THE INVESTIGATION OF PROTEINS BY DIFFUSION MEASUREMENTS[1]

HANS NEURATH

Department of Biochemistry, Duke University School of Medicine, Durham, North Carolina

Received January 12, 1942

CONTENTS

It is the object of the present contribution to the Symposium to review the method of diffusion as a tool for investigating the properties of proteins in solution. While the diffusion process has been known to the physical chemist for a long time, its importance for biochemical research was recognized only more recently. Modern developments in the field of protein chemistry, in particular, have contributed to a revival of the interest in diffusion measurements which, together with sedimentation and electrophoretic measurements, constitute the most revealing methods for determining the state of dispersion of protein solutions. It is the aim of the present review to demonstrate the usefulness of diffusion measurements for a study of proteins, and to consider the theoretical and practical limitations of the methods available at present.

I. Theory

It would take us too far afield to discuss here in detail the theory of diffusion. Reference may be made to the earlier publications on this subject, excellently reviewed by Williams and Cady in this journal (56), and to the more recent kinetic treatment of the diffusion problem by Eyring and coworkers (6, 11). In the following, we shall merely review briefly the elementary principles of the diffusion process as it is most frequently met in actual experimentation.

[1] Presented at the Symposium on Physicochemical Methods in Protein Chemistry, which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Biological Chemistry at the 102nd Meeting of the American Chemical Society, Atlantic City, New Jersey, September 8–12, 1941.

We consider a system in which a solution of concentration c_0 is in contact with its solvent, as shown in the top section of figure 1. The boundary between solution and solvent is at the point $x = 0$, and positive values are assigned to x in the direction of increasing concentration. It will be assumed that no external force other than that of osmotic pressure is acting on the solute molecules and that the solute concentration is in the "ideal solution" range. Diffusion is

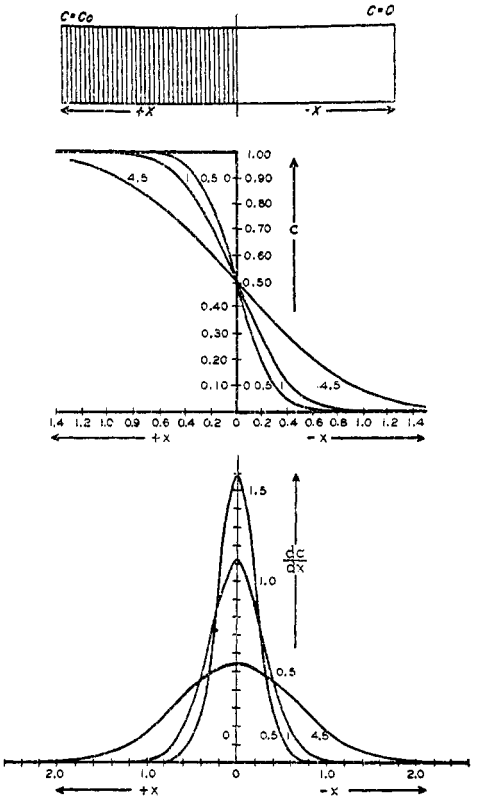

FIG. 1. Relation between concentration and distance of migration in a diffusion column. Top section: graphical illustration of the diffusion column. Center section: relation between concentration and distance of migration. Bottom section: relation between concentration gradient and distance of migration. The curves as drawn refer to diffusion times of 0.5, 1, and 4.5 hr., respectively. The whole diagram should be rotated counterclockwise about an angle of 90° in order to conform with the proper space directions.

confined to the vertical direction, while the force of gravity maintains the concentration constant within the horizontal plane. Under these conditions, the rate of diffusion is, according to Fick,

$$\frac{ds}{dt} = -DA\frac{dc}{dx} \tag{1}$$

where ds is the quantity of solute which in the time dt diffuses across a boundary of cross-sectional area A, under the influence of a concentration gradient dc/dx.

D is the diffusion constant, characteristic of the physical properties of the solute molecules in question. Its dimensions are square centimeters per second. Expressing mass in terms of concentrations, equation 1 may be reduced to the differential form

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} \tag{2}$$

The solution of this equation depends on the conditions imposed by the experimental methods applied for determining the diffusion rate.

For the conditions as stated above, the solution of equation 2 must fulfill the following boundary conditions: for $t = 0$,

$$c(x, 0) \begin{cases} = 0, & \text{for } x < 0 \\ = c_0, & \text{for } x > 0 \end{cases}$$

where c_0 is the initial solute concentration. Assuming that during the diffusion process no concentration changes occur at the extreme ends of the cell, integration of equation 2 leads to the following relation between c, x, and t:

$$c_x = \frac{c_0}{2}\left(1 - \frac{2}{\sqrt{\pi}} \int_0^y e^{-y^2}\, dy\right) \tag{3}$$

where

$$y^2 = \frac{x^2}{4DT}$$

Diffusion constants, D, can be calculated from either the c–t relation at constant x, or from the c–x relation at constant t. Both methods are used in actual experimentation.

The relation between c and x, for various constant values of t, is shown in the center section of figure 1. It will be noted that at the position of the original boundary between solution and solvent, $x = 0$, c_x is equal to $c_0/2$ during the entire experiment and that the curves flatten out as t increases. The light-absorption method, discussed below, furnishes a relation of this type. If the differential, dc/dx, of the curves shown in the center section of figure 1 is plotted against x, a family of curves is obtained as shown in the bottom section of figure 1. In the ideal case, these curves have the shapes of Gaussian distribution curves and are identical with one another with respect to their areas. Curves of this type are obtained with the refractometric diffusion method. They follow the equation

$$\frac{dc}{dx} = \frac{c}{2\sqrt{\pi Dt}}\, e^{-\frac{x^2}{4Dt}} \tag{4}$$

II. Experimental Methods

While a large number of methods have been designed from time to time to measure the diffusion rates in solution, we shall confine this discussion to those methods which have found general application to proteins.

A. THE POROUS-DISK METHOD

The first method to be described is based on the diffusion across a porous diaphragm which separates solution from solvent. It was introduced by Northrop and Anson (36) and by McBain and Liu (22). The diffusion cell consists of a bell-shaped Pyrex-glass vessel, closed at the narrow top end by a stopcock and at the wide bottom end by a sealed-in sintered-glass disk, as shown in figure 2. The cell is filled with the solution the diffusion rate of which is to be measured, and is immersed in a vertical position in the solvent, just touching its surface. The whole apparatus is mounted in a constant-temperature bath. After preliminary manipulations, necessary for establishing a diffusion gradient within the pores of the diaphragm, diffusion is allowed to proceed from the solution into the solvent, density differences and slight convection currents tending to keep the concentration uniform within each compartment. That is to say, diffusion

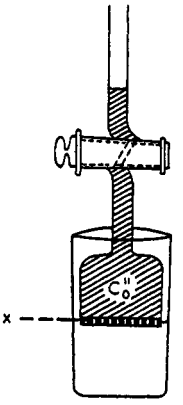

FIG. 2

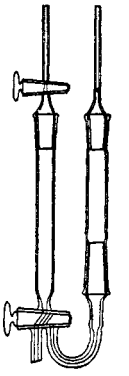

FIG. 3

FIG. 2. Diagram of the porous-disk diffusion cell (according to McBain and Liu (22)). x denotes the mean position of the boundary, c_0'' the initial solute concentration.

FIG. 3. Cylindrical diffusion cell, according to Svedberg (48) and Tiselius and Gross (53)

takes place only within the sintered-glass disk of thickness x, the concentration gradient dc/dx being constant within the pores, provided the diffusion rate is independent of solute concentration. Concentrations are determined after varying time intervals in aliquots of solution and solvent, and diffusion constants are calculated from the respective concentration changes by means of the equation

$$D_1 = \frac{1}{kt}\frac{V''V'}{V'' + V'}\left\{\ln c_0'' - \ln\left[c_0'' - \left(1 + \frac{V'}{V''}\right)c'\right]\right\} \qquad (5)$$

where k is a constant, equal to the mean effective surface area of the disk, t the time of diffusion, c_0'' the initial concentration of the solution, c' the concentration of the diffusate, and V'' and V' the volumes of the solution and diffusate, respectively (23). D_1 is the integral diffusion constant representing a mean value over a range of varying concentrations. It is identical with the true

diffusion constant D if the diffusion rate is independent of concentration. Otherwise, it can be related to it (23) by means of the equation

$$D = D_1 + c\frac{dD_1}{dc} \tag{6}$$

The constant k is determined by calibration with a solution of known diffusion constant, such as potassium chloride or sodium chloride. A thorough discussion of the method may be found in the papers by Mehl and Schmidt (23), Anson and Northrop (2), and McBain and Liu (22). The virtues of the porous-disk method are twofold. In the first place, it is simple experimentally and readily available. In the second place, it facilitates the determination of the diffusion constants of biologically active proteins, such as enzymes and viruses, in dilutions in which the concentration can hardly be measured by chemical methods but can be readily determined on the basis of biological activity. By comparing the rate of diffusion of the activity with that of a protein of known diffusion constant, the size of the smallest, biologically active, molecular kinetic unit can be estimated (10). The shortcomings of the method lie in the fact that it furnishes only a mean value for the diffusion constants if the solutions are polydisperse, and that only with the utmost care can satisfactory reproducibility of the results be obtained. The rather wide limits in experimental error that can be found in some of the published data (46) clearly emphasize this point. The initial decrease in diffusion rate observed by several investigators may represent a source of uncertainty worthy of further investigation. The method has been applied by Northrop and Anson (36) to hemoglobin, by Northrop to pepsin (35), by Scherp to trypsin[2] (45), by Kunitz and Northrop to chymotrypsinogen and chymotrypsin (14), by McBain, Dawson, and Barker to egg albumin (21), by Zeile (57) and by Stern (46) to catalase, and more recently by Hand to hemoglobin and catalase (10). A comparison of some of these results with those obtained by measurements of the absolute diffusion rate will be given in section III of this review.

B. THE LIGHT-ABSORPTION METHOD

The principle of this method is essentially the same as that employed in the earlier ultracentrifugal studies, and has been reviewed in detail by Svedberg and Pedersen (49). The method permits measurements of absolute diffusion rates and is based on the differences existing in the absorption spectra of solute and solvent. It may be described briefly as follows (53):

A transparent diffusion cell in which a sharp boundary between solution and solvent has been established is placed in a constant-temperature bath, between a light source and a photographic camera. Monochromatic light of such a wave length is used as will be absorbed by the solute but not by the solvent. With proteins containing no chromophoric groups, the absorption maximum lies in

[2] Scherp's value for the diffusion constant of trypsin is lower than that found for hemoglobin. As, however, the molecular weight of trypsin appears to be lower than that of hemoglobin, Scherp's data are probably erroneous.

the ultraviolet region of about 270 mμ wave length. With others, such as hemoglobin or cytochrome, the absorption maximum is in the visible range of the spectrum. The cell is photographed at the beginning of the diffusion experiment and at certain time intervals thereafter, care being taken to keep the light intensity and exposure time constant. Under such conditions, the concentration distribution in the cell is imaged on the plate by a corresponding distribution of light intensity. The relation between concentration and degree of blackening of the photographic plate is obtained by filling the cell with known dilutions of the solution and taking exposures with the same light intensity and exposure time as were applied in the diffusion run proper. The reference plate and experimental plate are then measured with a microphotometer, yielding a relation between c and x of the type pictured in the center section of figure 1.

The diffusion cell, as designed by Svedberg (48) and by Tiselius and Gross (53), is pictured in figure 3. It consists of a U-shaped glass tube, with the two limbs separated from each other by a three-way stopcock. The left limb is partly filled with the solution and the right limb with the solvent; a sharp boundary is formed by opening the stopcock after temperature equilibration in the constant-temperature bath. The boundary is then moved up in the solution compartment by means of a proper compensating arrangement until it reaches the center of the photographic field. The beginning of the diffusion process is made to coincide with the time of formation of the boundary.

