JOURNAL OF THE
AMERICAN CHEMICAL SOCIETY

VOLUME 76 OCTOBER 4, 1954 NUMBER 19

[CONTRIBUTION NO. 1225 FROM THE DEPARTMENT OF CHEMISTRY OF YALE UNIVERSITY]

Theory of the Self-diffusion of Water in Protein Solutions. A New Method for Studying the Hydration and Shape of Protein Molecules

BY JUI H. WANG

RECEIVED MAY 25, 1954

A theory is developed to treat quantitatively the self-diffusion of water in protein solutions. The result can be used to study the controversial problem of protein hydration from a new angle. The hydration of proteins computed from self-diffusion data by means of the present theory is independent of the molecular weight of the protein, and for proteins such as ovalbumin depends only slightly on molecular shape. The hydration of ovalbumin in its isoelectric salt-free solution at 10.0° is determined in the present work to be 0.18 ± 0.01 g. of water per g. of dry protein. By combining this value with the diffusion data for ovalbumin, the axial ratios of the equivalent ellipsoid of revolution for ovalbumin are calculated to be 2.6:1:1. These results are discussed and compared with some selected data from the dielectric absorption and low-angle X-ray scattering measurements for ovalbumin solutions. Existing data on the viscosity of ovalbumin solutions are also examined, and the usual interpretation of intrinsic viscosity is criticized.

The hydration of proteins in solution is a problem of fundamental importance in biochemistry and biophysics. The term "hydration" is usually defined as the average amount of water carried by unit weight of protein when the protein molecules migrate through solution (*e.g.*, in diffusion, ordinary electrophoresis and sedimentation experiments). Although a huge amount of research work has been reported in the literature, there still appears to be no single unambiguous method available for determining quantitatively the hydration of proteins in solution. Estimations of this hydration by different experimental methods have recently been comprehensively reviewed by Edsall.[1] In most of these methods the measured result is expressed as a function of hydration and the shape of protein molecule. Since neither the hydration nor the shape of the hydrated protein molecule is known, it is consequently not possible to determine accurately either of these by any of these methods. The only exception appears to be the dielectric dispersion method of Oncley[2] in which the dielectric absorption and dispersion are explained solely by the rotation of rigid ellipsoidal protein molecules with permanent dipole moments. However, according to Kirkwood and Shumaker[3] "the relaxation time spectrum of an ellipsoidal molecule will be determined not only by external rotatory diffusion but also by the diffusion of the mobile protons on the surface of the molecule." While evidences for the validity of the Kirkwood–Shumaker theory for dielectric dispersion are yet to be established, one cannot in the meantime accept the results of Oncley and co-workers without reservations. Moreover, even if we accept Oncley's interpretation of dielectric dispersion, the experimental uncertainties are usually such that it is difficult to draw quantitative conclusions from these measurements. In the present work the problem of protein hydration is studied from a new angle. This method is based experimentally on the measurement of the self-diffusion of water in protein solutions. Provided the self-diffusion coefficient of the protein is very small as compared to that of water in the same solution, the hydration of the protein determined by the present method is independent of the numerical value of the molecular weight of the protein. In some cases, *e.g.*, for prolate ellipsoids of revolution, the value of hydration so determined is also almost independent of the shape of the protein molecules.

(1) J. T. Edsall, Chap. 7, "The Proteins" (Edited by H. Neurath and K. Bailey), Vol. I, Part B, Academic Press, Inc., New York, N. Y., 1953.

(2) See reference 1.

(3) J. G. Kirkwood and J. B. Shumaker, *Proc. Natl. Acad. Sci.*, **38**, 855 (1952).

Theory of the Self-diffusion of Water in Protein Solutions

The self-diffusion coefficient of water in protein solutions is smaller than that in pure water for two reasons. Firstly, protein molecules have a much larger volume and a much smaller self-diffusion coefficient than the water molecules. These large

and almost stationary (compared to the Brownian motion of the water molecules) protein molecules obstruct the paths for water molecules; *i.e.*, the water molecules near a protein molecule have to diffuse along longer paths in order to get to the other side of the protein molecule. But since in experimental measurements we compute the self-diffusion coefficient of water by taking a macroscopic dimension of the diffusion apparatus as the length of the diffusion path irrespective of whether or not the diffusion path is blocked microscopically, the net effect of the above considerations will be to yield a measured self-diffusion coefficient of water in protein solution smaller than that in pure water. In what follows we shall refer to this effect as the "obstruction effect." The second reason is that a fraction of the water molecules are firmly attached to the protein molecules (hydration) and hence do not contribute to the rate of self-diffusion of water. For the term "hydration" we refer to the definition given at the beginning of this article. Thus, if the nth water molecule is so loosely attached to the protein molecule that it remains attached to the latter during only part of the time when the latter migrates, we shall count the former as only a fraction of a hydrated water molecule. Since in self-diffusion measurements we make no distinction between "bound" and "free" water molecules, the measured self-diffusion coefficient of total water in protein solution is consequently less than that in pure water. We shall refer to this second effect as the "direct hydration effect" to distinguish it from the effect of hydration on the volume of the hydrated protein molecules to be discussed in the following paragraphs under "obstruction effect."

For aqueous solutions of small molecules or ions there is a third effect due to the distortion or "breaking down" of water structure by the solute particles. Thus for some aqueous solutions of slightly hydrated electrolytes the distortion in water structure caused by these ions is large enough to make the self-diffusion coefficient of water in these solutions even greater than that in pure water.[4] Because of this kind of distortion effect the temperature coefficients of the limiting mobilities of these slightly hydrated ions are often appreciably smaller than the temperature coefficient of the fluidity of water.[5] Consequently, for these ions the product $D^\circ\eta_0/T$, where D° is the self-diffusion coefficient of the ion at infinite dilution and η_0 is the viscosity of pure water, has been found to decrease with increasing temperature.

For aqueous protein solutions, however, no appreciable amount of distortion of this kind has ever been detected. Indeed the general success of the approximate relationship $D_p^\circ\eta_0/T$ = constant used by numerous workers to correlate the limiting diffusion coefficient, D_p°, of protein molecules with the viscosity, η_0, of water at different temperatures T may be taken as evidence that no appreciable amount of distortion in the structure of "free" water exists in protein solutions. Consequently we shall neglect this third effect in the following treatment of the problem.

(4) J. H. Wang, *J. Phys. Chem.*, **58**, 686 (1954).
(5) J. H. Wang, THIS JOURNAL, **74**, 1612 (1952).

