

Estimation of Sedimentation Coefficients of Globular Proteins: An Application of Small-Angle X-Ray Scattering

THOMAS F. KUMOSINSKI AND HELMUT PESSEN

*Eastern Regional Research Center, Agricultural Research Service, U. S. Department of Agriculture,
Philadelphia, Pennsylvania 19118*

Received May 27, 1982

A new semiempirical procedure is presented which relates solution small-angle X-ray scattering parameters to sedimentation coefficients. With this method, sedimentation coefficients were calculated for a set of 20 globular macromolecules with molecular weights ranging from 1.3×10^4 to 7.0×10^6 ; all were in excellent agreement with experimental values. Best results were arrived at by obtaining: (1) the Stokes radius in Svedberg's equation by way of the scattering volume V of the macromolecule instead of the commonly used partial specific volume \bar{v} ; and (2) the structural frictional ratio $(f/f_0)_s$ from an axial ratio derived from $R_G S/V$ instead of the usual $3V/(4\pi R_G^3)$ relationship, where R_G is the radius of gyration and S is the external surface area of the molecule. This indicates that the frictional ratio is a function of the surface roughness of the macromolecule, in agreement with similar conclusions in the literature. In addition, structural parameters from the X-ray crystallographic structure are compared with those from small-angle X-ray scattering for a better insight into the contribution of hydration to the frictional coefficient.

The problems of measuring the interaction of water with biological macromolecules have long plagued biochemists. The importance of this interaction with respect to protein primary, secondary, and tertiary structure is well documented (1). Nevertheless, several basic aspects of this topic remain unresolved. Among these are a clear definition of the hydration of a protein, an understanding of the relationships between different experimental values of total hydration and the methods of measurement from which they are obtained, and whether the interactions are strong or weak.

Of the physical methods available for studying globular proteins, those based on hydrodynamics, in particular sedimentation, have been generally considered to provide the most information. For instance, sedimentation and diffusion coefficients of globular proteins are related to

the molecular weight, size, and shape, and, by way of the buoyancy term, to the hydration by the Svedberg relationship (2). However, throughout the years there has been much disagreement over the proper application of Svedberg's equation to rigid globular proteins. By the traditional method, devised by Oncley (3), a hydration range for a particular protein, corresponding to a range of plausible axial ratios, could be estimated. Various attempts at simultaneous solution of viscosity and sedimentation (or diffusion) equations have tended to give less than satisfactory results (1). It appeared that the best that could be accomplished was to assume a hydration of 0.25 g H₂O/g protein and, using an anhydrous molecular weight and a partial specific volume obtained from amino acid composition, to estimate translational frictional ratios; from these, axial ratios could be calculated by use of Perrin's equations

(4) on the assumption of either a prolate or oblate ellipsoid of revolution model. These axial ratios could then be compared only in a qualitative manner since they had no theoretical justification based on molecular structure.

Renewed interest in the frictional properties of macromolecules and, more particularly, in the contribution of hydration to the frictional coefficient, has emerged as a result of the availability of numerous X-ray crystallographic structures together with theoretical advances (5). Recently, Squire and Himmel (6) have attempted to separate the structural and the hydration contributions to the frictional ratio. The structural contribution was calculated from the unit cell parameters of the X-ray crystallographic structure. The hydration values calculated on this basis, however, were extremely high, probably because of the assumption that the protein can be modeled by a smooth-surfaced ellipsoid of revolution. Teller *et al.* (7), on the other hand, have performed a more rigorous calculation using a shell model to take into account the surface roughness (rugosity) of the molecule. These authors use the term "rugosity" without explicit definition to describe the wrinkled texture of the surface, a concept quantitated elsewhere in terms of an excess surface added to the smooth surface of the ellipsoid (8). Using the Kirkwood approximation (5, 7, 9, 10), these authors calculated frictional coefficients from the three-dimensional coordinates derived from X-ray diffraction. Agreement was reached between experimental and theoretical frictional coefficients only when a one-layer hydration shell was added to the crystallographic model. Also, they calculated theoretical frictional coefficients for several proteins, using a more rigorous theory than the Kirkwood approximation, and found agreement with experimental values when water was considered bound only to charged groups on the surface of the protein. They concluded that sedimentation of globular proteins depends upon the structural dimensions and rugosity of the protein, and on the hydration of the charged (and perhaps the polar) groups on the surface.

One problem remaining with the basic assumption is that sedimentation measures the dynamic hydrated structure, whereas X-ray crystallography observes a static structure (1); here, such factors as the amount of electrostriction due to charged groups upon crystallization of the protein, as well as the influence of protein breathing in solution, have been considered to have only negligible consequences. What is needed is a nonhydrodynamic method of analysis that measures the size and shape of the hydrated protein in solution to be compared with the sedimentation results. Such a method is small-angle X-ray scattering (SAXS)¹ (11). Although it does not yield the exact position of atoms within a protein molecule, SAXS does yield structural parameters such as the molecular weight, radius of gyration, hydrated volume, and external surface-to-total-volume ratio. It will be the purpose of this paper, therefore, to compare SAXS results with sedimentation behavior of globular proteins in an effort to link X-ray diffraction crystallographic results with solution sedimentation studies, as well as to attempt to quantitate the contribution of hydration to the frictional coefficient.

METHOD OF ANALYSIS AND SELECTION OF PROTEINS

The criteria for selection of proteins for this investigation were based, in the first place, on availability of two kinds of data in the literature: (1) the sedimentation coefficient $s_{0,w}^0$, or data which allow it to be calculated, and (2) requisite SAXS data. The latter refers to reported values of the radius of gyration and at least two other of the following three parameters: the hydrated volume V , the surface-to-volume ratio S/V (required for all proteins in the lower molecular weight range), and the axial ratio or other shape ratio, depending on the model. These are referred to hereafter as the primary parameters for the protein reported, as contrasted to the secondary parameters in a particular case, namely, those that can be derived from these if not reported independently. In addition, the proteins considered here were roughly globular, with no flexibility seen by SAXS. In the higher molecular weight region this was less important since the effect of rugosity on the frictional

¹ Abbreviation used: SAXS, small-angle X-ray scattering.

coefficient should become less significant, and finally negligible, as the size of the protein increases.