Diffusion constants are calculated from the microphotometer registration curves by means of the equation

$$D = \frac{1}{4y^2} \frac{x_2^2 - x_1^2}{t_2 - t_1} \tag{7}$$

where x_1 and x_2 are distances from the boundary at the times t_1 and t_2, respectively. The most convenient method of calculation consists in determining from two successive curves (t_1 and t_2) the values of x which correspond to a given value of c, such as $c/4$ or $3c/4$. In that case, $1/4y^2$ is a constant. The method has been used by Tiselius and Gross (53) with protein concentrations as low as 0.2 per cent. The light-absorption method has given way in recent years to the light-refraction method, which with both sedimentation and diffusion measurements permits a considerably higher degree of accuracy.

C. OTHER METHODS

Before considering in detail the refractometric-scale method, mention will be made at this place of a diffusion method which, however, is more of theoretical significance than of practical value. It is due to Svedberg (48) and is based on the measurement of the degree of boundary blurring occurring during sedimentation in the ultracentrifuge. In sedimentation, the centrifugal force acting on the solute molecules is opposed by their diffusion. While, with such high centrifugal fields as are employed in sedimentation-velocity measurements, the rate of diffusion is small in comparison with that of sedimentation, nevertheless it becomes perceptible by a blurring of the originally sharp, sedimenting boundary.

The calculation of the diffusion constant is based on a comparison of the experimental sedimentation curves with those to be expected if no diffusion were to occur. Basic to this method is the assumption that the solutions under investigation are strictly monodisperse. The method is, however, rather inaccurate, owing to the fact that (*1*) the extent of diffusion taking place during a sedimentation run is actually very small, (*2*) the temperature in the ultracentrifuge is rarely sufficiently constant, and (*3*) any small inhomogeneity of the solution will have a significant influence on the apparent diffusion constant calculated from the sedimentation diagrams.

The method as applied to both the light-absorption and the light-refraction systems has been described in detail in Svedberg and Pedersen's monograph (49).

Bourdillon (3) recently designed a simple, yet rather crude, method for estimating the diffusion constant of viruses.

D. THE REFRACTOMETRIC-SCALE METHOD

This method is based on Stefan's (16, *loc. cit.*) phenomenon of the curvature of light when passing through a medium of varying refractive index. The method was introduced by Wiener (55) and Thoevert (51, 52) and further developed by Lamm (15, 16). The basic principle is illustrated in figure 4

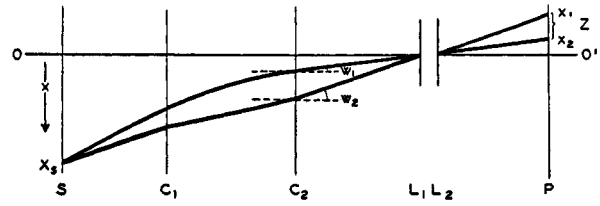

FIG. 4. Light path in the refractometric-scale method (not drawn to scale), according to Lamm (15, 16). *O–O'* is the optic axis, x the direction of increasing refractive-index gradient, S the plane of the transparent scale, C_1 and C_2 the planes enclosing the diffusion cell, L_1 and L_2 the principal planes of the camera lens, and P the plane of the photographic plate. W_1 and W_2 are the exit angles of the deviated and undeviated light pencils. For further explanations see text.

which, for purposes of demonstration, has not been drawn to scale. *O–O'* is the optic axis, S the plane of a transparent scale, and C_1 and C_2 the vertical planes enclosing the diffusion cell; L_1 and L_2 are the principal planes of a photographic lens, and P is the plane of a photographic plate. We shall first assume that the medium contained in the diffusion cell has a constant index of refraction. In this case, light emerging from the point x_s on the scale will be deflected linearly between the planes C_1 and C_2 and will be imaged on the plate at x_1. If, however, there is a refractive-index gradient in the medium, increasing in extent in the downward direction, light emerging from x_s will be deflected in a curved fashion and will be imaged at x_2. The scale-line displacement, $x_1 - x_2$, is proportional to the differences between the exit angles of the undeviated and deviated light pencils (W_2 and W_1), which in turn is proportional to the refractive-index

gradient, dn/dx, existing in a horizontal plane at a depth conjugate to about the position of the point x_2:[3]

$$x_1 - x_2 = Z = Gl(W_2 - W_1) = Gab\frac{dn}{dx} \tag{8}$$

Here G is the photographic magnification factor, l the optical distance $\overline{P - L_1}$, a the optical distance $\overline{C_1 - C_2}$, and b the optical distance from the scale to the center of the cell[4]. By determining the scale-line displacement for every point on the scale, the refractive-index gradient distribution can be determined over the whole depth of the cell. This method yields, therefore, a relation between the refractive-index gradient and the distance from the boundary. Since, however, in general, the index of refraction of a protein solution is a linear function of its concentration, of the type

$$n_1 - n_0 = kc \tag{9}$$

where n_1 and n_0 are the refractive indices of solution and solvent, respectively, one virtually obtains a relation between concentration gradient, dc/dx, and x as shown in the bottom part of figure 1.

The diffusion assembly consists of a transparent diffusion cell placed in a constant-temperature bath, a transparent scale placed behind it, and a photographic camera placed in front of it, the camera lens being focussed on the scale. The scale is illuminated uniformly from behind by a suitable light source. The image of the undistorted scale lines is obtained by photographing the scale when the cell is filled with a medium of constant refractive index, usually the solvent (reference scale). The image of the displaced scale lines is obtained by taking exposures of the diffusion column at suitable time intervals after the beginning of the diffusion process (experimental scale). The scale-line distances are then measured on the plate in a microcomparator for both the reference scale and the experimental scale, and the differences in reading (equal to the scale-line displacements Z) are plotted as ordinates against the corresponding readings on the experimental scale as abscissae (equal to the positions of the displaced lines, z).

The experimental details of the method have been described in full by Lamm (16) and summarized in brief by Polson (42), Saum (44), and Neurath and Saum (33). In view of the general importance of Lamm's method and the present difficulties of obtaining access to some of the European journals, it may be appropriate to review here certain experimental details of a description of the apparatus which was built by the author several years ago and reconstructed since.

[3] It is actually conjugate to the position of the point x_2 after the latter has been corrected for F (see equation 21).

[4] The optical distance is equal to the geometric distance divided by the refractive indices of the media traversed by the light.

1. Apparatus

An illustration of the diffusion assembly is given in figure 5. It is mounted on two horizontal I beams, bolted together in parallel position and resting on three concrete blocks. The light source (enclosed by a metal housing), light filters, and a simple condensing lens are mounted on an optic bench behind the constant-temperature bath. The light source is a Westinghouse high-intensity mercury lamp, H_3, 85 watts. Monochromatic light of 546 mμ wave length is obtained by a combination of Corning glass filters No. 551 and No. 512 and a Wratten filter No. 74.

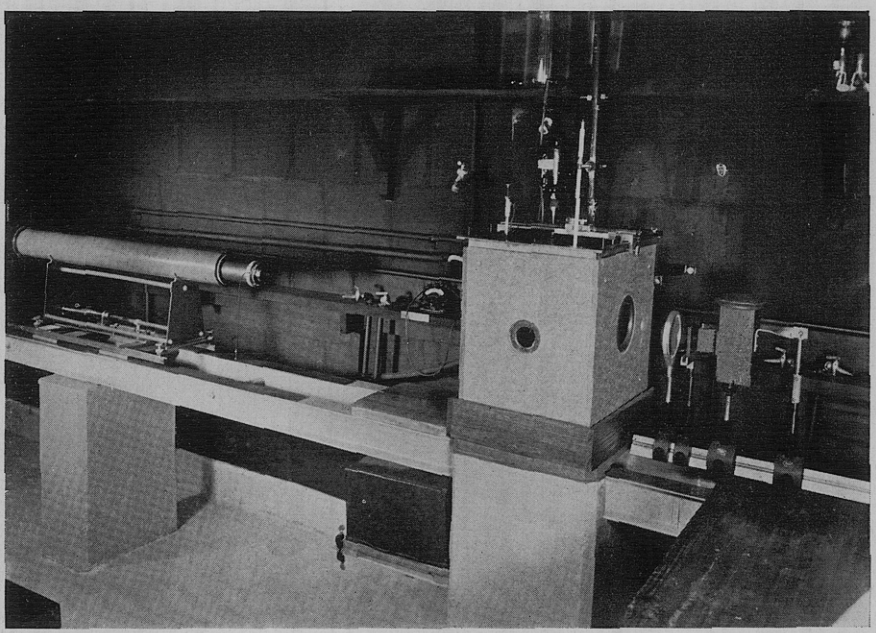

FIG. 5. Diffusion assembly for the refractometric-scale method, as used in the author's laboratory (for description see text).

A close-up view of the constant-temperature bath is shown in figure 6. The tank, 13 in. long in the direction of the optic axis, 16 in. wide, and 18 in. high, is made of brass and insulated by $\frac{1}{4}$-in. Celotex boards. Two circular glass windows, $4\frac{1}{2}$ in. in diameter and $\frac{1}{4}$ in. thick, and optically flat to within 0.5 wave length, are mounted on the walls perpendicular to the optic axis while a third, smaller window is mounted on the wall facing the observer. This latter window serves merely for the observation of the cell during an experiment. The tank is cushioned against mechanical vibrations by several alternate layers of sponge rubber and Celotex and is isolated from the optical assembly, since the I beams of the latter pass through a channel in the supporting concrete block without any direct contact with the tank.

A cross-piece, running parallel to the optic axis and mounted on the top edge

of the tank, holds the scale and the diffusion cell; the latter is attached to a movable rider which provides for variations in the cell–scale distance. This distance is measured on a graduated metal scale mounted on the cross-piece.

The transparent scale is ruled on glass and protected by a cover glass, held in place by transparent Lucite cement. It is 5 cm. in length and graduated in 0.02-cm. divisions[5].

A 1/16th H. P., vertical-shaft motor stirrer is mounted separately on a floor flange. Splashing is avoided by a hollow metal cylinder surrounding the shaft and suspended vertically from the top cover of the tank. The cylinder is 3 in.

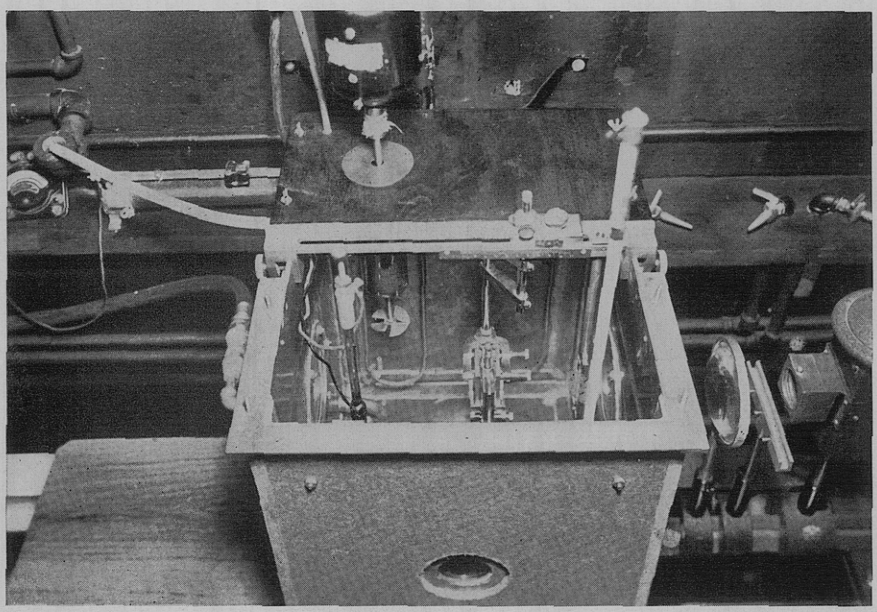

FIG. 6. Close-up view of constant-temperature bath of the diffusion assembly. The objects to be seen inside the tank are, from left to right, the thermoregulator, the stirrer with surrounding hollow cylinder, the cell, the scale, and the thermometer (for further description see text).

in diameter and terminates about 3 in. above the stirrer blades. A mercury thermoregulator with stainless-steel housing, in conjunction with a thyratron vacuum-tube relay, keeps the temperature constant to within a few thousandths of 1°C.