1. The Obstruction Effect.—To simplify the boundary conditions of the present problem let us consider two liquid baths of infinite capacity connected by a tube of small but uniform cross-sectional area. Let one bath be filled with a solution of protein of a given concentration in ordinary water, and the other bath be filled with a solution of protein in H_2O^{18}-labeled water at the same protein concentration. The connecting tube may initially be filled with the solution in either bath. Diffusion of H_2O^{18} through the connecting tube will take place. As $Dt/l^2 \to \infty$, where l is the length of the tube, a steady state will be reached at which the concentration at each point in the tube remains constant. Thus we have

$$\nabla^2 c = 0 \tag{1}$$

at each point in the tube, where ∇^2 is the Laplacian operator and c is the concentration of labeled water molecules. Let us assume that the shape of ovalbumin molecules can be approximated by ellipsoids with principal semi-axes a, b and c, respectively, and consider a particular ovalbumin molecule whose a-axis is for the moment parallel to the direction of the diffusion tube as the fundamental ellipsoid of our system of ellipsoidal coördinates λ, μ, ν. We shall refer to the direction of the diffusion tube as the x-axis with origin located at the center of our fundamental ellipsoid. If $a > b > c$, we have by definition $\lambda > -c^2 > \mu > -b^2 > \nu > -a^2$ and equation 1 can be written as

$$\begin{aligned}\Big\{&(\mu-\nu)\Big[(a^2+\lambda)^{1/2}(b^2+\lambda)^{1/2}(c^2+\lambda)^{1/2}\frac{\partial}{\partial\lambda}\Big]^2\\ &+(\nu-\lambda)\Big[(a^2+\mu)^{1/2}(b^2+\mu)^{1/2}(c^2+\mu)^{1/2}\frac{\partial}{\partial\mu}\Big]^2\\ &+(\lambda-\mu)\Big[(a^2+\nu)^{1/2}(b^2+\nu)^{1/2}(c^2+\nu)^{1/2}\frac{\partial}{\partial\nu}\Big]^2\Big\}c=0\end{aligned} \tag{2}$$

The equation of the fundamental ellipsoid is now simply $\lambda = 0$ and the new coördinates λ, μ, ν are related to the Cartesian coördinates x, y, z by

$$\frac{\partial x}{\partial\lambda} = \frac{x}{2(a^2+\lambda)} \tag{3}$$

etc.

A convenient solution of (2) for the present problem is of the form

$$c = K_1 + K_2 x\int_\lambda^{K_3}\frac{d\lambda}{\Omega_a(\lambda)} \tag{4}$$

where

$$\Omega_a(\lambda) = (a^2+\lambda)^{3/2}(b^2+\lambda)^{1/2}(c^2+\lambda)^{1/2}$$

and K_1, K_2 and K_3 are arbitrary constants. The boundary conditions are that at distances sufficiently far from the ellipsoid the "obstruction" effect should vanish, *i.e.*

$$c = \bar{c} + c'x \text{ at } \lambda = \infty \tag{5}$$

where $\bar{c}$ is the average concentration of labeled water at $x = 0$ and c' is equal to the difference in the concentration of labeled water in the two infinite baths divided by the length of the diffusion tube, and that the "free" water molecules in contact with the ellipsoid can only diffuse tangentially to the surface of the ellipsoid, *i.e.*

$$(\partial c/\partial\lambda)_{\lambda=0} = 0 \tag{6}$$

Combining equations 4, 5, 6, and 3 we obtain the final form of the solution of (2) as

$$c = \bar{c} + \frac{c'x}{\frac{2}{abc} - \omega_a}\left[\frac{2}{abc} - \int_0^\lambda \frac{d\lambda}{\Omega_a(\lambda)}\right] \qquad (7)$$

where

$$\omega_a = \int_0^\infty \frac{d\lambda}{\Omega_a(\lambda)}$$

To compute the obstruction effect let us consider the portion of liquid enclosed in the imaginary cylinder with dimensions much larger than those of the protein molecules and with axis parallel to the x-axis as depicted by the broken lines in Fig. 1. The rate of diffusional flow of labeled water molecules in the x-direction averaged over the entire volume of this imaginary cylinder is

$$q = \frac{-D_0}{A(x_2 - x_1)} \int_V \left(\frac{\partial c}{\partial x}\right) dv \qquad (8)$$

where A is the cross-sectional area, and $x_2 - x_1$ is the length of the imaginary cylinder, and the integration is to be carried out over the entire volume of the cylinder. Equation 8 can be transformed by means of Green's theorem to

$$q = \frac{-D_0}{A(x_2 - x_1)} \int\int_S c \cos\theta ds \qquad (9)$$

where the integration is to be carried out over the entire surface of the imaginary cylinder as well as the surfaces of all the ellipsoids enclosed in the cylinder, and θ is the angle between the normal of the surface and the x-axis.

Substituting the boundary values of c as given by (7) by putting $\lambda = \infty$ at the surface of the imaginary cylinder and $\lambda = 0$ at the surface of each of the ellipsoids and carrying out partial integration, we obtain

$$q = -D_0\left[c' - \frac{1}{A(x_2 - x_1)} \sum_i \int\int_{S_i} c \cos\theta ds\right] \qquad (10)$$

where the symbol $\sum_i$ represents summation over all ellipsoids enclosed in the imaginary cylinder.

Differentiating (7) with respect to x and putting $\lambda = 0$, we get

$$\left(\frac{\partial c}{\partial x}\right)_{\lambda=0} = \frac{2c'}{2 - abc\omega_a} \qquad (11)$$

Transforming the surface integrals in (10) back to volume integrals by means of Green's theorem and substituting the value of $(\partial c/\partial x)_{\lambda=0}$ as given by (11), we obtain

$$q = -D_0c'\left[1 - \frac{1}{A(x_2 - x_1)} \sum_i \left(\frac{2}{2 - abc\omega_a}\right) \int_{V_i} dv\right] \qquad (12)$$

$$= -D_0c'[1 - \alpha_a\phi]$$

where

$$\alpha_a = \frac{2}{2 - abc\,\omega_a}$$

and ϕ is the total volume fraction occupied by the hydrated protein molecules.

Let us define an effective self-diffusion coefficient, D_a', of the "free" water molecules in the protein solution by

$$q \equiv -D_a'\,c' \qquad (13)$$

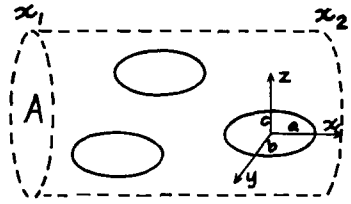

Fig. 1.—Diagrammatic representation of a section of the diffusion tube discussed in the treatment of the obstruction effect.