An extensive search of the literature produced no more than a total of 18 globular proteins and 2 spherical viruses that meet the above criteria (for references, see Table I). In view of the nearly three decades that the SAXS technique has been available, this is a small number. It would appear that most SAXS investigators do not determine the S/V ratio for solutions of biopolymers on account of severe experimental difficulties caused by the extremely small scattering signal from protein solutions. In fact, only 11 proteins in the data set actually have experimentally determined surface areas. Fortunately for our purpose, 10 of these proteins have molecular weights less than 100,000 and thus could be expected to show the effects of rugosity. The remaining ones have molecular weights greater than 100,000, where it could be expected that the rugosity would make relatively little contribution to the structural portion of the frictional ratio. (This assumption will be discussed further.) Shape information, however, is available for these proteins from SAXS results in the high-angle region. In fact, all the high-molecular-weight proteins have been found to be cylinders (either prolate or oblate), with the exception of β -lactoglobulin octamer and malate synthase; ~~the first has been found to be a prolate and the second an oblate ellipsoid of revolution.~~ *these are*

The form of Svedberg's equation to be used to calculate theoretical sedimentation coefficients $s_{20,w}^0$ from small-angle X-ray scattering structural parameters is (2)

$$s_{20,w}^0 = \frac{M(1 - \bar{v}\rho)}{(f/f_0)6\pi\eta Nr_0}, \quad [1a]$$

where M is the anhydrous molecular weight obtained from amino acid composition whenever possible, \bar{v} is the partial specific volume of the protein, ρ is the density and η is the viscosity of water at 20°C, and N is Avogadro's number. Here, r_0 , the Stokes radius (in centimeters) is related to the scattering volume V of the hydrated macromolecule in cubic centimeters instead of \bar{v} , by the relationship

$$r_0 = (3V/4\pi)^{1/3}, \quad [1b]$$

and the frictional ratio f/f_0 is the structural factor of the total ratio for the hydrated particle, based on a model of a prolate or oblate ellipsoid of revolution (4):

$$\frac{f}{f_0} = \frac{(p^2 - 1)^{1/2}}{p^{1/3} \ln [p + (p^2 - 1)^{1/2}]}, \quad (p > 1, \text{ prolate}), \quad [2a]$$

$$\frac{f}{f_0} = \frac{(1 - p^2)^{1/2}}{p^{1/3} \tan^{-1} [(1 - p^2)^{1/2}/p]}, \quad (p < 1, \text{ oblate}), \quad [2b]$$

where $p = a/b$, b is the equatorial radius, and a is the semiaxis of revolution of the ellipsoid. (The usage of $p = a/b$ is in agreement with that of Luzzati and co-workers (12); this p is the reciprocal of the p defined by Teller *et al.* (7).) All molecules were modeled by ellipsoids of revolution. The axial ratios p were determined from SAXS parameters by the method of Luzzati (12), with the use of either the ratio $3V/(4\pi R_G^3)$ or $R_G S/V$, where V is the volume of the macromolecule, R_G is the radius of gyration, and S is the external surface area, i.e.,

$$\frac{3V}{4\pi R_G^3} = \frac{p}{(2 + p^2)^{3/2}}, \quad (p \cong 1), \quad [3a]$$

$$R_G \frac{S}{V} = \frac{3}{2p} \left[1 + \frac{p^2}{(p^2 - 1)^{1/2}} \sin^{-1} \frac{(p^2 - 1)^{1/2}}{p} \right] \times \left(\frac{2 + p^2}{5} \right)^{1/2}, \quad (p > 1), \quad [3b]$$

or

$$R_G \frac{S}{V} = \frac{3}{2p} \left[1 + \frac{p^2}{(1 - p^2)^{1/2}} \tanh^{-1} (1 - p^2)^{1/2} \right] \times \left(\frac{2 + p^2}{5} \right)^{1/2}, \quad (p < 1). \quad [3c]$$

It should be noted that f/f_0 and r_0 are derived from solution structural parameters without any assumption regarding the contribution of hydration to the frictional ratio; also, no assumption is necessary concerning the symmetric or asymmetric placement of the water molecules, or concerning electrostriction effects, in contrast to the use of three-dimensional X-ray crystallographic structures for correlation with sedimentation data of globular proteins, where such assumptions cannot be avoided (13).

RESULTS AND DISCUSSION

Reliability of Data Bases

We first tested our set of 20 globular macromolecules against those of Squire and Himmel (6) and Teller *et al.* (7) (selected for a different purpose and according to different criteria) by use of the Svedberg relationship (2),

$$s_{20,w}^0 = M^{2/3}(1 - \bar{v}\rho) / [\bar{v}^{1/3}(3/4\pi)^{1/3}6\pi\eta N^{2/3}], \quad [4]$$

where all parameters have been previously defined. A plot of $s_{20,w}^0$ vs $(M^{2/3}/\bar{v}^{1/3}) \times (1 - \bar{v}\rho)$ is shown in Fig. 1 for all 20 macromolecules. Fitting to the 20 points

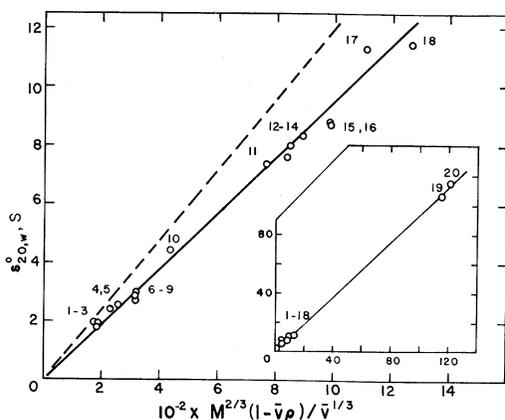


FIG. 1. Plot of $s_{20,w}^0$ vs the function $M^{2/3} (1 - \bar{v}\rho) / \bar{v}^{1/3}$ for the 18 proteins, as numbered in Table I. Solid line: linear least-squares fit, with slope 0.00943 ± 0.00013 . Dashed line: theoretical upper limit line expected for proteins considered spherical, with slope 0.0120. Inset: corresponding points and lines for 20 biopolymers, including two viruses in addition to the 18 proteins of the main figure; scales in same units; slope 0.00950 ± 0.00002 .

a least-squares straight line with zero intercept gives a slope of 0.00950 ± 0.00002 $S \text{ cm g}^{-1} \text{ mol}^{2/3}$. (The 18 proteins alone give a line of slope 0.00943 ± 0.00013 .) Also shown in Fig. 1 is the theoretical line for molecules considered as smooth spheres, which constitutes an upper limit of slope 0.0120 in the same units (7). Squire and Himmel (6) and Teller *et al.* (7) obtained slopes of 0.0108 and 0.010 for their respective sets of proteins. These values are not greatly different from ours. One may take it, therefore, that our set has approximately the same average rugosity as other globular proteins. This type of calculation, however, is purely empirical and has no structural foundation.