The camera lens, $1\frac{3}{5}$ in. in diameter, is a two-element projector lens of 24-in. focal length, corrected for spherical and chromatic aberration. It is focussed on the scale with a magnification of 1.1:1. A simple compur shutter provides for adjustment of the aperture (between *f*:50 and *f*:100) and exposure time. A slit in the back end of the camera and a movable plate-holder carriage allow a series of eight exposures to be taken on one plate. Eastman spectroscopic

[5] Scales of this type may be obtained from the Gaertner Scientific Corporation.

plates GV, sensitized for the green, proved to be best suited for sharp contrast and high light sensitivity. With an aperture of f:85, the exposure time is about 1 to 2 sec. The optical distances involved in the present diffusion assembly are as follows: plate to camera lens, 140.0 cm.; lens to scale, 113.9 cm.; b, the optical distance between the center of the cell and the scale, is variable between the limits of 3 and 11 cm.

The cell

In earlier studies the glass cell shown in figure 3 was employed. Its main disadvantage lies in the optical distortions produced by the lens effect of the cylindrical glass tube.

Lamm (16) designed a stainless-steel cell with plane-parallel windows, which eliminates most of the shortcomings of the Svedberg cell. A drawing of this cell is shown in figure 7. The cell consists of a stainless-steel disk with a rectangular, vertical slot in the center. Two optically flat, circular glass windows are pressed against the faces of the disk by threaded metal rings. A diaphragm, sliding in the horizontal slot and operated from the outside by a screw arrangement, serves to divide the vertical slot into two compartments.[6] The circular opening is filled with mercury through the right-hand side tube, in order to prevent leakage. The cell is filled with the solution to above the middle, with the diaphragm partly recessed. The lower half is then separated by advancing the diaphragm; the solution remaining above is rinsed out with solvent; and the top half of the cell is filled with the solvent. The boundary is then formed by slowly withdrawing the diaphragm. It is, of course, imperative that after each step of the above procedure the cell be brought to temperature equilibrium in the constant-temperature bath in order to eliminate thermal convection and volume changes.

Recently, the author (27) has described a diffusion cell[7] which is somewhat similar to that previously introduced by Loughborough and Stamm (19). This cell operates on the principle of the Tiselius electrophoresis cell; however, unlike the conditions with the latter, the boundary is formed directly in the photographic field and thus does not have to be moved by special compensating arrangements. The advantage of this cell over the Lamm cell lies in the fact that solution and solvent surfaces are in direct contact as the boundary is formed, thus eliminating the displacement of the upper column of liquid occurring in the Lamm cell when the diaphragm is withdrawn. The latest model of the cell, a detailed description of which has already been published (27), is pictured in figure 8.

The cell consists of two U-shaped stainless-steel blocks which can be slid

[6] With the cell used in the author's laboratory, the stainless-steel diaphragm originally in use has been replaced by one made of polystyrene, in order to avoid scratching of the glass windows by the hard steel edges. Lubriseal has been found to be a very satisfactory sealing material.

[7] The cell was designed in coöperation with Mr. H. S. Bush, Instrument Maker of the Department of Chemistry of Cornell University, from whom it can be obtained.

horizontally past each other by a screw arrangement. Two optically flat, rectangular glass windows are pressed against the vertical faces by the top and bottom clamps. The lower compartment, when separated from the top one, is filled with the solution through the left-hand tube, while the upper compartment is filled with the solvent through the right-hand tube. After temperature equilibration, the sharp boundary is formed by bringing the two compartments into vertical alignment. The dimensions of the cell proper are as follows: 1.7 cm. in the direction of the optic axis, 0.5 cm. wide, and 5 cm. high.

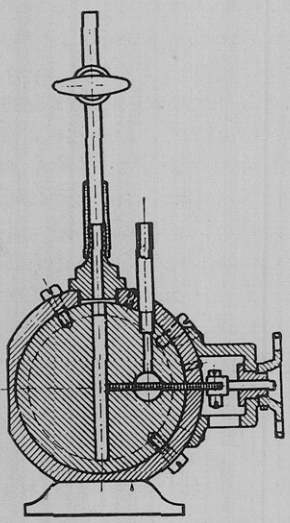

FIG. 7

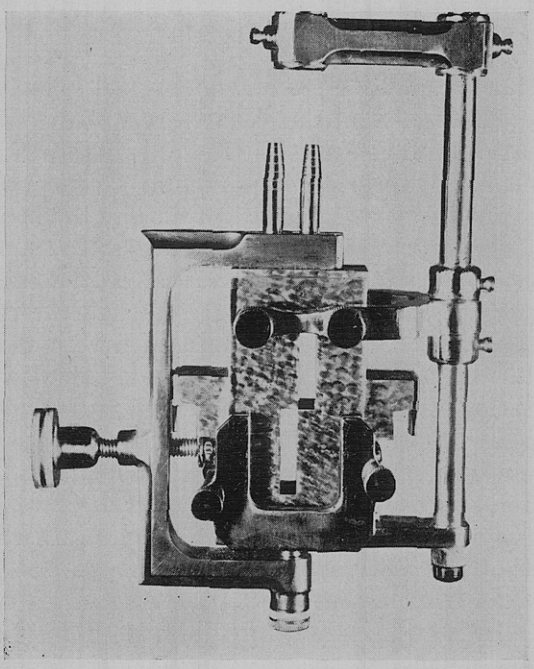

FIG. 8

FIG. 7. Schematic drawing of the Lamm diffusion cell (taken from reference 16)
FIG. 8. Photograph of the Neurath (27) diffusion cell (for description see text)

Lamm (16) has described a rectangular glass cell with a sliding glass diaphragm; however, no details of construction have been given.

Longsworth (18) and Rothen (43) have used the Tiselius electrophoresis apparatus for diffusion measurements. The diffusion curves were recorded by the schlieren-scanning method and the Philpot–Svensson (50) optical system, respectively. The results obtained in these studies will be referred to in section III of this review.

The earliest measurements carried out with the aid of a cylindrical lens and an inclined slit are those of Thoevert (52). It appears that they have not received the recognition they deserve; they are, therefore, reproduced in figure 9.

In figure 10 diffusion curves, as obtained by Longsworth[8] with the schlieren-scanning method, are reproduced.

2. Measurements

In order to measure the true diffusion rate of a solute, it is necessary to eliminate the influence of any external force other than that of osmotic pressure produced by the solute molecules in question. For this reason, the concentration gradient has to be negligible for all other ions and molecules that may be present; and,

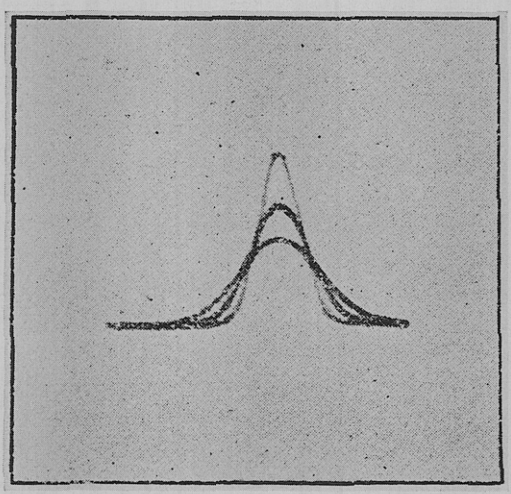

FIG. 9

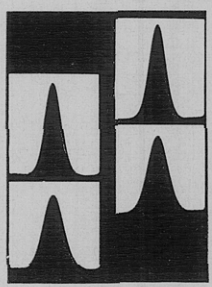

FIG. 10

FIG. 9. Refractive-index gradient curves for the diffusion of sodium chloride, as obtained by Thoevert (51) with the use of a cylindrical lens and an inclined slit. Reproduced from the original paper (52).

FIG. 10. Refractive-index gradient curves for the diffusion of a 1.4 per cent egg albumin solution, as obtained by the Longsworth schlieren-scanning method. The left- and right-hand pictures correspond, respectively, to the two boundaries in the Tiselius electrophoresis cell. The time of diffusion was 151,400 sec. in the top photographs and 235,800 in the bottom photographs. Unpublished experiments by Dr. L. G. Longsworth.

with solute molecules carrying electrical charges, such as the proteins, diffusion potentials have to be suppressed. Both requirements are met by carrying out the diffusion measurements in the presence of salts and by equilibrating the solution against the solvent, prior to the measurements, by dialysis through collodion or cellophane membranes. Salt concentrations of 0.1 ionic strength usually suffice to suppress diffusion potentials.

The diffusion cell is filled with solution and solvent as described above, and the time of diffusion is measured from the beginning of formation of the boundary. The length of the diffusion run, the time of the first exposure, and

[8] The author is indebted to Dr. L. G. Longsworth for placing these photographs at his disposal.

the time intervals between successive exposures depend on the diffusion rate of the protein under investigation and on the magnitude of the refractive-index gradients in the diffusion column.

Limitations on the over-all time of diffusion are imposed by both theoretical and practical considerations. The former have to do with the primary assumption involved in the derivation of equation 3, i.e., that the time of diffusion is short enough to avoid the occurrence of concentration changes at the extreme ends of the cell; the latter relates to the stability of the protein under investigation and requires that no chemical or physical changes occur during the diffusion process. Both factors tend to place an upper limit on the duration of the experiment.

The influence of the magnitude of the refractive-index gradient on the diffusion measurements can be best appreciated by considering certain approximations made in the application of the theory of the curvature of light to the present method (16).

The magnitude of Z, the scale-line displacement, depends on G, the optical magnification, b, the optical distance from the scale to the center of the cell, and ϑ, the angle of deviation of a light ray passing through the optically inhomogeneous medium contained in the diffusion cell (equation 8). The approximations involved in the derivation of this equation are as follows: (*1*) ϑ is sufficiently small to warrant the substitution of the arguments of the angles of inclination of the deviated light pencils with the optic axis for their tangents. Lamm has shown that for angles of 1.7°, this approximation introduces an error of 0.0003 radian. (*2*) n and dn/dx are constant in the region of the cell traversed by the light ray. This means, physically, that the vertical thickness of the light pencil is so small that the displacement Z is characteristic of a well-defined height in the diffusion column. Both conditions are more nearly fulfilled the longer the optical distance between camera lens and scale, the smaller the aperture of the lens, and the shorter the distance between cell and scale. A small aperture is also required on account of the lens action of the refractive-index gradient itself; however, a lower limit is placed on the aperture by the diffraction phenomena occurring when the aperture becomes too small. Values ranging between f:50 and f:80 are satisfactory. As a rule, b and the aperture will be made smaller for higher refractive gradients and larger for smaller gradients. From an extensive study of the influence of these various factors, Lamm (16) arrived at the conclusion that the theoretical requirements are fulfilled as long as the scale lines appear on the photographic plate undistorted and in focus, and this working rule can conveniently be used as a guide for actual experimentation. While too large differences in initial concentration must, therefore, be avoided, disturbances may also occur when the solute concentration is too low. In the latter case, the density difference between solution and solvent may become too small to keep the concentration gradients constant within the horizontal planes. The extreme limits in initial differences in concentration between solution and solvent lie in the neighborhood of 1.5 and 0.2 per cent protein, respectively.

After completion of the diffusion run the cell is filled with the solvent and

another exposure taken, yielding an image of the reference scale. The scale-line distances are then measured in a microcomparator, with an accuracy of $\pm 1\mu$. The most convenient procedure consists in recording the comparator readings as determined from the same starting point on the microcomparator and setting the cross-hair on the same starting line on each scale image. The differences between the readings on the reference scale and experimental scale are then plotted as ordinates, against the respective comparator readings on the experimental scale as abscissae, thus yielding a curve which relates the scale-line displacements, Z, to the positions of the displaced lines, z.