Comparison of (12) and (13) gives

$$D_a' = (1 - \alpha_a\phi)D_0 \qquad (14)$$

Equation 14 has been derived on the assumption that the ellipsoids are so oriented that their a-axes are parallel to the diffusion tube. Had we assumed that the b-axes were parallel to the diffusion tube, we would have obtained an expression similar to (14) but with a and b interchanged. Therefore, by symmetry we can immediately write down the general expression

$$D_i' = (1 - \alpha_i\phi)D_0$$

where

$$\left.\begin{aligned} \alpha_i &= \frac{2}{2 - abc\,\omega_i} \\ \omega_i &= \int_0^\infty \frac{d\lambda}{(i^2 + \lambda)(a^2 + \lambda)^{1/2}(b^2 + \lambda)^{1/2}(c^2 + \lambda)^{1/2}} \end{aligned}\right\} \qquad (15)$$

with $i = a, b, c$ in turn.

In order to simplify the computation, let us further assume that the ellipsoids have two of their principal axes equal, *i.e.*, ellipsoids of revolution. For prolate ellipsoids, we have

$$a = \rho b = \rho c, \text{ with } \rho > 1$$

Therefore

$$\omega_a = \int_0^\infty \frac{d\lambda}{(a^2 + \lambda)^{1/2}(b^2 + \lambda)} = \frac{1}{a^2 - b^2}\left[\frac{1}{\sqrt{a^2 - b^2}} \ln \frac{a + \sqrt{a^2 - b^2}}{a - \sqrt{a^2 - b^2}} - \frac{2}{a}\right] \qquad (16)$$

and consequently

$$\alpha_a = \frac{1}{\frac{\rho^2}{\rho^2 - 1} - \frac{\rho}{2(\rho^2 - 1)^{3/2}} \ln \frac{\rho + \sqrt{\rho^2 - 1}}{\rho - \sqrt{\rho^2 - 1}}} \qquad (17)$$

For the special case $a = b = c$, the ellipsoid becomes a sphere, and we have by (15)

$$\alpha_a = \alpha_b = \alpha_c = 1.5$$

Similarly

$$\alpha_b = \frac{2}{\frac{\rho^2 - 2}{\rho^2 - 1} + \frac{\rho}{2(\rho^2 - 1)^{3/2}} \ln \frac{\rho + \sqrt{\rho^2 - 1}}{\rho - \sqrt{\rho^2 - 1}}} \quad (18)$$

For oblate ellipsoids, we have

$$a = \rho b = \rho c, \text{ with } \rho < 1$$

Similar computation yields

$$\alpha_a = \frac{1}{1 - \frac{\rho}{1 - \rho^2}\left\{\frac{1}{\sqrt{1 - \rho^2}}\left[\tan^{-1}\left(\frac{\rho}{\sqrt{1 - \rho^2}}\right) - \frac{\pi}{2}\right] + \frac{1}{\rho}\right.} \quad (19)$$

and

$$\alpha_b = \frac{2}{\frac{2 - \rho^2}{1 - \rho^2} - \frac{\rho}{(1 - \rho^2)^{3/2}}\left[\frac{\pi}{2} - \tan^{-1}\left(\frac{\rho}{\sqrt{1 - \rho^2}}\right)\right]} \quad (20)$$

Since in a real solution these ellipsoids are oriented in all possible random directions, some average value, $\bar{\alpha}$, should be used instead of either α_a or α_b. To treat the general problem, let us consider an ellipsoid in solution so oriented that the direction cosines of the x-axis (parallel to the diffusion tube) referred to the principle axes of this ellipsoid are l_a, l_b and l_c. The rates of diffusional flow in the direction of the three principal axes are

$$q_a = -D'_a c' l_a,\ q_b = -D'_b c' l_b,\ q_c = -D'_c c' l_c \quad (21)$$

respectively. Thus the total rate of diffusional flow, q, in the x-direction is

$$q = q_a l_a + q_b l_b + q_c l_c \quad (22)$$

Substituting (21) in (22) and average over all possible orientations, we have

$$\bar{q} \equiv -D'c' = \frac{-c'}{4\pi}\int (D_a l_a^2 + D_b l_b^2 + D_c l_c^2)\, d\Omega \quad (23)$$

where Ω is the solid angle and the average effective self-diffusion coefficient, D', of the "free" water molecules is defined by (23). Since the average value of each direction cosine is equal to 1/3, we have from (23)

$$D' = \frac{1}{3}(D_a + D_b + D_c) \quad (24)$$

Equation 24 has been derived by Perrin[6] from the theory of Brownian motion. Defining $\bar{\alpha}$ by

$$D' = D^0(1 - \bar{\alpha}\phi) \quad (25)$$

and substituting values of D'_a, D'_b, D'_c given by (15) in (24), we obtain

$$\bar{\alpha} = \frac{1}{3}(\alpha_a + \alpha_b + \alpha_c) \quad (26)$$

Values of $\bar{\alpha}$ are computed from equations (26), (17), (18), (19) and (20) for ellipsoids of revolution with different axial ratio ρ. These are listed for prolate and oblate ellipsoids separately in Table I. The application and significance of these values will be discussed in later sections of this article.

It is clear from the above derivation that equation 23 is strictly valid only for dilute protein solutions. However, since in this treatment we have minimized the effect of interactions between protein molecules by disregarding the Brownian motion of the protein molecules, we may expect equation 25 to hold even for fairly concentrated protein solutions as long as the volume fraction of "free" water is still much larger than that of the hydrated protein. But for very concentrated protein solutions in which equation 5 can no longer be used as the correct boundary condition, the result (25) cannot be expected to hold. Thus one should be aware of the limitation of equation 25 in applying it to interpret experimental data. For example, it would obviously be absurd to predict from equation 25 that the effective self-diffusion coefficient of water be negative in protein solutions with $\phi > 1/\bar{\alpha}$.

(6) F. Perrin, *J. Phys.*, **7**, 1 (1936).