Use of Accessible Surface

Another approach would be to use the relationships developed by Teller (14) between accessible surface area A_s , volume V , radius R from the packing volume, and molecular weight M , derived by calculations based on the X-ray crystallographic three-dimensional structures of a set of proteins first used by Chothia (15). The

relationships are:

$$A_s = 11.12 \pm 0.16M^{2/3} \quad (\text{in } \text{\AA}^2), \quad [5a]$$

$$V = 1.273 \pm 0.006M \quad (\text{in } \text{\AA}^3), \quad [5b]$$

$$R = 0.672 \pm 0.001M^{1/3} \quad (\text{in } \text{\AA}). \quad [5c]$$

From these, axial ratios for prolate or oblate ellipsoids of revolution can be calculated by means of Eqs. [3a–c]. The molecular weight cancels out for both $3V/(4\pi R_G^3)$ (the smooth-surface model) and $R_G S/V$ (the rugose-surface model), as it must, these expressions being dimensionless. (It may be mentioned here that the smooth model is an equivalent ellipsoid of equal V , while the rugose model is an equivalent ellipsoid of equal S/V , and that the parameters derived from $R_G S/V$ thus are better adapted to account for the extra surface area attributable to rugosity (8).) Since Eq. [5c] is based on a spherical model (14), use of $3V/(4\pi R_G^3)$ will necessarily result in axial ratios of 1 for the smooth model. The information contained in A_s ($\equiv S$), however, is independent of the assumption of such a model and will, therefore, permit calculation of equivalent axial ratios from $R_G S/V$ (prolate: 3.96; oblate: 0.238), and thus frictional ratios from Eqs. [2a] and [2b] (prolate: 1.180; oblate: 1.178). Equations [1a] and [b], which use V instead of \bar{v} as a measure of the Stokes radius r_0 , then yield

$$s_{20,w}^0 = M^{2/3} (1 - \bar{v}\rho)k \times 10^{-13}, \quad [6]$$

where k , the collection of constants in Eq. [1a], becomes 0.01284 for all smooth-surface models, 0.01088 for prolate, and 0.01090 for oblate rugose-surface models.

Figure 2 is a plot of $s_{20,w}^0$ vs $M^{2/3}(1 - \bar{v}\rho)$ for our 20 macromolecules. A straight line fitted to the experimental data yields a value of 0.01079 ± 0.00003 for k in Eq. [6]. (For the 18 proteins alone, k equals 0.01043 ± 0.00014 .) Comparison between these and the above values derived from the X-ray crystallographic three-dimensional structures shows that a prolate or oblate ellipsoid of revolution with an equivalent S/V ratio (rugose-surface model) describes the hydrodynamic behavior of globular proteins to within about 1%, whereas the

smooth model is off by nearly 20%. This is in agreement with the conclusions concerning the rugosity of the surface reached by Teller *et al.* (7), who used a more exact calculation of the frictional coefficient.

Estimation of Sedimentation Coefficients from SAXS

With these considerations in mind, we turn to using SAXS results in an attempt to predict sedimentation coefficients. In Table I, the radius of gyration R_G , volume V , and surface-to-volume ratio S/V are listed for a set of 20 macromolecules. Also tabulated are the partial specific volumes \bar{v} and the anhydrous molecular weight M (obtained in most cases from the amino-acid composition), as well as indications of the geometric model which best describes the scattering particle as determined by SAXS. It should be noted, however, that experimental values of S/V are available only for proteins 1 through 11. Values for the other macromolecules had to be calculated from their smooth-surface model. As it happens, the last 10 molecules have molecular weights in excess of 100,000, so that their rugosity should make a relatively smaller contribution to the sedimentation coefficient. Axial ratios calculated for each protein from the SAXS results of Table I and Eq. [3a] for the $3V/(4\pi R_G^3)$ relationship, and Eqs. [3b] and [3c] for the $(R_G S/V)$ relationship, are given as $(a/b)_1$ and $(a/b)_2$, respectively.

For each of the first 10 proteins, $(a/b)_2$ is larger than $(a/b)_1$, although perhaps to a lesser extent as the molecular weight of the protein increases. This would be consistent with the notion that flow lines are influenced by the rugae (which presumably remain of about constant average dimensions) to a lesser extent as the volume of the particle increases. Frictional ratios for $(a/b)_1$ and $(a/b)_2$, calculated from Eqs. [2a] and [2b], are listed as $(f/f_0)_1$ and $(f/f_0)_2$, respectively. It should be noted that here the assumption is made that all proteins can be approximated by spherical, prolate, or oblate ellipsoidal models. This assumption is least exact for proteins 12, 13, 15, 16, and 18, which are more nearly cylin-

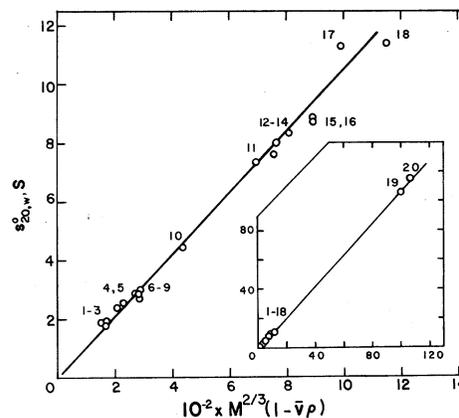


FIG. 2. Plot of $s_{20,w}^0$ vs the function $M^{2/3}(1 - \bar{v}\rho)$ for the 18 proteins, as numbered in Table I. Solid line: linear least-squares fit, with slope 0.01043 ± 0.00014 . Inset: corresponding plot and line including two viruses; scales in same units as main figure; slope 0.01079 ± 0.00003 .

ders; however, it is still a useful approximation which has been considered reasonable by other investigators (48). From the molecular weights, partial specific volumes, and the frictional ratios, one can obtain sedimentation coefficients for the smooth-surface (s_1) and rugose-surface (s_2) models by means of Svedberg's equation (Eq. [1a]), with the Stokes radius calculated from the scattering volume listed in the table. These values as well as the experimentally determined $s_{20,w}^0$ are given in the table.