3. Methods of calculation

(a) Monodisperse systems

The curves obtained in the manner described above are, except for a constant factor depending on the photographic enlargement and the optical cell–scale distance, of the same form as equation 4. The calculations of the diffusion

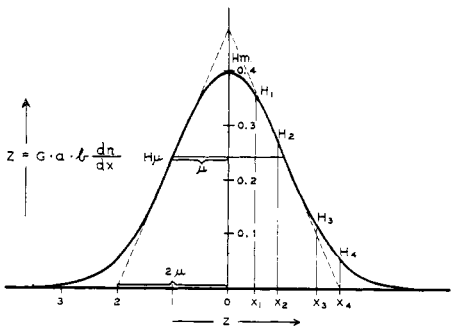

FIG. 11. Ideal scale-line displacement distribution curve (in normal coördinates). Z is the scale-line displacement, z the position of the displaced lines, H_m the maximum ordinate, μ half the distance between the inflection points, and H and x coördinates for points on the curve.

constant are based on the assumption that in the ideal case the curves have the properties of Gaussian distribution curves. We shall consider first the case of monodisperse solutions and ideal diffusion behavior. The following symbols will be used in the following equations which are most commonly used for calculating diffusion constants. Their meaning is also illustrated in figure 11.

H_m = maximum ordinate,
μ = half the distance between the inflection points,
x_i = abscissa of the point i,
H_i = ordinate of the point i,
A = area under the curve,
σ = standard deviation,
t = time of diffusion, and
D = diffusion constant.

(1) *Maximum ordinate method.* The maximum ordinate divided by $\sqrt{e}$ gives the ordinate of the inflection points, H_μ. The diffusion constant is

$$D = \frac{\mu^2}{2t}, \quad \text{where} \quad H_\mu = \frac{H_m}{\sqrt{e}} \tag{10}$$

(2) *Method of successive analysis.* Here the diffusion curve is divided into a series of vertical and horizontal chords, and the diffusion constant is calculated from a pair of successive values of x_i and H_i.

$$D = \frac{x_2^2 - x_1^2}{4t \ln \frac{H_1}{H_2}} \tag{11}$$

In practice, the diffusion curve is divided on either side of the bisecting ordinate into about three to four ordinates equally spaced from one another, and the diffusion constants are calculated as described. For a monodisperse solution, the D values should agree with one another to within 2 per cent or better.

(3) *Maximum ordinate-area method.* If, in equation 4, x is taken as zero, then

$$\left(\frac{dn}{dx}\right)_{x=0} = H_m = \frac{n_1 - n_0}{2\sqrt{\pi D t}} \tag{12}$$

and

$$D = \frac{(n_1 - n_0)^2}{4\pi t (H_m)^2} = \frac{A^2}{4\pi t (H_m)^2}$$

The area A is determined by graphical integration.

(4) *Statistical method.* This method, developed by Pearson (39) and applied to the present problem by Lamm (16), consists in treating the experimental curves as ideal displacement distribution curves. According to statistics, the standard deviation, σ, is equal to

$$\sigma = \sqrt{\mu_2^0} \tag{13}$$

where μ_2^0 is the second moment of the curve about the centroidal[9] ordinate. The diffusion constant D is related to σ by

$$D = \frac{\sigma^2}{2t} \tag{14}$$

In practice, the base line of the diffusion curve is divided into evenly spaced units of the breadth unity, numbered outward from an arbitrarily chosen origin near the center of the base line. If s_i is the respective number on the base line and S_i the corresponding ordinate, then the zero moment of the curve about the arbitrarily chosen central ordinate is

$$\mu_0' = \Sigma(S_i) = N \tag{15}$$

[9] The centroidal ordinate is that ordinate about which the first moment is equal to zero.

and is equal to area in the S units that were chosen. μ_1', the first moment about this ordinate, is

$$\mu_1' = \frac{\Sigma(s_i S_i)}{N} \tag{16}$$

and the second moment, μ_2', is

$$\mu_2' = \frac{\Sigma(s_i^2 S_i)}{N} \tag{17}$$

The position of the true centroidal ordinate is

$$x_0 = s_0 - \mu_1' \tag{18}$$

and the true second moment, μ_2, is

$$\mu_2 = \mu_2' - (\mu_1')^2 \tag{19}$$

The second moment calculated in this manner has yet to be converted into absolute units by the relation

$$\mu_2^0 = \mu_2 \omega^2 \tag{20}$$

where ω is the distance between successive vertical chords, S_i, in centimeters. In the interest of greater accuracy it is advisable to chose ω, the class breadth, not too large, and as a rule, 20 to 30 divisions of s on either side of the center of the base line will be most satisfactory.

The diffusion constant calculated by any one of the methods given has to be corrected for photographic magnification and optical distances between camera, cell, and scale. If G is the photographic enlargement factor, l the optical distances from the lens to the scale, and b the optical distance from the center of the cell to the scale, the correction factor by which the calculated D values have to be multiplied is

$$F = \left(\frac{l - b}{lG}\right)^2 \tag{21}$$

The extent to which the experimental diffusion curves deviate from an ideal displacement distribution curve can be studied by transforming the coördinates s and S into normal coördinates, ξ and ψ, in which case the time of diffusion disappears as parameter.

The relations between s, S, ξ, and ψ are

$$\xi = (s - \mu_1') \frac{\omega}{\sigma} \tag{22}$$

and

$$\psi = S \frac{\sigma}{\omega N} \tag{23}$$

where σ is the calculated standard deviation, as discussed previously. The curve obtained in this manner is, for the ideal case, identical with the normal form of the ideal distribution curve, the values of which can be found in mathe-

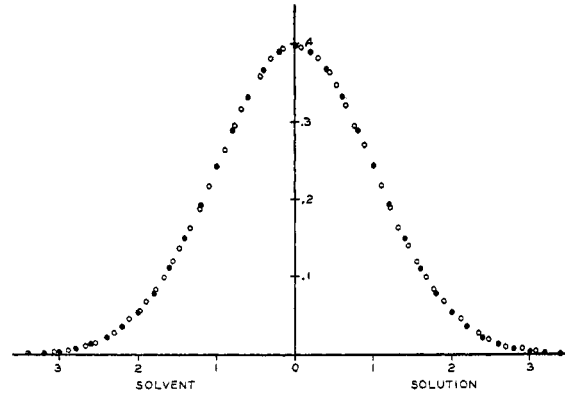

FIG. 12. Comparison of an ideal Gaussian distribution curve with the diffusion curve as obtained from measurements on a 0.8 per cent solution of urea-denatured serum albumin in 8 *M* urea (31). The open circles refer to the experimental points, plotted in normal coördinates. The full circles indicate the position of the ideal distribution curve.

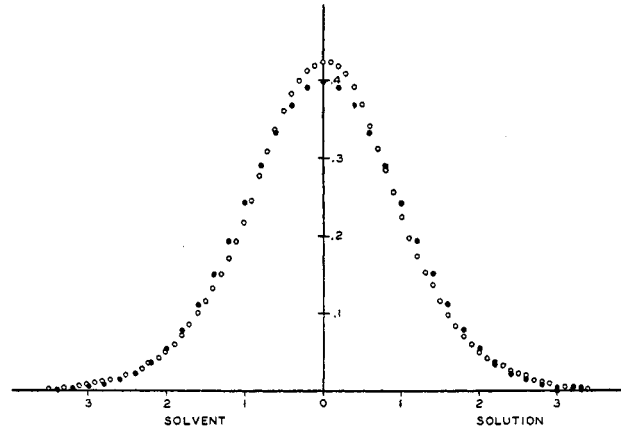

FIG. 13. Comparison of an ideal Gaussian distribution curve with the diffusion curve as obtained from measurements on a 0.8 per cent serum albumin solution at pH 7.6, denatured by heating for 30 min. at 70°C. Polydispersity is most clearly indicated by the difference in maximum ordinates of the ideal curve (full circles) and the experimental curve (open circles). Unpublished experiments by Neurath, Cooper, and Erickson.

matical tables. The "fit" of these two curves is a measure of the homogeneity of the solute under investigation, and is illustrated for a monodisperse solution in figure 12 and for a polydisperse solution in figure 13.

While, in the preceding cases, the diffusion constant has been calculated from a single curve, it is, of course, also possible to obtain values of D from a com-

parison of the properties of the curves obtained after consecutive time intervals. It will be noted that in the equations given, the diffusion constant is related to any two of the variables H_i, X_i, and A, t being a constant. By introducing t as a variable, another set of equations can be formulated if any one of the above-mentioned variables is chosen as constant. This method has been used recently in studies of the diffusion of the rabbit papilloma virus protein (32), with the particular object of differentiating between lack of monodispersity, on one hand, and restriction of free diffusion, on the other, as the origin of the

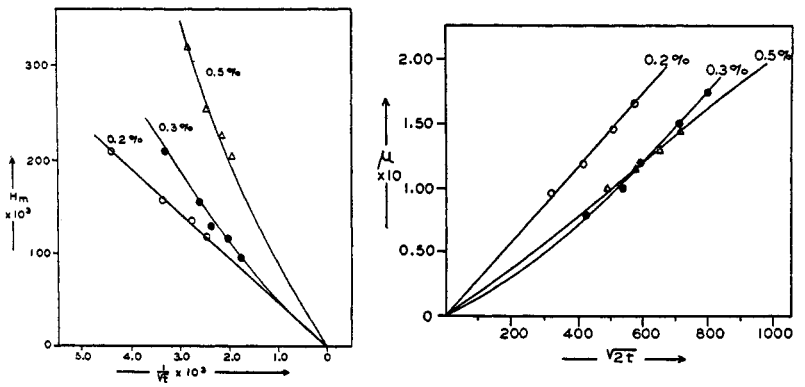

FIG. 14. Relation between time of diffusion and maximum ordinate, H_m (left), and half the distance between the inflection points, μ (right) of diffusion curves, as obtained from measurements on the rabbit papilloma virus protein (32). For further explanation see text and equations 24 and 25.

observed variations in apparent diffusion constant with protein concentration.

From equation 10, for instance, it follows that

$$\frac{\mu}{\sqrt{2t}} = \sqrt{D} \tag{24}$$

and

$$\frac{H_m}{\sqrt{t}} = \frac{K}{\sqrt{D}}, \quad \text{where} \quad K = \frac{A}{2\sqrt{\pi}} \tag{25}$$

The results obtained from a plot of these equations are shown in figure 14. While in the most dilute solutions the curves are linear in both plots, diffusion anomalies occurring in protein concentrations of 0.3 and 0.5 per cent are evidenced both by the variations in the slope and by the non-linear shape of the curves.

(b) Polydisperse systems

While, with monodisperse solutions, calculations by any one of the methods just given yields a value for the true diffusion constant of the solute in question, average values are obtained if the solutions are polydisperse.

An estimate of the diffusion constant of the fastest moving component can be obtained with the aid of equation 11, provided its diffusion rate, relative to that of the other components, is sufficiently fast to establish a concentration gradient of its own, at higher values of x. In that case a set of equations of the form given in equation 11 can be set up, and $D_{\text{lim.}}$, the limiting value of the diffusion constant, extrapolated from $x \to \infty$, $H_i \to 0$. A more detailed discussion of the limitations of this method will be given in section D4 (page 378).