TABLE I

VALUES OF $\bar{\alpha}$ IN EQUATION 25 FOR ELLIPSOIDS OF REVOLUTION WITH DIFFERENT AXIAL RATIO

Prolate ellipsoids $\rho = a/b$	$\bar{\alpha}$	Oblate ellipsoids $1/\rho = b/a$	$\bar{\alpha}$
1.0	1.500	1.0	1.500
1.5	1.516	1.5	1.525
2.0	1.539	2.0	1.577
2.5	1.561	2.5	1.649
3.0	1.576	3.0	1.730
4.0	1.601	4.0	1.909
5.0	1.616	5.0	2.101
6.0	1.627	6.0	2.295
7.0	1.634	7.0	2.489
8.0	1.640	8.0	2.694
9.0	1.644	9.0	2.916
10.0	1.647	10.0	3.107
∞	1.667	∞	∞

2. The Direct Hydration Effect.—The evaluation of the direct hydration effect is complicated by the rate of exchange of labeled water molecules between "bound" and "free" water, because this exchange introduces a time factor which is absent in equation 1. In order to analyze this complication it is expedient to consider the more general self-diffusion equation for non-stationary states. In the absence of protein hydration this equation may be written as

$$\frac{\partial c}{\partial t} = D' \frac{\partial^2 c}{\partial x^2}$$

for our experimental system. Let us define c_0^* and c_h^* as the concentration of total H_2O^{18} and the concentration of "bound" H_2O^{18}, respectively, in g. per cc. of solution, the concentration of "free" H_2O^{18} must then be $c_0^* - c_h^*$. Let c_0 and c_h be the concentrations of total water and total "bound" water, respectively. The general relationships between concentrations, time and coördinate in a system with both finite rate of self-diffusion and exchange may now be written as

$$\left.\begin{aligned} \frac{\partial(c_0^* - c_h^*)}{\partial t} &= D' \frac{\partial^2(c_0^* - c_h^*)}{\partial x^2} + k'\left[\frac{c_h^*}{c_h} - \left(\frac{c_0^* - c_h^*}{c_0 - c_h}\right)\right] \\ \frac{\partial c_h^*}{\partial t} &= D_p \frac{\partial^2 c_h^*}{\partial x^2} + k'\left[\left(\frac{c_0^* - c_h^*}{c_0 - c_h}\right) - \frac{c_h^*}{c_h}\right] \end{aligned}\right\} \quad (27)$$

where D_p is the self-diffusion coefficient of the protein and k' is defined as an effective rate constant for the exchange of labeled water molecules between "free" and "bound" water. Combining equations 27 we obtain

$$\frac{\partial c_0^*}{\partial t} = D' \frac{\partial^2(c_0^* - c_h^*)}{\partial x^2} + D_p \frac{\partial^2 c_h^*}{\partial x^2} \quad (28)$$

Let us now assume firstly that the term $D_p(\partial^2 c_h^*/\partial x^2)$ is negligibly small as compared to either of the other terms in equation 28, and secondly that the rate of exchange of labeled water molecules between "free" and "bound" water is instantaneously fast, *i.e.*, $k' \gg \frac{\partial(c_0^* - c_h^*)}{\partial t}$ or $\frac{\partial c_h^*}{\partial t}$ or $D' \frac{\partial^2(c_0^* - c_h^*)}{\partial x^2}$ or $D_p \frac{\partial^2 c_h^*}{\partial x^2}$ in equations 27. Consequently $\left|\frac{c_h^*}{c_h} - \left(\frac{c_0^* - c_h^*}{c_0 - c_h}\right)\right| \ll 1$, *i.e.*

$$\frac{c_h^*}{c_h} = \frac{c_0^* - c_h^*}{c_0 - c_h} = \frac{c_0^*}{c_0} \equiv \sigma \quad (29)$$

at all points along the diffusion path, where σ is the isotopic mole fraction of H_2O^{18} in water at given x and t. Here we have neglected the small isotope-effect for the case of H_2O^{18} as tracer.

Equation 28 can now be simplified to

$$\frac{\partial c_0^*}{\partial t} = c_0 \frac{\partial \sigma}{\partial t} = D'(c_0 - c_h) \frac{\partial^2 \sigma}{\partial x^2} \quad (30)$$

But by definition we have

$$\frac{\partial c_0^*}{\partial t} \equiv D \frac{\partial^2 c_0^*}{\partial x^2} = c_0 D \frac{\partial^2 \sigma}{\partial x^2} \quad (31)$$

where D is the experimentally measured apparent self-diffusion coefficient of water in protein solution. Hence

$$D = D'\left(1 - \frac{c_h}{c_0}\right) \quad (32)$$

Since the self-diffusion coefficient, D_p, of the protein is in general much smaller than the effective self-diffusion coefficient, D', of water and that for most protein solutions $(\partial^2 c_h^*/\partial x^2)$ is considerably smaller than $\partial^2(c_0^* - c_h^*)/\partial x^2$, our first assumption that the product $D_p(\partial^2 c_h^*/\partial x^2)$ can be neglected as compared to other terms in equation 28 is justified. The second assumption that the rate of exchange of labeled water molecules between "bound" and "free" water can be considered as instantaneously fast is supported by the failure of our attempt to detect a slow exchange of this kind. Thus by diluting a rapidly stirred concentrated solution of ovalbumin in ordinary water with labeled water and then immediately evaporating small fractions of the water in solution into a mass spectrometer, we found that the exchange rate was immeasurably fast at room temperatures. The self-diffusion of water in ovalbumin solutions at 10° has been measured by Wang, Anfinsen and Polestra.[7] Since each of their diffusion measurements lasted more than a day, we may for practical purposes consider the rate of exchange of labeled water molecules as instantaneous in interpreting their data by the present theory.

3. The Complete Theoretical Equation and Its Verification.—Let us define H as the hydration of the protein in solution expressed in g. of "bound" water per g. of anhydrous protein. Let c_p be the concentration of the protein in g. of anhydrous protein per cc. of solution, and w be the weight-fraction of anhydrous protein in solution (*i.e.*, $100w$ is the percentage by weight of dry protein in solution). Then we have

$$\frac{c_p}{c_0} = \frac{w}{1 - w} \quad (33)$$

and

$$c_p \bar{V}_p + c_0/d_0 = 1 \quad (34)$$

where $\bar{V}_p$ is the apparent specific volume of anhydrous protein in its aqueous solution and d_0 is the density of pure water. The volume-fraction, ϕ, of hydrated protein may now be written as

$$\phi = c_p(\bar{V}_p + H/d_0) = \frac{\bar{V}_p + H/d_0}{\bar{V}_p + (1/d_0)[(1 - w)/w]} \quad (35)$$