It is seen that whereas s_1 values are consistently larger than $s_{20,w}^0$, s_2 generally is very close to $s_{20,w}^0$, in agreement with Teller's conclusion that the hydrodynamic behavior of proteins is influenced by the rugose-accessible surface area. The agreement between s_2 and $s_{20,w}^0$ is particularly remarkable for the holo and apo forms of several proteins in this data set, viz., riboflavin-binding protein and glyceraldehyde-3-phosphate dehydrogenase. In these cases, $s_{20,w}^0$ values change owing to some configurational change in the protein, and the calculated s_2 values evidently follow these changes very closely.

It should be noted that in cases 12-20 the differences, if any, between $(a/b)_1$ and $(a/b)_2$ (and consequently between $(f/f_0)_1$

TABLE I
STRUCTURAL AND HYDRODYNAMIC PARAMETERS FROM SAXS

| Macromolecule ^a | Model ^b | Auxiliary parameters | | SAXS parameters | | | Calculated from SAXS | | Observed | From V and \bar{v} |
|---|--------------------|--|------|--|---|-----------------------------|--------------------------------|----------------------------|---------------------|------------------------|
| | | M^c \bar{v}^d (ml/g) | | R_G (Å) | V (Å ³) S/V (Å ⁻¹) | $(a/b)_1^e$ $(a/b)_2^f$ | $(f/f_0)_1^g$ $(f/f_0)_2^h$ | s_1^i (S) s_2^j (S) | $s_{20,w}^0$ (S) | A_1^k (g/g) |
| 1. Ribonuclease (bovine pancreas) (16) | PE | 13,690 ^t 0.696 (18) | 14.8 | 22,000 0.29 | 1.87 3.69 | 1.036 1.161 | 2.03 1.81 | 1.78 (19) | 0.272 | |
| 2. Lysozyme (hen's egg white) (16) | PE | 14,310 ^m 0.702 (18) | 14.3 | 24,200 0.25 | 1.42 2.92 | 1.011 1.107 | 2.07 1.89 | 1.91 (20) | 0.363 | |
| 3. α -Lactalbumin (bovine pancreas) (16) | PE | 14,180 ⁿ 0.704 (18) | 14.5 | 25,100 0.24 | 1.43 2.81 | 1.012 1.099 | 2.02 1.86 | 1.92 (21) | 0.362 | |
| 4. α -Chymotrypsin (bovine pancreas) (22) | PE | 22,000 ^o 0.736 | 18.0 | 37,170 ^p 0.157 | 2.0 2.02 | 1.044 1.045 | 2.36 2.36 | 2.40 (23, 24) | 0.282 | |
| 5. Chymotrypsinogen A (bovine pancreas) (22) | PE | 25,000 ^o 0.736 | 18.1 | 37,790 ^p 0.160 | 2.0 2.12 | 1.044 1.051 | 2.67 2.65 | 2.58 (25) | 0.175 | |
| 6. Pepsin ^q (26) | PE | 34,160 0.725 (27) (28) | 20.5 | 54,870 ^p 0.26 | 2.0 4.76 | 1.044 1.234 | 3.36 2.84 | 2.88 (29) | 0.243 | |
| 7. Riboflavin-binding protein, apo (hen's egg white) (8) | PE | 32,500 ^o 0.720 | 20.6 | 66,500 0.203 | 1.63 3.58 | 1.021 1.153 | 3.12 2.76 | 2.76 | 0.513 | |
| 8. Riboflavin-binding protein, holo (hen's egg white) (8) | PE | 32,500 ^o 0.720 ^r | 19.8 | 55,600 0.213 | 1.76 3.62 | 1.029 1.156 | 3.28 2.92 | 2.92 | 0.311 | |
| 9. β -Lactoglobulin A dimer (bovine milk) (30) | PE | 36,730 ^o 0.751 (18) | 21.6 | 60,250 ^t 0.166 ^t | 2.13 2.93 | 1.052 1.108 | 3.12 2.99 | 2.87 (31) | 0.237 | |
| 10. Bovine serum albumin (32) | PE | 66,300 ^u 0.735 (18) | 27.2 | 97,000 0.175 | 2.55 4.18 | 1.081 1.194 | 5.01 4.54 | 4.53 ^v | 0.146 | |
| 11. β -Lactoglobulin A octamer (bovine milk) (30) | OE | 146,940 0.751 ^w | 34.4 | 215,000 ^t 0.125 ^t | 0.347 0.255 | 1.097 1.162 | 7.89 7.45 | 7.38 (31) | 0.130 | |
| 12. Glyceraldehyde-3-phosphate dehydrogenase, apo (bakers' yeast) (34) | PC | 142,870 ^x 0.745 | 32.1 | 264,200 0.0907 ^p | 1.51 2.14 ^y | 1.015 1.053 ^y | 7.93 7.64 ^y | 7.60 (35) | 0.369 | |
| 13. Glyceraldehyde-3-phosphate dehydrogenase, holo (bakers' yeast) (34) | PC | 145,520 ^x 0.745 ^r | 31.7 | 245,500 0.0929 ^p | 1.60 2.18 ^y | 1.020 1.055 ^y | 8.24 7.96 ^y | 8.07 (35) | 0.271 | |
| 14. Malate synthase (bakers' yeast) (36) | OE | 170,000 ^o 0.735 | 39.6 | 338,000 0.0843 ^p | 0.363 — | 1.089 — | 8.40 — | 8.25 (37) | 0.463 | |
| 15. Pyruvate kinase, apo (brewers' yeast) (38) | OEC | 190,800 ^o 0.734 (39) | 43.5 | 406,000 0.0879 ^p | 0.321 0.298 | 1.112 1.127 | 8.70 8.62 | 8.70 (39) | 0.548 | |
| 16. Pyruvate kinase, holo (brewers' yeast) (38) | OEC | 192,160 ^o 0.734 ^r | 42.5 | 406,000 0.0855 ^p | 0.349 0.320 | 1.096 1.113 | 8.92 8.80 | 8.81 (39) | 0.539 | |