The problem of the resolution of a compound diffusion curve of a polydisperse system into its component curves has received attention recently by Dr. L. W. Nordheim and the author. Assuming a solution to be composed of two components of different diffusion constants, the individual constants and the concentrations of the two components may be calculated as follows:

The diffusion equation is as stated previously,

$$\frac{dn}{dx} = \frac{n_1 - n_0}{2\sqrt{\pi Dt}} e^{-\frac{x^2}{4Dt}} \tag{26}$$

Substituting for $2Dt = \sigma^2$ and for $n_1 - n_0 = kc$,

$$\frac{dn}{dx} = \frac{c}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \tag{27}$$

For a solution[10] containing i components with concentrations c and standard deviations σ, the general diffusion equation is (16)

$$g(x) = \int_{\sigma=0}^{\infty} \frac{c(\sigma)}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} d\sigma \tag{28}$$

and for $i = 2$

$$g(x) = \frac{c_1}{\sigma_1\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_1^2}} + \frac{c_2}{\sigma_2\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_2^2}} \tag{29}$$

Calculation of higher moments and integration with respect to x yields the following set of equations:

$$\int_0^{\infty} g(x)\,dx = \nu_0 = \frac{c_1}{\sigma_1\sqrt{2\pi}}\sigma_1\sqrt{\frac{\pi}{2}} + \frac{c_2}{\sigma_2\sqrt{2\pi}}\sigma_2\sqrt{\frac{\pi}{2}} = \frac{1}{2}(c_1 + c_2) \tag{30}$$

$$\int_0^{\infty} xg(x)\,dx = \nu_1 = \frac{c_1}{\sigma_1\sqrt{2\pi}}\sigma_1^2 + \frac{c_2}{\sigma_2\sqrt{2\pi}}\sigma_2^2 = \frac{1}{\sqrt{2\pi}}(c_1\sigma_1 + c_2\sigma_2) \tag{31}$$

$$\int_0^{\infty} x^2 g(x)\,dx = \nu_2 = \frac{c_1}{\sigma_1\sqrt{2\pi}}\sqrt{\frac{\pi}{2}}\sigma_1^3 + \frac{c_2}{\sigma_2\sqrt{2\pi}}\sqrt{\frac{\pi}{2}}\sigma_2^3 = \frac{1}{2}(c_1\sigma_1^2 + c_2\sigma_2^2) \tag{32}$$

$$\int_0^{\infty} x^3 g(x)\,dx = \nu_3 = \frac{c_1}{\sigma_1\sqrt{2\pi}}2\sigma_1^4 + \frac{c_2}{\sigma_2\sqrt{2\pi}}2\sigma_2^4 = \frac{2}{\sqrt{2\pi}}(c_1\sigma_1^3 + c_2\sigma_2^3) \tag{33}$$

[10] Here it is assumed that dn/dc is constant not only for the whole solution but also for the individual components. If the specific refractive-index increments of the components are different, refractive indices have to be used in the calculations in place of concentrations. For proteins of closely related properties such as those derived from a mixture by fractional precipitation, the refractive-index increments may be considered to be the same.

Here ν_0 is equal to one-half the area under the curve, and ν_1, ν_2, and ν_3 to the first three moments calculated for either one of the halves of the curve, with the centroidal ordinate as origin.

Setting

$$a = \frac{1}{\nu_0} 2\nu_0 \tag{34}$$

$$b = \frac{\sqrt{2\pi}}{\nu_0} \nu_1 \tag{35}$$

$$c = \frac{1}{\nu_0} 2\nu_2 \tag{36}$$

$$d = \frac{\sqrt{2\pi}}{2\nu_0} \nu_3 \tag{37}$$

we obtain

$$\sigma_{1,2} = -\frac{1}{2} \frac{bc - ad}{ac - b^2} \pm \sqrt{\frac{1}{4}\left(\frac{bc - ad}{ac - b^2}\right) - \frac{bd - c^2}{ac - b^2}} \tag{38}$$

where σ_1 and σ_2 are the standard deviations of the two components, related to their respective diffusion constants by equation 14. Substitution of σ_1 and σ_2 into equations 31 and 32 yields the relative and absolute concentrations of the two components[11].

The method as presented here is strictly applicable only if the experimental diffusion curves are truly symmetrical and if the solutions under investigation are known to be composed of only two components. If there are more than two, this method yields only mean values which, however, are probably closer approximations than those obtained from a single standard deviation. The method should prove particularly useful when applied to protein solutions for which the presence of two components has been established by sedimentation analysis, since it facilitates the determination of the respective diffusion constants without previous separation and isolation of the individual components. A discussion of the limitations of the method will be given in section D4 (page 378).

The problem of the diffusion of polydisperse systems has also been considered recently by Gralén (9). His calculations are directed primarily toward an evaluation of the average diffusion constant. The type of average has been shown by him to depend on the method of calculation employed. Thus with equation 12 the average diffusion constant D_0 is

$$D_0^{A,H_m} = \left(\frac{\Sigma c_i}{\Sigma \frac{c_i}{\sqrt{D_i}}}\right)^2 \tag{39}$$

[11] Equation 33 may also be replaced by the relation

$$\frac{c_1}{\sigma_1} + \frac{c_2}{\sigma_2} = H_m$$

where H_m is the maximum ordinate. This method may be more convenient in that it eliminates the calculation of the third moments.

and with equation 14

$$D_0^{\sigma} = \frac{\Sigma c_i D_i}{\Sigma c_i} \tag{40}$$

where i is the number of diffusing components, c their respective concentration, and D the respective diffusion constant. The ratio of these two average diffusion constants

$$\frac{D_0^{\sigma}}{D_0^{A,H_m}} = \frac{\Sigma c_i D_i \left(\Sigma \frac{c_i}{\sqrt{D_i}}\right)^2}{(\Sigma c_i)^3} \tag{41}$$

is considered to be a measure of the degree of polydispersity. Assuming a logarithmic distribution function, Gralén calculates the distribution of diffusion constants in relation to the relative concentrations of the diffusing components.

4. *Resolving power and accuracy*

We have stated in the preceding discussion that, for a monodisperse solution, the diffusion constants calculated with the aid of the various equations should agree with one another to within 2 per cent or better. The question arises, however, whether such a close agreement is necessarily an indication of monodispersity or whether it may not merely simulate monodisperse behavior of an actually polydisperse system. This problem has, as far as we are aware, not been considered previously. Its importance for determinations of the state of dispersion of a protein solution and the determination of the molecular weights of proteins is obvious and can hardly be overemphasized.

The following discussion is limited to a comparison of a monodisperse solution, on one hand, with one composed of two diffusing entities, on the other. If in equation 29 one of the two right-hand members is small enough to be neglected in comparison with the other, the equation can be reduced to the logarithmic form given in equation 11. This is possible if (*1*) $\sigma_2 = 0$, in which case one is dealing with a monodisperse solution; (*2*) $\sigma_2 \ll \sigma_1$; or (*3*) $c_2 \ll c_1$. Case (*2*) implies that, as mentioned, at a given distance, x, the observed refractive-index gradient is due only to the first component and is zero for the second component. Case (*3*) implies merely that the concentration of the second component is altogether too small to give rise to a measurable refractive-index gradient. If, therefore, $\sigma_1 = \sigma_2$, or $\sigma_2 = 0$, values of D calculated with equation 11 should be constant over the whole region of the diffusion curves and the solution may be said to be monodisperse. If D is found to increase with increasing values of x, the limiting value of the diffusion constant of the fastest moving component can be determined by extrapolation to $x \rightarrow \infty$, provided condition (*2*) is known to be fulfilled. Otherwise, equation 11 is invalid and the method of "successive analysis" is not applicable.

The most readily detectable difference between the observed diffusion curves

and normal Gaussian distribution curves exists with respect to their maximum ordinates. If, in equation 29, x is taken as zero, this equation reduces to

$$\left(\frac{dn}{dx}\right)_{x=0} = H_m^{1,2} = \frac{c_1}{\sigma_1\sqrt{2\pi}} + \frac{c_2}{\sigma_2\sqrt{2\pi}} \tag{42}$$

where $H_m^{1,2}$ is the maximum ordinate, and σ_1 and σ_2 are the standard deviations of the two components present in concentrations c_1 and c_2. σ, it will be recalled, is proportional to the square root of the diffusion constant.

For a monodisperse substance we have

$$H_m = \frac{c}{\sigma\sqrt{2\pi}} \tag{43}$$

Setting

$$c = c_1 + c_2$$

the maximum ordinate of a diffusion curve of a monodisperse solution can be compared with that of a curve produced by the simultaneous diffusion of two components. The quantity $\left(\frac{H_m^{1,2}}{H_m} - 1\right)100$ gives the percentage change in the maximum ordinate of the diffusion curve of a mondisperse solution. In table 1 values are given for this quantity, calculated for various ratios of D_2/D_1 in relation to assumed values of c_2/c_1. σ has been assumed to be equal to σ_1, i.e., we compare the maximum ordinate of the diffusion curve of a monodisperse solution of the faster diffusing component with that of a solution containing both components in varying relative concentrations.

It may be seen from table 1 that the maximum ordinate of the compound diffusion curve is always higher than that of the diffusion curve of component 1. The percentage increase is smaller the more closely the two diffusion constants approach each other and, for a given value of D_2/D_1, the lower the relative concentration of the second component. The accuracy with which the maximum ordinate can be determined experimentally is about 1 per cent (about $\pm$ 2 μ for a maximum scale-line displacement of 300 μ). Hence, it is practically impossible to detect with *this method* of calculation the presence of about 20 per cent of material having a diffusion constant 10 to 20 per cent lower than that of the major component; the resolving power increases, however, with increasing differences in the two diffusion constants, and with ratios of D_2/D_1 approaching a value of 0.5, the resolution becomes sufficiently distinct to detect the presence of the second component[12] in a concentration as low as about 5 per cent of that of the total protein.

[12] We have stated in a recent paper (30) that the diffusion measurements on serum albumin denatured by urea, or by guanidine hydrochloride, indicate the solutions to be monodisperse. The ratios of the diffusion constants of native and denatured protein are, in these cases, about 0.5, and the agreement between the observed and calculated maximum ordinates is within the limits of the experimental error (see figure 12). This conclusion is in accord with the present ideas.

The situation becomes somewhat different if the compound diffusion curve is compared with that of a single substance, the diffusion constant of which lies

TABLE 1

Comparison of the maximum ordinate of a compound diffusion curve of a mixture of two components of diffusion constants D_1 and D_2 with the maximum ordinate of the diffusion curve of a single component of diffusion constant D_1

c_1 and c_2 are the relative concentrations of components 1 and 2, D_2/D_1 is the ratio of their diffusion constants, and $\left(\frac{H_m^{1,2}}{H_m^1} - 1\right) \times 100$ is the percentage change in the maximum ordinate, H_m^1, of the diffusion curve of a monodisperse solution of component 1

c_1	c_2	PER CENT CHANGE IN H_m^1 FOR $\frac{D_2}{D_1}$ =					
		0.9	0.79	0.70	0.50	0.25	0.10
1	0	0	0	0	0	0	0
0.9	0.1	0.5	0.6	2	4.1	10	126
0.75	0.25	1.3	3.3	4.9	10.4	25	142
0.50	0.50	2.5	7.0	9.7	20.8	50	166
0.25	0.75	4.0	10.0	14.8	31.2	75	191
0.10	0.90	4.8	11.2	17.8	37.5	90	206
0	1.0	5.2	13.7	19.8	41.6	100	216

TABLE 2

Comparison of the maximum ordinate of a compound diffusion curve of a mixture of two components of diffusion constants D_1 and D_2 with the maximum ordinate of a diffusion curve for a single component of diffusion constant D_0, where $\sqrt{D_0} = \frac{\sqrt{D_1} + \sqrt{D_2}}{2}$

c_1 and c_2 are the relative concentrations of components 1 and 2; D_1 and D_2 are their diffusion constants; $\left(\frac{H_m^{1,2}}{H_m^0} - 1\right) \times 100$ is the percentage change in the maximum ordinate, H_m^0, of the diffusion curve of a monodisperse solution of a component with diffusion constant D_0 and concentration $c_0 = c_1 + c_2$

c_1	c_2	$\left(\frac{H_m^{1,2}}{H_m^0} - 1\right) \times 100$
		per cent
1	0	−14.7
0.90	0.10	−11.2
0.75	0.25	−5.9
0.50	0.50	+3.0
0.25	0.75	+11.9
0.10	0.90	+17.2
0	1.0	+20.8

between those of the two components. Such a comparison is made in table 2 for the case in which the diffusion constant of the single substance, $\sqrt{D_0}$, is equal to $\frac{\sqrt{D_1} + \sqrt{D_2}}{2}$. The change in maximum ordinate is again expressed

in terms of percentage of that of the diffusion curve for the single substance, $\left(\frac{H_m^{1,2}}{H_m^0} - 1\right)100$. Here, it will be noted, the maximum ordinate of the compound curve is lower than that of the single curve for values of c_2 below 0.5 and higher for values above 0.5. The region near $c_2 = 0.5$ is of interest because the per cent deviation approaches zero, implying that it is impossible to decide, on the basis of the maximum ordinate alone, whether one is dealing with a mixture of two components or with a single component, the diffusion constant of which is about half-way between that of the two components.