The use of the expression $c_p(\bar{V}_p + H/d_0)$ for the volume-fraction of the hydrated protein has been a controversial subject in the literature. For example, Scheraga and Mandelkern[8] hold that this neglects the electrostriction, selective adsorption, other hydrodynamic effects, etc., making it impossible to identify $\bar{V}_p$ with the specific volume of the anhydrous protein. The fact is, however, that $\bar{V}_p$ is in general different from the specific volume of the dry protein, and it would be a mistake to identify these two quantities. Had we used the specific volume of the dry protein instead of $\bar{V}_p$ in equation 35, we would indeed have introduced an error there. But we used the apparent specific volume, $\bar{V}_p$, of the anhydrous protein in solution which already includes the effect of electrostriction. The only approximation we made in equation 35 is that the effect of electrostriction is appreciable only in the hydration layer of each protein molecule. If this is true, the quantity $(\bar{V}_p + H/d_0)$ should represent the true specific volume of the hydrated protein, because then $\bar{V}_p$ will be smaller than the "true specific volume" of the anhydrous protein in solution and H/d_0 will be greater than the "true volume" of "bound" water by exactly the same amount. There is some experimental evidence in favor of this approximate assumption. For example, it is generally agreed that electrostriction in protein solutions is caused by the local collapse or distortion of the structure of water surrounding the protein due to the strong attraction of the latter for the former molecules. It has already been pointed above that the approximate constancy of the quantity $D_p^0\eta_0/T$ for a given protein at different temperatures indicates the absence of appreciable distortion in the structure of "free" water in protein solutions. Consequently we may infer that electrostriction in protein solutions occurs only in "bound" water.

The influence of selective adsorption and other hydrodynamic effects as pointed out by Scheraga and Mandelkern may be important in interpreting protein diffusion, sedimentation velocity and viscosity data. But inasmuch as we are here concerned only with salt-free protein in water solutions, there can be no selective adsorption phenomenon other than simple hydration. The hydrodynamic effects are irrelevant to the present problem, because we have made no use of hydrodynamic theories in this work. Indeed, the above treatment bears more mathematical resemblance to the

(7) J. H. Wang, C. B. Anfinsen and F. M. Polestra, THIS JOURNAL, **76**, 4763 (1954).

(8) H. A. Scheraga and L. Mandelkern, *ibid.*, **75**, 179 (1953).

dynamics of perfect fluids than hydrodynamics. While deviations of the real hydrated protein molecules from ellipsoids would cause some uncertainty in the above estimated magnitude of the obstruction effect, this uncertainty is much less serious than the corresponding uncertainties in the hydrodynamic theories of protein diffusion or viscosity. This point will be further examined in a later section of this article.

By combining equations 25, 32, 33 and 35, we obtain the complete equation

$$D = D_0\left\{1 - \bar{\alpha}\left[\frac{\bar{V}_p + H/d_0}{\bar{V}_p + \left(\frac{1}{d_0}\right)\left(\frac{1-w}{w}\right)}\right]\right\}\left[1 - \left(\frac{w}{1-w}\right)H\right] \quad (36)$$

For the convenience of numerical computations, equation 36 may be rearranged to the following form

$$\frac{D}{D_0} - \Delta_1 = 1 - [\bar{\alpha}(\bar{V}_p d_0 + H) + H]w + \Delta_2 \quad (37)$$

where

$$\Delta_1 = \frac{\bar{\alpha}\bar{V}_p d_0(\bar{V}_p d_0 - 1)w^2}{1 + (\bar{V}_p d_0 - 1)w}$$

$$\Delta_2 = \left[\frac{\bar{\alpha}H(\bar{V}_p d_0 - 1)}{1 + (\bar{V}_p d_0 - 1)w} - \frac{H}{1-w} + \frac{\bar{\alpha}(\bar{V}_p d_0 + H)H}{1 + (\bar{V}_p d_0 - 1)w(1-w)}\right]w^2$$

Equation 37 will be used as the basis of the computation for H from the self-diffusion data of Wang, Anfinsen and Polestra[7] for ovalbumin solutions. It may be noticed from (37) that the magnitudes of both Δ_1 and Δ_2 increase with protein concentration. However, even for the most concentrated ovalbumin solution in the above mentioned work the magnitudes of Δ_1 and Δ_2 are still very small as compared to the other terms in (37). Thus for their most concentrated solution $w \approx 0.25$. Using the approximate values $\bar{\alpha} \approx 1.6$, $H \approx 0.2$, $\bar{V}_p = 0.75$, $d_0 = 1$ for ovalbumin solution, we can compute the approximate value of Δ_1 and Δ_2 to be

$$\Delta_1 \approx -0.02;\ \Delta_2 \approx -0.003$$

It may be noticed from Table II that the measured value of D/D_0 varies from 1 at infinite dilution to about 0.6 for 25% ovalbumin solution with average experimental uncertainty of about 1%. Hence even for the most concentrated solution in the above mentioned work the term Δ_2 is still negligible. Consequently if we plot $(D/D_0) - \Delta_1$ *vs.* w we should get a straight line with slope equal to $\bar{\alpha}(\bar{V}_p d_0 + H) + H$ from which the hydration of protein can be computed. Values of $(D/D_0) - \Delta_1$ computed from the experimental data[7] are listed in Table II.

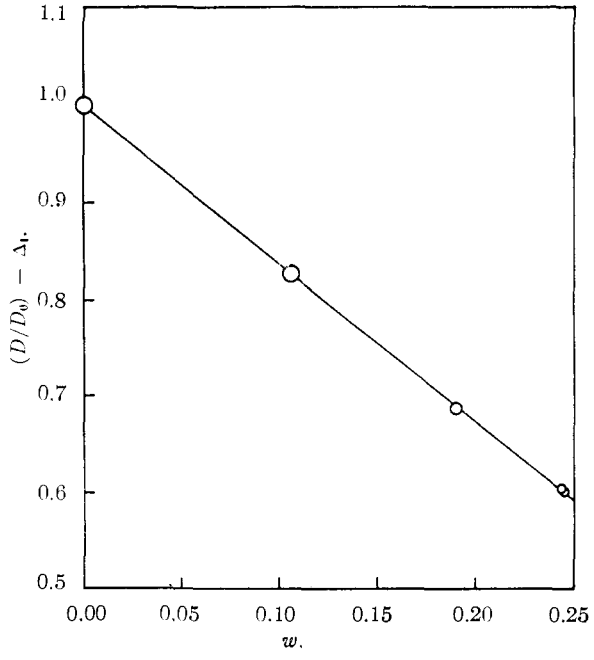

Fig. 2.—Experimental test of equation 37 for ovalbumin solutions.

