within the cell. In systems where conservation of solute cannot be guaranteed one needs a precise concentration reference at some point in the cell in order to determine the effective initial concentration to be used in equation (1) or its analogs. This paper deals with the establishment of such a concentration reference and the application of the resulting system.

In equilibrium ultracentrifugation the speed is usually chosen (Svedberg and Pedersen, 1940) so that the concentration at the base of the cell is about three times that at the meniscus. In such a procedure the concentration increases gradually throughout the cell. If the speed is now increased by a factor of about three, the concentration ratio across the cell becomes about $2 \times 10^4$. The concentration near the meniscus becomes virtually independent of position and may be neglected in comparison with the initial concentration. Thus it becomes easy to establish a known reference concentration in the cell, *provided* the speed is chosen so that the meniscus concentration of the smallest component of the system is negligible. Concentrations in the cell may then be determined simply as being proportional to the readily observable refractive index differences between the region near the meniscus and the points of interest.[1]

There are a number of consequences of such high speed equilibrium experiments. The marked concentration of the solution near the base of the cell allows measurements on more dilute initial solutions than is customary. The degree of concentration is a strong function of the effective molecular weight of the species making up the solution and therefore leads to a partial size fractionation. In favorable cases this fractionation is sufficient that estimates of the size of the smallest species present may be obtained, even though the initial solution is paucidisperse. This appears to be a useful feature in the characterization of systems where the species of interest is present along with its polymers, for example in studies of protein subunits.

Since the number of solute molecules at the meniscus under these conditions is a negligible fraction of the total present in the cell, one can obtain estimates of the number-average molecular weight, along with the usual determination of the other moments of the molecular weight distribution. The form of many of the equations for these higher molecular weight averages and their evaluation become simpler when the meniscus concentration approaches zero. The large concentration range spanned in a single experiment proves useful in studying concentration-dependent systems. A further result of high speed operation is a significant decrease of the time required for attainment of equilibrium.

Most of the work with this method has been carried out with the Rayleigh interference optics rather than the schlieren optics, primarily because measurement of the interference patterns appears less subjective at the lower concentrations and because the interference optics allow a direct correction for small optical distortions in the cells. This work has been limited primarily to the use of 3-mm columns of solution, i.e., 3 mm between the meniscus and base of the solutions. Such a column height appears convenient for many purposes, giving moderate resolution with convenient equilibrium times. Further, limiting the column height to 3 mm has allowed the construction of cells for observing three solution-solvent pairs simultaneously. However, longer columns offer some distinct advantages. For example, higher initial concentrations may be employed with the interference optics. Also lower molecular weights may be studied by lengthening the solution column.

[1] A somewhat similar method, usually employing lower speeds, has been used by Wales *et al.* (1951) for the characterization of polydisperse polymer systems. They calculated the meniscus concentrations essentially by successive approximations, assuming particular types of molecular weight distributions.

## Theoretical

It is convenient to introduce a limiting parameter $\sigma$ called the *effective reduced molecular weight* (Yphantis and Waugh, 1956). The value of $\sigma$ is defined as $\omega^2 M(1 - \bar{V}\rho)_o/RT$, where $(1 - \bar{V}\rho)_o$ is evaluated at infinite dilution and at a pressure of 1 atm. This parameter is a convenient abbreviated index of the behavior of a monodisperse solute at infinite dilution at the speed of the experiment in question.

The usual differential equation, ignoring pressure dependence, for two component systems[2] (Van Holde and Baldwin, 1958) is

$$\frac{1}{rc}\frac{dc}{dr} = \frac{d \ln c}{d(r^2/2)} = \frac{M(1 - \bar{V}\rho)}{1 + c\left(\frac{\partial \ln y}{\partial c}\right)_{P,T}}\frac{\omega^2}{RT} = \frac{\sigma\frac{(1 - \bar{V}\rho)}{(1 - \bar{V}\rho)_o}}{1 + c\left(\frac{\partial \ln y}{\partial c}\right)_{P,T}} = \sigma_w(r) \quad (2)$$

where $y$ is the activity coefficient of the solute on the $c$ scale and $P$ is the pressure. Strictly speaking, values of $\sigma$ (and therefore of molecular weight) may be evaluated from experimental data only in the limit of zero concentration unless the values of the nonideality term, $c(\partial \ln y/\partial c)_{P,T}$, are known. However, a number of parameters, essentially ratios of various derivatives of the concentration distributions as a function of position, may always be directly evaluated from the experimental data. One of these is indicated in equation (2) as $\sigma_w(r)$; it is the familiar ratio $\omega^2 s/D$, where $s$ and $D$ are the sedimentation and diffusion coefficients (La Bar and Baldwin, 1963). The value of $\sigma_w(r)$ is, in general, a function of position and concentration. In the limit as the pressure approaches 1 atm. and as the concentration approaches zero, the value of $\sigma_w(r)$ obviously approaches the value of $\sigma$ for a monodisperse system; for a disperse system the limit would be the weight-average value of $\sigma$ existing at $r$. The other observable factors will be defined after a short digression.

With aqueous solutions of macromolecules it is usual to assume that the solute and solvent have the same compressibility. Further, at the low concentrations of the present method the density increment from the solute may be neglected. Under these two assump-

[2] Systems of biological interest are generally composed of at least three components: water, the macromolecular species, and the supporting electrolyte(s). For convenience of interpretation these systems are usually treated as two-component systems: the aqueous solution of electrolyte(s) is considered the solvent component and the macromolecular species the solute component. Such a treatment appears valid when there is no preferential interaction between the macromolecular species and the solvent components. An appropriate definition of components permits description of systems with concentration-independent, preferential interaction as two-component systems (Casassa and Eisenberg, 1961).

tions the value of the term $(1 - \bar{V}\rho)$ is independent of pressure and of concentration and may be taken equal to its value at infinite dilution and at atmospheric pressure. The only remaining difference between $\sigma_w(r)$ and $\sigma$ in equation (2) is then the nonideality term $c(\partial \ln y/\partial c)_{P,T}$. In many cases this term also may be neglected, leading to significant simplification. With this additional assumption, that $c(\partial \ln y/\partial c)_{P,T} \ll 1$, equation (2) then may be integrated to give the familiar exponential expression $c = A \exp(\sigma r^2/2)$ where the constant $A$ has the value $0.5c_o\sigma(b^2 - a^2)/[\exp(\sigma b^2/2) - \exp(\sigma a^2/2)]$ if conservation of mass is invoked in a sectoral cell.

*Molecular Weight Averages.*—Disperse systems may be treated simply if they satisfy the above criteria and are ideal in showing no interactions between the various components and no preferential interactions with the solvent. The total concentration at any point $r$ may then be written as the sum of the exponential expressions for the $n$ macromolecular components of the system:

$$c(r) = \sum_{i=1}^{n} c_i(r) = \sum_{i=1}^{n} A_i \exp(\sigma_i r^2/2) \qquad (3)$$

From this relation and the definitions of the various averages (Svedberg and Pedersen, 1940) the following may readily be shown for these *ideal* systems: the number average $\sigma$ is:

$$\begin{aligned}\sigma_n(r) &= \sum_{i=1}^{n} c_i(r) \Big/ \sum_{i=1}^{n} c_i(r)/\sigma_i \\ &= \sum_{i=1}^{n} A_i \exp(\sigma_i r^2/2) \Big/ \sum_{i=1}^{n} \frac{A_i}{\sigma_i} \exp(\sigma_i r^2/2) \qquad (4a) \\ &= \frac{c(r)}{\int_{r=a}^{r} c(r)\, d(r^2/2) + \sum_{i=1}^{n} \frac{A_i}{\sigma_i} \exp(\sigma_i a^2/2)}\end{aligned}$$

where the sum in the last term is proportional to the number of solute molecules per unit volume at the meniscus $a$. The weight average $\sigma$ is:

$$\begin{aligned}\sigma_w(r) &= \sum_{i=1}^{n} \sigma_i c_i(r) \Big/ \sum_{i=1}^{n} c_i(r) \\ &= \sum_{i=1}^{n} A_i \sigma_i \exp(\sigma_i r^2/2) \Big/ \sum_{i=1}^{n} A_i \exp(\sigma_i r^2/2) \qquad (4b) \\ &= \frac{1}{c(r)} \frac{dc}{d(r^2/2)} = \frac{d \ln c}{d(r^2/2)}\end{aligned}$$

The $z$-average $\sigma$ is:

$$\begin{aligned}\sigma_z(r) &= \sum_{i=1}^{n} \sigma_i^2 c_i(r) \Big/ \sum_{i=1}^{n} \sigma_i c_i(r) \\ &= \sum_{i=1}^{n} A_i \sigma_i^2 \exp(\sigma_i r^2/2) \Big/ \sum_{i=1}^{n} A_i \sigma_i \exp(\sigma_i r^2/2) \\ &= \frac{d^2c}{d(r^2/2)^2} \Big/ \frac{dc}{d(r^2/2)} \qquad (4c)\end{aligned}$$

and the $(z + q)$ average $\sigma$ is:

$$\sigma_{z+q}(r) = \frac{d^{q+2}c}{d(r^2/2)^{q+2}} \Big/ \frac{d^{q+1}}{d(r^2/2)^{q+1}} \qquad (4d)$$

where $q$ is any positive integer. These $\sigma_k(r)$ are the various averages of the effective reduced molecular weights existing at the point $r$. If all the ideal macromolecular solutes have the same values of $\bar{V}_i = \bar{V}$, then these $\sigma_k(r)$ correspond, after multiplication by the factor $RT/\omega^2(1 - \bar{V}\rho)$, to the various actual molecular weight averages at the point $r$. Note that if concentrations are determined by refractometric techniques, then the concentrations used in evaluating equations (4) will not be on a weight basis but on a refractive index basis and the $\sigma_k(r)$ will reflect any differences in the specific refractive increments of the components (Van Holde and Baldwin, 1958).

In general the number-average effective molecular weight cannot be evaluated directly since the sum $\sum_{i=1}^{n} \exp(\sigma_i a^2/2) A_i/\sigma_i$, which is essentially proportional to the number of molecules per unit volume at the meniscus, is unknown. However, if the meniscus concentration may be neglected with respect to the concentration at some point $r'$ further toward the base of the cell and if the molecular weight of the species of lowest $\sigma$ is large enough so that the number of molecules per unit volume at the meniscus is also negligible, then the sum may be set equal to zero and equation (5) used to determine the number average effective molecular weights, $\sigma_n(r)$, for the points centrifugal to the point $r'$.[3] If the effective molecular weight of the smallest species, $\sigma_1$, is large enough the point $r'$ will not be far from the meniscus $a$. However, if $\sigma_1$ is so small that there is an appreciable number of molecules at $a$, then use of equation (5) becomes somewhat similar to an osmotic pressure experiment with a membrane partially permeable to some component(s) of the system.

$$\sigma_n(r) \cong c(r) \Big/ \int_{r=a}^{r} c(r)\, d(r^2/2) \qquad (r > r') \qquad (5)$$

The relations of equations (4) and (5) are derived only for ideal systems. Their derivation is not valid for concentration-dependent systems. It is useful, however, to *define* the $\sigma_k(r)$, for the general case, as the experimental observables suggested by equations (4b,c,d) and (5), i.e., as $\sigma_k(r) \equiv D^k c(r)/D^{k-1}c(r)$ where $D^k$ is the operator $d^k/d(r^2/2)^k$ for $k \geq 0$ and $\int_a^r c(r)\, d(r^2/2)$ for $k = -1$ and where $k = 0, 1, 2$ for the number, weight, and $z$-averages. With a monodisperse nonideal solute the various $\sigma_k(r)$ will reflect the effects of nonideality in different ways. In the general case of dispersity and concentration dependence these $\sigma_k(r)$ will have no simple direct interpretations but appropriate extrapolations will be useful for extracting information.

$$\overline{\sigma_n} = \int_{r=a}^{b} c(r) d(r^2/2) \Big/ \int_{r=a}^{b} c(r)/\sigma_n(r)\, d(r^2/2) \qquad (6a)$$

$$\begin{aligned}\overline{\sigma_w} &= \frac{\int_{r=a}^{b} c(r)\sigma_w(r)\, d(r^2/2)}{\int_{r=a}^{b} c(r)\, d(r^2/2)} \\ &= \frac{\int_{r=a}^{b} c(r)\left[\frac{1}{c(r)} \frac{dc(r)}{d(r^2/2)}\right] d(r^2/2)}{\int_{r=a}^{b} c(r)\, d(r^2/2)} \\ &= \frac{c(b) - c(a)}{\int_{r=a}^{b} c(r)\, d(r^2/2)} \qquad (6b)\end{aligned}$$

[3] Wales (1951) and Wales *et al.* (1951) calculated number-average molecular weights by approximation of the sum, $\sum_{i=1}^{n} \exp(\sigma_i a^2/2) A_i/\sigma_i$, in equation (4a) by a number of methods, recommending the use of high speeds so as to minimize the effects of the approximation. With paucidisperse systems it is often feasible to use speeds high enough to permit ignoring the sum completely.

Usually the quantities of interest are not these point values of the molecular weight averages, but rather averages of these quantities over all the contents of the cell so as to characterize the initial solution as a whole. These averages over the whole cell will be distinguished by a superscript bar. Their values are given in equations (6) for sector-shaped cells (Svedberg and Pedersen, 1940; Lansing and Kraemer, 1935). If the meniscus concentration is negligible it may be seen from equation (5) that equation (6b) reduces to

$$\overline{\sigma_w} = \sigma_n(b) \qquad (6b')$$

i.e., the weight-average effective molecular weight of the complete solution is simply the number average obtained at the base of the cell. The $z$-average effective molecular weight of the entire solution is

$$\overline{\sigma_z} = \frac{\int_{r=a}^{b} \frac{dc}{d(r^2/2)}\,\sigma_z(r)\,d(r^2/2)}{\int_{r=a}^{b} \frac{dc}{d(r^2/2)}\,d(r^2/2)} = \frac{\int_{r=a}^{b} \frac{dc}{d(r^2/2)}\left[\frac{d^2c}{d(r^2/2)^2}\Big/\frac{dc}{d(r^2/2)}\right] d(r^2/2)}{\int_{r=a}^{b} dc} = \frac{\left[\frac{dc}{d(r^2/2)}\right]_{r=b} - \left[\frac{dc}{d(r^2/2)}\right]_{r=a}}{c(b) - c(a)} \qquad (6c)$$

If both the meniscus concentration and its gradient may be neglected, this reduces to:

$$\overline{\sigma_z} = \sigma_w(b) \qquad (6c')$$

Similarly it may be shown that

$$\overline{\sigma_{z+q}} = \sigma_{z+q-1}(b) \qquad (6d')$$

Direct determinations of concentration and especially of concentration gradients are not feasible at the end points of the cell. Some form of extrapolation of these quantities is usually required to estimate their values for use in equations (6). However, equations (6′) suggest the extrapolation of the derived quantities $\sigma_k(r)$ which show much less variation with position than the primary data.

Note that the $\overline{\sigma_k}$ will refer to the initial solution only if there is no selective removal of components from solution, for example, if adsorption or precipitation may be neglected or if the adsorption or precipitation affects all components equally on a weight basis. It should be pointed out that the $\overline{\sigma_k}$ are only apparent values without a simple direct meaning if the disperse systems exhibit concentration dependence. An extrapolation to infinite dilution is then required so that the $\overline{\sigma_k}$ will characterize the effective reduced molecular weight distributions (Williams *et al.*, 1958; Fujita, 1962).

*Nonideality or Dispersity?*—Examination of an experiment at a single initial concentration and at a single speed will not distinguish between an association equilibrium or the presence of polymers *not* in chemical equilibrium with each other (Svedberg and Pedersen, 1940). For example, the distribution

$$c(r) = A_1 \exp(\sigma_1 r^2/2) + A_2 \exp(\sigma_1 r^2) \qquad (7)$$

would be given by two independent ideal components with $\sigma_2 = 2\sigma_1$. Exactly the same distribution would be given by a monomer-dimer system in chemical equilibrium with no volume change on dimerization, provided the components were otherwise ideal and provided the association constant for the dimerization were equal to $M_1 A_2/2A_1^2$ where $M_1$ is the monomer molecular weight and the concentrations in equation (7) are given in g/liter. Simple nonideality may compensate for the effects of dispersity so that a preparation may even appear ideal and homogeneous when in fact it is both disperse and nonideal. Dispersity and nonideality may be differentiated if the same solution is examined at a number of sufficiently different speeds or if more than one initial concentration is run at a common speed. Both approaches have their advantages, but the simplicity of running more than one concentration at the same time and the possibility of instability and time dependence with many biological systems favor the latter approach.

