

[11] Structural Interpretation of Hydrodynamic Measurements of Proteins in Solution through Correlations with X-Ray Data

Introduction

Analysis of protein structure by X-ray crystallography has revealed irregular surfaces consisting of many clefts, grooves, and protuberances. To understand the overall contribution of these features to protein properties, one must bear in mind that globular proteins generally carry out their biological function in aqueous solution, where they are both fully solvated and in a dynamic state, whereas crystallography observes a static structure, and one in which bound water is nearly undetectable. Significant questions from a biochemical perspective are, therefore, whether the solution structure differs from the crystallographic structure, what the possible differences are, and how they may affect biological activity. Hydrodynamic parameters, which are sensitive to surface characteristics, have been calculated from X-ray crystallographic coordinates and compared with solution values in attempts to answer such questions, but with only marginal success. There appears to be a need for structural information from other nonhydrodynamic sources for use in conjunction with hydrodynamic parameters, so that the combined results may be compared with those from X-ray crystallography. Small-angle X-ray scattering (SAXS) is a method particularly suited to meet that need.

The calculation of hydrodynamic parameters from SAXS data can afford various insights. Comparison with observed values can remove ambiguities regarding the relative contributions of shape and hydration to the frictional ratio. X-Ray data from either diffraction or SAXS can provide a measure of molecular surface roughness. Comparison of diffraction results with those from SAXS can indicate differences between a technique that gives detailed structural information, albeit without regard to hydration, and one that gives less geometric detail, but furnishes information on hydration as well as molecular shape and surface area. This chapter will deal with each of these aspects in turn to show how they can be used to advantage.

At other times it may be of value to have a means for deriving structural parameters from available hydrodynamic data which would not permit this unambiguously by traditional methods. It will be shown how from sedimentation coefficients (the best suited among hydrodynamic param-

eters for this purpose), in conjunction with general expressions and a knowledge only of molecular weights and partial specific volumes, one can obtain reasonable estimates of a number of useful characteristics. Among these are molecular surface areas, hydrated volumes, axial ratios, and radii of gyration of globular proteins, both in the native state and on undergoing structural changes.

Approaches to the Problem Area

The three hydrodynamic parameters under consideration here are obtained experimentally by the observation of flow under the influence of an applied force. The proportionality factors relating the flow rates to the respective forces are the sedimentation coefficient, s , for a gravitational field; the diffusion coefficient, D , for a concentration gradient; and the viscosity, η , for a shearing force. In solutions, these coefficients will vary with concentration. In the cases of sedimentation (the major focus of this chapter) and diffusion, quantities more characteristic of the solute are the respective parameters s^0 and D^0 , obtained by extrapolation to infinite dilution; in the case of viscosity, similar extrapolation of a derived quantity, the reduced viscosity, gives the intrinsic viscosity, $[\eta]$. Each of the first two coefficients can be used individually to characterize a macromolecule with respect to either its molecular weight or its frictional properties, provided the other of these is already known from independent measurement or is held constant. Used jointly, sedimentation and diffusion give molecular weight and frictional information simultaneously. The intrinsic viscosity is closely related to the frictional coefficient (i.e., the ratio of frictional force to relative particle velocity) but, in contrast to flexible and extended molecules, for the globular proteins which are our concern here it is not significantly related to molecular weight.

It thus becomes evident that an understanding of frictional coefficients is central to an interpretation of the three hydrodynamic parameters. Frictional coefficients, however, are not directly accessible by experiment. The experimentally accessible frictional ratio (the ratio of the frictional coefficient of the actual—hydrated and nonspherical—molecule to that of a corresponding theoretical—nonhydrated and spherical—molecule of equal dry volume) by definition combines information on two kinds of properties, not readily separated: hydration and shape (the latter in terms of anisotropy, expressed for convenience as the axial ratio of a hypothetical ellipsoid of revolution). Although neither may be obtained explicitly, a range of reasonable assumptions for one will give a range of possible values for the other. This procedure is feasible in the case of globular proteins, for which both ranges are relatively limited, and has led to the practice of balancing the relative contributions of hydration and

shape according to some particular criterion, or else of assuming an average hydration value, such as 0.2 or 0.25 g of water per gram of dry protein, in order to arrive at approximate axial ratios.¹⁻³ Results have often appeared dubious.^{4,5} Such procedures have been criticized on the basis that some of the hydration model assumptions are not appropriate and, instead, simultaneous use of parameters from sedimentation, diffusion, and viscosity measurements on the same solutions has been advocated.^{4,6,7} The improvements are still not altogether convincing. The problem may be that frictional coefficients derived from the various hydrodynamic processes are not, in principle, identical, because the different types of forces involved require different hydrodynamic models.⁵

A reason for less than satisfactory results from both of the above approaches, which is worth examining in some detail, is suggested by the accumulated X-ray crystallographic evidence relating to the surface structure of the protein molecule. The assumption of smoothness inherent in the models, like the use of ellipsoidal modeling in the first place, has been no more than a convenience, adopted because these were the only geometrical bodies for which a complete theory predicting frictional ratios was available.⁸⁻¹¹ However, the severe surface irregularities revealed by X-ray diffraction data for globular proteins are certain to affect frictional properties. Earlier calculations of structural frictional coefficients from sedimentation coefficients, based simply on unit-cell parameters from X-ray diffraction¹² and leading to excessively high hydration values, have been improved by rigorous calculations for a number of proteins based on a shell model¹³ to take into account the dependence of frictional coefficients on surface roughness, or "rugosity," i.e., wrinkledness. (This concept was introduced without explicit definition; it is quantitated elsewhere in terms of a surface area in excess of the smooth surface of the model.¹⁴)

¹ J. L. Oncley, *Ann. N.Y. Acad. Sci.* 41, 121 (1941).

² E. J. Cohn and J. T. Edsall, "Proteins, Amino Acids and Peptides," pp. 428ff. Van Nostrand-Reinhold, Princeton, New Jersey, 1943.

³ C. Tanford, "Physical Chemistry of Macromolecules," pp. 358ff. Wiley, New York, 1961.