| | | | | | | | | | | |
|---|----|---------------------------------|--------------|------------------|---|--------------|----------------|----------------|-----------|-------|
| 17. Catalase (bovine liver) (40) | PC | 248,000 0.730 | (41) (18) | 39.8 | 420,000 0.0752 ^p | 1.91 2.24 | 1.038 1.060 | 12.20 11.96 | 11.3 (41) | 0.290 |
| 18. Glutamate dehydrogenase (bovine liver) (42) | PC | 312,000 0.749 | (43) (43) | 47.0 | 668,000 0.0648 ^p | 1.98 2.30 | 1.043 1.064 | 12.18 11.93 | 11.4 (43) | 0.541 |
| 19. Turnip yellow mosaic virus (44) | S | 4.97 × 10 ⁶ 0.666 | (45) (45) | 118 ^p | 11.49 × 10 ^{6p} 0.0214 ^p | 1.0 — | 1.0 — | 104 — | 106 (45) | 0.727 |
| 20. Southern bean mosaic virus (46) | S | 6.63 × 10 ⁶ 0.696 | (47) (47) | 111 ^p | 12.25 × 10 ^{6p} 0.0210 ^p | 1.0 — | 1.0 — | 124 — | 115 (47) | 0.417 |

^a Numbers in parentheses following entries indicate references. Tabulated data were taken from the references thus designated in this column, unless noted otherwise for a particular parameter.

^b Geometric model used to describe scattering particle: PE, prolate ellipsoid; OE, oblate ellipsoid; PC, prolate cylinder; OEC, oblate elliptical cylinder; S, sphere.

^c Molecular weights, by preference, were based on amino acid compositions and sequences wherever available, except in some cases where the cited authors' values appeared more reliable or consistent with the other parameters under the conditions of measurement.

^d Partial specific volumes were the cited authors' values or, in some cases, more accurate values found in the literature. Corrections for temperature differences between 25 and 20°C were not in general made for \bar{v} because resulting differences in $s_{20,w}^0$ are minimal and do not affect comparisons between the different s values.

^e From Eq. [3a]. Prolate or oblate cylinders were modeled as equivalent prolate or oblate ellipsoids, respectively.

^f From Eq. [3b] or [3c]. Cylinders modeled as in footnote *e*.

^g From Eq. [2a] or [2b], based on Eq. [3a].

^h From Eq. [2a], based on Eq. [3b]; or Eq. [2b], based on Eq. [3c].

ⁱ From Eqs. [1a] and [1b], based on Eq. [3a].

^j From Eqs. [1a] and [1b], based on Eq. [3b] or [3c].

^k From Eq. [9c].

^l From Dayhoff (17), p. D-130.

^m From Dayhoff (17), p. D-138.

ⁿ From Dayhoff (17), p. D-136.

^o Value reported by cited authors (see footnote *c*). For 15, molecular weight calculated from value for subunits by same authors.

^p Secondary parameter, calculated with use of indicated model from values of primary parameters of cited authors (see Selection of Proteins).

^q Origin of preparation not stated.

^r Value for apoenzyme used, since \bar{v} for holoenzyme was not available.

^s From Dayhoff (17), Suppl. 1 (1973), p. S-83.

^t Unpublished data of authors of Ref. (29) (S. N. Timasheff, personal communication).

^u From Dayhoff (17), Suppl. 2 (1976), p. 267.

^v From Baldwin (33), corrected to 20°C following Pedersen (2), p. 36, and Appendices I-III.

^w Value for 9 (dimer) used, since \bar{v} for octamer not available.

^x From Dayhoff (17), pp. D-147-D-148.

^y On basis of solid cylinder; hollow core of model neglected.

^z Calculated from value for apoenzyme.

^{z'} Molecular weight calculated from value for subunits reported by Bischofberger *et al.* (39).

and $(f/f_0)_2$, and between s_1 and s_2) are not due to rugosity. Clearly, in the absence of experimental S/V values the rugosity could not be taken into account. In cases 14, 19, and 20, S/V was calculated from models of smooth ellipsoids or spheres, so that the information content of $R_G S/V$ must be identical to that of $3V/(4\pi R_G^3)$, and only one axial ratio and one s is calculated and listed (designated here as s_1 , since the designation s_2 would incorrectly imply that an independent S/V was involved). In cases of other smooth bodies, such as cylinders (12, 13, 15–18), there will be a difference between $(a/b)_1$ and $(a/b)_2$, and thus between s_1 and s_2 , since these bodies have been represented by ellipsoids of equal V or equal S/V , for which the frictional ratios can be readily calculated by means of Perrin's equations. These differences will not, however, reflect surface rugosity, but the excess surface due to difference in model (elsewhere (8) termed S_B , the excess surface due to body shape other than ellipsoidal, as distinguished from S_X , the additional contribution to surface area due to rugose texture). To the extent that this additional surface affects hydrodynamic properties, s_2 in these cases also should afford the better estimate of $s_{20,w}^0$.

In a few instances, the agreement between s_2 and $s_{20,w}^0$, while still satisfactory, is less striking than in the majority of the cases. In 4 and 5, the molecular weights reported by the authors were substantially lower than values from known amino acid composition, so that the possibility of partial autolysis cannot be excluded, with unknown consequences for the SAXS values. In 9, we are dealing with a known dimer, which might be more accurately represented by an elongated, rounded cylinder than by a prolate ellipsoid. Altogether, however, the agreement shown in Table I is all the more remarkable when it is borne in mind that these SAXS data, compiled from scattered and sometimes fragmentary sources ranging over a period of nearly

three decades, were obtained by a variety of observers, of varying familiarity with the technique, using different instruments of several different types, and different methods of data evaluation.