Consider first a disperse solution of $n$ ideal macromolecular components: the concentration distribution for such a system is given by equation (3) which is rewritten below after insertion of the values of the $A_i$:

$$c(r) = \frac{c_0(b^2 - a^2)}{2} \sum_{i=1}^{n} \frac{\sigma_i f_i \exp(\sigma_i r^2/2)}{\exp(\sigma_i b^2/2) - \exp(\sigma_i a^2/2)} \cong \frac{c_0(b^2 - a^2)}{2} \sum_{i=1}^{n} \sigma_i f_i \exp[\sigma_i(r^2 - b^2)/2] \qquad (3')$$

Here $f_i = (c_i/\sum_{i=1}^{n} c_i)_0$ is the (weight) fractional concentration of the $i$th component in the original solution and the approximation on the right-hand side is valid only if the meniscus concentrations for all the $n$ components may be neglected. If the initial concentrations $c_0$ are varied keeping the same values of $a$ and $b$ then the point values of the various molecular weight averages $\sigma_k(r)$ will be identical functions of position. In general, values of the $\sigma_k(r)$ for experiments at the same speed may be compared directly only if the ratio $a/b$ is maintained the same for all runs and if the $\sigma_k(r)$ are expressed as functions of the variable $\xi = (b^2 - r^2)/(b^2 - a^2)$. However, if the meniscus concentrations are negligible the $\sigma_k(r)$ for a purely disperse system are identical functions of the simpler variable $r^2 - b^2$ for any cell loading, i.e., for any values of $a$, $b$, and $c_0$.

In purely concentration-dependent systems, whether associating or nonideal, the apparent point-average molecular weights $\sigma_k(r)$ calculated from equations (4) are identical functions only of $c(r)$, the concentrations at $r$. These $\sigma_k(r)$ show a completely different behavior when plotted against $r^2 - b^2$. The curves for different initial concentrations, although similar, are shifted toward the base of the cell as the initial concentration is lowered. If these $c$-dependent $\sigma_k(r)$ are plotted as a function of the concentration $c(r)$ at the points $r$, then the various curves will be coincident as far as they extend. The same plot for the purely disperse systems will give a different curve for each initial concentration. For a purely disperse system these curves will be similar to each other insofar as they extend, except that the scale factor for concentration in each case will be proportional to the quantity loaded for each run. Systems that are both disperse and concentration dependent will exhibit behavior between these two extremes.

*Estimation of the Size of the Smallest Species Present.*—Consider a paucidisperse ideal system of $n$ components for which the total concentration at any point is given by equation (3′). The ratio of the concentration of the component of lowest $\sigma$ ($= \sigma_1$) to that of any other component $j$ is given by

$$c_1(r)/c_j(r) = (A_1/A_j) \exp[(\sigma_1 - \sigma_j)r^2/2] \cong \frac{\sigma_1 f_1}{\sigma_j f_j} \exp[(\sigma_j - \sigma_1)(b^2 - r^2)/2] \qquad (8)$$

where the approximation on the right is valid when the meniscus concentrations are negligible. It is obvious from equation (8) that the regions closest to the meniscus will be enriched in the amount of the smallest component relative to any of the others, the extent of enrichment depending strongly on the difference $\sigma_j - \sigma_1$. If this difference were great enough and if the relative amount of the smallest component were high enough in the starting solution, one could then observe the smallest component virtually uncontaminated by the larger components in the centripetal region of the useful portion of the distribution. This may be realized in many instances, especially with preparations where a monomer is contaminated with polymers higher than dimer or trimer. However, there is often a significant overlap of the various components and a *direct* examination of the logarithmic concentration graphs then allows only rough estimation of the size of the smallest component. With systems made up of a component and its polymers, as in many biochemical preparations, one can often still obtain reliable estimates of the size of the monomer species from the point-average molecular weights. The feasibility of such a procedure depends on the presence of significant ratios of monomer to only one of its polymers in a region of the cell. Perhaps the most direct way of estimating monomer size would be to express the experimental concentrations in a form such as equation (7) for the monomer-dimer case, i.e., as the sum of two exponential curves. When the type of polymer, whether dimer, trimer, etc., is known this procedure requires the determination of three constants, $A_1$, $A_n$, and $\sigma_1$, which is laborious to do properly because of the exponential nature of the functions and which may be misleading if done improperly. Such a procedure is even more hazardous if the type of polymer is unknown, requiring a four-constant curve-fitting exercise.

Another possible procedure is the solution for $\sigma_1$ of the simultaneous equations prescribed by the various point-average molecular weights. This appears feasible for many cases, but suffers from the large multiplication of errors entailed in the solution of such simultaneous equations. The procedures that have been adopted are simple to apply and in the most commonly encountered type of system, the monomer contaminated with dimer, are unaffected by a moderate linear concentration dependence.

Let us assume that we have a region where only the two components 1 and $j$ are present in significant amounts and that these components behave ideally with $\sigma_j = j\sigma_1$. Then the total concentration at a point $r$ will be given by equation (3′) and the effective average molecular weights will be given by:

$$\sigma_n(r) = \frac{f_1\sigma_1 \exp[\sigma_1(r^2 - b^2)/2] + f_j\sigma_j \exp[\sigma_j(r^2 - b^2)/2]}{f_1 \exp[\sigma_1(r^2 - b^2)/2] + f_j \exp[\sigma_j(r^2 - b^2)/2]}$$

$$= \sigma_1 \frac{1 + \frac{jf_j}{f_1} \exp[(j-1)\sigma_1(r^2 - b^2)/2]}{1 + \frac{f_j}{f_1} \exp[(j-1)\sigma_1(r^2 - b^2)/2]}$$

$$= \sigma_1 \frac{1 + j\mu\lambda}{1 + \mu\lambda} \cong \sigma_1[1 + (j-1)\mu\lambda - (j+1)\mu^2\lambda^2 + \cdots] \quad (9a)$$

and similarly

$$\sigma_w(r) = \sigma_1 \frac{1 + j^2\mu\lambda}{1 + j\mu\lambda} \cong \sigma_1[1 + j(j-1)\mu\lambda - j^2(j+1)\mu^2\lambda^2 + \cdots] \quad (9b)$$

$$\sigma_z(r) = \sigma_1 \frac{1 + j^3\mu\lambda}{1 + j^2\mu\lambda} \cong \sigma_1[1 + j^2(j-1)\mu\lambda - j^4(j+1)\mu^2\lambda^2 + \cdots] \quad (9c)$$

where for convenience the quantity $\exp[(j-1)\sigma_1(r^2 - b^2)/2]$ has been set equal to $\lambda$ and $\mu$ represents $f_j/f_1$, the ratio of the concentrations of the two components in the initial solution. Note that $\lambda \leq 1$ and that the product $\mu\lambda$ is the ratio $c_j(r)\sigma_1/c_1(r)\sigma_j$ of the number of $j$-mer molecules to those of monomer at the point $r$. If $j^2\mu\lambda$ is small compared to unity then these $\sigma_k(r)$ may be expanded in the form of the series given at the right of equations (9). The total concentration at a point $r$, equation (3′), is given in this notation by

$$c(r) = B\lambda^{1/j-1}(1 + j\mu\lambda) \quad (10)$$

where $B = c_o(b^2 - a^2)/2$. Comparison of equations (9a, b, c) and (10) suggests expressing the $\sigma_k(r)$ as a power series in the total concentration to the $(j-1)$ power so that

$$\sigma_k(r) = \alpha_{1k} + \alpha_{2k}c(r)^{j-1} + \alpha_{3k}c(r)^{2j-2} + \cdots$$
$$= \alpha_{1k} + \alpha_{2k}B^{j-1}\lambda + \alpha_{2k}(j^2 - j)B^{j-1}\mu\lambda^2 + \cdots + \alpha_{3k}B^{2j-2}\lambda^2 + \cdots \quad (11)$$

where values of $c(r)$ have been inserted from equation (10) and terms of order higher than $\lambda^2$ have been omitted. The coefficients $\alpha_{ik}$ may be determined by equating the coefficients of the same powers of $\lambda$ in equations (9) and (11). This yields the following values for the $\alpha_{ik}$: The $\alpha_{1k}$ are all equal to $\sigma_1$, the $\alpha_{2k}$ are equal to $(j-1)j^k\sigma_1\mu/B^{j-1}$ and the $\alpha_{3k}$ are equal to $-(j^3 - 2j^2 + j + j^{k+1} + j^k)j^k\sigma_1\mu^2/B^{2j-2}$, where $k = 0$, 1, and 2 for the number, weight and $z$-averages. Thus at sufficiently low values of $c(r)/c_o$ all the point-average molecular weights are linear functions of $c(r)^{j-1}$ and all the point-average molecular weights extrapolate to the weight of the monomer species at zero concentration, allowing determination of the monomer molecular weight.

Another useful procedure is suggested by the simple relations between the limiting slopes of the $\sigma_k(r)$ as a function of $c(r)^{j-1}$. Consider the relations

$$\sigma_1 \cong \frac{j\sigma_n(r) - \sigma_w(r)}{j - 1} \quad (12a)$$

$$\sigma_1 \cong \frac{j^2\sigma_n(r) - \sigma_z(r)}{j^2 - 1} \quad (12b)$$

$$\sigma_1 \cong \frac{j\sigma_w(r) - \sigma_z(r)}{j - 1} \quad (12c)$$

All these relations give estimates of $\sigma_1$ without requiring extrapolation. They are valid only in a restricted range: for example with a monomer-dimer system equation (12a) is valid to within $0.05\sigma_1$ if the weight fraction of the dimer at the points considered is less than about 29% and with a monomer-trimer system if the weight fraction of trimer is less than ~25%; equation (12c) is valid to within $0.05\sigma_1$ if the dimer weight fraction is less than ~17% or if the trimer weight fraction is less than ~10%, again at the points considered. Note that reversibly associating systems which yield identical forms of concentration distributions as fixed monomer-polymer systems may be treated in this way.

It may be shown that either the extrapolations of the $\sigma_k(r)$ or the use of equations (12) will yield good estimates of $\sigma_1$ in nonideal monomer-dimer systems if the term $c(\partial \ln y/\partial c)_{P,T}$ for the monomer component is proportional to concentration and is not too large, say, less than ~0.2. Such nonideality could lead to incorrect estimates of $\sigma_1$ in other systems, however,

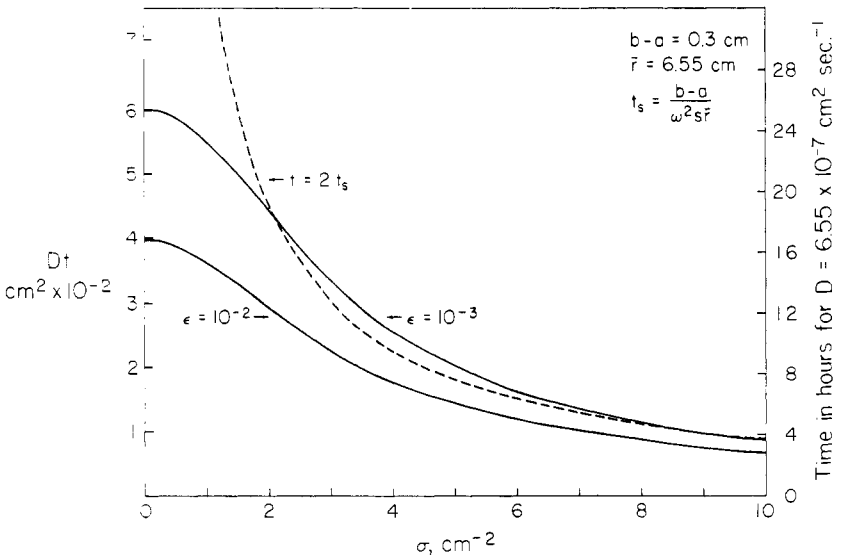


FIG. 1.—Calculated equilibrium times as a function of $\sigma$ ($= \omega^2 s/D$) for a 3-mm solution column. The dashed curve corresponds to the classical estimate of twice the time required for a molecule to sediment across the cell. See text for the definition of $\epsilon$.

unless guarded against.[4] It is strongly recommended that more than one initial concentration be examined and that the $\sigma_k(r)$ obtained be plotted against $c(r)^{j-1}$ and against $r^2 - b^2$ in each case. The value of $j$ may be estimated from the power of $c(r)$ giving the best fit. Other methods of determining $j$ and $\sigma_1$ will be discussed in a future publication.

*Equilibrium Times.*—Van Holde and Baldwin (1958) presented equations for the time required for the concentration difference across the cell, $c(b) - c(a)$, to be within a fraction $\epsilon$ of its equilibrium value. Their relations are valid even for the ranges of parameters considered here, as shown by comparison with concentration distributions calculated from the rectangular approximation they used. Figure 1 presents equilibrium times calculated from their equations (11 and 12) for $\epsilon = 10^{-2}$ and $10^{-3}$ for a 3-mm column as a function of $\sigma$ with the times presented in generalized form as $Dt$ and for a typical case with $D = 6.55 \times 10^{-7}$ cm² sec$^{-1}$. It is obvious that considerably less time is required for equilibrium at the higher speeds, with a reduction of about 60% in the equilibrium time on increasing $\sigma$ from zero to the values used in this study ($\sim$5 cm$^{-2}$).

There is a question as to what criteria should be used in estimating equilibrium times. In experiments with comparatively high initial concentrations the concentration difference across the cell, $c(b) - c(a) = c_o\sigma(b^2 - a^2)/2$, is large and $10^{-3}$ times $c(b) - c(a) = c(b)$ is still beyond experimental error. For example, an extremely high initial concentration of 10 fringes, when $(b^2 - a^2) = 4$ and $\sigma = 5$ cm$^{-2}$ gives $10^{-3}c(b) = 0.1$ fringe, a quantity about four times the experimental error of measurement in the section of the cell where the fringes are clearly resolved. An examination of Archibald's (1942) exact solution of the differential equation for transient ultracentrifugation shows that the first eigenfunction, which is primarily responsible for the deviations from equilibrium at large times, has the following characteristics (see Table 2.2 in Fujita, 1962): the eigenfunction has a large negative value near the base of the cell and goes to zero at the so-called hinge point (where the equilibrium concentration equals the initial concentration), then it has a smaller positive maximum and decreases to a small positive value near the meniscus. Examination of the corresponding eigenfunctions in the rectangular approximations shows that they too have the same behavior, and that the ratio of the value of the positive maximum to that at the base is less than $^1/_5$ if $\sigma \geqslant 3$ cm$^{-2}$. In the experiment given as an example the observations would be restricted to regions where the concentrations are less than the initial concentration (10 fringes), and thus to regions centripetal to the hinge point. Therefore the observable deviations from equilibrium will be less than one-fifth those at the base, and the criterion of $\epsilon = 10^{-3}$ is valid for such an experiment.

On the other hand consider an experiment where all of the solution is to be observed, e.g., for determination of the $\bar{\sigma}_k$. The initial concentration should then be below $\sim$0.4 fringe so that the concentration gradients near the base of the cell will be low enough to allow measurements there. For the same $\sigma$ and $b^2 - a^2$ as in the previous example, the base concentrations will be below $\sim$4 fringes. For this case a value of $\epsilon = 10^{-2}$ gives deviations which have a maximum value at the base of the cell of less than 1.6 times the observational error. One can conclude, therefore, that the criterion $\epsilon = 10^{-3}$ is in general adequate, and that if necessary one could use a criterion of $\epsilon = 10^{-2}$ for experiments at the lowest initial concentrations.