⁴ H. Scheraga and L. Mandelkern, *J. Am. Chem. Soc.* 75, 179 (1953).

⁵ I. D. Kuntz and W. Kauzmann, *Adv. Protein Chem.* 28, 239 (1974).

⁶ M. Stern, *Biochemistry* 5, 2558 (1966).

⁷ P. G. Squire, P. Moser, and C. T. O'Konski, *Biochemistry* 7, 4261 (1968).

⁸ A. Overbeck, *J. Reine Angew. Math.* 81, 62 (1876).

⁹ D. Edwardes, *Q. J. Pure Appl. Math.* 26, 70 (1893).

¹⁰ R. O. Herzog, R. Illig, and H. Kudar, *Z. Phys. Chem., Abt. A* 167, 329 (1934).

¹¹ F. Perrin, *J. Phys. Radium* [7] 7, (1936).

¹² P. G. Squire and M. E. Himmel, *Arch. Biochem. Biophys.* 192, 165 (1979).

¹³ D. C. Teller, E. Swanson, and C. De Haën, this series, Vol. 61, p. 103.

¹⁴ T. F. Kumosinski, H. Pessen, and H. M. Farrell, Jr., *Arch. Biochem. Biophys.* 214, 714 (1982).

Using first an approximation due to Kirkwood¹⁵⁻¹⁷ and later a more rigorous theory, Teller *et al.*¹³ calculated frictional coefficients from X-ray coordinates. They found that agreement with experimental values was reached only when a single-layer hydration shell was added to the crystallographic model in the first case, and only to charged groups on the protein surface in the second.

This illustrates a further reason for disagreements: X-ray diffraction does not observe quite the same entities as do solution methods. Proteins, as mentioned, function in an environment where they are expected to be hydrated, but X-ray structural analysis can show only a fraction of even the most tightly bound water. Model calculations to derive frictional ratios directly from data-bank X-ray coordinates, therefore, can take no account of bound water, although its presence must moderate the effects of the surface irregularities. In addition, X-ray diffraction observes a static structure, disregarding protein breathing¹⁸ in solution. Yet other differences may be due to electrostriction in the crystal resulting from charged groups in the protein.

To resolve some of these ambiguities, an ideal complement to the hydrodynamic methods would be a method which (1) is not hydrodynamic, and therefore not dependent on frictional ratios; (2) independently gives hydration information, as well as structural information from which frictional ratios can be obtained; and (3) unlike X-ray diffraction, does examine proteins in solution. Small-angle X-ray scattering meets all these requirements.¹⁹ It has the additional qualification that it allows the determination of the surface-to-volume ratio, an exceedingly useful parameter in this context. A comparison of hydrodynamic coefficients obtained by SAXS with empirical values thus will provide an approach to interpreting the latter in a meaningful way.

There are two different expressions that can be used as a starting point for the calculation of axial ratios (and thus frictional ratios¹¹) of scattering-equivalent ellipsoids from SAXS data²⁰; they give different answers. Method 1 gives an estimate of the overall molecular shape, without regard to rugosity. Method 2 makes use of the surface-to-volume ratio obtained from SAXS and translates it into a hypothetical axial ratio descriptive of the surface area instead of the overall geometry of the molecule. In effect, it provides the molecular model of Method 1 with the required additional surface by stretching or flattening it (depending on whether one deals with

a prolate or an oblate ellipsoid) and arrives at a frictional ratio reflecting the extra surface presented by the rugosities. One can use either of these axial ratios (denoted 1 and 2), by way of frictional ratios derived from them, to estimate hydrodynamic coefficients, in particular sedimentation coefficients.²⁰ It turns out that those predicted from the surface-sensitive axial ratios (Method 2) are in excellent agreement with experimental sedimentation coefficients, whereas the more conventional (Method 1) axial ratios are poor predictors.²¹

Establishing Correlations between Hydrodynamic and X-Ray Data

The approach described in the following is largely adapted from the work of the present authors.²¹ Its main thrust will be directed toward the utilization of sedimentation coefficients because these are found to furnish the best correlations; diffusion and viscosity correlations will be discussed to the extent the requisite data were available.

The criterion for selection of proteins for the present purpose was the availability of two kinds of data in the literature: (1) the sedimentation coefficient $s_{20,w}^0$, or data which allow it to be calculated, and (2) requisite SAXS data. The latter refers to reported values of (a) the radius of gyration, R_G , and (b) at least two others 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 some other shape ratio, depending on the model. These SAXS data are referred to hereafter as the primary parameters for the protein reported, as contrasted to secondary parameters in a particular case, namely, those that could be derived from the primary ones if not reported independently. In addition, the proteins considered here are roughly globular, with no flexibility, as seen by SAXS.

An extensive search of the literature to date has produced a total of 19 globular proteins and 2 spherical viruses that meet the stated criteria (for references, see Table I; this differs from Table I of Ref. 21 by an additional protein, 10a, and a number of corrections and recalculated values, but the original sequence of numbers has been retained). In view of the nearly three decades that the SAXS technique has been available, this is a comparatively small number. It would appear that most SAXS investigators do not determine S/V for solutions of bipolymers on account of severe experimental difficulties caused by the extremely small scattering signal from protein solutions as well as certain instrumental limitations. In fact, only 11 proteins in the data set actually have experimentally determined surface areas. Fortunately for our purpose, 10 of these proteins