Such an apparent uncertainty can be partly dispelled if, from the history of the experiment, predictions can be made regarding the values of the diffusion constants to be expected. However, the calculations given in tables 1 and 2 refer to particular cases, and it can be readily seen that equation 29 permits of an infinite number of solutions depending on the choice of the values for c_1, c_2, σ_1, and σ_2.

The problem becomes definite upon application of the statistical method of analysis. We have seen that the four equations (30 to 33) contain four unknowns and, therefore, are capable of one solution only. As a test for monodispersity, the following relations, which follow directly from these equations, may be applied:

$$\nu_0 \nu_1 \geqq (\nu_1)^2 \frac{\pi}{2} \tag{44}$$

and

$$\nu_0 \nu_3 \geqq 2\nu_1 \nu_2 \tag{45}$$

The right-hand members of these equations should be equal to the left-hand members for a monodisperse solution, and smaller for a two-component solution. Moreover, the absolute values of the moments will be different for a monodisperse solution, on one hand, and for a polydisperse solution, on the other. An illustration of these relations is given in table 3. The symbols have the same meaning as in the corresponding equations. Column A refers to a 1:1 mixture of materials the diffusion constants of which are assumed to be 7×10^{-7} and 5×10^{-7}, column B to a 1:1 mixture, diffusion constants 7×10^{-7} and 3.5×10^{-7}, and column C to a single substance, diffusion constant 7×10^{-7}. In the first case, the per cent difference is 0.8 for equation 44, and 1.5 for equation 45; in the second case, the corresponding differences are 3.5 per cent and 5.6 per cent. While the former values are probably too small to be detected experimentally, the latter are outside the limits of the experimental error. The differences between the absolute values calculated for the three cases are very marked and may be used as a guide when other evidence is available to indicate an expected value for D or c.

In summarizing this discussion, it may be stated that application of diffusion measurements to the determination of the state of dispersion of a protein solution is limited by two factors, one of theoretical, the other of practical nature. The

former has to do with the insensitivity of Gaussian distribution curves to small variations in the standard deviation; the latter is related to the accuracy with which the curves can be obtained experimentally, even when the greatest care is exercised. There is, however, little doubt that the application of the above relations to experimental data should place diffusion measurements on a much more reliable basis than they appear to be at present.

TABLE 3

Comparison of statistical properties of diffusion curves of monodisperse and polydisperse solutions

D_1, D_2 = diffusion constants of components 1 and 2; c_1, c_2 = relative concentrations of both components; t = time of diffusion in seconds; ν_0, ν_1, ν_2, ν_3 = moments of diffusion curves (for their meaning see text); Δ_I = difference between values listed in rows 10 and 11, expressed in per cent of value listed in row 10; Δ_{II} = difference between values listed in rows 13 and 14, expressed in per cent of value listed in row 13

	PROPERTIES	A	B	C
(1)	$D_1 \times 10^7$	7	7	7
(2)	$D_2 \times 10^7$	5	3.5	
(3)	t in seconds	86,400	86,400	86,400
(4)	c_1	0.5	0.5	1.0
(5)	$c_2 = 1 - c_1$	0.5	0.5	0
(6)	ν_0	0.5	0.5	0.5
(7)	ν_1	0.1280	0.1184	0.1387
(8)	ν_2	0.0519	0.0454	0.0605
(9)	ν_3	0.0270	0.0228	0.0337
(10)	$\nu_0\nu_1$	0.0260	0.0227	0.0302
(11)	$(\nu_1)^2\frac{\pi}{2}$	0.0258	0.0220	0.0302
(12)	Δ_I	0.8	3.5	0
(13)	$\nu_0\nu_3$	0.0135	0.0114	0.0168
(14)	$2\nu_1\nu_2$	0.0133	0.0107	0.0168
(15)	Δ_{II}	1.5	5.6	0

III. Results

The diffusion constants of a large number of proteins have been measured in recent years. The impetus to these determinations came largely from sedimentation-velocity measurements, which require a knowledge of the diffusion constants for the determination of molecular weights. Measurements with the light-absorption and the light-refraction methods have been carried out most extensively in Svedberg's laboratory, where these methods were originated, in the laboratory of J. W. Williams, and in that of the present author. Before considering these results it may be of interest to compare the values of diffusion constants as obtained for certain proteins by the various experimental methods previously discussed. This is done in table 4.

The results obtained with the porous-disk method are more susceptible to experimental error than those obtained from measurements of the absolute diffusion rate, for the reasons already given. The discrepancies to be noted in

table 4 are probably due to a large extent, to differences in the preparation of the materials employed by the various investigators. It would be of interest to see how well the results would agree if one and the same material were to be used for comparative measurements.

Sufficient material has accumulated to make it possible to estimate the precision of the refractometric-scale method. Lamm's (16) measurements on potassium chloride and pentaerythritol, analyzed by means of the methods of calculation considered previously (section D3), reveal a standard deviation from

TABLE 4

Comparison of the diffusion constants of certain proteins as determined by various methods of measurement

$D^{25°}_{H_2O}$ = diffusion constant calculated for 25°C. in water

PROTEIN	T	METHOD	$D^{25°}_{H_2O}$	REFERENCE
	°C.		*cm.² per second* $\times 10^7$	
Pepsin	8	Porous disk	8.9*	(35)
	20	Scale	10.3†	(42)
	25	Scale	10.0, 9.2‡	(29)
Hemoglobin (carbon monoxide)	5	Porous disk	6.9*	(36)
	0	Porous disk	8.0	(10)
	20	Light absorption	7.2	(53)
	20	Scale	7.9	(42)
Egg albumin	25	Porous disk	11.1§	(21)
	20	Light absorption	8.7	(53)
	20	Scale	8.7	(42)
	0	Schlieren scanning	8.7¶	(18)

* Corrected for erroneous value of the diffusion constant of 0.1 *N* hydrochloric acid originally used in the calibration of the cell (2).

† Unfractionated crystalline preparation used by the investigator.

‡ The two values refer to two crystalline fractions differing from one another in solubility.

§ This value may be in error, since the measurements were carried out in the absence of salts.

¶ Extrapolated to infinite dilution according to Polson's data (42).

the mean not exceeding 1 per cent. With monodisperse protein preparations such as egg albumin (42), excelsin (42), serum albumin (29), serum pseudoglobulin (29), and others, the standard deviation from the mean lies close to 2 to 3 per cent.

In the following, we shall examine the factors which have been found to influence the diffusion rate of proteins and which, unfortunately, have not always received the attention they deserve. The most important of these factors are the time of diffusion, the temperature, the solvent viscosity, and the solute concentration.

A. TIME OF DIFFUSION

While the majority of measurements that have been published do not reveal any significant drift of the apparent diffusion constant with time, there has been noted in several instances a gradual decrease of the diffusion rate, approaching asymptotically a constant value after longer time intervals.

This effect has been observed with both the porous-disk method (10, 57) and the refractometric-scale method. Of the monodisperse proteins studied by Polson (42), about one-fourth revealed a downward drift of D with time, whereas the remainder did not. Of the experiments carried out in the author's laboratory, about one-half exhibited this effect. With secalin, studied by Andrews (1), D decreased to as much as one-fourth of the initial value. Longsworth's recent measurements on egg albumin[13] show a downward drift at pH 4.64 but not at pH 11.81. The origin of this effect is rather obscure and its elucidation must await further detailed measurements on monodisperse solutions; it also remains to be seen whether in experiments where this effect is observed the mean value or the limiting value should be taken as the best approximation to the true diffusion constant.

B. TEMPERATURE

The influence of temperature on the diffusion constant is given by the equation

$$\frac{D_x}{D_y} = \frac{T_x}{T_y}\frac{\eta_y}{\eta_x} \tag{46}$$

where D_x and D_y are the diffusion constants at the absolute temperatures T_x and T_y, respectively, and η_x and η_y the corresponding viscosities of the solvent. The equation has been verified experimentally by Tiselius and Gross for phycoerythrin (53) (measured at 20° and 30°C.) and by Polson for egg albumin (42) (measured at 15°, 20°, and 25.1°C.). Longsworth's measurements on egg albumin (18), carried out at 0°C., likewise agree with Polson's data when corrected for the temperature by means of equation 46.

C. VISCOSITY

While this factor will be considered more fully in the following paragraph, it may be said here that, by and large, the viscosity correction

$$D_x = D_y\frac{\eta_y}{\eta_x} \tag{47}$$

holds, for values of η_y/η_x not exceeding 1.1. Here D_x and D_y are the diffusion constants measured at a constant temperature in solvents of viscosities η_x and η_y, respectively. Diffusion constants are usually referred to water as a solvent, in which case η_x is the water viscosity at a given temperature (20° or 25°C.) and η_y that of the solvent into which diffusion takes place.

Polson's measurements on egg albumin (42), in ammonium sulfate solutions,

[13] The author is indebted to Dr. L. G. Longsworth for placing these data at his disposal.

varying in relative viscosity (with respect to water) between 1.090 and 1.437, indicate, however, a slight increase in the apparent diffusion constant when corrected for the solvent viscosity by means of equation 47; recent measurements on serum albumin in solutions of sucrose, as well as in solutions of urea and guanidine hydrochloride (30) (η/η_0 varying between 1.15 and 1.30), cast some doubt on the strict validity of equation 47. In these instances, the relative viscosity influential in retarding the diffusion rate appears to be about 10 per cent higher than that actually measured. This problem is being further investigated in the author's laboratory.

D. THE INFLUENCE OF CONCENTRATION ON THE DIFFUSION RATE

We have mentioned at the outset that the validity of the diffusion equation 3 requires the independent motion of the diffusing entities; under these conditions the diffusion rate is independent of the solute concentration. While on a molar basis, protein solutions of 1 to 5 per cent concentration are still in the "infinite dilution" range, yet the diffusion rate is not independent of concentration. This non-ideal behavior is reflected also by sedimentation and viscosity and is frequently met with in solutions of high-molecular-weight compounds.

The influence of concentration on the diffusion rate may be studied by either one of the following methods: (*1*) the initial concentration difference is varied by increasing the concentration of the solution and keeping that of the solvent at zero; (*2*) the initial concentration difference is kept constant, but the absolute concentration of both solution and "solvent" is increased. The first method is limited in scope by the optical disturbances occurring with high refractive-index gradients, as discussed in section D2 (page 22). It can be used up to protein concentrations of about 1.5 per cent.

Where a concentration dependency of the diffusion constant, within the limits of about 0.2 and 1.5 per cent, is observed, the true diffusion constant can be determined by extrapolation to zero concentration. This is analogous to the procedure used in other physical measurements, such as those of osmotic pressure or sedimentation rate. Of course, this method is restricted in applicability if, for instance, particle dissociation occurs in very dilute solutions, as has been observed with hemoglobin (40). A protein the diffusion rate of which is fairly independent of the concentration up to about 1.4 per cent is egg albumin, as can be seen from table 5. With the tobacco mosaic virus protein (33), on the other hand, the diffusion constant is concentration dependent in concentrations as low as 0.2 per cent. This is evidenced by the variations in the calculated values of the diffusion constant, and also by the shape of the individual diffusion curves. They are skewed, as shown in figure 15, and the position of the maximum ordinate shifts with time toward regions of lower concentration. These phenomena have been related to a restriction of the diffusion rate in the regions of relatively high concentration, owing to interparticle attraction, and to a corresponding acceleration in regions of relatively lower protein concentration (33).

In the second method for studying the influence of concentration on the diffusion rate, the absolute solute concentrations are varied in both solute and

"solvent," while their relative concentrations remain constant. Thus, the diffusion of a 1 against a 0 per cent solution, of a 2 against a 1 per cent solution, or of a 6 against a 5 per cent solution should proceed with equal rate if the diffusion constant is concentration independent. Such measurements have been carried out by Polson (42) on six different proteins, varying in molecular weight between 17,000 and 750,000, and in concentrations ranging from 1 to 9 per cent.