It is useful to know the variation of equilibrium time with the column height. Equilibrium times calculated for $\sigma = 5$ cm$^{-2}$, $\epsilon = 10^{-2}$ and $10^{-3}$, and for column heights up to 1 cm are given in Figure 2. The times for $b - a \sim 0.3$ cm are seen to be essentially proportional to $(b - a)$ instead of to $(b - a)^2$ as they are for experiments at low speeds (Van Holde and Baldwin, 1958).

[4] The use of $\sigma_1 = 2\sigma_n(r) - \sigma_w(r)$ appears feasible in many cases even if the true value of $j$ is not known or if there is a moderate amount of linear concentration dependence. For example, if the larger component has a $\sigma$ less than $2.7\sigma_1$ this expression will give estimates of $\sigma_1$ valid to $0.05\sigma_1$ provided no more than 22% of the distribution at the points considered is composed of the heavier component. If the $\sigma$ of the larger component is greater than $2.7\sigma_1$ the amount of polymer that may be present and still permit estimation of $\sigma_1$ in this manner is much less; e.g., if $\sigma_j = 3\sigma_1$ only $\sim$10% and if $\sigma_j = 4\sigma_1$ only $\sim$3% of the heavier component may be present and still obtain values of $\sigma_1$ valid to 5%.

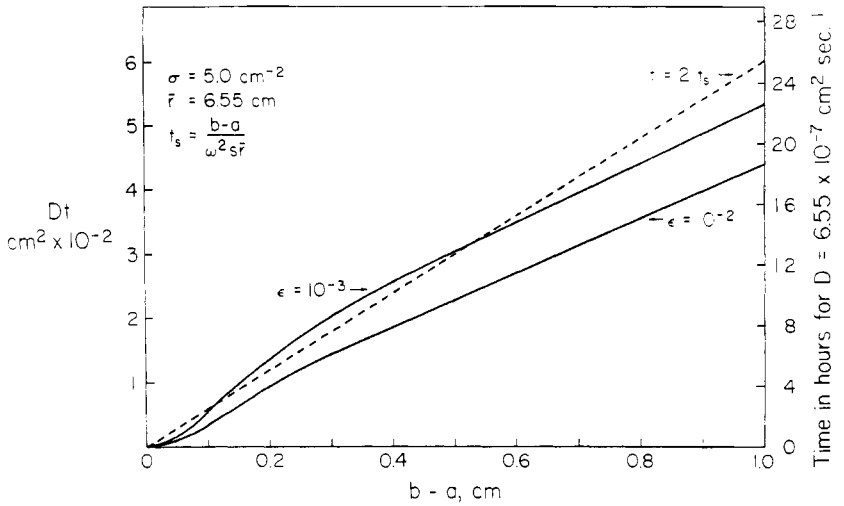


FIG. 2.—Calculated equilibrium times as a function of column height for $\sigma = 5$ cm.$^{-2}$

The classical estimate of equilibrium times (Weaver, 1926; Svedberg and Pedersen, 1940) was $t_{eq} \sim 2(b - a)/\omega^2 sr$), i.e., twice the time required for a boundary to sediment across the cell. These times are indicated on both Figures 1 and 2 by dotted lines and appear to approximate the $\epsilon = 10^{-3}$ curves over an appreciable range. This suggests a simple rule of thumb for conservatively estimating equilibrium times for high speed experiments. $t_{eq} \sim 2.3(b - a)/(\omega^2 sr)$ which on both Figures 1 and 2 would lie above the curves for $\epsilon = 10^{-3}$.

The most satisfactory criterion of equilibrium is, of course, the experimental one: are the distributions time independent within observational error? The preceding discussion should be taken only as a guide of what to expect. Actual determination is to be preferred whenever possible.

## EXPERIMENTAL

*Equipment.*—The Spinco Model E ultracentrifuge used was equipped with Rayleigh interference optics using coated lenses, and a temperature-control unit calibrated against a National Bureau of Standards certified thermometer. The optics were adjusted so that the cell was observed in collimated light parallel to the axis of rotation and so that the mid-point of the solution, with quartz windows, was in focus at the camera. The alignment procedure has been described previously (Yphantis, 1960) with the following exceptions: The slits of the double-slit aperture, which was mounted directly over the lower collimating lens, were narrowed to 0.020 in. (Goodman *et al.*, 1962; Schachman, 1963). The front-to-rear positioning of the light source was determined by defocusing the camera lens by about 1″ and adjusting the light-source position so that the shadow of a meniscus at about 7000 rpm and the light reflected from the meniscus would be coincident (L. Gropper, private communication, 1961). For this procedure the light source was rotated to the interference position and a carefully centered ~0.5-mm aperture was temporarily substituted for the usual interference mask over the light source.

A few minor modifications of the ultracentrifuge proved helpful with these equilibrium experiments. A gentle stream of dry air was directed over the outside of the lower collimating lens to minimize moisture condensation during prolonged experiments. Automatic photography at intervals up to 800 minutes was made possible by connection of a Microflex counter (Eagle Signal Co., No. HZ40A6) into the interval-timing circuit. Bothersome line-voltage fluctuations were eliminated by connection of a voltage regulator (Sorenson No. 1001, 1000 va, 0.01% regulation) to the drive "Variac" input after acceleration to operating speed.

The An-D and An-E rotors were used at the higher speeds, while at speeds below 12,000 rpm the heavy An-J rotor proved more stable. Care was taken to insure that the drive was level for each rotor used. Drives were selected for minimum vibration and precession. For this purpose a rigidly mounted telescope was focused through the view-port to a point just above the cell. The swingout mirror was temporarily replaced by a rigidly mounted mirror to avoid picking up vibrations through the mirror mount. With the light source in the schlieren position and displaced forward a few millimeters, a series of fine interference fringes were seen, similar to those of the Lloyds mirror experiment (Jenkins and White, 1957). The motion of these fringes relative to the telescope cross hairs was a sensitive indication of precession and of slow vibrations. These fringes were blurred at certain speeds with some drives, indicating rapid vibrations.

The Rayleigh patterns were recorded on Kodak spectrographic plates, emulsion Type II-G, with a Kodak Type 77-A filter directly over the light source. This combination of filter and plate response served to isolate the broadened 5461 A Hg line with typical exposure time of 5–16 seconds. Quartz windows were used for most experiments at speeds under 40,000 rpm. Sapphire windows proved invaluable at the higher speeds, and are currently being used routinely at all speeds.

A number of different types of cell centerpieces have been used. The initial experiments were performed with standard 12- and 30-mm-thick duraluminum-filled epoxy double-sector centerpieces. A few experiments employed the recent double-sector Kel-F coated aluminum centerpieces. Most of the work, including all reported here, was performed with centerpieces similar to that illustrated in Figure 3a. These centerpieces allow the simultaneous interferometric observations of three solution-solvent pairs. Figure 3b presents an interference pattern obtained with three initial concentrations of a good sample of β-lactoglobulin in one of these centerpieces. Such centerpieces have been milled or broached in 12- and 30-mm thicknesses with blanks of several materials. Duraluminum-filled epoxy blanks were purchased from Spinco. These

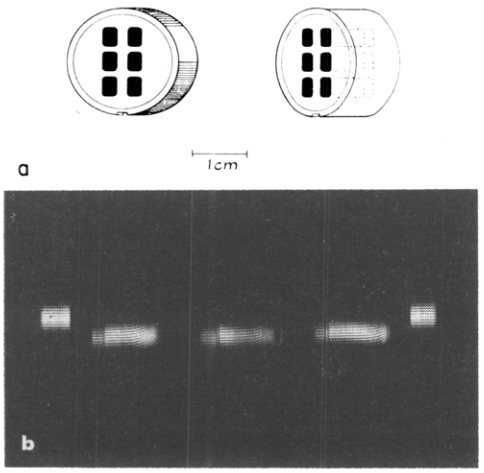


FIG. 3.—(a) Centerpiece for observing three solution-solvent pairs (channels). The centrifugal direction is toward the foot of the page. (b) Rayleigh interference pattern obtained with such a centerpiece in an equilibrium experiment on β-lactoglobulin at 35,600 rpm. The initial concentrations loaded were 0.2% for the first, or most centripetal channel, 0.05% for the second, and 0.01% for the third channel. See Table I for the remaining experimental conditions.

furnished the most stable and reproducible cells for the highest speeds, up to 59,780 rpm, but they are not suitable for use at extremes of *p*H or where trace-metal contamination may be deleterious. Centerpieces of Kel-F press-fitted into duraluminum rings have been most useful. These give cells that operate satisfactorily up to ~35,000 rpm; at higher speeds they begin to show permanent deformations. They have the distinct advantage of chemical inertness and are adequate for studies of proteins of molecular weight above ~30,000. More recently centerpieces have been machined from pure epoxy blanks supplied by Spinco. These appear useful for extremes of *p*H and for speeds up to 59,780 rpm, although they do not appear to be quite as stable mechanically as the duraluminum-filled epoxy cells. Note that the channels in these cells (Fig. 3a) are essentially rectangular rather than sector shaped. Such a rectangular shape is easier to fabricate and is practically equivalent to a sector shape if the distribution extends only 0.3 cm at a radius of 6 cm. The rounded corners are essential for maintaining the strength of the centerpieces. Sharp corners lead to permanent deformation of the centerpieces at significantly lower speeds.

*Materials.*—Crystalline bovine plasma albumin, Lot T-68304, was purchased from Armour and Co. It was examined in the following isoelectric buffer (Kegeles and Gutter, 1951): 0.15 M sodium chloride, 0.02 M sodium acetate, 0.03 M acetic acid, *p*H $4.4_0$.

The ribonuclease-A was the gift of Dr. D. L. Eaker who prepared it by countercurrent distribution (King and Craig, 1958) of an Armour sample, Lot 381-062. The lyophilized preparation was stored at $\sim -10°$ in a tightly sealed container for 15 months previous to the experiment reported. This sample was dissolved in a *p*H $7.7_5$ buffer composed of 0.1 M NaCl, 0.035 M $K_2HPO_4$, and 0.004 M $KH_2PO_4$.

The β-lactoglobulin was prepared by Dr. T. P. King who chromatographed a commercial sample (Pentex, Lot F23) on a 600 × 2.5-cm column of G-75 Sephadex in a *p*H 5.5, 0.5 M ammonium acetate buffer. An aliquot of the effluent containing the major component was examined directly and the fractions containing the remainder of the major component were pooled. The β-lactoglobulin was recovered by salting out with solid $(NH_4)_2SO_4$ (45 g/100 ml). The precipitate was redissolved, dialyzed against distilled water, and stored as a frozen ~1.6% solution. Samples for study were thawed and dialyzed against 0.5 M ammonium acetate containing enough acetic acid to attain *p*H 5.5.

Samples of southern bean mosaic virus were donated by Dr. B. Magdoff. They were examined after dialysis against isoelectric buffers: strain S-I in 0.1 M sodium acetate, *p*H 3.9, and strain S-III in 0.1 M sodium acetate, *p*H 5.6.

The partial specific volumes used in the calculations for these materials were: bovine plasma albumin, $\bar{V} = 0.734_8$ (Dayhoff *et al.*, 1952); β-lactoglobulin, $\bar{V} = 0.751_4$ (Pedersen, 1936); ribonuclease, $\bar{V} = 0.695$ (Harrington and Schellman, 1956); and southern bean mosaic virus, $\bar{V} = 0.696$ (Miller and Price, 1946). Solvent densities were either measured or estimated from data in the *International Critical Tables*. No correction was applied for density increments due to the solutes of interest.

*Procedure.*—The operating speed must be chosen so that the meniscus concentration of the smallest component is essentially zero. An approximate molecular weight should be known or determined. Assuming an estimate of the size, say, from a velocity experiment, which is strongly recommended as a general preliminary, the speed is then chosen so that the effective reduced molecular weight, $\sigma = \omega^2 M(1 - \bar{V}\rho)RT = \omega^2 s/D$ will be about 5 $cm^{-2}$ when using the 3-mm-high columns.

The concentrations to be run depend on the type of information desired and on the centerpiece thickness. With 12-mm centerpieces an initial concentration of about 0.01% gives a displacement of about four fringes at the base of the cell and usually allows essentially all of the solution to be viewed. Initial concentrations of 0.1% or occasionally higher are useful for determination of the size of the smallest species present. Intermediate concentrations have frequently been used in studies of disperse or interacting systems. The lowest concentrations that may be run depend on the desired precision and on the nature of the sample. Useful results have been obtained with initial concentrations as low as 0.003%. These concentrations should be scaled appropriately with other centerpiece thicknesses, as should the loading volumes.

In the standard double-channel centerpieces the solution channel should be filled with 0.1–0.11 ml of solution and with 0.01 ml of a base fluid[5] so that as much of the solution column as possible may be visible. The loading of the solvent-reference channel is critical only if the buffer components sediment appreciably. In general it is more important that the concentrations of the solvent components be matched rather than that the geometrical filling of the solution and solvent channels be identical. The criterion for acceptable matching has been that the refractive-index gradients from the solvent components in the solution and solvent channels should be the same within experimental error: a maximum difference of 0.01 fringe/mm (~0.03 fringe from meniscus to base) has been chosen as an index of matching. Higher values of this difference may be tolerated if corrected for. Consider first the case of identical geometrical loading of the two channels and assume for convenience that the solvent is an ideal two-component system of pure solvent and a

[5] The fluorocarbon FC-43, perfluorotributyl amine (Minnesota Mining and Mfg. Co., Minneapolis, Minn.) has been used as a base fluid since it is dense, volatile, and extremely inert.

solute b, that dissociates into $\nu$ ions, with an effective reduced molecular weight, $\sigma_b = \omega^2 M_b(1 - \bar{V}_b\rho)/\nu RT$. An initial concentration difference, $\Delta c_o$, between the two channels gives rise to a difference across the two channels of $\Delta[c(b) - c(a)] = \sigma_b \Delta c_o (b^2 - a^2)\ 2 \cong 2\sigma_b \Delta c_o$, where $b^2 - a^2$ has been evaluated for a typical 3-mm column. The criterion chosen corresponds to $\Delta c_o \simeq 0.015/\sigma_b$ fringe. For NaCl, $M(1 - \bar{V}\rho)/\nu \cong 17$ and $\sigma_b \simeq 0.007$ cm$^{-2}$ at 30,000 rpm; the allowable missmatch in initial concentrations at this speed is $\sim$2 fringes or $\sim$0.05 g/100 ml for a 12-mm cell thickness. At 60,000 rpm the allowable missmatch is one-fourth of this, while at 15,000 rpm it is four times this. Similar estimations apply to similar salts, e.g., KCl, $NaNO_3$, sodium acetate. On the other hand many commonly used salts must be matched more precisely, for example $KH_2PO_4$, with $M(1 - \bar{V}\rho)\ \nu \cong 72$, may be missmatched at 30,000 rpm by only $\sim$0.5 fringe ($\sim$0.012 g/100 ml for a 12-mm cell).

Consider now a cell loaded with the same concentration, $c_o$, of the solvent component but with different volumes in the two channels, assuming for simplicity that the menisci are aligned but that the base of the solvent channel extends $\Delta b$ cm further from the center of rotation than the base of the solution channel. When $\sigma_b$ is small it is easy to show that the concentration difference between the two channels at any point $r$ is approximately $\Delta c(r) = \Delta b \sigma_b \bar{r} c_o/2$ where $r = (a + b)/2$, and $c_o$ refers to the initial concentration of solvent component b. The concentration difference between the channels from meniscus to base is then $\Delta[c(b) - c(a)] = \Delta b \sigma_b^2 \bar{r}^2 c_o (b - a)/2$, and the allowable values of $\Delta b$ are functions of both $c_o$ and $\sigma_b$. For NaCl at 30,000 rpm the criterion used corresponds to a missmatch of 0.5 mm if $c_o \simeq 400$ fringes ($\sim$10% NaCl) while at 60,000 rpm, $c_o$ must be less than $\sim$25 fringes for the same value of $\Delta b$. The allowable values of $\Delta b \cdot c_o$ are inversely proportional to the square of the effective reduced molecular weight of the buffer solute. At low speeds, accordingly, where $\sigma_b < \sim 0.05$ cm$^{-2}$, there is a considerable leeway in the geometry of loading. At higher speeds or with heavier solvent components one must beware of unequal loading.