TABLE I
STRUCTURAL AND HYDRODYNAMIC PARAMETERS FROM SAXS

Macromolecule ^a	Model ^b	Auxiliary parameters		SAXS parameters			Calculated from SAXS		From sedimentation $s_{20,w}^0$, S
		M θ , ml/g ^c	R_G , Å	V , Å ³ S/V , Å ⁻¹	$(alb)_1^d$ $(alb)_2^d$	$(f/f_0)^e$ $(f/f_0)^e$	s_1, S' s_2, S'		
1. Ribonuclease (bovine pancreas) ^{1,11}	PE	13,690 ¹ 0.696 ¹	14.8	22,000 0.29	1.87 3.69	1.036 1.161	2.03 1.81	1.78 ^a	
2. Lysozyme (chicken egg white) ¹	PE	14,310 ¹¹ 0.702 ¹	14.3	24,200 0.25	1.42 2.92	1.011 1.107	2.07 1.89	1.91 ^b	
3. α -Lactalbumin (bovine milk) ¹	PE	14,180 ^a 0.704 ¹	14.5	25,100 0.24	1.43 2.81	1.012 1.099	2.02 1.86	1.92 ^b	
4. α -Chymotrypsin (bovine pancreas) ⁷	PE	22,000 ^a 0.736	18.0	37,170 ^a 0.157	2.0 2.02	1.044 1.045	2.36 2.36	2.40 ^{b,2}	
5. Chymotrypsinogen A (bovine pancreas) ⁷	PE	25,000 ^a 0.736	18.1	37,790 ^a 0.160	2.0 2.12	1.044 1.051	2.67 2.65	2.58 ¹⁰	
6. Pepsin ^{11,4}	PE	34,160 ¹² 0.725 ¹³	20.5	54,870 ^a 0.26	2.0 4.76	1.044 1.234	3.36 2.84	2.88 ¹⁴	
7. Riboflavin-binding protein, apo (pH 3.0) (chicken egg white) ¹⁵	PE	32,500 ^a 0.720	20.6	66,500 0.203	1.63 3.58	1.021 1.153	3.12 2.76	2.76	
8. Riboflavin-binding protein, holo (pH 7.0) (chicken egg white) ¹⁵	PE	32,500 ^a 0.720 ^a	19.8	55,600 0.213	1.76 3.62	1.029 1.156	3.28 2.92	2.92	
9. β -Lactoglobulin A dimer (bovine milk) ¹⁶	PE	36,730 ¹ 0.751 ¹	21.6	60,250 ¹ 0.166 ¹	2.13 2.93	1.052 1.108	3.12 2.99	2.87 ¹⁷	
10. Bovine serum albumin ¹⁸	PE	66,300 ^a 0.735 ¹	30.6	142,000 ^a 0.146	2.90 3.88	1.105 1.174	4.62 4.34	4.30 ^a	
10a. Lactate dehydrogenase, M ₄ (dogfish) ¹	OE	138,320 0.741	34.7 ¹	253,300 ¹ 0.0893 ¹	0.409	1.069	7.50	7.54 ¹⁹	
11. β -Lactoglobulin A octamer (bovine milk) ¹⁶	OE	146,940 0.751 ¹	34.4	215,000 ¹ 0.125 ¹	0.347 0.255	1.097 1.162	7.89 7.45	7.38 ¹⁷	

TABLE I (continued)

Macromolecule ^a	Model ^b	Auxiliary parameters	SAXS parameters			Calculated from SAXS		From sedimentation
		M^c \bar{v} , ml/g ^d	R_G , Å	V , Å ³ S/V , Å ⁻¹	$(alb)_1^e$ $(alb)_2^f$	$(\beta/\alpha)_1^g$ $(\beta/\alpha)_2^h$	s_1, S^i s_2, S^j	$s_{20,w}^k, S$
12. Glyceraldehyde-3-phosphate dehydrogenase, apo (bakers' yeast) ²¹	HOC	142,870 ^c 0.737 ^d	32.1	264,200 0.0995 ^e	0.636 0.389	1.018 1.078	8.15 7.70	7.6 ²³
13. Glyceraldehyde-3-phosphate dehydrogenase, holo (bakers' yeast) ²¹	HOC	145,520 ^{mm} 0.737 ^e	31.7	250,000 0.1016 ^e	0.614 0.384	1.024 1.080	8.46 7.97	8.0 ²³
14. Malate synthase (bakers' yeast) ²⁴	OE	170,000 ^{u,25} 0.735	39.6	338,000 0.0843 ^e	0.363 —	1.089 —	8.40 —	8.25 ²⁵
15. Pyruvate kinase, apo ²⁶ (brewers' yeast) ²⁶	OEC	190,800 ¹¹ 0.734 ²⁷	43.5	406,000 0.0879 ^e	0.321 0.298	1.112 1.127	8.70 8.62	8.70 ²⁷
16. Pyruvate kinase, holo ²⁶ (brewers' yeast) ²⁶	OEC	192,160 ^{mm} 0.734 ^e	42.5	406,000 0.0855 ^e	0.349 0.320	1.096 1.113	8.92 8.80	8.81 ²⁷
17. Catalase (bovine liver) ²⁸	PC	248,000 ²⁹ 0.730 ³	39.8	420,000 0.0752 ^e	1.91 2.24	1.038 1.060	12.20 11.96	11.3 ²⁹
18. Glutamate dehydrogenase (bovine liver) ³⁰	PC	312,000 ³¹ 0.749 ³¹	47.0	668,000 0.0648 ^e	1.98 2.30	1.043 1.064	12.18 11.93	11.4 ³¹
19. Turnip yellow mosaic virus ³²	S	4.97 × 10 ⁶ ³³ 0.666 ³³	108 ^e	11.49 × 10 ⁶ ^e 0.0214 ^e	1.0 —	1.0 —	104 —	106 ³³
20. Southern bean mosaic virus ³⁴	S	6.63 × 10 ⁶ ³⁵ 0.696 ³⁵	111 ^e	12.25 × 10 ⁶ ^e 0.0210 ^e	1.0 —	1.0 —	124 —	115 ³⁵

^a Superscript numerals following entries indicate references as listed below. Tabulated data were taken from the references thus designated in the first column, unless noted otherwise for a particular parameter.