TABLE 5

Diffusion constant of egg albumin in relation to protein concentration (Polson (42))

PROTEIN CONCENTRATION	$D_{H_2O}^{20°}$
per cent	$\times 10^7$
1.4	7.64
0.91	7.71
0.88	7.76
0.83	7.73
0.7	7.71

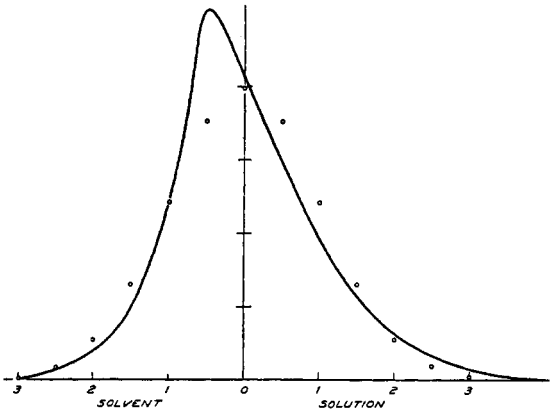

FIG. 15. Comparison of an ideal Gaussian distribution curve with the diffusion curve as obtained from measurements on a 1 per cent solution of tobacco mosaic virus protein. The open circles indicate the position of the ideal curve; the solid line refers to the experimental curve plotted in normal coördinates (33).

In evaluating his results, Polson assumes that the frictional resistance effective in retarding the diffusion rate is only that of the buffer solution in which the proteins were dispersed, whereas the protein molecules contained in the "solvent" do not contribute to the friction. Thus, in systems in which, for instance, a buffered solution containing 6 per cent of protein diffuses against a buffered 5 per cent solution of the same protein, the viscosity correction is assumed to be $\eta_{buffer}/\eta_{H_2O}$, i.e., the same as that employed in cases where the protein diffuses into the buffer itself. On this basis, Polson arrived at the conclusion that with some proteins, such as hemocyanin *Homarus*, the diffusion constant increases very slightly with increasing protein concentration, and that with amandin it remains

constant, whereas with serum globulin, egg albumin, lactoglobulin, and erythrocruorin *Lampetra* it decreases as the concentration is increased. Polson's data are tabulated in column 3 of table 6 and are compared with the respective values of D_∞, the limiting value of the diffusion constant obtained by extrapolation to zero concentration[14] (column 7 of table 6).

If, however, one applies the full correction for the viscosity of the solution into which diffusion takes place, $\left(\frac{\eta_{x \text{ per cent protein}}}{\eta_{H_2O}}\right)$, values are obtained as given

TABLE 6

Diffusion in concentrated solutions

(1) SOLUTE	(2) CONCENTRATION Solution	(3) CONCENTRATION Solvent	(4) $D^{20} \times 10^7$	(5) η/η_0 SOLVENT	(6) $D^{20} \times \frac{\eta}{\eta_0} \times 10^7$	(7) D_∞^{20} (OBSD.) $\times 10^7$	(8) $\frac{D^{20}}{D_\infty^{20}}$	(9) D_∞^{20} (CALCD.)
Egg albumin (42)*	3	2	7.43	1.089	8.14	7.76	0.958	7.74
	6	5	6.94	1.240	8.62	7.76	0.884	7.77
	9	8	6.11	1.48†	9.05	7.76	0.787	7.78
Lactoglobulin (42)	5.5	5	6.46	1.246	8.05	7.30	0.885	7.26
Serum albumin (5)	5	4	5.54	1.213	6.72	6.11	0.907	6.11
	6.7	5.7	5.20	1.348†	7.00	6.11	0.850	6.15
Sucrose (42)	5	4.5	42.8	1.116‡	47.8	45.6	0.939	45.1
	10	9	39.9	1.280‡	51.0	45.6	0.875	45.6
Sucrose in 4 per cent serum albumin (5)	1	0	41.4	1.213	52.2	45.6	0.907	45.7
Amandin (42)	4.5	4	3.55	1.23	4.37	3.62	0.98	3.95
	8.66	8	3.57	1.64	5.83	3.62	0.98	5.00
Hemocyanin *Homarus* (42)	3.4	2.5	2.80	1.13	3.16	2.78	1.00	2.98
	6.8	5.9	2.83	1.36	3.84	2.78	1.02	3.36
Hemoglobin *Lampetra* (42)	1.66	1.08	10.17	1.05§	10.68	10.65	0.955	10.4
	3.33	2.71	9.34	1.13§	10.58	10.65	1.02	9.95

* The numbers given in parentheses refer to the observer as listed in the bibliography. † Extrapolated values. ‡ Interpolated value. § Calculated from the dissymmetry constant with the aid of the Perrin and Simha equations (29).

in column 6 of table 6. It will be noted that, according to these calculations, the apparent diffusion constant increases with increasing protein concentra-

[14] While Polson believes that the concentration dependency of a protein is directly related to its molecular volume (the larger the volume, the smaller the interparticle distance and the greater the concentration dependency), it appears to us that such a viewpoint neglects many other factors which are known to influence the interaction between molecules of colloidal dimensions. Thus, for instance, the shape of the molecules and the nature of their surface undoubtedly play an important rôle. This can be seen from a comparison of the diffusion behavior of the tobacco mosaic virus protein (33), on one hand, and of the rabbit papilloma virus protein on the other (32). Both have molecular weights of comparable magnitude, but the concentration dependency of the diffusion constant of the rod-shaped tobacco mosaic virus protein molecules is considerably greater than that of the more nearly symmetrically shaped molecules of the rabbit papilloma virus protein.

tion. Such could only be the case if the molecules either dissociate into smaller units, or become more nearly spherical in shape. The fact, however, that a simple substance like sucrose follows the same trend, together with the known stability of some of these proteins, discredits such a hypothesis and calls for a closer examination of the problem. In its broadest form, it can be stated as follows: What is the frictional resistance encountered by a molecule when it diffuses through a medium containing molecules of its own kind or molecules of comparable size? We have seen in the preceding discussion that the solvent can, by and large, be considered as continuum as long as the solvent molecules are small in comparison with those of the solute; in that case, the simple viscosity correction as given in equation 47 is applicable. The data given in table 6 demonstrate that, in the present case, the calculated values for the diffusion constant are too high if the full viscosity correction is applied, and too low, if the contribution of the large molecules to the frictional resistance is neglected.

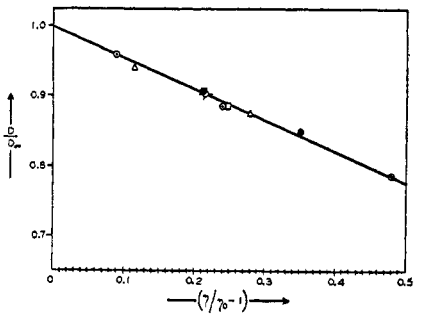

FIG. 16. Diffusion in concentrated solutions. A plot of D/D_∞ against the specific viscosity of the "solvent" $\left(\frac{\eta}{\eta_0} - 1\right)$, according to equation 48. ⊙, egg albumin; □, lactoglobulin; ●, serum albumin, △, sucrose; ◇, sucrose in serum albumin. Compare also table 6.

A thorough discussion of the problem would greatly exceed the scope of this review; however, it may be of interest to present here an empirical relation which, in the course of a preliminary investigation, Dr. G. R. Cooper and the author have found to be applicable to the diffusion of protein into protein, of sucrose into sucrose, and of sucrose into protein (5). In all these cases, the diffusing molecules are either of *comparable* size or *smaller* than those contained in the medium into which diffusion occurs. The experimental data employed were those of Polson (42), as well as unpublished measurements carried out in this laboratory on the diffusion of concentrated serum albumin solutions, and on the diffusion of 1 per cent sucrose dispersed in a 4 per cent serum albumin solution and diffusing into the 4 per cent albumin solution.

A linear relation was obtained when D/D_∞ was plotted against $\left(\frac{\eta}{\eta_0} - 1\right)$, where D is the measured diffusion constant, D_∞ its limiting value for infinite dilution, and $\left(\frac{\eta}{\eta_0} - 1\right)$ the specific viscosity of the medium into which diffusion took place. The data are plotted in figure 16. In this plot, only those mono-

disperse proteins have been considered for which the observed diffusion constant decreases with increasing concentration. Polson's data for serum globulin were omitted on account of the polydisperse nature of the preparation employed by him. The straight line drawn through the points in figure 16 follows the equation

$$\frac{D}{D_\infty} = 1 - k\left(\frac{\eta}{\eta_0} - 1\right) \tag{48}$$

where k turned out to be 0.446. In column 9 of table 6, there are given values of D/D_∞, calculated by means of equation 48. Comparison with the D_∞ values given in column 7 of this table indicates a very satisfactory agreement between these two sets of data. The physical meaning of equation 48 is as yet obscure. It will be recalled that $\left(\frac{\eta}{\eta_0} - 1\right)$, the specific viscosity of a solution, is proportional to its volume concentration; hence, the equation may also be written as

$$\frac{D}{D_\infty} = 1 - k'\phi \tag{49}$$

where ϕ is the ratio of the volume of the dispersed phase to that of the dispersion medium[15].

While no theoretical explanation will be offered at this time, nevertheless it is remarkable that this equation is equally valid for the high-molecular-weight proteins as well as for the low-molecular-weight sucrose. This fact suggests the equation to be of universal application. When Polson's measurements on hemocyanin *Homarus* and amandin are evaluated with our equation, one finds an increase of the apparent diffusion constant with increasing protein concentration. This might be due either to a splitting of the molecules, or to a decrease in molecular asymmetry without concomitant changes in size. The former interpretation appears to be more likely; Pedersen has found protein–protein interaction to cause, with certain proteins, dissociation into smaller units (40), and the hemocyanins in general, and amandin in particular, are known to be susceptible to dissociation. Moreover, the dissymmetry constants of both proteins in the native state are relatively low and could not decrease to the extent required by the increase in diffusion constant if the effects were ascribed entirely to changes in molecular shape.

With *Lampetra* hemoglobin, the observed decrease in diffusion constant is larger than that predicted by equation 48, suggesting, on the basis of the present considerations, an increase either in molecular asymmetry, or in molecular size, or both.

The empirical equation 48 may conceivably find application also to sedimentation-velocity measurements. Here, too, the sedimenting molecules are moving through a medium containing molecules of their own kind and should thus be affected by their frictional resistance. While in very dilute solutions this effect can be neglected, it should become significant in more concentrated solutions.

[15] For solutions of spherical particles, k should be $1/2.5 = 0.4$. The close agreement between that value, and the value of 0.446 obtained here, is probably fortuitous.

The absolute values for the diffusion constants of proteins vary between about 1×10^{-6} and 1×10^{-7}. With the large virus proteins, considerably lower values have been found, about 3×10^{-8} for tobacco mosaic virus protein and 6.4×10^{-8} for the rabbit papilloma virus protein, for instance. In table 7 are summarized

TABLE 7

*Observed diffusion constants and calculated molecular weights and dissymmetry constants of certain proteins**

PROTEIN	$D^{25°}_{H_2O} \times 10^7$	M	f/f_0	REFERENCE
Ribonuclease	13.6†	12,700	1.04	(43)
Pepsin B	10.0			(29)
A	9.2			(29)
Oxytocic pressure hormone	8.5	30,000	1.18	(54)
Horse globin	7.5	37,000	1.47	(8)
Horse serum albumin (McMeekin)	7.4	69,000‡	1.21	(29)
Horse hemoglobin (recombined)	7.2	69,000	1.23	(8)
Diphtheria toxin	7.1	70,000	1.22	(41)
Horse serum albumin (Hewitt)	7.0	73,000‡	1.25	(29)
Digested antitoxin (diphtheria)	6.7	98,000	1.14	(41)
Myogen	5.5	150,000	1.26	(7)
Catalase	5.2			(47)
Secalin	4.8			(1)
Horse pseudoglobulin GI	4.6	174,000‡	1.42	(29)
GII	4.8	170,000‡	1.39	(29)
Antitoxin (diphtheria)	4.4	150,000	1.4	(38)
Horse serum albumin denatured				
by 8 *M* urea	4.1	78,000‡	1.86	(30)
by 8 *M* guanidine hydrochloride	3.7	81,000‡	2.03	(30)
Horse pseudoglobulin GII denatured				
by 5.6 *M* guanidine hydrochloride	4.0	75,000‡	1.95	(31)
by 8 *M* urea	3.0	170,000‡	1.95	(31)
Myogen denatured by 6 *M* urea	2.7	72,000	3.2	(7)
Tomato bushy stunt virus	1.2	10,000,000	1.27	(28)
Calf thymus nucleohistone	1.2	2,300,000	2.5	(4)
Rabbit papilloma virus	0.59	47,000,000	1.49	(32)
Tobacco mosaic virus	0.3§	60,000,000§	2.5§	(33)

* Including only those data that have been reported since the publication of Svedberg and Pedersen's monograph (49). For previous work see Table 48 of the monograph.