The loading of the standard double-sector cells presents no problem. The following procedure was adopted for filling the cells of Figure 3a since they lack the customary filling holes: The empty cell was completely assembled and tightened with the torque wrench, then one of the usual filling-hole screws was tightened against the centerpiece to hold it in place. The cell was loosened with the torque wrench and the top window holder and its window were removed. This was facilitated by having the two holes of the top window holder tapped for 4/36 screws which could be inserted to lift the window holder. The solutions, solvents, and base fluid were measured in their channels using standard 0.5-ml syringes with Kel-F or polyethylene tubing attached instead of metal needles. The top window and its holder were then reinserted and the cell was tightened with the torque wrench both before and after removing the filling-hole screws. Care was taken during these operations to prevent dust or lint from falling between the windows and the centerpiece as this appears to be a potential source of variability of optical distortions in these cells.

Photographic exposures were taken at low speeds, usually $\sim$5000 rpm, and also shortly after reaching full speed, as a check on the blanks. Periodic exposures were made, usually at 3- to 6-hour intervals, until equilibrium was attained. These exposures have proved valuable not only for ascertaining equilibrium but also as partial checks on the stability of the preparations. From the discussion on equilibrium times it is apparent that the most sensitive criterion of equilibrium for a given initial concentration will be given by measurements near the base of the cell. However this region is often inaccessible and the next most sensitive indication is given by measurements near the positive maximum of the first eigenfunction. This occurs at the position where the equilibrium concentration is approximately $0.3c_o$ if $3 < \sigma < 10$ cm$^{-2}$. For the higher initial concentrations the fringe displacements were measured at 1-mm (on the plate) intervals in the region from the meniscus to the point where the fringe displacement was about 2 fringes so as to cover this sensitive region. After a brief transient period the displacements in this region decrease exponentially with time to constant values at equilibrium. Occasionally, unstable materials generate low molecular weight substances at such a rate that these displacements no longer show this characteristic drop but may even increase with time. A series of interference photographs were customarily taken at equilibrium and the best exposure was selected for measurement.

The photographic plates were measured with a two-coordinate comparator (Gaertner Toolmakers Microscope, M2001-RS) in either of two ways. After alignment of the radial direction along the $x$-coordinate of the comparator, the average of the $y$-coordinates of 5 fringes, usually 3 white and 2 black, was determined as a function of the $x$-coordinate, taking care to keep the measurements in the central region of the diffraction envelope. Readings were taken at intervals of 0.5 mm on the plate until the displacements of the fringes relative to the meniscus region were about 100$\mu$. From this point on, measurements were taken at 0.1-mm intervals until the fringes could no longer be resolved clearly, generally up to a fringe-density of about 13 fringes/mm on a good plate. The reproducibility of these measurements depends greatly on the contrast and quality of the photographed pattern. The standard deviations of measurements on good plates are about 5 $\mu$ ($\sim 1/60$ fringe) in regions where the fringes have moderate slope and increase to about 1% of the net displacement near the maximum observable gradients.

Initially, and still for occasional runs, the patterns have been measured by the quicker procedure of determining the $x$-coordinate for given fringe displacements, usually multiples of 0.5 fringe (along with measurements of the displacements near the meniscus). This procedure appears less precise at the lower concentration region of the distributions, where the fringes exhibit little slope, but is adequate at higher displacements.

Blanks have been used, not to correct for buffer redistribution, but rather to take care of residual optical inequalities between solution and solvent channels. The following procedure was used with the standard double-sector cells to obtain blanks without taking the cell apart: FC-43 fluorocarbon and water were added and the cell was taken up to speed to displace any solution trapped between the windows and the centerpiece. The cell was thoroughly rinsed again and loaded with water to approximately the same total volumes as in the experiment, and then taken up to the speed of the run. The only transient phenomena when the cells are loaded with water are from thermal gradients that are rapidly dissipated so that blank exposures may be taken after a few minutes.

The present design of the multichannel cells requires that they be disassembled each time they are loaded.[6]

[6] Multichannel cell assemblies that may be filled without dismantling are currently being tested.

Accordingly, greater care is required for blanks with these cells than with the usual double-channel cells. The blanks for these multichannel cells appear to be independent of speed and time and are usually reproducible to better than 10$\mu$ if certain precautions are taken. This reproducibility appears to be a long-term reproducibility lasting over dozens of experiments and periods of months if a given centerpiece is assembled with only one specific set of cell components, i.e., cell casing, windows, and window holders, but with replacement of the rubber window liners for each experiment and for each blank, and also if care is taken during assembly of the cells to prevent foreign matter from lodging between the cell windows and the centerpieces. An occasional blank will exhibit a deviation from the others. Most often such a deviation from the preceding blanks is temporary and probably reflects a failure to assemble the cell reproducibly. Infrequently the deviations are semipermanent with subsequent blanks clustering about a new average. Such changes in blanks usually occur at speeds near those showing readily visible permanent deformations of the centerpieces.

In general the blank corrections are small, with displacements of less than $^1/_7$ fringe and often considerably less, across a channel for centerpieces that have been ground and lapped flat. Deviations from the average values of the blanks are infrequent, occurring less than 5% of the time if the cells are used at speeds below 80% of their maximum useful speeds. Additionally the experiments themselves provide checks on the blanks for any suspicious cases: since the optical inequalities between the channels appear independent of speed as long as the cell components are not overstressed, the low-speed exposures at the start of an experiment may be used as blanks. This is feasible provided the solutions loaded are uniform at the time of observation, but this procedure will fail if the solutions loaded have not been thoroughly mixed or if there has been significant evaporation, especially from concentrated buffers. The procedure is more difficult to apply with the 30-mm-thick centerpieces because of the increased sensitivity to small refractive-index differences. Caution should also be exercised to allow the decay of any thermal gradients. Exposures taken shortly after operating speed is reached may also be compared with the blanks. The same precautions apply here and care must also be taken to avoid using the end points of a channel, where there is solute redistribution. A further check on the blanks is provided with "clean" samples, i.e., samples not contaminated with small solutes; subtraction of an appropriate blank from the experiment should yield virtually constant values of the net fringe displacement near the meniscus if the speed has been chosen properly and if the concentration being processed is not excessive.

Deposition of oil on the lenses or condensation of moisture may affect the patterns observed. In such cases it appears best to use blanks taken immediately after a run without any cleaning or wiping of the lenses so that the blank and the experiment will tend to be affected equally. The use of a closed-system oil drain appears helpful in minimizing drive-oil deposits, while deposition of diffusion-pump oil, identified by infrared adsorption, was virtually eliminated after water-cooling the elbow connecting the diffusion pump to the chamber. Generally there was no significant oil deposition, even with prolonged runs, unless the drive started to consume an excessive amount of oil.

*Calculations.*—The $y$-displacements of the blank are subtracted from the measured $y$-displacements of the experiment. Then the initial region, up to about 50 $\mu$ net displacement from the meniscus region, is plotted as a function of position (Fig. 4a). Inspection of such a graph allows a choice of the zero-concentration reference level. Generally this choice is defined to within a 5-$\mu$ displacement[7] if the blank corrected displacements are constant over several points. If the graph of these initial displacements is not flat then the reasons must be found. The following have accounted for all deviations to date: choice of an improper blank, choice of too high a loading concentration or too low an operating speed, missmatch of solvent and solution, convection, the presence of unsuspected low molecular weight component(s), or a continuous generation of small quantities of low molecular weight species from the main component that sets up a concentration gradient. If the speed has been chosen too low, or if too high an initial concentration has been run, one can often make a guess as to the zero-concentration level, provided something is known about the sample, and accept the results provisionally. It is better, of course, to repeat the experiment under proper conditions. A solvent missmatch may be corrected for, either by calculation or preferably by centrifuging an appropriate blank, with the same missmatch, to equilibrium.

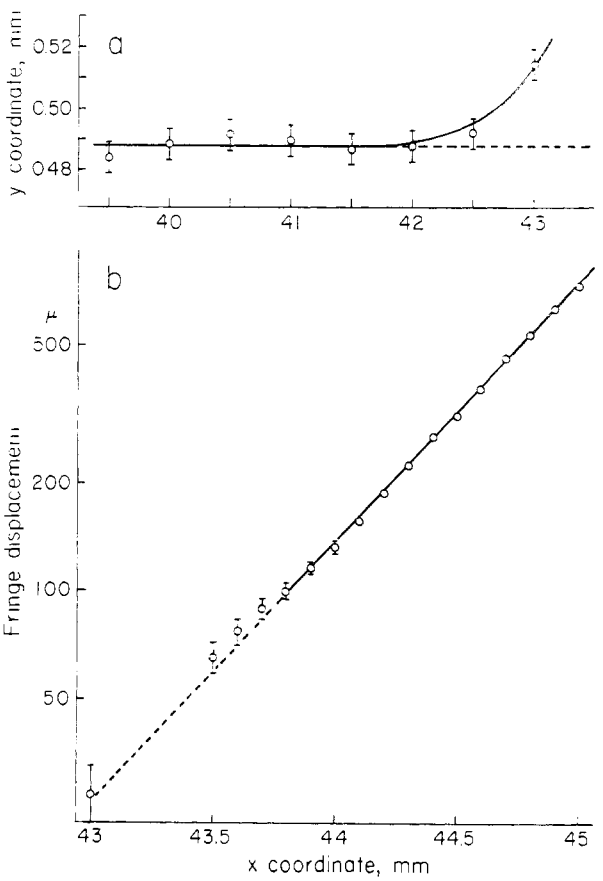


FIG. 4.—Measurements on the 0.01% $\beta$-lactoglobulin pattern presented in the previous figure. (a) Blank-corrected fringe displacements as a function of comparator $x$-coordinate for the centripetal region of the channel. The dotted line indicates the zero-concentration level chosen. (b) Fringe displacements, on a logarithmic scale, after subtraction of the zero-concentration level, are presented as a function of comparator $x$-coordinate. The indicated ranges correspond to $\pm 5$ $\mu$ fringe displacement.

[7] With the ultracentrifuge used one fringe corresponds to a displacement of 289 $\mu$.

TABLE I
MOLECULAR WEIGHTS AND STANDARD DEVIATIONS FOR β-LACTOGLOBULIN AS GIVEN BY THE LINEAR LEAST SQUARES ESTIMATES[a]

| Initial Concentration (%) | Mol Wt from Measurements Every 0.1 mm | Mol Wt from Measurements Every 0.5 Fringe |
|---|---|---|
| 0.2 | 35,100 ± 170 | 35,400 ± 100 |
| 0.09[b] | 36,560 ± 140 | |
| 0.05 | 36,500 ± 140 | 36,600 ± 170 |
| 0.01 | 36,400 ± 330 | 35,500 ± 360 |

[a] Experimental conditions: *p*H 5.5, 0.5 M ammonium acetate; 25°, 16 hours at 35,600 rpm; 12-mm-thick, three-channel, duraluminum-filled epoxy centerpiece. [b] From measurements on the direct column eluate; the other values are from an experiment after storage for 17 months as a frozen 1.6% aqueous solution.

The natural logarithms of the differences, $y(r_i) - y_0$, between the blank corrected fringe displacements; $y(r_i)$ and the zero-concentration level chosen, $y_0$, are then plotted against a position variable. Strictly speaking this variable should be $r^2$ and the slope of the resulting curve at any point will then give $\sigma_w/2$ at that point. However it is simpler to plot the logarithms of the net displacements against the $x$-coordinates read directly from the comparator. The value of $\sigma_w(r)$ at any point is then given by $\{d \ln [y(r) - y_0]/dx\}m/r$, where $x$ is the comparator $x$-coordinate, $m$ is the magnification factor from centimeters in the cell to comparator $x$-units, i.e., $dx\ dr$, and $r$ is the radius of the point considered. In these runs the useful portions of the logarithm plots are usually contained within a 1-mm (cell coordinate) space. The slope then given by a single ideal component differs by less than 1% from the mean values even though plotted against a quantity proportioned to $r$ instead of $r^2$. Figure 4b presents such a graph of the logarithm of the net fringe displacement versus $x$-coordinate for data obtained from the pattern of the lowest concentration channel of Figure 3b. Values of fringe displacements are presented from ~25 μ upward. Generally not too much significance should be attached to values of $\sigma$ determined by displacements smaller than 100–150 μ since experimental error introduces a large uncertainty in the logarithms of these fringe displacements. These smaller displacements are useful, however, in indicating the possible presence of lower species or of convection (see later).

"Averaged" values of $\sigma$ have been estimated by a weighted least squares treatment. Each point was assigned a (statistical) weight inversely proportional to its estimated error in ln $c$ and a linear dependence of ln $[y(r) - y_0]$ on $r^2$ was assumed.[8] Only points with a net displacement greater than 100 μ have been used. The molecular weights resulting from this least squares treatment are listed in Table I. This table also includes the values found for the other two channels of the experiment pictured in Figure 3b as measured by both procedures (every 0.1 mm and every 0.5 fringe) along with a value from the sample before recovery from the column effluent and storage. The differences between the values presented are not considered significant, even though they are much larger than the standard deviations from the least squares estimations: when using points with net displacements as small as 100 μ a variation of 5 μ in the choice of $y_0$, the zero concentration level, results in about a 2% variation in the least squares slope without a significant increase of the standard deviation. The lower values of the molecular weights estimated for the highest concentration (0.2%) in Table I probably result from having only a limited low concentration region near the meniscus for use in estimating the zero level. In general such a least squares treatment does not appear to offer a significant advantage over graphical estimations: the same values, within experimental uncertainty, may be obtained from the graphs of log $[y(r) - y_0)$ vs. $x$ when data going down to ~100 μ are included.

[8] A "Sicom" program for a Control Data 160-A computer with a plotter has been written for routine reduction of data and for performing this and most of the subsequent calculations. This program will be described in a later publication.

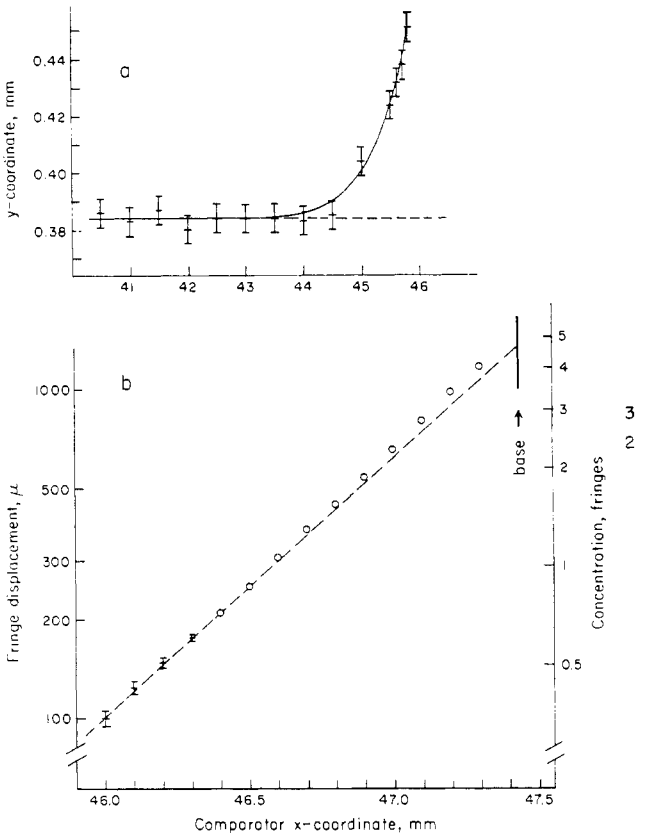


FIG. 5.—Measurements on an 0.01% solution of bovine plasma albumin in a *p*H 4.4 buffer. The data are from one channel of a 12-mm, three-channel, duraluminum-filled epoxy cell after 17 hours at 25,980 rpm and at 19.6°. (a) Blank-corrected fringe displacements as a function of the comparator $x$-coordinate for determination of the zero-concentration level. (b) Net fringe displacement on a logarithmic scale as a function of comparator $x$-coordinate. The dotted line indicates the slope found with an 0.1% solution in the same experiment when corrected to the average radius of these points. The discrepancy between the points and the slope indicates heterogeneity. See the text for details.