^b Geometric model used to describe scattering particle: PE, prolate ellipsoid; OE, oblate ellipsoid; PC, prolate (elongated) cylinder; OEC, oblate (flattened) elliptical cylinder; HOC, hollow oblate 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° were not in general made for \bar{v} because resulting differences in $s_{20,w}$ 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). Cylinder modeled as in Note 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 Key to references:

- ¹ H. Pessen, T. F. Kumosinski, and S. N. Timasheff, *J. Agric. Food Chem.* **19**, 698 (1971).
- ² M. O. Dayhoff, ed., "Atlas of Protein Sequence and Structure," Vol. 5. Natl. Biomed. Res. Found., Washington, D.C., 1972.
- ³ J. C. Lee and S. N. Timasheff, *Biochemistry* **13**, 257 (1974).
- ⁴ D. A. Yphantis, *J. Phys. Chem.* **63**, 1742 (1959).
- ⁵ A. J. Sophianopoulos, C. K. Rhodes, D. N. Holcomb, and K. E. Van Holde, *J. Biol. Chem.* **237**, 1107 (1962).
- ⁶ M. J. Kronman and R. E. Andreotti, *Biochemistry* **3**, 1145 (1964).
- ⁷ W. R. Krigbaum and R. W. Godwin, *Biochemistry* **7**, 3126 (1968).
- ⁸ G. W. Schwert, *J. Biol. Chem.* **179**, 655 (1949).
- ⁹ G. W. Schwert and S. Kaufman, *J. Biol. Chem.* **190**, 807 (1951).
- ¹⁰ P. E. Wilcox, J. Kraut, R. D. Wade, and H. Neurath, *Biochim. Biophys. Acta* **24**, 72 (1957).
- ¹¹ A. A. Vazina, V. V. Lednev, and B. K. Lemazhikin, *Biokhimiya (Moscow)* **31**, 629 (1966).
- ¹² T. G. Rajagopalan, S. Moore, and W. H. Stein, *J. Biol. Chem.* **241**, 1000 (1966).
- ¹³ T. L. McMeekin, M. Wilensky, and M. L. Groves, *Biochem. Biophys. Res. Commun.* **7**, 151 (1962).
- ¹⁴ R. C. Williams, Jr. and T. G. Rajagopalan, *J. Biol. Chem.* **241**, 4951 (1966).
- ¹⁵ T. F. Kumosinski, H. Pessen, and H. M. Farrell, Jr., *Arch. Biochem. Biophys.* **214**, 714 (1982).
- ¹⁶ J. Witz, S. N. Timasheff, and V. Luzzati, *J. Am. Chem. Soc.* **86**, 168 (1963).
- ¹⁷ T. F. Kumosinski and S. N. Timasheff, *J. Am. Chem. Soc.* **88**, 5635 (1966).
- ¹⁸ V. Luzzati, J. Witz, and A. Nicolajeff, *J. Mol. Biol.* **3**, 379 (1961).
- ¹⁹ G. L. Miller and R. H. Golder, *Arch. Biochem. Biophys.* **36**, 249 (1952).
- ²⁰ A. Pesce, T. P. Fondy, F. Stolzenbach, F. Castillo, and N. O. Kaplan, *J. Biol. Chem.* **242**, 2151 (1967).
- ²¹ H. Durchschlag, G. Puchwein, O. Kratky, I. Schuster, and K. Kirschner, *Eur. J. Biochem.* **19**, 9 (1971).
- ²² R. Jaenicke, D. Schmid, and S. Knof, *Biochemistry* **7**, 919 (1968).
- ²³ R. Jaenicke and W. B. Gratzer, *Eur. J. Biochem.* **10**, 158 (1969).
- ²⁴ D. Zipper and H. Durchschlag, *Eur. J. Biochem.* **87**, 85 (1978).
- ²⁵ G. Schmid, H. Durchschlag, G. Biedermann, H. Eggerer, and R. Jaenicke, *Biochem. Biophys. Res. Commun.* **58**, 419 (1974).

References to TABLE I (continued)

- 26 K. Müller, O. Kratky, P. Röschlau, and B. Hess, *Hoppe-Seyler's Z. Physiol. Chem.* **353**, 803 (1972).
- 27 H. Bischofberger, B. Hess, and P. Röschlau, *Hoppe-Seyler's Z. Physiol. Chem.* **352**, 1139 (1971).
- 28 A. G. Malmom, *Biochim. Biophys. Acta* **26**, 233 (1957).
- 29 J. B. Sumner and N. Gralén, *J. Biol. Chem.* **125**, 33 (1938).
- 30 I. Pilz and H. Sund, *Eur. J. Biochem.* **20**, 561 (1971).
- 31 E. Reisler, J. Pouyet, and H. Eisenberg, *Biochemistry* **9**, 3095 (1970).
- 32 P. W. Schmidt, P. Kaesberg, and W. W. Beeman, *Biochim. Biophys. Acta* **14**, 1 (1954).
- 33 R. Markham, *Discuss. Faraday Soc.* **11**, 221 (1951).
- 34 B. R. Leonard, Jr., J. W. Anderegg, S. Shulman, P. Kaesberg, and W. W. Beeman, *Biochim. Biophys. Acta* **12**, 499 (1953).
- 35 G. L. Miller and W. C. Price, *Arch. Biochem. Biophys.* **10**, 467 (1946).
- 36 From Dayhoff (Ref. 2, this table), p. D-130.
- 37 From Dayhoff (Ref. 2, this table), p. D-138.
- 38 From Dayhoff (Ref. 2, this table), p. D-136.
- 39 Value reported by cited authors (see Note c).
- 40 Secondary parameter, calculated with use of indicated model from values of primary parameters of cited authors (see under Selection of Proteins).
- 41 Origin of preparation not stated.
- 42 Value for apoenzyme used, since \bar{v} for holoenzyme not available.
- 43 From M. O. Dayhoff, ed., "Atlas of Protein Sequence and Structure," Suppl. 1, p. S-83. Natl. Biomed. Res. Found., Washington, D.C., 1973.
- 44 Unpublished data of authors of Ref. (16), this table (S. N. Timasheff, personal communication).
- 45 From M. O. Dayhoff, ed., "Atlas of Protein Sequence and Structure," Suppl. 2, p. 267. Natl. Biomed. Res. Found., Washington, D.C., 1976.
- 46 This value of molecular volume appears to be high, as was the molecular weight of 81,200 reported by the listed authors, pointing to the possible presence of aggregation products. For s_1 and s_2 listed in the table, the amino acid sequence molecular weight was used, together with a proportionally adjusted volume of 115,940. The inconsistent use of $V = 142,000$ with $M = 66,300$ would result in $s_1 = 4.43$ and $s_2 = 4.06$.
- 47 From Miller and Golder (Ref. 19, this table). The value reported in Ref. 18, this table, is 4.1 at 0.75%. Allowing for the concentration dependence according to Ref. 19, the two values are equivalent.
- 48 Unpublished data of H. Pessen, T. F. Kumosinski, G. S. Fosmire, and S. N. Timasheff.
- 49 Value for 9 (dimer) used, since \bar{v} for octamer not available.
- 50 From Dayhoff (Ref. 2, this table), pp. D-147, D-148.
- 51 Calculated from value for apoenzyme.
- 52 The designations "apo" and "holo," although not strictly correct in this case, are used for brevity. They refer to "native" and "fructose diphosphate liganded," respectively.
- 53 Molecular weight calculated from value for subunits reported by Bischofberger *et al.* (Ref. 27, this table).