TABLE II

DATA FOR TESTING EQUATION (37)

w	D/D_0	Δ_1	$(D/D_0) - \Delta_1$
0.000	1.000	0.000	1.000
.106	0.824	− .003	0.827
.190	.674	− .012	.686
.244	.585	− .019	.604
.245	.584	− .019	.603

In computing the values of Δ_1 listed in Table II, the value of $\bar{\alpha}$ has been assumed to be 1.6. It may be noticed from Table I that the exact value of $\bar{\alpha}$ for prolate ellipsoids varies from 1.5 to 1.67 depending on the axial ratio. Therefore this assumed value of $\bar{\alpha}$ may have a maximum error of 7%. But since the magnitude of Δ_1 is less than 4% of the magnitude of $(D/D_0) - \Delta_1$, the error in $(D/D_0) - \Delta_1$ due to error in this assumed value of $\bar{\alpha}$ must be less than 0.3% and is consequently entirely negligible as compared to other experimental uncertainties.

Values in the last column of Table II are plotted *vs.* w in Fig. 2. The plot is, within experimental uncertainties, a very good straight line as predicted by equation 37. Indeed this agreement is even better than what we expected, for in deriving (37), we neglected the perturbing effect of other protein molecules on the concentration gradients in the vicinity of the fundamental ellipsoid. This kind of perturbation may have appreciable effect on the measured self-diffusion coefficient of water in concentrated solutions. The rather surprising agreement between theory and experiment as depicted in Fig. 2 may either be due to the particular case of the ovalbumin as partly fortuitous, or to the fact that the effect of these interactions on the measured self-diffusion coefficient of water is small. More experimental work in this direction is necessary before a definite conclusion on this point can be drawn. But even if the $(D/D_0) - \Delta_1$ *vs.* w plot for some protein solutions deviates from a straight line at high concentrations, we should not have much difficulty in determining the hydration of the protein from such a plot, because it is only necessary to plot the values of $(D/D_0) - \Delta_1$ in dilute solutions and get the limiting slope.

Computation of the Hydration of Ovalbumin

The slope of the $(D/D_0) - \Delta_1$ *vs.* w plot in Fig. 2 is −1.63. Thus by (37) we have

$$1.63 = \bar{\alpha}(\bar{V}_p d_0 + H) + H \quad (38)$$

If we assume that the shape of hydrated ovalbumin molecules can be approximated by prolate ellipsoids, Table I shows that the value of $\bar{\alpha}$ varies only slightly when the axial ratio changes by a factor of 2 or 3. Thus as a first approximation we may guess that the axial ratio of ovalbumin lies between 2 and 4 corresponding to $\bar{\alpha} = 1.539$ and 1.601, respectively. If we take the average value 1.57 for $\bar{\alpha}$ in (38) with $\bar{V}_p = 0.746$ and $d_0 = 1$, we obtain

$$H = 0.18 \pm 0.01$$

where the uncertainty in H is estimated from the experimental errors of the self-diffusion measurements. This value of hydration, together with the experimentally determined limiting diffusion coefficient of ovalbumin, enable us to compute the axial ratio of the equivalent ellipsoid for ovalbumin. It will be shown in the next section that this axial ratio is 2.6, corresponding to $\bar{\alpha} = 1.56$. If we now substitute this new value of $\bar{\alpha}$ in (38) to make a second approximation in computing H, we will find that the new value of H differs from that given above by an amount less than the experimental uncertainties, and consequently the second approximation is quite unnecessary.

We may also recall that equation 37 was derived on the assumption that the shape of hydrated protein molecules can be approximated by compact ellipsoids. The actual shape of these molecules may of course deviate more or less from an ellipsoid. However, Table I shows that $\bar{\alpha}$ changes only little even when the shape of the ellipsoid changes continually from that of a sphere, through a series of stages of elongated ellipsoids, and finally into that of an infinite thin rod. This indicates that our method is not shape-sensitive, and that the value of hydration so obtained should remain valid even if the actual shape of the protein molecules deviates slightly from that of a compact prolate ellipsoid.

For oblate ellipsoids, Table I shows that $\bar{\alpha}$ is quite sensitive to the axial ratio, and consequently the above mentioned advantage disappears. This difference in property between prolate and oblate ellipsoids is easily understandable from simple geometric considerations. Thus as $\rho \rightarrow \infty$, the prolate ellipsoids degenerate into infinite thin rods. The water molecules can still diffuse through the network of thin rods no matter how these rods are oriented. But as $1/\rho \rightarrow \infty$, the oblate ellipsoids degenerate into infinite sheets. Some of these sheets will be so oriented as to block the diffusion path completely, *i.e.*, to exhibit an infinite obstruction effect.

The hydration of oblate protein molecules may be computed by successive approximations similar to those described above. But inasmuch as $\bar{\alpha}$ for oblate ellipsoids with $(1/\rho) > 1.5$ is quite shape-sensitive, the numerical result so obtained may include appreciable error due to deviation of the shape of real molecules from an oblate ellipsoid. Had we assumed an oblate ellipsoid for the shape of ovalbumin and tried to compute the hydration and axial ratio from the diffusion data for both water and protein by means of this procedure, we would obtain $H = 0.13$, $1/\rho = 2.9$ as our final results. It will be shown in a later section of this article that these values are not consistent with low angle X-ray scattering data. Consequently we may take this observation as evidence against the oblate ellipsoid as a possible shape for ovalbumin if we accept the X-ray data as correct.

Estimation of the Shape of the Hydrated Ovalbumin Molecule

If we accept 5.80×10^{-7} cm.^{2}/sec. as the limiting diffusion coefficient of ovalbumin in its infinitely dilute aqueous solution at 10°7, we may compute the axial ratio of the equivalent ellipsoid of ovalbumin by the usual procedure[9], using $H = 0.18$ obtained above and 44000 as the molecular weight of ovalbumin. The axial ratio, $\rho = a/b$, so obtained is 2.6 ± 0.1. But since the real hydrated ovalbumin molecules may deviate considerably from ellipsoids of revolution, the exact relationship between this axial ratio and the shape of the real molecules is somewhat uncertain.[8]

Comparison with Low Angle X-Ray Scattering Data

Guinier[10] has shown that the scattering of X-ray at low angles by protein solutions can be interpreted by a single constant R characteristic of the shape of the protein molecules. According to Guinier this constant R, which he called the "radius of gyration," is related to the axial ratio $\rho = a/b$ and the semi-axis b by

$$R = \left(\sqrt{\frac{2 + \rho^2}{5}}\right) b \qquad (39)$$

for ellipsoids of revolution. Thus if either ρ or b is known, the other quantity can be computed from the value of R determined from X-ray scattering measurements according to (39). By combining the values of R obtained from X-ray scattering measurements with diffusion and sedimentation data, Ritland, Kaesberg and Beeman[11] have estimated the hydration and axial ratio of a number of proteins. The values estimated by these workers appear reasonable, although their method of computation which is based on the implicit assumption that ρ has the same value in both X-ray scattering and diffusion measurements is in general not strictly valid. In some special cases, such as the uniform distribution of hydration throughout the protein molecule, ρ will indeed have the same value in these two sets of measurements. However, in order to explain the high solubility of proteins such as ovalbumin, the weight of opinion favors the assumption that a major part of the hydration is bound externally to the protein. Thus if most of the hydration is distributed in the form of a thin layer of uniform thickness δ wrapped around a more or less anhydrous protein core, the axial ratio for X-ray scattering would be $(a - \delta)/(b - \delta)$ which is in general not equal to a/b, although the difference is small for weakly hydrated proteins such as ovalbumin. Consequently inasmuch as the distribution of hydration is unknown, the exact meaning of the numerical results obtained in this way remains somewhat uncertain.