This sample of β-lactoglobulin appears homogenous under the experimental conditions used. All the "log" plots are linear and the values of $\overline{M}_w$ and $\overline{M}_z$ as determined by extrapolation of the point-average molecular weights are the same: 36,500 ± 1,100. It is a rare sample that exhibits such ideal behavior, more commonly the plots of ln $[y(r) - y_0]$ exhibit curvature as illustrated by the data for a typical sample of commercial crystalline bovine plasma albumin. The meniscus region graph is shown in Figure 5a and the "log" plot in Figure 5b. The curvature of this latter graph is slight but the slope calculated for this radius from the data of another channel of the same experiment at $c_o$ = 0.1% is indicated by the dotted line. Com-

parison of the apparent slopes (or "averaged" $\sigma$) of identical samples at different initial concentrations is a simple method of detecting heterogeneity if the measured displacements span the same concentration range. In the present case the average slope from the data of Figure 5b gives $\sigma = 5.73$ cm$^{-2}$ whereas the data of the channel at $c_0 = 0.1\%$ gives $\sigma = 5.35$ cm$^{-2}$. Since the data span essentially the same concentration range the conclusion of heterogeneity appears inescapable, if the possibility of pressure dependence be discounted. In such a case as this more elaborate treatments are called for to describe the sample in terms of its average properties and, if possible, in terms of its principal or smallest component. As pointed out in the theoretical section one convenient way of doing this is through the point average effective molecular weights defined by equations (4). There are numerous specific ways of calculating these, each with its own advantages. Two types of approaches appear generally applicable: one is based on the application of curve fitting to all of the primary experimental data[9] with derivation of the particular values of interest from the fitted curves. The other approach rests on direct and preferably simple treatment of the data to obtain the requisite information. The first has the feature of "smoothing" the data, which is often an advantage, but the procedure may also lead to loss of significant detail. The procedure adopted here is the classical one of using the raw unsmoothed data directly. This leads to "noisy" results where the individual experimental fluctuations show up in the derived values but has the advantage of keeping direct contact with the data.

The simplest and most directly computed point average is the effective reduced weight-average molecular weight, $\sigma_w(r)$, which from equation (4b) is simply the slope of ln $c$ vs. $r^2/2$ at each point. Evaluation of this quantity by direct differencing of adjacent points is excessively noisy for closely spaced data. The procedure used for manual computations has been to take five equally spaced (on the $x$-coordinate and hence on the $r$-coordinate) adjacent data points and to determine the least square straight line through these points. The slope of this line is equal to the slope at the central point given by fitting a least quadratic through the same five points. A straightforward application of the unweighted least squares procedure to five equidistant points gives the following as the value of the slope at the central point:

$$\left(\frac{du}{dx}\right)_0 = \frac{0.1}{\Delta x}(2u_2 + u_1 - u_{-1} - 2u_{-2}) \qquad (13)$$

where the $u_j$ are the five values of the logarithm of the net displacement; the values of $j$ are in sequence, with $j = -2$ for the first point, $j = 0$ for the central point, etc., and $\Delta x$ is the common $x$-spacing. Equation (13) is simple and rapid to apply, allowing evaluation of the derivative in one simple operation on many calculators. The value of $\sigma_w(r)$ is then simply $(du/dx)m/r$.

The averaging involved in the use of equation (13) does not seem to obscure any significant detail in the bulk of the solution. However, the use of equation (13) does leave a gap at the base of the cell without any values of $\sigma_w(r)$. Occasional experiments appear "noisier" than others, usually because of oil deposition on the lenses or from photographic difficulties. For these it is useful to take seven instead of five points for a similar least squares treatment but at the expense of a larger "dead" region.

Values of $\sigma_n(r)$, the effective reduced number average molecular weight, have been obtained through the use of equation (5), with the integral $\int_a^r c(r)\, d(r^2/2) = \int_a^r r\, c(r)dr$ evaluated by straightforward trapezoidal summation. This appears precise enough if points are taken every 0.1 mm on the photographic plates. In the region where measurements have been made only every 0.5 mm, usually with net displacements less than 100 $\mu$, it may be advisable to obtain points by interpolation on the ln $(y - y_0)$ vs. $x$ graphs. However, if these graphs are linear within experimental error in this region[10] one may then make the assumption that the solute is effectively monodisperse in and centripetal to this initial region with $\sigma_n(r) = \sigma_w(r) = \sigma(r_1)$ for $r = r_1$, where $r_1$ is any point of this initial region. This allows the ready approximation of the integral up to $r_1$ as

$$I \equiv \int_{r=a}^{r_1} rc(r)\, dr \cong \frac{c(r_1)}{\sigma_n(r_1)} \cong \frac{c(r_1)}{\sigma_w(r_1)} = \frac{y(r_1) - y_0}{\sigma_w(r_1)} \qquad (14)$$

where $\sigma_w(r_1)$ is obtained from the best slope for this initial region and where concentrations have been expressed in the units of the fringe displacements. For points centrifugal to $r_1$ the values of $\sigma_n(r)$ are then given by

$$\sigma_n(r_k) = \frac{c(r_k)}{\int_{r=a}^{r_1} rc(r)dr + \int_{r=r_1}^{r_k} rc(r)dr} \cong \frac{y(r_k) - y_0}{I + \frac{\Delta x}{2m_i}\sum_{i=2}^{k} r_i[y(r_i) - y_0] + r_{i-1}[y(r_{i-1}) - y_0]} \qquad (15)$$

where the points $r_i$ are all equally spaced with $x$ interval $\Delta x$ and with the first of these equally spaced points ($i = 1$) taken within the initial linear region so as to define the value of $r_1$. Since there may be moderate uncertainty in the values of $\sigma_w(r_1)$ the initial values of $\sigma_n(r)$ will also be uncertain, but this uncertainty will decay rapidly with increasing fringe displacements. For example, with a monodisperse system an uncertainty of 20% (a rather extreme value) in $\sigma_w(r_1)$ at $y_1 - y_0 = 100$ $\mu$ would decrease to an uncertainty of about 4% in $\sigma_n(r_k)$ at $y_k - y_0 = 500$ $\mu$. In general, values of $\sigma_n(r)$ for displacements under one fringe ($\sim$300 $\mu$) are not considered significant when computed in this manner (see the section on the analysis of errors).

Values of $\sigma_n(r)$ and $\sigma_w(r)$ calculated from the data of Figure 5 are given in Figure 6a as a function of concentration. Sedimentation velocity experiments showed the presence of no component smaller than the major component. It is feasible therefore to attempt determinations of the size of the major component from these values of $\sigma_n(r)$ and $\sigma_w(r)$. The linear extrapolation of these values to zero concentration gives an estimate of the monomer $\sigma$ of 5.3 $\pm$ 0.2 cm$^{-2}$, corresponding to a monomer molecular weight of 66,600 $\pm$ 2,600.[11] Values of $2\sigma_n(r) - \sigma_w(r)$ also provide an estimate of $\sigma_1$ as 5.3 $\pm$ 0.1 cm$^{-2}$ with a total uncertainty estimated as $\sim$5% in this case. Note that values of $\sigma_n(r)$ are not presented for concentra-

[9] See Donnelly (1960) for an example of such curve fitting with subsequent estimation of the average properties of the sample.

[10] In most systems log $[y(r) - y_0]$ is linear in $r$ for fringe displacements less than 100 $\mu$.

[11] Unless otherwise specified the ranges indicate the estimated scatter or standard deviation of the calculated values. A discussion of the total uncertainty of derived value is to be found in the section on errors.

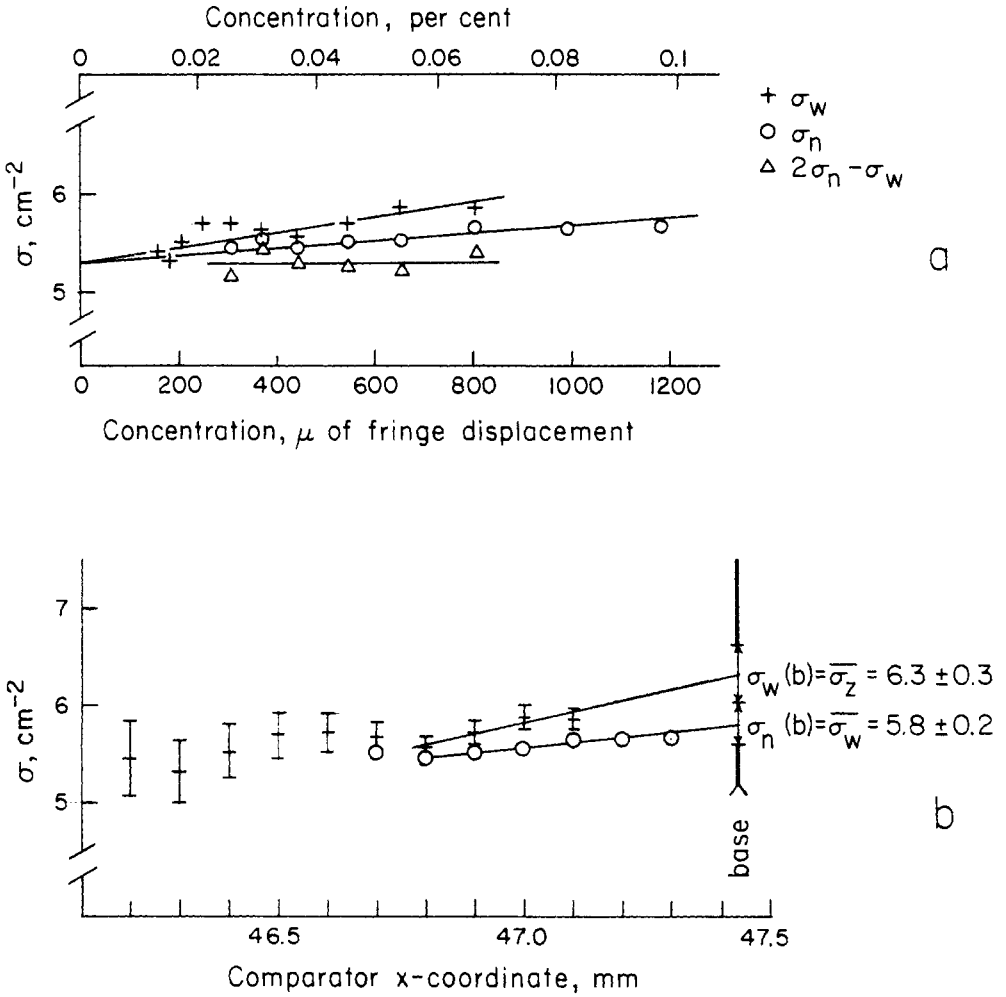


FIG. 6.—Use of the point-average effective molecular weights $\sigma_k(r)$ for characterization of the bovine plasma albumin. The original data are presented in the previous figure. (a) Determination of the monomer molecular weight. The effective reduced weight and number-average molecular weights determined at various points in a single channel are presented as a function of the concentration at those points. Extrapolation to zero concentration gives an estimate of the "monomer" size. The values of $2\sigma_n(r) - \sigma_w(r)$ also give an estimate of the size of the smallest significant species without requiring extrapolation. (b) Determination of the weight- and $z$-average effective molecular weights of the complete sample by linear extrapolation of the more centrifugal points to the base of the column. Note that the abscissa in this graph is the comparator $x$-coordinate.

tions corresponding to displacements of less than one fringe and that the scatter of the values of $\sigma_w(r)$ decreases with concentration. The total estimated uncertainties of the same $\sigma_w(r)$ are shown in Figure 6b; uncertainties in the $\sigma_n(r)$ are expected to be about 1.5–2 times these estimates. In general, the values of $2\sigma_n(r) - \sigma_w(r)$ are considered significant only for fringe displacements greater than 500 $\mu$, although, as here, they often appear well behaved for smaller displacements.

The values of $\sigma_n(r)$ and $\sigma_w(r)$ may be used to estimate the average molecular weights of the entire sample using equations (6′). This requires extrapolation of the values of $\sigma_n(r)$ and $\sigma_w(r)$ to the base of the solution column. Using calculated distributions it has been found that linear extrapolation of the values for the most centrifugal points available as a function of position appears to give estimates valid to within 3 or 4% provided the extrapolations involved span less than 0.2 or 0.3 mm in the cell and provided there is no significant quantity of very large components, say, greater than three times the size of the main component present, that would not be fully observed because of the centrifugal fractionation. The position of the base has been taken as that of the mid-point of the base meniscus. Any significant increase of apparent thickness of a base meniscus over about 0.007 cm (cell coordinate) has been taken as an indication of precipitation or of the presence of components too large to be properly included in the molecular weight averages.

The values of $\sigma_n(r)$ and $\sigma_w(r)$ closest the base have been linearly extrapolated to the cell base in Figure 6b, and the maximum plausible range of the linearly extrapolated values is indicated. These extrapolated values correspond to $\bar{M}_w$ = 72,900 ± 2,500 and $\bar{M}_z$ = 79,200 ± 3,800, corresponding to 9.5 and 10.5% dimer for this sample on the assumption that only monomer and dimer molecules are present. Velocity experiments on the same preparation showed the presence of 8 ± 2% dimer and of a trace amount of higher polymers.

A useful, "quick and dirty" method of estimating the average properties of a sample consists of subtracting the contribution of the monomer or lowest component from the experimental net displacements and plotting the logarithm of the remaining displacements against position. This procedure is illustrated for the data of Figure 5. The dotted curve in that figure corresponds to the (~ monomer) $\sigma$ found with

TABLE II
MOLECULAR WEIGHTS FROM A SINGLE RUN ON A STORED SAMPLE OF RIBONUCLEASE-A[a]

| (1) Initial Concentration (%) | (2) Weight-Average Mol Wt of Entire Sample | (3) *z*-Average Mol Wt of Entire Sample | (4) Graphical Estimate of Smallest Mol Wt Present | (5) $2M_n(r) - M_w(r)$ |
|---|---|---|---|---|
| 0.07 | | | 13,900 ± 600 | 13,500 ± 500[b] |
| 0.021 | 17,500 ± 900 | 20,000 ± 1500 | 14,800 ± 400 | 13,950 ± 470 |
| 0.007 | 17,500 ± 600 | 20,300 ± 600 | 15,300 ± 800 | (14,300 ± 1400) |

[a] Experimental conditions: *p*H 7.8; 25°; 50,740 rpm; 12-mm-thick, three-channel, unfilled epoxy centerpiece. [b] Standard deviations from the mean are indicated in this column. The other columns present the estimated uncertainty.