Solvation Effects

Up to this point, we have made no assumptions concerning the hydration of the protein. We have merely used SAXS results, which implicitly contain the hydration term, to calculate sedimentation constants. In order to deal with problems of hydration, we use a multicomponent expression for the sedimentation coefficient adapted from Schachman (49):

$$\mu_{123} = \mu_2 + (km_1 + \alpha)\mu_1 + km_3\mu_3, \quad [7a]$$

where μ_{123} is the total chemical potential of the sedimenting unit containing component 1 (water), 2 (macromolecule), and 3 (salt); μ_i and m_i are the chemical potential and the molality of component $i = 1, 2, 3$; k is a proportionality constant equal to the ratio of the fraction of salt bound to the molality of the protein; and α is the preferential hydration of the protein, i.e., the hydration beyond that corresponding to the bulk ratio of water to salt, in moles water bound preferentially per mole protein. It is readily seen that there are $(km_1 + \alpha)$ moles of water and km_3 moles of salt per mole of protein bound to the macromolecule. We have used this expression since now the term α can be related to the preferential salt binding $(\partial m_3/\partial m_2)_\mu$, used in investigation with other experimental techniques, by the expression

$$\alpha = -\left(\frac{m_1}{m_3}\right)\left(\frac{\partial m_3}{\partial m_2}\right)_\mu. \quad [7b]$$

Differentiating Eq. [7a] at constant temperature with respect to pressure and combining the result with the transport equation, Eq. [1a], with due regard for the makeup of the sedimenting unit, gives

$$s_{20,w}^0 = \frac{M_2(1 - \bar{v}_2\rho) + \alpha M_1(1 - \bar{v}_1\rho) + km_1 M_1(1 - \bar{v}_1\rho) + km_3 M_3(1 - \bar{v}_3\rho)}{f_{123}N}, \quad [8a]$$

where M_i and \bar{v} are the molecular weight and partial specific volume of component $i = 1, 2$, and 3 , respectively, and f_{123} is the frictional coefficient of the sedimenting unit. But sedimentation coefficients regularly are extrapolated to zero protein concentration and the experiments were performed in 0.1 M salt solutions, hence, $\bar{v}_1 \cong 1$, $\rho \cong 1$, $m_3 \ll 1$, and the second, third, and fourth term in the above expression should be negligibly small, so that

$$s_{20,W}^0 \cong \frac{M_2(1 - \bar{v}_2\rho)}{f_{123}N} \quad [8b]$$

From this, it may appear at first that the solvation of the protein should have no effect on the sedimentation coefficient. This, however, would be losing sight of the variability of the term f_{123} , which indeed has been commonly neglected in the context of sedimentation coefficients, even though the variability of the analogous term f_{12} has long been acknowledged (3, 48). As a matter of principle it would, therefore, be an error not to consider this effect here. Since in our approach we evaluate f_{123} from

$$f_{123} = (f/f_0)_{123} 6\pi\eta N(r_0)_{123}, \quad [8c]$$

where $(r_0)_{123}$, the Stokes radius of the sedimenting unit, equals $(3V_{123}/4\pi)^{1/3}$, there are two terms that depend upon the binding of salt and water to the protein, namely, $(f/f_0)_{123}$ and V_{123} , the hydrated volume from SAXS. The binding contribution to the term $(f/f_0)_{123}$ cannot be obtained from the present study, but V_{123} is readily related to the binding as follows. Differentiating Eq. [7a] as before (49), rearranging, replacing the molal units by concentrations in grams per gram of water, and noting that the partial specific volume of the sedimenting unit $\bar{v}_{123} = V_{123}N/M_{123}$, yields

$$\begin{aligned} V_{123}N/M_2 \\ = \bar{v}_2 + (k'g_1 + \xi_1)\bar{v}_1 + k'g_3\bar{v}_3, \quad [9a] \end{aligned}$$

where g_i is the concentration of component i in grams per gram of water (i.e., $g_1 \equiv 1$); $k' = 1000 k/M_2$ equals the amount of component 1 or 3 bound in proportion to its

concentration in the bulk of the solution, in grams of component so bound per gram of protein; and $\xi_1 = \alpha M_1/M_2$ is the preferential binding, in grams of water so bound per gram of protein. As before, under the conditions of these studies $\bar{v}_1 \cong 1$, $g_1 \equiv 1$ and, in 0.1 M salt solution, $g_3 \cong 0.006 \ll 1$, and Eq. [9a] then reduces to

$$V_{123}N/M_2 \cong \bar{v}_2 + k' + \xi_1. \quad [9b]$$

But the sum of the hydration proportional to bulk concentration, $k'g_1$, and the preferential hydration ξ_1 , by the definition of the latter, equals the total hydration A_1 (i.e., $\xi_1 \equiv A_1 - k'g_1$; cf. Ref. (50), where A_3/g_3 evidently is identical to our k'). Thus,

$$V_{123}N/M_2 \cong \bar{v}_2 + A_1, \quad [9c]$$

and it is seen that it is essentially the total hydration which lowers $s_{20,W}^0$ by way of the hydrated-volume term in Eq. [8c]. The effect of salt binding in this respect is negligible as long as salt concentrations are of the order indicated above. In solutions of high salt concentration, or even moderate concentration when salt binding is strong (i.e., when the preferential salt binding is positive and the preferential hydration in consequence is negative (cf. Eq. [7b])), salt will contribute to the solvated volume by way of the third term in Eq. [9a]. Apart from this, salt enters into Eqs. [8b] and [8c] only through its effect on $(f/f_0)_{123}$, as mentioned above, with lowered $s_{20,W}^0$ again the likely result.