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. 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 approximated best by cylinders (either prolate or oblate), with the exception of lactate dehydrogenase, β -lactoglobulin octamer, and malate synthase; these resemble oblate ellipsoids of revolution.

Theory

For the frictional ratio we make use of a decomposition, due to Oncley,¹ of the total ratio, f/f_0 , into shape- and hydration-dependent factors, f_e/f_0 and f/f_e , respectively. The form of Svedberg's equation to be used to calculate theoretical sedimentation coefficients $s_{20,w}^0$ from SAXS structural parameter is²²

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

where the subscript "20,w" denotes reference to water at 20°, M is the anhydrous molecular weight obtained from the amino acid sequence or composition whenever possible, \bar{v} is the partial specific volume of the protein (calculated or, preferably, experimental), ρ is the density and η is the viscosity of water at 20°, and N is Avogadro's number. It should be noted that for our calculations r_0 , the Stokes radius (in cm), will be related to the scattering volume V of the hydrated macromolecule (in cm³) instead of the more customary \bar{v} , by the relationship

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

Since the scattering volume, in contrast to \bar{v} , already reflects the hydrated molecule, the corresponding frictional ratio is really f_e/f_0 , although it was written above (and for simplicity will continue to be written in the following) as f/f_0 .

The frictional ratio f/f_0 , then, is here the structural factor of the total frictional ratio for the hydrated particle. We model all molecules as prolate or oblate ellipsoids of revolution¹¹:

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

²² T. Svedberg and K. O. Pedersen, "The Ultracentrifuge," p. 22. Oxford Univ. Press (Clarendon), London and New York, 1940.

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

where p equals a/b , b is the equatorial radius, and a is the semi-axis of revolution of the ellipsoid. (The usage of $p = a/b$ is in agreement with that of Luzzati and co-workers²⁰; note that this p is the reciprocal of the p defined by Teller *et al.*¹³) The axial ratios p were determined from SAXS parameters by the method of Luzzati,²⁰ with the use of either of two dimensionless ratios, r_1 and r_2 , defined as follows (where V is the volume of the macromolecule, R_G is the radius of gyration, and S is the external surface area):

$$r_1 = \frac{3V}{4\pi R_G^3} = p / \left(\frac{p^2 + 2}{5} \right)^{3/2}, \quad (p \geq 1) \quad (3a)$$

and

$$r_2 = R_G \frac{S}{V} = \frac{3}{2p} \left[1 + \frac{p^2}{(p^2 - 1)^{1/2}} \sin^{-1} \frac{(p^2 - 1)^{1/2}}{p} \right] \left(\frac{p^2 + 2}{5} \right)^{1/2}, \quad (p > 1) \quad (3b)$$

or

$$r_2 = R_G \frac{S}{V} = \frac{3}{2p} \left[1 + \frac{p^2}{(1 - p^2)^{1/2}} \tanh^{-1} (1 - p^2)^{1/2} \right] \left(\frac{p^2 + 2}{5} \right)^{1/2}, \quad (p < 1) \quad (3c)$$

These equations incorporate the geometric relationships $V = (4/3)\pi ab^2$, $R_G = [(a^2 + 2b^2)/5]^{1/2}$, and the expressions for S in terms of a and b for prolate and oblate ellipsoids of revolution, respectively. The ratios r_1 and r_2 may be seen from the limiting case of a sphere to be subject to the constraints $r_1 \leq 2.152$ and $r_2 \geq 2.324$.

Regarding the evaluation of p from Eq. (2), when f/f_0 is known, or from Eq. (3), when r_1 or r_2 is known, it may be noted that no closed expression for p is available. For this reason the earlier literature (p. 326 of Ref. 3; 12) made use of plots of these functions (readily calculated for p as the independent variable), which then permitted graphical evaluation to a limited precision. Greater precision could be obtained from compilations of these values, involving somewhat voluminous tables or, if more compact (cf. p. 41 of Ref. 22), requiring interpolation. With the availability of computers, or even programmable desk calculators, it is a relatively simple matter to program an iterative algorithm which can rapidly evaluate p to the desired precision for any value of f/f_0 , r_1 , or r_2 within the domains of these variables.

It is to be emphasized 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.²³

Correlations between Sedimentation and X-Ray Diffraction Data

To ensure that our special selection criteria have resulted in a set of data not very different from those for globular proteins in general, we first test our set of 21 globular macromolecules against those of Squire and Himmel¹² and Teller *et al.*¹³ (selected for a different purpose and according to different criteria), as suggested by the Svedberg relationship²² for spherical molecules in the form

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

where all parameters have been previously defined. A plot of $s_{20,w}^0$ vs $M^{2/3}(1 - \bar{v}\rho)/\bar{v}^{1/3}$ is shown in Fig. 1 for all 21 macromolecules. Fitting a least-squares straight line with zero intercept to all points gives a slope of 0.00950 ± 0.00002 S cm g⁻¹ mol^{2/3}. (The 19 proteins alone give a line of slope 0.00931 ± 0.00009 .) 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, obtained from Eq. (4) by evaluation of the collection of constants.¹³ Squire and Himmel¹² and Teller *et al.*¹³ obtained the equivalents of slopes of 0.0108 and 0.010 for their respective sets of proteins. These values are not greatly different from those above. One may assume, therefore, that our set has approximately the same average rugosity as other globular proteins. This statistical correlation is purely empirical and has no structural foundation; frictional coefficients are not explicitly considered.