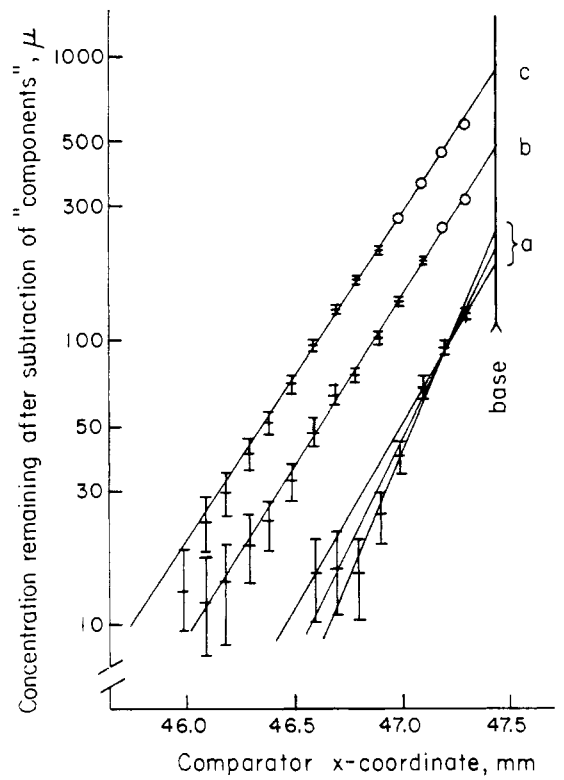


FIG. 7.—Difference curves for the data of Figure 4 used in an approximate method of determining the weight- and *z*-average molecular weights of the sample. See the text for details.

the 0.1% solution. This curve was subtracted from the points and the logarithm of this difference is presented as curves *a* of Figure 7. The indicated ranges correspond to ± 5 μ. The slope of these curves give an "average" $\sigma$ of 10.7 ± 1.7 $cm^{-2}$ for the components remaining after subtraction of the monomer, indicating that they are probably dimers. Values of the $\bar{\sigma}_k$ for the whole sample may be estimated by calculation of the $\sigma_{k-1}(b)$ (see equations 6'). The values of $\sigma$ and the concentrations found by extrapolation to the cell base for the difference curve and for the subtracted curve are inserted into equations (4), assuming that each of the curves represents but a single ideal component. Values of $\bar{\sigma}_w$ and $\bar{\sigma}_z$ calculated in this way appear moderately insensitive to any reasonable choice of a curve to be subtracted, i.e., if care is taken to avoid systematic deviations from a description of the experimental curve as the sum of two exponential curves. In choosing the lowest $\sigma$ curve care should be taken to best account for the lowest concentration region, and in drawing the "log" curve for the differences most reliance should be placed on the higher concentration points. The average values of $\bar{\sigma}_w$ and $\bar{\sigma}_z$ determined from each of the three curves marked *a* in Figure 7 are $5.7_2$ ± $0.1_2$ and $6.1_0$ ± 0.3 $cm^{-2}$ compared to the values of 5.8 ± 0.2 and 6.3 ± 0.3 found by extrapolation. Comparatively large variations of the subtracted curves are tolerable: choice of the subtracted $\sigma$ as 5.11 $cm^{-2}$ (giving curve *b*) or 4.43 $cm^{-2}$ (curve *c*) gave $\bar{\sigma}_w$ = 5.8 or 5.7 $cm^{-2}$ and $\bar{\sigma}_z$ = 6.1 or 6.3 $cm^{-2}$, respectively. In these last two cases the individual curves into which the experimental curve was decomposed have no intrinsic physical significance in themselves. Obviously, unless the subtracted curve corresponds to a real component the individual curves are only useful ways of expressing the experimental data. It appears that this procedure is useful for rapidly obtaining estimates of $\bar{\sigma}_w$ and $\bar{\sigma}_z$ good to 5% or so.[12] If the value of $\sigma_1$ is known reliably, the decomposition of the experimental curves is often useful in estimating the properties of the remaining material. However, caution should be exercised in ascribing physical existence to components when using such a procedure, since there is usually much freedom in fitting experimental curves by sums of exponentials. Information from other sources, especially from sedimentation velocity experiments, may be invaluable in such cases.

*Results.*—This method has been used to examine many preparations. The following results have been selected to illustrate the accessible range of molecular weights and concentrations. The experiments chosen are on preparations contaminated with heavier components than the ones of interest so as to exhibit the ability to extract information in such cases.

The requirement that $\sigma$ be about 5 $cm^{-2}$ and the speed limitations of the ultracentrifuge set limits to the observable range of molecular weights. Ribonuclease is a well-characterized solute that is small enough to lie near the limit imposed by the maximum available speeds. Table II presents the values found for a stored lyophilized preparation of ribonuclease-A. This preparation contained less than 2% polymer before storage; as has been noted previously (Yphantis, 1960), ribonuclease preparations often appear unstable on storage in the solid state, with considerable polymer formation. This is strikingly illustrated here by the values of the weight- and *z*-average molecular weights (Table II, columns 2 and 3). These values were obtained by extrapolation; in the case of the 0.021% solution a longer extrapolation was required than for the more dilute solution, accounting for the larger estimated uncertainty in spite of the higher concen-

[12] The mathematical basis of this method appears sound if the sum of the curves used depicts the data within experimental error and if one does not try to extract molecular weights above the *z*-average. The accuracy of the procedure depends on the quality of the fit; if needed, more than two exponential terms could be used to describe the experimental data more closely and the values of the averages determined would then approach the true values even more closely. The procedure is equivalent in many ways to the curve fitting used by Donnelly (1960).

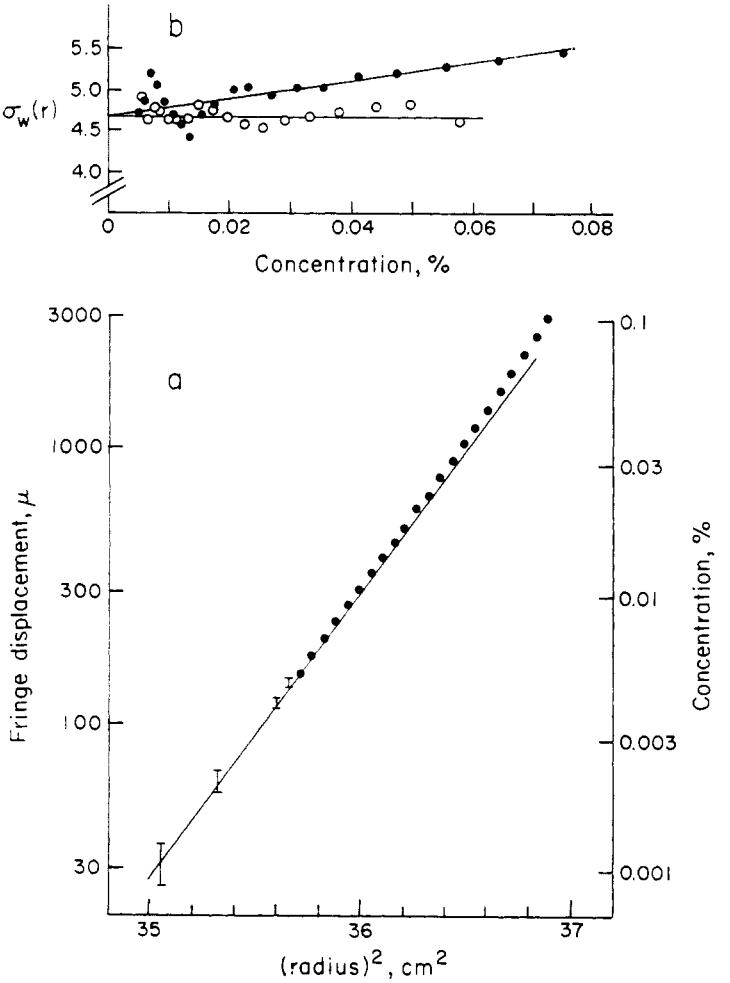


FIG. 8.—Results from an experiment on bovine plasma albumin at an initial concentration of 0.05% at $p$H 4.4$_0$. The data are from a one channel of an experiment in a 30-mm-thick, three-channel, Kel-F centerpiece, run for 14 hours at 24,630 rpm and at 24.7°. (a) The observed fringe displacements and the corresponding concentrations are presented (on a logarithmic scale) as a function of the square of the radius. The indicated points were obtained with the camera lens focused on the middle of the cell and the line represents the data obtained with the camera lens focused on the $^2/_3$ plane. The indicated ranges correspond to fringe displacements of $\pm 5$ $\mu$. (b) Comparison of the $\sigma_w(r)$ obtained with the $^1/_2$ focus (●) and the $^2/_3$ focus (○) of the camera lens. The line drawn for the points with the $^2/_3$ focus corresponds to the least-squared averaged value of $\sigma$.

trations available for measurement. Assuming that all the polymer is in the form of a dimer, the values of the weight-average molecular weights correspond to 27% dimer, and the values of the $z$-average molecular weights to 31%.

Simple graphical estimates of the smallest component present (Table II, column 4) were made by visually fitting the initial slopes, placing most reliance on the data from ~0.5 to ~2 fringes displacement. Such estimates are subjective, but do give reasonable values if the centrifugal fractionation is high enough, as in the most concentrated solution studies. More reliable estimates are obtained for poorer centrifugal fractionations by using equations (12). Column 5 of Table II presents the values (and their standard deviations from the mean) obtained through equation (12a) with $j = 2$ using values of $\sigma_n(r)$ and $\sigma_w(r)$ for concentrations corresponding to fringe displacement of more than 500 $\mu$. The values in parentheses are an exception; in this particular channel only about 600 $\mu$ of fringe displacement was measurable since the initial concentration was low. The values presented for this channel cover the range 260–370 $\mu$ displacement and are subject to more uncertainty than those obtained at higher concentrations. The monomer molecular weight estimates may be compared to the formula weight of 13,683 (Hirs *et al.*, 1956). Application of equations (12) to this sample would be difficult to justify on the basis of velocity runs because of the poor resolution with such low molecular weights. The justification in this case is provided by knowledge of similar samples that do not exhibit smaller molecules than the main component on Sephadex chromatography and by the absence of any indication of lower $\sigma$ species at the higher initial concentrations of this experiment.

Some use has been made of 30-mm-thick centerpieces. These allow the observation of more dilute regions of the concentration distributions but the results are subject to the effects of extensive Wiener skewing if the camera lens is focused on the middle of the cell (see the section on the consideration of errors). Figure 8 presents the results of an experiment on the same bovine plasma albumin as before but in a 30-mm cell. The initial concentration was 0.05%, corresponding to an initial concentration of ~0.125% in a 12-mm cell. According to the results for the 0.1% solution in the 12-mm cell, there should have been essentially no polymer in the region where the fringes were measurable. However, when the camera lens was focused on the mid-plane of the cell

the logarithmic plot (indicated points on Fig. 8a) was curved and the values of $\sigma_w(r)$ (Fig. 8b) increased with concentration, indicating the apparent presence of material larger than monomer. This anomaly appears to be an optical artifact and may be eliminated by the procedure that Svensson (1956) recommended: The camera lens was focused on a plane two-thirds of the distance from the front to the rear of the solution and the magnification factor used for this focus was determined from photographs taken, with collimated light, of a scale in the usual mid-plane position. Under these conditions the logarithmic plot was no longer curved (line of Fig. 8a) and the values of $\sigma_w(r)$ were independent of concentration (Fig. 8b). There are difficulties, however, with this procedure that may interfere with determinations of the $\bar{\sigma}_k$; for example the end points of the solution columns are not sharply imaged and the fringe contrast at the higher fringe displacements is poorer than with the mid-plane focus.

Obviously, direct estimates of polymer content or of concentration dependence in such thick cells will be subject to serious systematic error when using points with large fringe displacements unless the camera lens is focused on the $^2/_3$ plane. However, it is often possible to obtain useful results with the usual mid-plane focus: The effects of the Wiener skewing are negligible in regions with small refractive index gradients, while at the higher gradients the effects often, but not always, mimic the presence of a few per cent of dimer (for details see the section on the consideration of errors). In the case illustrated here the optical aberrations have little effect on the determination of monomer size: the values of $\sigma_w(r)$ extrapolate, at zero concentration, to (approximately) the least squares value of $\sigma$ found for the $^2/_3$ focus (4.68 $\pm$ 0.01 $cm^{-2}$) and the average (and standard deviation from the mean) of the values of $2\ \sigma_n(r) - \sigma_w(r)$ for fringe displacements $> 500\ \mu$ is 4.67 $\pm$ 0.07 $cm^{-2}$ with focus on the mid-plane and 4.66 $\pm$ 0.11 $cm^{-2}$ with focus on the $^2/_3$ plane. Under the experimental conditions used the value of $\sigma = 4.67\ cm^{-2}$ corresponds to a molecular weight of 66,100 for the albumin monomer.

Experiments on southern bean mosaic virus provide an example of a large solute run at a speed close to the lower limit of the commercially available ultracentrifuges. Figure 9 presents values of $\sigma_w(r)$ as a function of $b^2 - r^2$ to allow intercomparison of three different initial concentrations of the same experiment. The values of $\sigma_w(r)$ show an increased scatter at the lowest fringe displacement ($\sim$200 $\mu$). Larger deviations, both systematic and random, also appear for points near the maximum observable fringe displacements ($\sim$2000 $\mu$ for the two highest initial concentrations). This sample appears to be heterogenous but no estimate of average properties was attempted because of the apparent thickening of the bases (and menisci) at the low speed used. An analysis of the sedimentation coefficient distribution, $g(s)$ (see for example Williams *et al.*, 1952) was performed by Mr. P. Englund of this Institute. Using the Rayleigh interference optics at an initial concentration of 0.06% he found that the sample contained three sedimentation coefficient classes with about 87% of the material corresponding to $s$ $\sim$115 Svedbergs, $\sim$8% to $s$ $\sim$122 S, and the remainder to $s \sim$ 146 S. No smaller species were evident in this analysis nor in experiments at higher concentrations (with the schlieren optics). Since no components smaller than the main component were seen either in velocity or equilibrium experiments equations (12) may be safely applied to estimate the size of the main component.

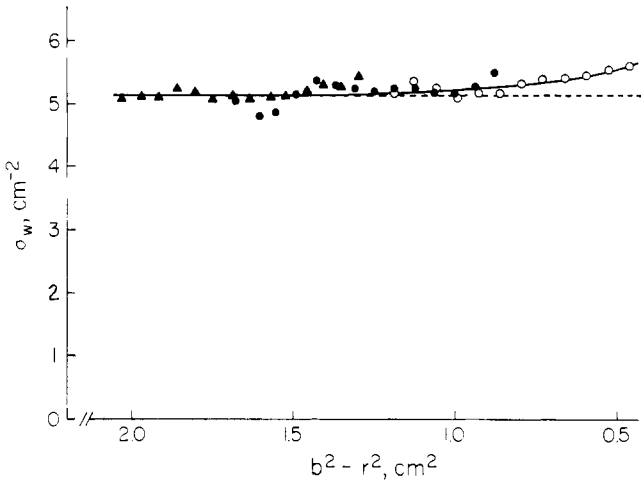


FIG. 9.—Direct comparison of the effective reduced weight-average molecular weights observed in three channels of an equilibrium experiment on southern bean mosaic virus, strain S-I. The values of the $\sigma_w(r)$ should fall on one continuous curve if the concentration dependence is negligible and this line should be flat if the sample is homogeneous. With significant concentration dependence the curves for each initial concentration should be different whether or not the sample is homogeneous. The initial concentrations were: (▲) 0.14%, (●) 0.042%, and (○) 0.14%. See Table III for the remaining experimental conditions.

Values of the molecular weights of the species of interest may be obtained in this case directly from the data of the highest concentration channel (Fig. 9). Essentially the same values were found by application of equation (12a) with $j = 2$ to all the concentrations (column 3, Table III). Reasonable estimates were also obtained from the initial slopes of the graphs of the logarithms of the net fringe displacements (column 4). Other strains of southern bean mosaic virus also have been examined, e.g., a preparation of strain S-III appeared homogeneous with a molecular weight of $6.5_2 \pm 0.2_6 \times 10^6$. The values found may be compared with values in the literature of $6.63 \times 10^6$ from the combination of sedimentation and diffusion coefficients (Miller and Price, 1946) and of $6.3 \times 10^6$ from sedimentation equilibrium in a magnetically suspended rotor (Hexner *et al.*, 1962).

*Consideration of Errors.*—Random variations $\delta y_i$ in the individual measured $y$-displacements will introduce fluctuations in the calculated results. In addition, an uncertainty $\delta y_0$ in the choice of the zero concentration level will introduce a systematic error in any particular set of calculations. The total uncertainty in a derived value may be taken as the root mean square sum of the effects of these two types of uncertainty. The value

TABLE III
MOLECULAR WEIGHTS OF SOUTHERN BEAN MOSAIC VIRUS, STRAIN S-I[a]

| (1) Initial Concentration (%) | (2) Mol Wt from the Linear Least Squares Slope | (3) "Monomer" Mol Wt from Eq. (12a) | (4) Graphical Estimate of "Monomer" Mol Wt |
|---|---|---|---|
| 0.14 | 6.53 ± 0.01 × $10^6$ | 6.52 ± 0.16 × $10^6$ | 6.49 ± 0.18 × $10^6$ |
| 0.042 | 6.53 ± 0.02 | 6.57 ± 0.16 | 6.43 ± 0.13 |
| 0.014 | 6.71 ± 0.02 | 6.43 ± 0.06 | 6.56 ± 0.17 |

[a] Experimental conditions: 62 hours at 2378 rpm; pH 3.9, 0.1 M acetate; 6.5°; 12-mm three-channel Kel-F centerpiece.