In the last column of Table I are listed the values of A_1 calculated from the SAXS volume. Here, the first 9 proteins have an average value of 0.283 g of components 1 and 3 per gram of protein, in surprising accord with the previous generally assumed value of 0.25 g water/g protein (3). The last 10, which have higher molecular weights ($>100,000$) and are actually oligomeric structures, have a higher value of 0.430 g of components 1 and 3 per gram of protein. This higher value might be expected since the phenomenon of trapped solvent (internal solvation) has been observed by other investigators in such multisubunit structures as casein micelles, vi-

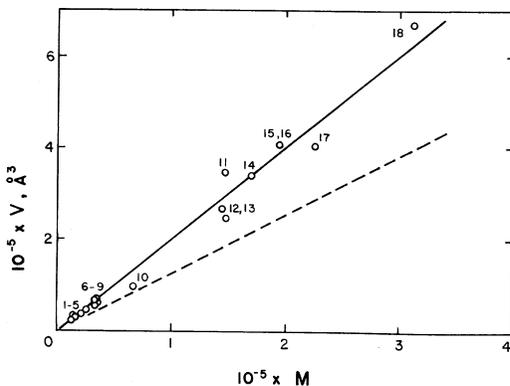


FIG. 3. Plot of scattering volume vs molecular weight for the 18 proteins in Table I. Solid line (SAXS): linear least-squares fit, with slope 1.95 ± 0.05 . Dashed line: from X-ray crystallographic structure data (cf. Eq. [5b]), with slope 1.27.

rules, and aspartate transcarbamylase (1).

With regard to the solvation effects on $(f/f_0)_{123}$, while it is not possible from the present study to determine exactly where the solvent (water and salt) binding sites on a protein are located, it is possible to compare the (f/f_0) evaluated from X-ray diffraction data with the $(f/f_0)_{123}$ from SAXS.

Structural Comparisons of SAXS and X-Ray Diffraction

A comparison of the volumes derived from SAXS with theoretical volumes derived from the X-ray diffraction three-dimensional structure according to Teller (7) is shown in Fig. 3. The SAXS solution volume is seen to be consistently higher than the volume from the X-ray crystallographic structure. Fitting a least-squares straight line with zero intercept to the SAXS volume as a function of molecular weight gives a slope of 1.95 ± 0.05 , while the corresponding slope for the diffraction data is 1.27 (7). Further, the SAXS surface area (Fig. 4) can be compared with the accessible surface area according to Teller (14). Here, the SAXS surface area is slightly lower, and fitting a straight line to the data as a function of $M^{2/3}$ gives a slope of 9.42 ± 0.136 , while Teller's value

is 11.12. (It may be added that each of the above calculations was also attempted with a polynomial of degree 2, i.e., with extra terms in M^2 for the volume and $M^{4/3}$ for the surface area, but the extra terms were found to result in no statistically significant differences.) The volume of a protein in solution, therefore, is found to be larger than the volume from the X-ray crystallographic results, whereas the surface area is slightly lower than the crystallographic accessible surface area. The increase in volume can be expected owing to solvation effects; other factors being equal, such an increase would be expected also to yield a correspondingly increased surface area. The decreased surface area observed appears to indicate that the binding of solvent to the macromolecule results in less asymmetry, less rugosity, or a combination of both. In fact, the binding sites should lie within some of the rugae or deeper clefts or grooves of the macromolecule. Calculation of $(a/b)_2$ from the fitted SAXS results (i.e., $A_s = 9.42M^{2/3}$ and $V = 1.95M$), along with the spherical assumption used for X-ray crystallographic data (i.e., $R_G = (3/5)^{1/2}(3V/4\pi)^{1/3}$), yields an average axial ratio for a prolate ellipsoid of revolution of 2, as compared with 3.96 from the X-ray diffraction results. Although this calculation cannot be entirely

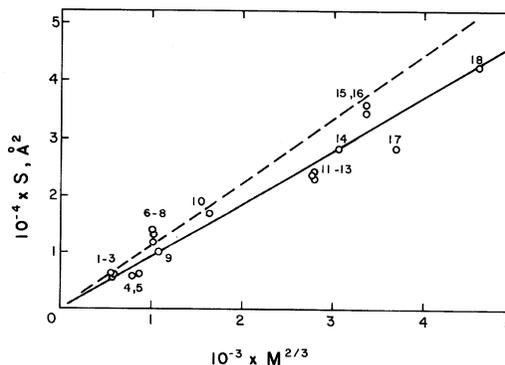


FIG. 4. Plot of surface area from small-angle X-ray scattering vs $2/3$ power of molecular weight for the 18 proteins in Table I. Solid line (SAXS): linear least-squares fit, with slope 9.42 ± 0.30 . Dashed line (XRD): accessible surface area computed from three-dimensional X-ray structure (cf. Eq. [5a]), with slope 11.12.

correct since it can be seen from the table that the average smooth-surface axial ratio is 1.8 rather than 1, this is the only type of comparison available in view of the lack of literature values for R_G calculated from the X-ray crystallographic structures.

However, the above results with respect to the increase in SAXS volume over the X-ray crystallographic volume could be due also to electrostriction of the protein upon crystallization. In fact, the concept of a dynamic alteration of protein conformation in solution ("breathing") has been previously introduced (51). Whether the observed increase in volume is due to binding of solvent components or to the breathing of the macromolecule cannot be resolved here. Needed would be extensive additional studies, including sedimentation in $H_2^{17}O$ and $H_2^{18}O$ for increased solvent density, along with small-angle neutron scattering experiments in $H_2^{17}O$, since D_2O has been shown to increase hydrophobic interactions. Only such work, by investigators with access to a neutron reactor, could settle this question.