(9) E. J. Cohn and J. T. Edsall, Chapter 18, "Proteins, Amino Acids and Peptides," Reinhold Publ. Corp., New York, N. Y., 1943.

(10) A. Guinier, *Ann. phys.*, **12**, 161 (1939).

(11) H. N. Ritland, P. Kaesberg and W. W. Beeman, *J. Chem. Phys.*, **18**, 1237 (1950).

As an alternative approach to this problem, we may compute the values of R corresponding to different distributions of hydration from the values $\rho = 2.6$ for diffusion and $H = 0.18$ obtained above.

If we assume that the molecular weight of ovalbumin is 44000, the volume of each hydrated ovalbumin molecule can be computed readily from the above values of ρ and H to be 6.76×10^{-20} cm.^{3} Equating this to $(4\pi/3)(2.6b)b^2$, we obtain $b = 18.4$ Å. Thus if the hydration is distributed uniformly throughout the ovalbumin molecule, we have by (39)

$$R = \left(\sqrt{\frac{2 + (2.6)^2}{5}}\right)(18.4) = 24.4 \text{ Å}.$$

On the other hand if the hydration is distributed as a thin layer of uniform thickness δ wrapped around an anhydrous protein core, we have $(4\pi/3)(2.6b - \delta)(b - \delta)^2(6.023 \times 10^{23}) = 44000(0.746 + \Delta')$ where Δ' is a positive correction term to be added to the apparent specific volume, $\bar{V}_p = 0.746$, of anhydrous ovalbumin in solution to give the "true specific volume" of the anhydrous protein. For approximate calculations, we may solve this equation for δ by neglecting Δ'. This gives $\delta = 1.65$ Å. corresponding to the above value of b. Consequently

$$R = \left(\sqrt{\frac{2 + (2.6b - \delta)^2/(b - \delta)^2}{5}}\right)(b - \delta) = 23.2 \text{ Å}.$$

Inclusion of Δ' in this computation would yield a value of R slightly larger than 23.2 Å.

The most reliable measurement on low angle X-ray scattering in ovalbumin solutions appears to be that carried out by Ritland, Kaesberg and Beeman.[11] According to these workers R is equal to 24.0 for ovalbumin. Thus if this value is correct, the actual distribution of hydration for ovalbumin must be somewhat intermediate between the two extreme cases mentioned above.

It may also be recalled that had we assumed an oblate ellipsoid as the general shape of the hydrated ovalbumin molecule, we would obtain the values $H = 0.13(1/\rho) = 2.9$ instead of the values used above. The value of R computed from these values of H and ρ is 23.0 Å. for uniform distribution of hydration throughout the ovalbumin molecule and 22.2 Å. for completely external hydration. If the experimental value $R = 24.0$ Å. can be trusted to within $\pm 5\%$, we may take the deviations of these calculated values from the experimental value as evidence against the oblate ellipsoid as a possible shape for ovalbumin molecules.

Comparison with Dielectric Absorption Data

By assuming that the measured dielectric relaxation times are due to the rotations of rigid ellipsoidal molecules with permanent dipole moments, Oncley[12] estimated that for ovalbumin the axial ratio lies between 3 and 7, and the hydration lies between 0.08 and 0.26. These values are not conspicuously inconsistent with the values obtained in this work. However, it is difficult to decide whether this comparison can be used as evidence in favor of Oncley's interpretation of dielectric relaxation, because preliminary computations based on the Kirkwood–Shumaker theory yield relaxation times of the same order of magnitude.[13]

Valuable information on the hydration of proteins recently has been obtained by Buchanan and co-workers[14] from dielectric measurements at high frequencies at which the measured dielectric absorption is entirely due to the rotation of the water molecules. Estimates on both the amount of "irrotationally bound water" and total hydration were made by these workers as a function of axial ratio, and the possible values so estimated were presented in the form of diagrams for several proteins. For example, if we accept $\rho = 2.6$ determined in this work for ovalbumin, the possible values predicted by their diagram lie between 0 and 0.1 for "irrotationally bound water" and between 0 and 0.28 for total hydration. The latter range of possible values is not inconsistent with the present result $H = 0.18 \pm 0.01$ g. of water per g. of anhydrous ovalbumin.

Examination of the Axial Ratios Obtained from Viscosity Data

The axial ratio of the equivalent ellipsoid for ovalbumin has been computed from the viscosity data of Polson to be 4.4 and 3.3 for the assumed hydration of $H = 0.1$ and $H = 0.3$, respectively.[15] Since the actual hydration is $H = 0.18$, the corresponding value of axial ratio must be about 3.9 according to the above mentioned computations. This value differs from our new value by an amount much greater than the experimental uncertainties. Indeed, examination of existing data indicates that the axial ratios for different proteins computed from viscosity measurements are often considerably greater than those from diffusion or X-ray measurements.[16] This general trend in discrepancies suggests that there may be some theoretical inadequacy in the usual interpretation of intrinsic viscosity.

A possible explanation of these discrepancies is that the effect of molecular interactions has been neglected completely in the Einstein–Simha theory of viscosity.[16]

Obviously, molecular interactions in protein solutions, if present, will contribute to the measured specific viscosity. Thus the expression for specific viscosity of a protein solution in general should be of the form

$$\frac{\eta}{\eta_0} - 1 = \nu_1\phi + \nu_2\phi^2 + \ldots\ldots + f_1\phi + f_2\phi^2 + \ldots\ldots \quad (40)$$

where the terms $\nu_1\phi$, $\nu_2\phi^2$, etc., are due to the volume effect treated in the Einstein–Simha theory, and the terms $f_1\phi$, $f_2\phi^2$, etc., are due to molecular interactions, *e.g.*, protein–ion interactions, protein–protein interactions, etc. If we neglect the molecular interaction terms, we obtain the usual expression at infinite dilution

$$\lim_{\phi \to 0} \frac{(\eta/\eta_0) - 1}{\phi} \equiv 100[\eta] / \left(\bar{V}_p + \frac{H}{d_0}\right) = \nu_1 \quad (41)$$

(12) See p. 562 of reference 9.