5 $\mu$, typical of the standard deviation of an individual measurement from the mean, has been taken as an estimate of the standard deviation of the blank corrected $y$-displacements.[13] The uncertainty $\delta y_0$ in the choice of the zero-concentration level depends on the particular experiment considered. For good runs $\delta y_0$ appears to be less than 5 $\mu$. However, if there is only a small "flat" region near the meniscus, or if the blanks are poor, then this uncertainty may be larger. Expressions derived from a straightforward propagation of errors analysis are presented for estimation of the resultant errors in any particular case.

Values of $\sigma_w(r)$ calculated through equation (13) are subject to a fractional error given by

$$\frac{\delta\sigma_w(r)}{\sigma_w(r)} = \delta \ln \sigma_w(r) = \left(\frac{0.1\overline{\delta y^2}}{\sigma_w^2(r)r^2(\Delta r)^2} + \delta y_0^2\right)^{1/2} \Big/ (y(r) - y_0) \quad (16)$$

where $\Delta r = \Delta x/m$ is the $r$ interval between successive measurements and $\sqrt{\overline{\delta y^2}}$ is the standard deviation of the net deflections. The fractional error in values of $\sigma_n(r)$ calculated from equation (15) has been split up into two expressions: that part arising from uncertainty in $y_0$ is given by

$$\delta \ln \sigma_n(r_k) = \left[-1 + \frac{2\sigma_n(r_k)}{\sigma_w(r_1)} + \sigma_n(r_k)(r_k^2 - r_1^2)/2\right] \times \delta y_0/[y(r_k) - y_0] \cong (1 + \ln\{[y(r_k) - y_0]/[y(r_1) - y_0]\})\delta y_0/[y(r_k - y_0)] \quad (17)$$

where the approximation is exact if $\sigma_n(r_1) = \sigma_w(r_1) = \sigma$ is constant. Using the same approximation, the fractional error in $\sigma_n(r)$ from the random fluctuations of the $y$-displacements is given by equation (18), which is not valid for points too close to $r_1$.

$$\delta \ln \sigma_n(r_k) \cong \left\langle 2 + \frac{0.1}{\sigma^2 r_1^2(\Delta r)^2} + \frac{\ln\{[y(r_k) - y_0]/[y(r_1) - y_0]\}}{k - 1} \right\rangle^{1/2} \sqrt{\overline{\delta y^2}}/[y(r_k) - y_0] \quad (18)$$

The term containing the logarithm of the net displacements is usually insignificant since it reflects the contribution of the fluctuations to the sum of equation (15) where they are effectively averaged. For practical cases the major contribution appears to be from the second term which reflects the effects of the random errors on the choice of $\sigma_w(r_1)$, which in this derivation was assumed to have been determined by application of equation (13) at $r_1$.

The systematic error in the values of $2\sigma_n(r) - \sigma_w(r)$ caused by an uncertainty in $y_0$ is approximated by

$$\delta \ln [2\sigma_n(r) - \sigma_w(r)] \cong (1 + 2 \ln\{[y(r_k) - y_0]/[y(r_1) - y_0]\})\delta y_0/[y(r_k) - y_0] \quad (19)$$

and the random error may be obtained by root mean square summation of the random error in $\sigma_w(r)$ and twice the random error in $\sigma_n(r)$.

Values of the total error in $\sigma_w(r)$, $\sigma_n(r)$, and $2\sigma_n(r) - \sigma_w(r)$ are presented in Figure 10 as a function of the net fringe displacement, assuming typical values of the values: $\sigma = 5$ cm$^{-2}$, $\Delta r = 5 \times 10^{-3}$ cm, $y(r_1) - y_0 = 100$ $\mu$, and $\sqrt{\overline{\delta y^2}} = \delta y_0 = 5$ $\mu$. The errors are strong functions of concentration, tending to zero at the higher concentrations with the present assumption of 5 $\mu$ random error in the $y(r_1)$. The actual random errors, however, increase to, say, 0.01 of the net deflection as the fringes become too closely spaced for adequate resolution, leading to limiting values of the errors of about 2–3% with negligible effects from variations in the choice of the zero concentration level. Inspection of Figure 10 shows that under the assumptions made the uncertainty of a value of $\sigma_n(r)$ is about 1.5–2 times that of the corresponding $\sigma_w(r)$ and that uncertainty in a value of $2\sigma_n(r) - \sigma_w(r)$ is about 3–4 times that of the $\sigma_w(r)$. It should be noted that the choice of $r_1$ will influence the estimated errors: from equations (18) and (19) $r_1$ should be chosen as large as possible, consistent with a complete description of the data for points centripetal to $r_1$ by a single exponential curve. Other evaluations of the value of the integral $I$ are also possible, e.g., in special cases the value of $\sigma_n(r_1)$ may be known, as with a well-characterized sample. In such cases the error in the $\sigma_n(r)$ may be decreased by an appropriate formulation.

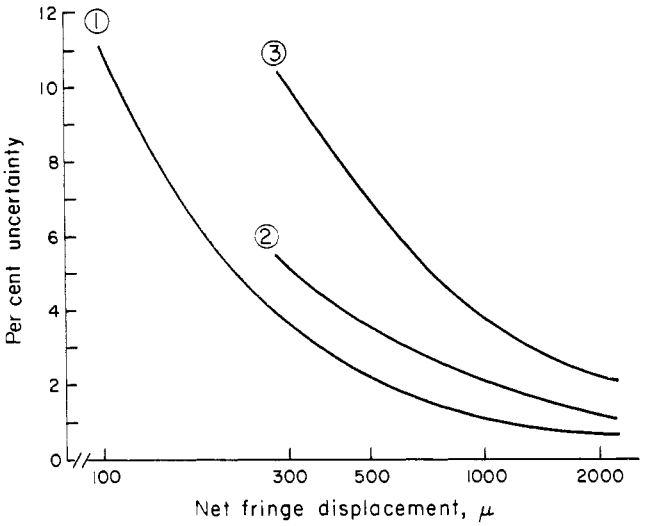


FIG. 10.—The total estimated uncertainty in calculated values of (1) the weight-average, (2) the number-average, and (3) the lowest molecular weights is presented as a function of the net fringe displacement (logarithmic scale). See the text for details.

There are other systematic errors that may be present with this method. These errors stem from the enhancement of the optical aberrations by the steep refractive index gradients used. Such aberrations have been examined both theoretically (Svensson, 1954a, 1956) and experimentally (Svensson, 1954b) and include as the most pronounced component in this study the familiar Wiener skewing. The contribution of these errors may be estimated from equation (60) of Svensson's (1956) consideration of the third-order aberrations of interferometric systems. This equation is reproduced below with slight changes in notation, omitting terms that vanish as a result of illuminating the cell with collimated light parallel to the axis of rotation:

$$\Delta S = h(n - m) + \frac{h^3(n - m)(dn/dr)^2}{6m^2} + \frac{h^3(dn/dr)^2}{6m}(2 - 3q) + \frac{h^5(dn/dr)^2 d^2n/dr^2}{30m^2}(15q^2 - 20q + 7) \quad (20)$$

Here $\Delta S$ is the optical path difference (fringe shift in terms of refractive index times distance), $h$ is the

[13] Least squares fitting of experimental distributions by sums of exponentials indicates typical standard deviations from the fitted curves of 3–5 $\mu$. More scatter than this is found when there are experimental difficulties, such as poor photography, oil deposition on the lenses, upset technicians, etc.

TABLE IV
TOTAL THIRD-ORDER OPTICAL ABERRATIONS AS PER CENT INCREASE OF ln $c$[a]

| Cell Thickness | $(dn_0/dr)$ | Fringe Displacement[b] (mm) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\sigma$ = 5 cm$^{-2}$ | | | | $\sigma$ = 7.1 cm$^{-2}$ | | | |
| | | 0 | 0.5 | 1.0 | 2.0 | 0 | 0.5 | 1.0 | 2.0 |
| 12 mm | 0 | 0 | 0.8 | 1.6 | 3.4 | 0 | 1.6 | 3.4 | — |
| | | *0* | *<0.1* | *<0.1* | *0.2* | *0* | *<0.1* | *0.2* | — |
| 30 mm | 0 | 0 | 2.0 | 4.4 | (9.9) | 0 | 4.4 | (9.9) | — |
| | | *0* | *<0.1* | *0.25* | *(1.1)* | *0* | *0.25* | *(1.1)* | — |
| 12 mm | 1.1 × 10$^{-3}$ cm$^{-1}$[c] | 0.6 | 1.4 | 2.3 | 4.4 | 0.9 | 2.6 | 3.8 | — |
| | | *0* | *<0.1* | *<0.1* | *<0.2* | *0* | *<0.1* | *0.25* | — |
| 30 mm | 1.1 × 10$^{-3}$ cm$^{-1}$ | 3.3 | 6.0 | 8.8 | — | 5.4 | 10.4 | (17.6) | — |
| | | *0* | *0.2* | *0.5* | — | *0* | *0.6* | *(1.7)* | — |

[a] Values given in italic type pertain to focus on the middle of the solution, all others to focus on a plane $^2/_3$ of the cell thickness from the front wall. ( ) denotes conditions where it is difficult to resolve the interference fringes; — indicates unobservable conditions. [b] One-mm fringe displacement corresponds to 3.46 fringes or a path length of 1.89 × 10$^{-3}$ cm for the particular ultracentrifuge used. [c] The value of 1.1 × 10$^{-3}$ cm$^{-1}$ corresponds to a 5 M guanidinium chloride solution at 50,740 rpm.

cell thickness, $m$ is the refractive index of the solvent, assumed to be constant in Svensson's treatment, $n$ is the refractive index of the solution at the point $r$ imaged by the camera lens, and $q$, the cell-focusing parameter, is the distance, expressed as a fraction of the cell thickness, from the front wall to the position in the solution that is exactly imaged onto the photographic plane, i.e, $q = {}^1/_2$ if the camera lens is focused on the middle of the solution as in these studies.

Equation (20) was derived on the assumption that the reference channel exhibits no change in refractive index with position. This is the case in the usual diffusion experiments but not in ultracentrifuge experiments. Here the refractive index gradients of the macromolecular species are superposed on refractive index gradients from solvent redistribution, solvent compression, and deformation of the windows. These gradients are assumed to be the same in both solvent and solution and to a first approximation are subtracted out by the interference optics. (Subtraction of the blank corrects for inequalities of cell path or differences in optical distortions of the windows between the two channels.) However, the total refractive index gradients should be considered in the evaluation of the aberrations. It will be assumed that all these common refractive index gradients have the same constant value $dn_0/dr$ over the (small) regions of interest in both solvent and solutions. In order to apply equation (20) we introduce a hypothetical reference channel with constant refractive index $m$ and consider the path difference $\Delta S_c$ between this hypothetical reference channel and the solution channel. The refractive index in the solution channel is $n = m + n_c$, where $n_c$ is the refractive increment of the solution at the point considered and the refractive index gradient in the solution channel is $dn/dr = dn_c/dr + dn_0/dr$. The actual path difference $\Delta S_{tot}$ may be obtained by subtracting from $\Delta S_c$ the path difference $\Delta S_0$ between the same hypothetical reference channel and the actual solvent channel with $n = m$ but with $dn/dr = dn_0/dr$. Dividing $\Delta S_{tot}$ by $hn_c$, the ideal path difference, and subtracting unity from both sides we have an estimate of the error in ln $\Delta S_{tot}$ as

$$\Delta \ln \Delta S_{tot} \cong \frac{\Delta S_{tot} - hn_c}{hn_c} = \frac{h^2\left(\frac{dn_c}{dr} + \frac{dn_0}{dr}\right)^2}{6m^2} + \frac{h^2\left[\left(\frac{dn_c}{dr} + \frac{dn_0}{dr}\right)^2 - \left(\frac{dn_0}{dr}\right)^2\right]}{6mn_c}(2 - 3q) + \frac{h^4\left(\frac{dn_c}{dr} + \frac{dn_0}{dr}\right)^2 \frac{d^2n_c}{dr^2}}{30m^2n_c}(15q^2 - 20q + 7) \quad (21)$$

This equation gives estimates for the fractional error in path difference for ultracentrifuge experiments in general when observed with collimated light parallel to the axis of rotation. The second term on the right represents the Wiener skewing which may be removed by focusing on the $^2/_3$ plane in the cell,[14] a procedure which also minimizes the last term (Svensson, 1956). The various derivatives of $n_c$ may readily be evaluated for equilibrium of a monodisperse ideal solute as $dn_c/dr = r\sigma n_c$ and $d^2n_c/dr^2 = r^2\sigma^2 n_c$ $(1 + 1/\sigma r^2)$ giving the following estimate for the error in ln $c$:

$$\Delta \ln c \cong \Delta \ln S_{tot} = \frac{h^2(dn_0/dr)^2 + 2h^2\sigma r(dn_0/dr)n_c + h^2\sigma^2r^2n_c^2}{6m^2} + \frac{h^2\sigma^2r^2n_c + 2h^2\sigma r(dn_0/dr)}{6m}(2 - 3q) + \frac{h^4\sigma^4r^4n_c^2 + 2h^4\sigma^3r^3(dn_0/dr)n_c + h^4(dn_0/dr)^2\sigma^2r^2}{30m^2}(1 + 1/r^2\sigma)(15q^2 - 20q + 7) \quad (22)$$

Examination of this equation shows that the errors are strongly dependent on the value of $\sigma$ as well as on the choice of the plane imaged on the photographic plate and that these errors may be reduced to quite small values by focusing on the $^2/_3$ plane (Table IV). In these studies the camera lens was focused on the middle[15] of the solution column. Although this appears acceptable with the most common types of experiments (12-mm cell thickness, low values of $dn_0/dr$, and $\sigma \sim 5$ cm$^{-2}$) a choice of the $^2/_3$ plane would have been superior from the standpoint of these errors and appears essential with the longer cells.

The value of 1.1 × 10$^{-3}$ cm$^{-1}$ for $dn_0/dr$ used in Table IV is an extreme; it corresponds to a total gradient from a typical cell with sapphire windows, filled with 5 M guanidinium chloride and spun at 50,740 rpm. Much lower values of $dn_0/dr$ are the rule, e.g., ~0.14 × 10$^{-3}$ cm$^{-1}$ for 0.15 M NaCl at 24,630 rpm, so that the terms in $dn_0/dr$ may usually be neglected. The errors then remaining are predominantly from the second set of terms of equation (22) and are either negligible when $q = {}^2/_3$ or else essentially proportional to the concentration, a behavior simulating linear concentration dependence—as though there were a small tendency towards dimerization. Although the apparent molecular weight averages $\sigma_k(r)$ and $\bar{\sigma}_k$

[14] Provided the magnification factor is determined using a calibration scale at the mid-plane of the cell.

[15] In centrifugations with sapphire windows and 12-mm cells the value of $q$ was ~0.59, tending to give somewhat smaller aberrations than with the quartz windows.

are affected these errors should have a negligible effect on the determination of $\sigma_1$ from equations (12) if $j$ is taken as 2.

On the other hand, runs in the longer cells at high values of $dn_0/dr$, as in concentrated urea or guanidinium chloride solutions or even in moderately concentrated phosphate buffers, appear subject to considerable error even in the limit of zero solute concentrations unless the $^2/_3$ focus is used. This zero-concentration error will persist in the determination of $\sigma_1$.

*Pitfalls.*—There appear to be two major dangers with this method of determining molecular weight. The first of these lies in the possibility of failure to satisfy the primary premise on which the method depends, i.e., that the meniscus concentration be negligible with respect to the concentrations used for the calculations. In general such a failure will be made obvious by the presence of a distinct refractive index gradient at the meniscus. However there are situations when these gradients are below the limits of detection. In such cases the average molecular weights determined in the usual manner may include only a partial contribution from the smallest species. In most such cases the apparent point-average molecular weights will vary significantly with position and the dispersity of the sample will be evident. Under a few conditions however, i.e., when less than 20% of the sample has a molecular weight of one-fourth or so of the main component, the sample will appear to be effectively monodisperse but with an apparent molecular weight less than that of the major component if the sample is examined as a dilute solution ($c_0 \cong 0.01\%$). Centrifugations of such samples at higher initial concentrations (say, 10 times higher) will, however, demonstrate the heterogeneity of the sample if there is enough of the smallest solute to significantly lower the apparent molecular weight observed at the lower initial concentration from the value corresponding to the main component.