Frictional ratios may be introduced into this approach by means of the relationships developed by Teller²⁴ between accessible surface area A_s , packing volume V_p , radius R_p from the packing volume, and molecular weight M , which were derived by calculations based on the X-ray crystallographic structures of a set of proteins first used by Chothia.²⁵ The relationships are

²³ F. M. Richards, *Annu. Rev. Biophys. Bioeng.* 6, 151 (1977).

²⁴ D. C. Teller, *Nature (London)* 260, 729 (1976).

²⁵ C. Chothia, *Nature (London)* 254, 304 (1975).

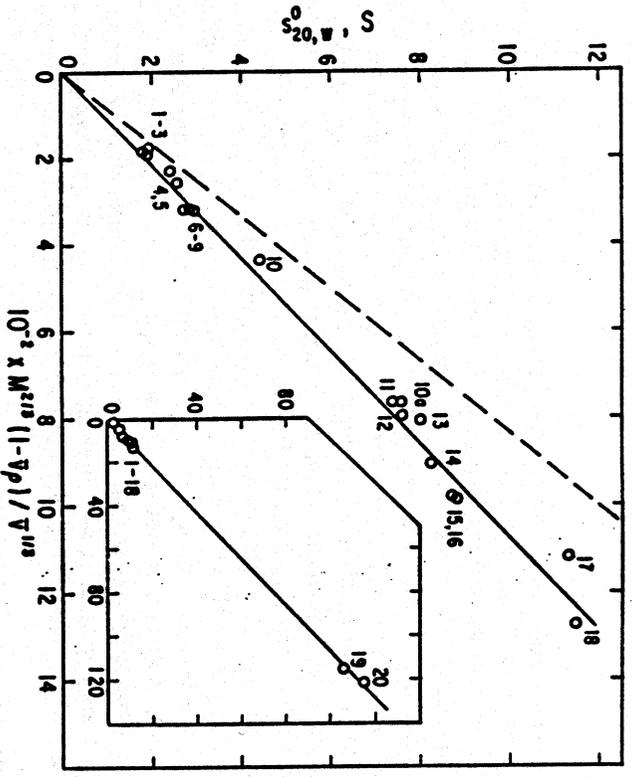


FIG. 1. Plot of $s_{20,w}^0$ vs the function $M^{2/3} (1 - \bar{v}p)/\bar{v}^{1/3}$ for the 19 proteins, as numbered in Table 1. Solid line: linear least-squares fit, with slope 0.00931 ± 0.00009 . Dashed line: theoretical upper limit line expected for proteins considered spherical, with slope 0.0120. Inset: corresponding points and lines for 21 biopolymers, including two viruses in addition to the 19 proteins of the main figure; scales in same units; slope 0.00950 ± 0.00002 . (Adapted from Fig. 1 of Ref. 21.)

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

$$V_p = 1.273 \pm 0.006 M \quad (\text{in } \text{\AA}^3) \quad (5b)$$

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

by reason of $V_p = (4/3)\pi R_p^3$. R_p is related to the radius of gyration by $R_G = (3/5)^{1/2} R_p$, as may be verified from Eq. (3a), with $p = 1$. From these expressions, axial ratios for prolate or oblate ellipsoids of revolution can be calculated by means of Eqs. (3a-c). (S and V here are represented by A_s and V_p , respectively, although it should be realized that these are rough approximations only, and that V_p , in particular, is not a hydrated volume.) 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

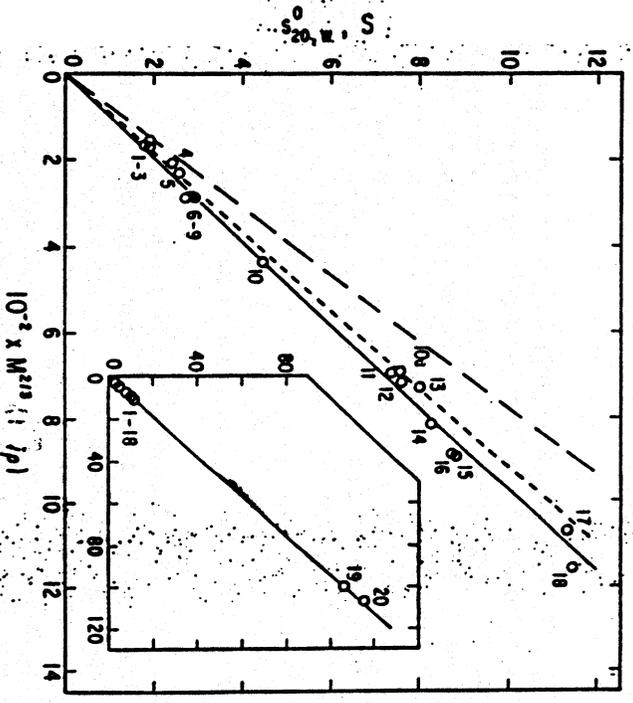


FIG. 2. Plot of $s_{20,w}^0$ vs the function $M^{2/3} (1 - \bar{v}p)$ for the 19 proteins, as numbered in Table 1. Solid line: linear least-squares fit, with slope 0.01030 ± 0.00010 ; dashed line: theoretical for smooth-surface models; dotted line: theoretical for rugose-surface models. Inset: corrected plot, and line including two viruses; scales in same units as main figure; slope 0.01079 ± 0.00003 . (Adapted from Fig. 2 of Ref. 21.)

must, these expressions being dimensionless. Since Eq. (5c) is based on spherical model,²⁴ use of $3V/(4\pi R_G^3)$ here will necessarily result in axial ratios of 1 for the smooth model. The information contained in 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,b), with V again used in place of \bar{v} , as a measure of the Stokes radius r_0 , then yield

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

where k , the collection of constants in Eq. (4), becomes 0.01284 for a smooth-surface models and about 0.0109 (prolate: 0.01088; oblate 0.01090) for rugose-surface models.