(13) J. G. Kirkwood, private communication.

(14) T. J. Buchanan, G. H. Haggis, J. B. Hasted and B. G. Robinson, *Proc. Roy. Soc.* (*London*), **213**, 379 (1952).

(15) See, for example, p. 692 of reference 1.

(16) See, for example, reference 1 for references to the literature.

where $[\eta]$ is the intrinsic viscosity and ν_1 is the viscosity increment of the protein. But if we retain the molecular interaction terms in (40) we obtain the new limiting expression

$$\lim_{\phi \to 0} \frac{(\eta/\eta_0) - 1}{\phi} \equiv 100[\eta] \Big/ \left(\bar{V}_p + \frac{H}{d_0} \right) = \nu_1 + f_1 \quad (42)$$

Equation 42 shows that the measured value of $100[\eta]/[\bar{V}_p + (H/d_0)]$ is in general greater than the viscosity increment ν_1 by an amount equal to the coefficient f_1 in (40). The usual heuristic explanation that the effect of molecular interactions on $[\eta]$ disappears at infinite dilution is based on the implicit assumption that the contribution due to these interactions involves only terms of the type $f_n\phi^n$ with $n > 1$. There is no definite experimental support to this assumption. Indeed, for solutions of simple electrolytes we should even include an interaction term of the form $f_{1/2}\phi^{1/2}$ according to Falkenhagen.[17]

For protein solutions, $(\eta/\eta_0 - 1)/\phi$ approaches a constant value as $\phi \to 0$, consequently we have $f_{1/2} = 0$. But there seems to be no reason for us to neglect the term $f_1\phi$.

The evaluation of the coefficients f_1, f_2, etc., would require a theoretical treatment of molecular interaction in protein solutions, but the simple formal relationship considered above should suffice to show the heuristic nature of the usual interpretation of intrinsic viscosity.

While for macromolecules of a high degree of asymmetry we may expect that f_1 is indeed negligible as compared to ν_1, for molecules with low axial ratios the existing structural information obtained from viscosity data may need considerable revision in view of (42). Furthermore (42) also suggests a new experimental approach for studying molecular interactions in protein solutions.

Acknowledgments.—The author is grateful to Professor Lars Onsager for his valuable advice regarding the present work. He also wants to thank Professors H. S. Harned and J. M. Sturtevant who read and corrected the manuscript.

NEW HAVEN, CONNECTICUT

(17) See, for example, Harned and Owen, "The Physical Chemistry of Electrolytic Solutions," Reinhold Publ. Corp., New York, N. Y., 1950, p. 67.

[CONTRIBUTION NO. 1138 FROM THE DEPARTMENT OF CHEMISTRY OF YALE UNIVERSITY AND THE LABORATORY OF CELLULAR PHYSIOLOGY OF NATIONAL HEART INSTITUTE]

The Self-diffusion Coefficients of Water and Ovalbumin in Aqueous Ovalbumin Solutions at 10°

BY JUI H. WANG, CHRISTIAN B. ANFINSEN AND FRANCESCO M. POLESTRA

RECEIVED MAY 25, 1954

The self-diffusion coefficients of water and ovalbumin in aqueous ovalbumin solutions at 10° have been determined with H_2O^{18} and C^{14}-labeled ovalbumin, respectively, as the tracers.

One of the problems that interests and puzzles many experimenters engaged in sedimentation and ordinary diffusion studies on proteins is the quantitative relationship between the mobility of the protein molecules and the measured viscosity of the solution. Theoretical investigation of this problem is complicated by the inhomogeneity of the liquid phase in the direction of sedimentation or diffusion. For self-diffusion, however, there is neither net back-flow of the solvent nor variation in the activity coefficients of the components along the diffusion path. Consequently the situation becomes much simpler theoretically, and one may expect to get a better understanding of the problem by careful examination of the self-diffusion coefficients of proteins in solution. Unfortunately, no such data exist in the literature. In the present work the self-diffusion coefficients of water and ovalbumin in aqueous ovalbumin solutions at 10° have been determined. The results of similar measurements on several other protein solutions will be reported in later communications.

Experimental

Diffusion Measurements.—The improved capillary method[1] was used in the present work. The rate of stirring in the diffusion bath was so adjusted that $2\Delta l/l$ is negligible as compared to other experimental error. Consequently Δl could be neglected and the simple relationship

$$\frac{c_{av}}{c_0} = \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2\eta + 1)^2} \exp\left[-(2n + 1)^2\pi^2 Dt/4l^2\right] \quad (1)$$

could be used to compute the self-diffusion coefficient D from the measured values of t, l and c_{av}/c_0. In the present work the self-diffusion coefficients of both water and ovalbumin in aqueous ovalbumin solutions at 10.00 ± 0.01° were determined as a function of protein concentration. In the measurements on the self-diffusion of water, capillaries of about 0.002 cm.² in cross-sectional area and with length between 2 and 3 cm. were used. The diffusion time t was between 1 and 1.5 days. Under these conditions Dt/l^2 was almost always greater than 0.2 so that it is sufficiently accurate to omit all terms after the first on the right-hand side of equation 1. Thus we have

$$\frac{Dt}{l^2} = \frac{4}{\pi^2} \ln\left(\frac{8}{\pi^2} \times \frac{c_0}{c_{av}}\right) \quad (2)$$

Equation 2 was used to compute all the self-diffusion coefficients of water from the experimental data. For measurements on the self-diffusion of ovalbumin, capillaries with the same cross-sectional area as those described above but with length between 0.7 and 0.9 cm. were used. The use of these short capillaries for the protein is necessary because of the small self-diffusion coefficient of ovalbumin and the ease with which ovalbumin denatures. By using these extremely short capillaries it was found possible to complete the diffusion measurements in from six to 14 days without excessive amounts of denaturation. Despite these long diffusion times, the value of Dt/l^2 for the protein was often much less than 0.2 so that it was necessary to evaluate D from the experimental data by means of equation 1. This

(1) J. H. Wang, C. V. Robinson and I. S. Edelman, THIS JOURNAL, 75, 466 (1953).