REFERENCES

- KUNTZ, I. D., AND KAUZMANN, W. (1974) *Advan. Protein Chem.* **28**, 239-345.
- SVEDBERG, T., AND PEDERSEN, K. D. (1940) *The Ultracentrifuge*, p. 22, Clarendon Press, Oxford.
- ONCLEY, J. L. (1941) *Ann. N. Y. Acad. Sci.* **41**, 121-150.
- PERRIN, F. (1936) *J. Phys. Radium* **7**, 1-11.
- GARCIA DE LA TORRE, J., AND BLOOMFIELD, V. A. (1981) *Q. Rev. Biophys.* **14**, 81-139.
- SQUIRE, P. G., AND HIMMEL, M. E. (1979) *Arch. Biochem. Biophys.* **196**, 165-177.
- TELLER, D. C., SWANSON, E., AND DE HAËN, C. (1979) in *Methods in Enzymology* (Hirs, C. H. W., and Timasheff, S. N., eds.), Vol. 61, pp. 103-124, Academic Press, New York.
- KUMOSINSKI, T. F., PESSEN, H., AND FARRELL, H., JR. (1982) *Arch. Biochem. Biophys.* **214**, 714-725.
- KIRKWOOD, J. G. (1949) *Rec. Trav. Chim. Pays-Bas* **68**, 649-660.
- KIRKWOOD, J. G. (1954) *J. Polym. Sci.* **12**, 1-14.
- PESSEN, H., KUMOSINSKI, T. F., AND TIMASHEFF, S. N. (1973) in *Methods in Enzymology* (Colowick, S. P., and Kaplan, N. O., eds.), Vol. 27D, pp. 151-209, Academic Press, New York.
- LUZZATI, V., WITZ, J., AND NICOLAIEFF, A. (1961) *J. Mol. Biol.* **3**, 367-378.
- RICHARDS, F. M. (1977) *Annu. Rev. Biophys. Bioeng.* **6**, 151-176.
- TELLER, D. C. (1976) *Nature (London)* **260**, 729-731.
- CHOTHIA, C. (1975) *Nature (London)* **254**, 304-308.
- PESSEN, H., KUMOSINSKI, T. F., AND TIMASHEFF, S. N. (1971) *J. Agr. Food Chem.* **19**, 698-702.
- DAYHOFF, M. O. (ED.) (1972) *Atlas of Protein Sequence and Structure*, Vol. 5, Nat. Biomed. Res. Found., Washington, D. C.
- LEE, J. C., AND TIMASHEFF, S. N. (1974) *Biochemistry* **13**, 257-265.
- YPHANTIS, D. A. (1959) *J. Phys. Chem.* **63**, 1742-1747.
- SOPHIANOPOULIS, A. J., RHODES, C. K., HOLCOMB, D. N., AND VAN HOLDE, K. E. (1962) *J. Biol. Chem.* **237**, 1107-1112.
- KRONMAN, M. J., AND ANDREOTTI, R. E. (1964) *Biochemistry* **3**, 1145-1160.
- KRIGBAUM, W. R., AND GODWIN, R. W. (1968) *Biochemistry* **7**, 3126-3131.
- SCHWERT, G. W. (1949) *J. Biol. Chem.* **179**, 655-664.
- SCHWERT, G. W., AND KAUFMAN, S. (1951) *J. Biol. Chem.* **190**, 807-810.
- WILCOX, P. E., KRAUT, J., WADE, R. D., AND NEURATH, H. (1957) *Biochim. Biophys. Acta* **24**, 72-78.
- VAZINA, A. A., LEDNEV, V. V., AND LEMAZHIKIN, B. K. (1966) *Biokhimiya (Moscow)* **31**, 629-633.
- RAJAGOPALAN, T. G., MOORE, S., AND STEIN, W. H. (1966) *J. Biol. Chem.* **241**, 4940-4950.
- MCMEEKIN, T. L., WILENSKY, M., AND GROVES, M. L. (1962) *Biochem. Biophys. Res. Commun.* **7**, 151-156.
- WILLIAMS, R. C., JR., AND RAJAGOPALAN, T. G. (1966) *J. Biol. Chem.* **241**, 4951-4954.
- WITZ, J., TIMASHEFF, S. N., AND LUZZATI, V. (1963) *J. Amer. Chem. Soc.* **86**, 168-173.
- KUMOSINSKI, T. F., AND TIMASHEFF, S. N. (1966) *J. Amer. Chem. Soc.* **88**, 5635-5642.
- LUZZATI, V., WITZ, J., AND NICOLAIEFF, A. (1961) *J. Mol. Biol.* **3**, 379-392.
- BALDWIN, R. L. (1957) *Biochem. J.* **35**, 503-512.
- DURCHSCHLAG, H., PUCHWEIN, G., KRATKY, O., SCHUSTER, I., AND KIRSCHNER, K. (1971) *Eur. J. Biochem.* **19**, 9-22.
- JAENICKE, R., AND GRATZER, W. B. (1969) *Eur. J. Biochem.* **10**, 158-164.
- ZIPPER, D., AND DURCHSCHLAG, H. (1978) *Eur. J. Biochem.* **87**, 85-99.
- SCHMID, G., DURCHSCHLAG, H., BIEDERMANN, G., EGGERER, H., AND JAENICKE, R. (1974) *Biochem. Biophys. Res. Commun.* **58**, 419-426.

38. MÜLLER, K., KRATKY, O., RÖSCHLAU, P., AND HESS, B. (1972) *Hoppe-Seyler's Z. Physiol. Chem.* **353**, 803-809.
39. BISCHOFBERGER, H., HESS, B., AND RÖSCHLAU, P. (1971) *Hoppe-Seyler's Z. Physiol. Chem.* **352**, 1139-1150.
40. MALMON, A. G. (1957) *Biochim. Biophys. Acta* **26**, 233-240.
41. SUMNER, J. B., AND GRALÉN, N. (1938) *J. Biol. Chem.* **125**, 33-36.
42. PILZ, I., AND SUND, H. (1971) *Eur. J. Biochem.* **20**, 561-568.
43. REISLER, E., POUYET, J., AND EISENBERG, H. (1970) *Biochemistry* **9**, 3095-3102.
44. SCHMIDT, P. W., KAESBERG, P., AND BEEMAN, W. W. (1954) *Biochim. Biophys. Acta* **14**, 1-11.
45. MARKHAM, R. (1951) *Discuss. Faraday Soc.* **11**, 221-227.
46. LEONARD, B. R., JR., ANDEREGG, J. W., SHULMAN, S., KAESBERG, P., AND BEEMAN, W. W. (1953) *Biochim. Biophys. Acta* **12**, 499-507.
47. MILLER, G. L., AND PRICE, W. C. (1946) *Arch. Biochem. Biophys.* **10**, 467-477.
48. BLOOMFIELD, V., DALTON, W. C., AND VAN HOLDE, K. E. (1967) *Biopolymers* **5**, 135-148.
49. SCHACHMAN, H. K. (1959) *Ultracentrifugation in Biochemistry*, p. 229, Academic Press, New York.
50. NA, G. C., AND TIMASHEFF, S. N. (1981) *J. Mol. Biol.* **151**, 165-178.
51. CREIGHTON, T. E. (1978) *Progr. Biophys. Mol. Biol.* **33**, 231-297.