Figure 2 is a plot of $s_{20,w}^0$ vs $M^{2/3} (1 - \bar{v}p)$ for our 21 macromolecules. A straight line with zero intercept fitted to the experimental data yields a value of 0.01078 ± 0.00003 for k in Eq. (6). (For the 19 proteins alone,

equals 0.01030 ± 0.00010 .) Comparison between these experimental values for k and those above, derived from the X-ray crystallographic 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.*,¹³ who, as mentioned in the Introduction, used a more exact calculation of the frictional coefficient, with data from a different set of proteins.

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 our set of 21 macromolecules. Also tabulated are the partial specific volumes, \bar{v} , and the anhydrous molecular weights, M (obtained in most cases from the amino acid sequence or 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. 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) or (3c) (as the case may be) for the $R_G S/V$ relationship, are given as $(a/b)_1$ and $(a/b)_2$, respectively.

For each of the proteins for which both values are available, $(a/b)_2$ is larger than $(a/b)_1$ when the model is prolate, the reverse if oblate. The differences become somewhat less 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 recalled that the assumption was made that all proteins can be approximated by spherical, prolate, or oblate ellipsoidal models. This assumption is least exact for proteins 12, 13, and 15–18, which are more nearly cylinders; however, it is still a useful approximation and generally considered reasonable.¹⁸ From the molecular weights, partial specific volumes, and 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 in this equation calculated from the scatter-

ing volume listed in the table. These values as well as the experimentally determined $s_{20,w}^0$ are given in Table I.

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.¹³ 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 quite faithfully.

It must be noted that in cases 12–20 the differences, if any, between $(a/b)_1$ and $(a/b)_2$ [and consequently between $(f/f_0)_1$ and $(f/f_0)_2$, and between s_1 and s_2] are not due to rugosity since, in the absence of experimental S/V values, the rugosity could not be taken into account. Instead, S/V in cases 10a, 14, 19, and 20 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 —as the designation s_2 would incorrectly imply that an independent S/V was involved). In cases of other smooth bodies, such as cylinders (Nos. 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 , because these bodies have been represented by ellipsoids of equal volume, for the sole reason that frictional ratios for ellipsoids can be readily calculated by means of Perrin's equations. These differences will not, therefore, reflect rugosity but the excess surface due to the difference in model (elsewhere¹⁴ 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 surface 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 very satisfactory, is less striking than in the majority of the cases. In 4 and 5, the molecular weights reported by the authors were somewhat lower than values from known amino acid composition, so that the possibility of partial autoproteolysis 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, when it is considered that these SAXS data were compiled from scattered and sometimes fragmentary sources ranging over a period of nearly three decades—obtained by a variety of observers, of varying familiarity with the technique, and

TABLE II
VISCOSITY AND DIFFUSION

Protein	$[\eta]$, ml/g	$D_{20,w}^0 \times 10^7$	$(a/b)_1$	$(a/b)_{\text{visc}}$	$(a/b)_{\text{diff}}$	$(a/b)_2$
1. Ribonuclease	3.30 ^a	10.68 ⁱ	1.87	2.56	3.46	3.69
2. Lysozyme	2.5 ^b	10.4 ^j	1.42	1.55	3.41	2.92
3. α -Lactalbumin	3.01 ^c	10.57 ^c	1.43	1.82	3.21	2.81
5. Chymotrypsinogen A	2.5 ^d	9.5 ^k	2.0	1.7	2.47	2.12
6. Pepsin	3.93 ^e	9.0 ^e	2.0	3.4	1.4	4.76
9. β -Lactoglobulin dimer	3.4 ^f	7.82 ^l	2.13	2.7	3.1	2.92
10. Bovine serum albumin	3.69 ^g	6.16 ^m	2.55	3.8	3.97	4.18
17. Catalase	3.9 ^h	4.1 ⁿ	1.91	3.1	3.05	2.24

- ^a J. G. Buzzell and C. Tanford, *J. Phys. Chem.* **60**, 1204 (1956).
^b J. Léonis, *Arch. Biochem. Biophys.* **65**, 182 (1956).
^c D. B. Wetlaufer, *C. R. Trav. Lab. Carlsberg* **32**, 125 (1961).
^d C. Tanford, K. Kawahara, S. Lapanje, T. M. Hooker, Jr., M. H. Zarlengo, A. Salahuddin, K. C. Aune, and T. Tagahaki, *J. Am. Chem. Soc.* **89**, 5023 (1967).
^e A. Polson, *Kolloid-Z.* **88**, 51 (1939).
^f L. G. Bunville, Ph.D. Thesis, State University of Iowa, Ames (1959).
^g C. Tanford and J. G. Buzzell, *J. Phys. Chem.* **60**, 225 (1956).
^h R. E. Lovrien, Ph.D. Thesis, State University of Iowa, Ames (1958).
ⁱ J. M. Creeth, *J. Phys. Chem.* **62**, 66 (1958).
^j J. R. Colvin, *Can. J. Chem.* **30**, 831 (1952).
^k G. W. Schwert, *J. Biol. Chem.* **190**, 799 (1951).
^l R. Cecil and A. G. Ogston, *Biochem. J.* **44**, 33 (1949).
^m R. L. Baldwin, L. J. Gosting, J. W. Williams, and R. A. Alberty, *Discuss. Faraday Soc.* **20**, 13 (1955).
ⁿ J. B. Sumner and N. Gralén, *J. Biol. Chem.* **125**, 33 (1938).

using different instruments of several different types and different methods of data evaluation—the agreement shown in Table I is all the more remarkable.