The limit of detection of small molecules depends very much on the particular experimental conditions: the quality of the measurements, the extent of the flat region near the meniscus, and the size and concentration of the low and high $\sigma$ species all are significant factors. Estimates of the limits of detection of smaller components have been made by treating theoretically calculated distributions containing a smaller component exactly as though these distributions were experimental data subject to the usual uncertainties. Such estimates are necessarily somewhat subjective but do indicate the *approximate* range of detection of the presence of smaller species. The criteria of detection were the presence of a significant gradient near the meniscus or the variation beyond likely experimental error of the point-average molecular weights with position. If the $\sigma$ of the contaminating species is one-tenth the $\sigma$ of the main component, then according to these estimates one can detect the presence of about 40%, 20%, and 10% by weight of small component with respective initial concentration of 0.01%, 0.03%, and 0.1%. For contaminants of smaller effective size than this the sensitivity of detection is approximately proportional to the $\sigma$ of the small species. If the $\sigma$ of the contaminating species is one-fourth that of the major component the limits of detection are about 20%, 7%, and 5% for total initial concentrations of 0.01, 0.03, and 0.1%. Slightly less sensitivity of detection than this was estimated when the contaminating species is one-half the size of the major component. These estimates are predicated on the following assumptions: 12-mm cell thickness, equal refractive increments for the two ideal components, and $\sigma = 5$ cm$^{-2}$ for the major component. These are estimates from the data of only a single channel. More discrimination is possible by combining data from other initial concentrations. Larger experimental error than usual will of course make detection more difficult. Higher sensitivity in detecting smaller species may be provided by increasing both the operating speed and the initial concentrations of the runs but at the sacrifice of adequate characterization of the main component.

The other major danger originates from the mechanical limitations of the available ultracentrifuges. Vibration, precession, and temperature fluctuations will tend to induce convective disturbances unless the distributions are stabilized by a sufficient density gradient from the redistribution of the solute or solvent component (Pickels, 1942; Schachman, 1959; Yphantis, 1963). Conditions in these high speed equilibrium runs are particularly conducive to such convection: as the solute concentration approaches zero near the meniscus, the density gradient contribution from the solute becomes negligible and, in the absence of other stabilization, one can expect to find convection. Fortunately, biological preparations often are examined in solvents that automatically ensure adequate stabilization, but the margin of safety is not large. In the early runs, before precautions were taken to select drives for minimum vibration and precession, there were a number of cases in typical buffers where the concentration distributions were distinctly abnormal. At times there was even a complete absence of any observable concentration gradient in the channels with the lowest initial concentrations. Yet a change of speed by a few per cent, either up or down, showed the presence of solute in these channels by the rapid establishment of concentration gradients which, on prolonged centrifugation at the new speeds, often gave well-behaved equilibrium distributions. The conditions for instability were highly variable, the same speed giving both stable and unstable distributions from run to run. Usually the effects of convection are more subtle, consisting of an abnormally steep initial region (generally in the range below 100–200 $\mu$ net fringe displacement) on the graphs of log $c$ vs. $x$. At higher displacements this steep line bends abruptly and assumes a more moderate slope, in keeping with the expected slope. Sometimes these abnormal distributions appear to be steady states, with convection taking place below a certain concentration gradient and "eroding" the foot of the concentration gradient.[16] At other times the convection appears to be intermittent, with the concentration distribution not approaching a steady state, but varying with time, even to the extent of a recurrent boundary formation at the meniscus.

Convection has not been detected in the more recent experiments with the selected drives, *provided* the speed of the centrifuge was above 17,000 rpm and *provided* there was about 0.1 M or more of a typical salt in the buffer used for the experiment. Experiments with lower salt concentrations are feasible if a little sucrose (0.5–2%) is added to the buffer to provide a stabilizing density gradient. Experiments attempted without significant density gradients from solvent redistribution have invariably exhibited convection with the unexplained exception of centrifugations at speeds below 5000 rpm; centrifugations that appear convection free at such low speeds have been made

[16] Such behavior has often been observed with low-speed synthetic-boundary cell runs and leads to too low estimates of the refractive increment for the solution.

with several preparations that did not contain any small molecules to set up effective density gradients. (Since the density gradients from the solvent components are essentially proportional to $[\omega(1 - \bar{V}\rho)]^2$ the salt present with such preparations should have had a negligible stabilizing effect in comparison to that at the higher speeds.) Centrifugations at speeds between 5000 and 17,000 rpm have shown variable degrees of stability with the usual safe concentration of ~0.1 M, even though some experiments were performed with the heavy An-J rotor used for the lower speeds. Limited experiments in the upper part of this speed range showed (with one particular drive, an An-E rotor, and three-channel 30-mm Kel-F cells) that salt distributions were unstable if the predicted equilibrium density gradients were below $\sim 1.8 \times 10^{-4}$ g cm$^{-4}$. However, insufficient experiments have been carried out to determine the amount of stabilization required for routine convection-free centrifugations in this range of speeds. For higher speeds, salt concentrations of ~0.1 M appear sufficient if the value of $1 - \bar{V}\rho$ for the salts is $\gtrsim 0.4$. Salts with low values of $1 - \bar{V}\rho$, such as triethylammonium chloride ($1 - \bar{V}\rho \sim 0$), should not be depended on to provide density stabilization.

## Discussion

The method described is of moderate precision. It is not capable of the precision attained at higher concentrations with the usual approach using either the schlieren (Van Holde and Baldwin, 1958) or the Rayleigh interference optics (Richards and Schachman, 1959). The precision at these high speeds seems to be about 1–2% at best, even with ideal monodisperse preparations and after focusing so as to remove the Wiener skewing. Higher precision appears possible only at lower values of $\sigma$. Lower speeds and longer columns (required so that the meniscus concentration be negligible) would allow an increase of precision with this method, but only with a considerable increase in equilibrium time. However, the combination of a high speed and a lower speed experiment on the same solution in 3-mm columns can yield a higher precision with only a moderate increase in equilibrium time (to be published).

Offsetting this lack of precision are the low concentrations that may be observed and the centrifugal fractionation that is obtained at high speed. The low initial concentrations required suggest applications where only small quantities of solute are available. For example, it is feasible to examine many column effluents directly. About 0.1 ml of a protein solution of $OD_{280} \sim 0.1$ in a 12-mm cell (or ~0.25 ml of $OD_{280} \sim 0.04$ in a 30-mm cell) appears adequate for a partial study or indication of molecular weight. Careful studies, however, require runs at more than one initial concentration, with the lower initial concentration giving information about the larger species present and the more concentrated solutions about the smaller species.

The centrifugal fractionation imposes limitations on the application of this method. Polydisperse systems with broad distributions will be difficult to handle adequately. Two conditions must be met in order to obtain the molecular weight averages usually employed to characterize such systems: the meniscus concentration must be negligible and the refractive index gradients must not be so high as to preclude measurements near the base of the cell. It will not be possible to satisfy these requirements simultaneously at useful initial concentrations if the molecular weight distributions are too broad. Most biochemical preparations are paucidisperse systems with only a moderate size range. For these the method will usually provide estimates of molecular weight averages that are of moderate accuracy. However, the possibility of missing contributions from the smallest and largest species should be kept in mind. Often only a very limited size range is of interest and gel-filtration chromatography (e.g., on Sephadex) may be useful for removal of components outside this range.

Centrifugal fractionation increases the sensitivity of detection of heterogeneity. The larger the concentration ratio (from the meniscus to the base of the cell) of an equilibrium distribution, the more the effects of heterogeneity are accentuated. At the higher speeds used here each equilibrium distribution spans several decades in concentration, of which about one decade may be reliably observed in a single channel. In comparison, only a 2- to 4-fold range is available in the usual ultracentrifugation at lower speeds. In ideal systems, the reliable measurement range may be extended to two decades by combining measurements from a 10-fold higher initial concentration (e.g., Fig. 9). Even larger spans of the equilibrium distributions may be examined by running higher initial concentrations at higher speeds.

Under certain conditions the centrifugal fractionation also permits estimation of the size of the main component. In order to do this it must be shown, from velocity experiments or other knowledge about the sample, that there is no significant quantity of a smaller species and that the amount of the heavier species is not too large. Many biochemical preparations are unstable and, although perhaps initially monodisperse, they aggregate readily to give paucidisperse solutions, with the species of interest contaminated by its dimer or higher polymers. In such a case the present technique may prove useful for characterization of the monomer. Another obvious application is in the determination of the existence and size of protein subunits. The ability to study dilute solutions should prove valuable when the dissociation to subunits is concentration dependent.

Studies of reversibly aggregating systems can yield information not only about the size of the species involved but also about the values of the equilibrium constants and the derived quantities $\Delta F°$, $\Delta H°$, and $\Delta S°$. Such studies are currently under way on human hemoglobin. In certain cases the equilibria may be studied successfully even though there is a significant amount of heavier species contaminating the system but not involved in the reactions.

The concentrations used for the measurements are low, under 0.1–0.2%, so that nonideality is usually not a serious disturbing factor. Small amounts of the usual linear concentration dependence will be taken care of automatically during application of equations (12) when $j$ is taken as 2. For these reasons no extensive discussion of the effects of concentration dependence has been made. Systems exhibiting large nonideality, both from charge effects and from association, will be considered in forthcoming publications.

## Acknowledgments

This work was made possible by the patient and painstaking technical assistance of Owen Griffith and Raymond Kikas. I am also grateful to L. G. Longsworth and A. T. Ansevin for many invaluable discussions, to T. P. King, D. L. Eaker, and B. Magdoff for generous gifts of preparations, and to N. Jernberg and

his associates for their skillfull help in the design and fabrication of cell components.


REFERENCES

Archibald, W. J. (1942), *Ann. N. Y. Acad. Sci. 43*, 211.
Casassa, E. F., and Eisenberg, H. (1961), *J. Phys. Chem. 65*, 427.
Dayhoff, M. O., Perlmann, G. E., and MacInnes, D. A. (1952), *J. Am. Chem. Soc. 74*, 2515.
Donnelly, T. H. (1960), *J. Phys. Chem. 64*, 1830.
Fujita, H. (1962), Mathematical Theory of Sedimentation Analysis, New York, Academic.
Goodman, M.. Schmitt, E. E., and Yphantis, D. A. (1962), *J. Am. Chem. Soc. 84*, 1283.
Harrington, W. F., and Schellman, J. A. (1956), *Compt. Rend. Trav. Lab. Carlsberg, Ser. Chim. 30*, 21.
Hexner, P. E., Kupke, D. W., Kim, H. G., Weber, F. N., Jr., Bunting, R. F., and Beams, J. W. (1962), *J. Am. Chem. Soc. 84*, 2457.
Hirs, C. H. W., Stein, W. H., and Moore, S. (1956), *J. Biol. Chem. 221*, 151.
Jenkins, F. A., and White, H. E. (1957), Fundamentals of Optics, 3rd ed., New York, McGraw-Hill.
Kegeles, G., and Gutter, F. J. (1951), *J. Am. Chem. Soc. 73*, 3770.
King, T. P., and Craig, L. C. (1958), *J. Am. Chem. Soc. 80*, 3366.
LaBar, F. E., and Baldwin, R. L. (1963), *J. Am. Chem. Soc. 85*, 3105.
Lansing, W. D., and Kraemer, E. O. (1935), *J. Am. Chem. Soc. 57*, 1369.
Miller, G. L., and Price, W. C. (1946), *Arch. Biochem. 10*, 467.
Pedersen, K. O. (1936), *Biochem. J. 30*, 961.
Pickels, E. G. (1942), *Chem. Rev. 30*, 341.
Richards, E. G., and Schachman, H. K. (1959), *J. Phys. Chem. 63*, 1578.
Schachman, H. K. (1959), Ultracentrifugation in Biochemistry, New York, Academic.
Schachman, H. K. (1963), *Biochemistry 2*, 887.
Schachman, H. K., Gropper, L., Hanlon, S., and Putney, F. (1962), *Arch. Biochem. Biophys. 99*, 175.
Svedberg, T., and Pedersen, K. O. (1940), The Ultracentrifuge, London, Oxford University Press (Johnson Reprint Corp., New York).
Svensson, H. (1954a), *Optica Acta 1*, 25.
Svensson, H. (1954b), *Optica Acta 1*, 90.
Svensson, H. (1956), *Optica Acta 3*, 164.
Trautman, R., and Burns, V. W. (1954), *Biochim. Biophys. Acta 14*, 26.
Van Holde, K. E., and Baldwin, R. L. (1958), *J. Phys. Chem. 62*, 734.
Wales, M. (1951), *J. Phys. Chem. 55*, 235.
Wales, M., Adler, F. T., and Van Holde, K. E. (1951), *J. Phys. Chem. 55*, 145.
Weaver, W. (1926), *Phys. Rev. 27*, 499.
Williams, J. W., Baldwin, R. L., Saunders, W. M., and Squire, P. G. (1952), *J. Am. Chem. Soc. 74*, 1542.
Williams, J. W., Van Holde, K. E., Baldwin, R. L., and Fujita, H. (1958), *Chem. Rev. 58*, 715.
Yphantis, D. A. (1960), *Ann. N. Y. Acad. Sci. 88*, 586.
Yphantis, D. A. (1963), *in* Ultracentrifugal Analysis in Theory and Experiment, Williams, J. W., ed., New York, Academic.
Yphantis, D. A., and Waugh, D. F. (1956), *J. Phys. Chem. 60*, 623.


# Chemical Studies on Mixed Soluble Ribonucleic Acids from Yeast

DAVID BELL,* R. V. TOMLINSON, AND G. M. TENER†

*From the Department of Biochemistry, University of British Columbia, Vancouver, British Columbia*

*Received September 20, 1963*

The products of alkaline digestion of a mixed s-RNA preparation, from growing yeast, were analyzed by DEAE-cellulose and paper chromatography, and by electrophoresis. The isolated components were estimated spectrophotometrically and the composition of the s-RNA was determined. Eighty per cent of the s-RNA chains were shown to have adenosine in the 3′-terminal position and the remainder had cytidine. More than 90% of the 5′- termini were phosphorylated and consisted of 78% pGp..., 10% pUp..., 7% pAp..., and 5% pCp.... A pancreatic ribonuclease digest of the same s-RNA preparation was subjected to chromatography on DEAE-cellulose in the presence of 7 M urea, thereby fractionating the components on the basis of net negative charge. Phosphorylated 5′-terminal sequences were eluted along with nonterminal fragments containing two more purine residues. They were then isolated as previously described by R. V. Tomlinson and G. M. Tener (1962, *J. Am. Chem. Soc. 84*, 2644; 1963, *Biochemistry 2*, 697, 703). The distribution of the major and minor bases in the mono-, di-, tri-, ... nucleotides was determined and the composition of the s-RNA was compiled from the results. Some of the sequences arising from phosphorylated 5′- termini were identified: 11.5% of the molecules terminated in pUp..., 6.5% in pCp..., 26% in pGpCp..., 0.7% in pApUp..., 0.4% in pGpUp..., and less than 0.1% in pApCp. Mixed trinucleoside tetraphosphates and tetranucleoside pentaphosphates from 5′- ends were also isolated and some evidence for the existence of the terminal sequences $pGp(Gp)_4Up...$ and $pGp(Ap)(Gp)_3Up$ was also obtained.

Soluble ribonucleic acid (s-RNA) has been the subject of intensive investigation ever since the discovery by Hoagland *et al.*, in 1957, that this RNA fraction, in the presence of the appropriate enzyme systems, is capable of accepting acyl-activated amino acids and transferring them to protein.

Present evidence indicates that s-RNA is a mixture of short-chain polyribonucleotides of molecular weight 20,000–30,000 (Osawa, 1960; Luborsky and Cantoni, 1962), each probably able to accept only one amino acid (Hoagland *et al.*, 1957, 1958; Berg and Offengand, 1958). Active s-RNA molecules are believed to ter-

* Present address: Department of Biochemistry, University of Glasgow, Glasgow, Scotland.
† Medical Research Associate, Medical Research Council of Canada. This work has been supported by a United States Public Health Service grant (CA-05342-03).