† Determined with the Tiselius electrophoresis apparatus.

‡ Calculated from diffusion and viscosity data.

§ This value is tentative, in view of the nature of the preparation employed in diffusion measurements, and in view of the observed diffusion anomalies.

the results of the most important diffusion measurements that have been reported since the publication of Svedberg and Pedersen's monograph (49). Unless otherwise indicated, the refractometric-scale method was used for these measurements. There are given also in this table values for the molecular weights and dissymmetry constants of the proteins in question.

IV. Applications

The information that can be obtained from diffusion measurements on proteins is threefold. In the first place, the size of the protein molecules may be obtained directly from diffusion data if the molecules are known to be spherical in shape; if they are non-spherical, diffusion measurements have to be supplemented by some other physical or chemical method which furnishes a measure of a second parameter of the dimensions of the molecules. In the second place, it is possible to determine from changes in the diffusion rate any changes in the size or shape of the molecules which may have occurred as a result of chemical or physical reactions. Lastly, diffusion measurements yield an estimate of the degree of homogeneity of a protein solution.

The relation between diffusion constant and molecular dimensions has been reviewed on various recent occasions (20, 25, 26, 29, 37) and need not be considered here in any detail. The relation between diffusion constant and frictional coefficient is, generally,

$$D = \frac{RT}{f} \tag{50}$$

and for spherical molecules,

$$D_0 = \frac{RT}{f_0} \tag{51}$$

where

$$f_0 = 6\pi\eta N r$$

N is the Avogadro number, η is the viscosity of the solvent, and r is the molecular radius. For spherical molecules, the diffusion constant is, therefore, related to the molecular weight by

$$D = \frac{RT}{6\pi\eta N}\left(\frac{3MV}{4\pi N}\right)^{-1/3} \tag{52}$$

If the molecules are non-spherical, we have the relation

$$f = \frac{M(1 - V\rho)}{s} = \frac{RT}{D} \tag{53}$$

and

$$M = \frac{RTs}{D(1 - V\rho)} \tag{54}$$

This equation is usually employed for molecular-weight determinations from combined diffusion and sedimentation-rate measurements. Here s is the sedimentation constant, V the partial specific volume of the solute, and ρ the density of the solvent. Molecular weights can also be determined from diffusion

measurements combined with any method which yields a value for f/f_0, the dissymmetry constant. The corresponding relation is

$$M = K\frac{1}{D^3 V}\left(\frac{f_0}{f}\right)^{1/3} \tag{55}$$

where

$$K = \frac{R^3 T^3}{162\pi^2 \eta^3 N^2}$$

The dissymmetry constant is related to the apparent molecular shape, as expressed by the ratio of a major to a minor axis of a prolate or oblate ellipsoid of revolution, by the Perrin equation; hence, any method that relates to the apparent molecular shape can be used, in conjunction with diffusion measurements, for molecular-weight determinations. It has already been pointed out (29) that values obtained in this manner are essentially independent of the degree of hydration.

Combined diffusion and viscosity measurements have been carried out in this laboratory on a large number of native and denatured proteins, and the agreement between the values obtained in this manner and those obtained by other methods, such as those of osmotic pressure or combined diffusion and sedimentation, is rather satisfactory (29, 30, 31). Measurements of the dielectric dispersion or double refraction of flow (37) can also be used for an estimation of the molecular shape and, in combination with diffusion data, of the molecular weight.

If the molecular weight of a protein is known, diffusion measurements lend themselves to an estimation of the dissymmetry constant, and hence of the *apparent* molecular shape. This term has been used by us to denote the shape that is calculated if the influence of hydration is neglected (25). It should be emphasized here that it was never inferred that the influence of hydration is of a negligibly small order of magnitude; rather, these calculations were made merely to show what kind of result is obtained if the dissymmetry factor is interpreted solely in terms of molecular asymmetry. The estimate of about 33 per cent hydration made in more recent calculations (29) appears to represent a good approximation. The relation between hydration and molecular asymmetry has been considered by Kraemer (12, 13) and others (29, 37) and need not be discussed here. It is only in the case of spherical shape that the degree of hydration can be estimated from diffusion data, in which case (12, 13)

$$\frac{D_0}{D} = \left(\frac{V_a + rV_1}{V_1}\right)^{1/3} \tag{56}$$

where D is the observed diffusion constant, D_0 that calculated for a spherical anhydrous molecule, V_a the partial specific volume of the anhydrous protein, V_1 that of the solvent, and r the amount of solvent combined with 1 g. of anhydrous protein. This equation assumes the density of the adsorbed solvent to be the same as that of the liquid in bulk.

This approach to the general problem of the relation between diffusion constant and molecular dimensions contains a number of theoretical approximations and uncertainties which call for further attention. The brief discussion has been given merely to illustrate the importance of diffusion measurements for an estimation of the dimensions of protein molecules.

Many of the ideas expressed in this paper have evolved as a result of discussions with Dr. L. W. Nordheim of the Department of Physics, Duke University, and with Dr. G. R. Cooper and Mr. John O. Erickson of this Department. The author is also indebted to the Rockefeller Foundation and to the Lederle Laboratories, Inc., for financial support which has made the work carried out in this laboratory possible.

REFERENCES

(1) ANDREWS, A. C.: J. Am. Chem. Soc. **62,** 942 (1940).
(2) ANSON, M. L., AND NORTHROP, J.: J. Gen. Physiol. **20,** 575 (1936–37).
(3) BOURDILLON, J.: J. Gen. Physiol. **24,** 459 (1941); **25,** 263 (1941).
(4) CARTER, R. O.: J. Am. Chem. Soc. **63,** 1960 (1941).
(5) COOPER, G. R., AND NEURATH, H.: Unpublished data.
(6) GLASSTONE, S., LAIDLER, K. J., AND EYRING, H.: *The Theory of Rate Processes.* McGraw-Hill Book Company, Inc., New York (1941).
(7) GRALÉN, N.: Biochem. J. **33,** 1342 (1939).
(8) GRALÉN, N.: Biochem. J. **33,** 1907 (1939).
(9) GRALÉN, N.: Kolloid-Z. **95,** 188 (1941).
(10) HAND, D. B.: J. Am. Chem. Soc. **61,** 3180 (1939).
(11) KINCAID, J. F., EYRING, H., AND STEARN, A. E.: Chem. Rev. **28,** 301 (1941).
(12) KRAEMER, E. O.: J. Franklin Inst. **229,** 393, 680 (1940).
(13) KRAEMER, E. O.: In T. Svedberg and K. O. Pedersen's *The Ultracentrifuge,* p. 57, ff. Oxford University Press, London (1940).
(14) KUNITZ, M., AND NORTHROP, J.: J. Gen. Physiol. **18,** 433 (1934–35).
(15) LAMM, O.: Z. physik. Chem. **A138,** 313 (1928); **A143,** 177 (1929).
(16) LAMM, O.: Nova Acta Regiae Soc. Sci. Upsaliensis [4] **10,** No. 6 (1937).
(17) LAMM, O., AND POLSON, A.: Biochem. J. **30,** 538 (1936).
(18) LONGSWORTH, L. G.: Ann. N. Y. Acad. Sci. **41,** 267 (1941).
(19) LOUGHBOROUGH, D. L., AND STAMM, A. J.: J. Phys. Chem. **40,** 1113 (1936).
(20) LUNDGREN, H. P., AND WILLIAMS, J. W.: J. Phys. Chem. **43,** 988 (1939).
(21) MCBAIN, J. W., DAWSON, C. R., AND BARKER, H. A.: J. Am. Chem. Soc. **56,** 1021 (1934).
(22) MCBAIN, J. W., AND LIU, T. H.: J. Am. Chem. Soc. **53,** 59 (1931).
(23) MEHL, J. W., AND SCHMIDT, C. L. A.: University of California Publications in Physiology **8,** 165 (1937).
(24) NEURATH, H.: Cold Spring Harbor Symposia Quant. Biol. **6,** 196 (1938).
(25) NEURATH, H.: J. Am. Chem. Soc. **61,** 1841 (1939).
(26) NEURATH, H.: Cold Spring Harbor Symposia Quant. Biol. **8,** 80 (1940).
(27) NEURATH, H.: Science **93,** 431 (1941).
(28) NEURATH, H., AND COOPER, G. R.: J. Biol. Chem. **135,** 455 (1940).
(29) NEURATH, H., COOPER, G. R., AND ERICKSON, J. O.: J. Biol. Chem. **138,** 411 (1941).
(30) NEURATH, H., COOPER, G. R., AND ERICKSON, J. O.: J. Biol. Chem. **142,** 249 (1942).
(31) NEURATH, H., COOPER, G. R., AND ERICKSON, J. O.: J. Biol. Chem. **142,** 265 (1942).
(32) NEURATH, H., COOPER, G. R., SHARP, D. G., TAYLOR, A. R., BEARD, D., AND BEARD, J. W.: J. Biol. Chem. **140,** 293 (1941).

(33) NEURATH, H., AND SAUM, A. M.: J. Biol. Chem. **126,** 435 (1938).
(34) NEURATH, H., AND SAUM, A. M.: J. Biol. Chem. **128,** 347 (1939).
(35) NORTHROP, J.: J. Gen. Physiol. **13,** 739 (1929–30).
(36) NORTHROP, J., AND ANSON, M. L.: J. Gen. Physiol. **12,** 543 (1928–29).
(37) ONCLEY, J. L.: Ann. N. Y. Acad. Sci. **49,** 121 (1941).
(38) PAPPENHEIMER, A. M., JR., LUNDGREN, H. P., AND WILLIAMS, J. W.: J. Exptl. Med. **71,** 247 (1940).
(39) PEARSON, K.: Phil. Trans. London **185A,** 71 (1894).
(40) PEDERSEN, K. O.: In T. Svedberg and K. O. Pedersen's *The Ultracentrifuge*, p. 356. Oxford University Press, London (1940).
(41) PETERMANN, M. L., AND PAPPENHEIMER, A. M., JR.: J. Phys. Chem. **45,** 1 (1941).
(42) POLSON, A.: Kolloid-Z. **87,** 149 (1939); **88,** 51 (1939).
(43) ROTHEN, A.: J. Gen. Physiol. **24,** 203 (1940).
(44) SAUM, A. M.: Thesis, Cornell University, 1938.
(45) SCHERP, H. W.: J. Gen. Physiol. **16,** 795 (1932–33).
(46) STERN, K. G.: Z. physiol. Chem. **217,** 237 (1933).
(47) SUMNER, J. B., DOUNCE, A. L., AND FRAMPTON, V. L.: J. Biol. Chem. **136,** 343 (1940).
(48) SVEDBERG, T.: Kolloid-Z. (Ergänzungsband) **36,** 53 (1925).
(49) SVEDBERG, T., AND PEDERSEN, K. O.: *The Ultracentrifuge.* Oxford University Press, London (1940).
(50) SVENSSON, H.: Kolloid-Z. **81,** 181 (1939); **90,** 141 (1940).
(51) THOEVERT, J.: Ann. chim. phys. [7] **26,** 366 (1902).
(52) THOEVERT, J.: Ann. phys. [9] **2,** 369 (1914).
(53) TISELIUS, A., AND GROSS, D.: Kolloid-Z. **66,** 12 (1934).
(54) VAN DYKE, H. B., CHOW, B. F., GREEP, R. O., AND ROTHEN, A.: Am. J. Physiol. **133,** 473 (1941).
(55) WIENER, O.: Ann. Physik Chem. [N.F.] **49,** 105 (1893).
(56) WILLIAMS, J. W., AND CADY, L. C.: Chem. Rev. **14,** 171 (1934).
(57) ZEILE, K.: Biochem. Z. **258,** 347 (1938).