Correlation of Diffusion Coefficients and Intrinsic Viscosities with SAXS Data

The above method for calculation of sedimentation coefficients from SAXS results may be useful in the calculation of other hydrodynamic quantities. Table II presents values of axial ratios $(a/b)_{\text{diff}}$ derived from experimental diffusion coefficients and the use of SAXS volumes from Table I, along with the smooth-surface axial ratio, $(a/b)_1$, and rugose-surface axial ratio, $(a/b)_2$, of Table I. Even though large errors frequently exist in experimental diffusion coefficients, the axial ratios derived from them are seen to be mostly closer in magnitude to the rugose-surface than

to the smooth-surface axial ratio. [In fact, some are larger than the rugose-surface axial ratio, most likely because of experimental error in the diffusion coefficient. The value for pepsin is very low, probably for the same reason: judging from its molecular weight relative to those of chymotrypsinogen A and β -lactoglobulin dimer (see Table I), the single value of the diffusion coefficient recorded for pepsin in the literature, 9.0, appears much too high; a value nearer 8.2 would be more in accord with those for the other two proteins. This would lead to 2.96 for $(a/b)_{\text{diff}}$.] The indication is that linear diffusion depends on the surface characteristics of a particle, as is the case with sedimentation. Determination of axial ratios from scattering volume and the intrinsic viscosity are less straightforward, as seen also in Table II. Here the experimentally derived axial ratio $(a/b)_{\text{visc}}$ is closer to the $(a/b)_2$ of Table I for only β -lactoglobulin A (dimer), bovine serum albumin, and catalase. The $(a/b)_{\text{visc}}$ for lysozyme, α -lactalbumin, and chymotrypsinogen A is closer to $(a/b)_1$, while ribonuclease and pepsin have $(a/b)_{\text{visc}}$ values approximately equidistant between the smooth-surface and rugose-surface axial ratios.

No clear reason for these discrepancies is apparent. However, Kuntz and Kauzmann⁵ (pp. 289–306) have also observed that hydration values derived from intrinsic viscosity and sedimentation coefficients show inconsistencies beyond those expected from experimental error. They suggest that discrepancies arise because the hydrodynamic volumes for diffusion and viscous flow are inherently different. While the results presented here do suggest such a difference, resolution of these matters will require the investigation of a larger number of proteins for which SAXS, diffusion, sedimentation, and viscosity data may become available. Meanwhile, when attempting to use sedimentation coefficients in conjunction with intrinsic viscosities to arrive at estimates of axial ratios, surface areas, or hydrated volumes for a particular protein, one should be aware of the possibility that the two methods do not actually measure identical geometric parameters. The applications section of this chapter will present an alternative procedure for estimating the contribution of hydration to the total frictional coefficient in order to obtain the contribution of the axial ratio by itself, and thus estimates of geometric parameters from sedimentation coefficients.

Structural Comparisons of SAXS and X-Ray Diffraction

A comparison of volumes from SAXS with theoretical volumes derived from the X-ray diffraction structure according to Teller²⁴ is shown in Fig. 3. The SAXS solution volume is seen to be consistently higher than the volume from the crystallographic structure. Fitting a least-squares

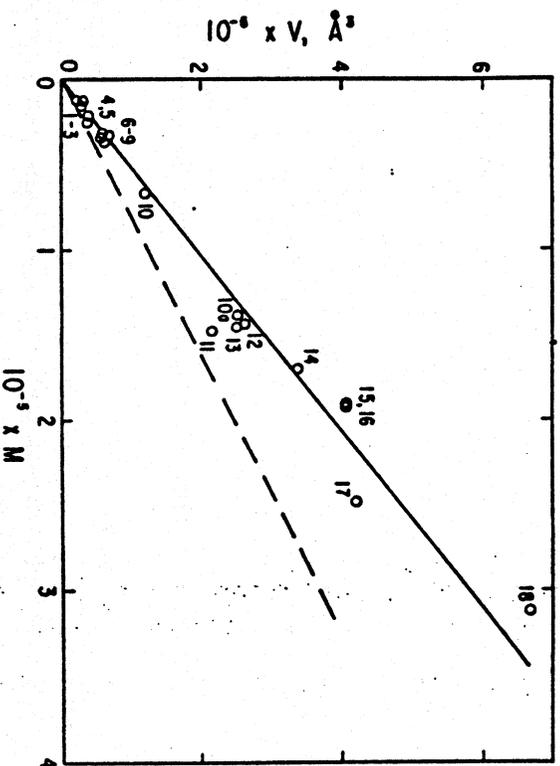


Fig. 3. Plot of scattering volume vs molecular weight for the 19 proteins in Table 1. Solid line (SAXS): linear least-squares fit, with slope 1.964 ± 0.045 . Dashed line: from X-ray crystallographic structure data [cf. Eq. (5b)], with slope 1.27. (Adapted from Fig. 3 of Ref. 21.)

straight line with zero intercept to the SAXS volume vs molecular weight plot gives a slope of 1.964 ± 0.045 , while the corresponding slope for the diffraction data is 1.27.²⁴

Further, the SAXS surface area (Fig. 4) can be compared with the accessible surface area according to Teller.²⁴ Here, the SAXS surface area is slightly lower, and fitting a straight line with zero intercept to the data as a function of $M^{2/3}$ gives a slope of 9.49 ± 0.25 , while Teller's value is 11.12. (It may be added that each of the above calculations can also be attempted with a polynomial of degree 2, i.e., with extra terms in M^2 for the volume, and in $M^{4/3}$ for the surface area, but the extra terms are found to result in no statistically significant differences.) The volume of a protein in solution from SAXS, therefore, is found to be larger than the volume from the X-ray crystallographic results, whereas the surface area in solution is slightly lower than the crystallographic accessible surface area. The increase in volume can be expected owing to solvation effects (see this volume [14]); other factors being equal, such an increase would be expected also to yield a correspondingly increased surface area. The contrary decrease in surface area actually observed appears to indicate that the binding of solvent to the macromolecule results in less anisotropy, less rugosity, or a combination of both these effects. In fact, the

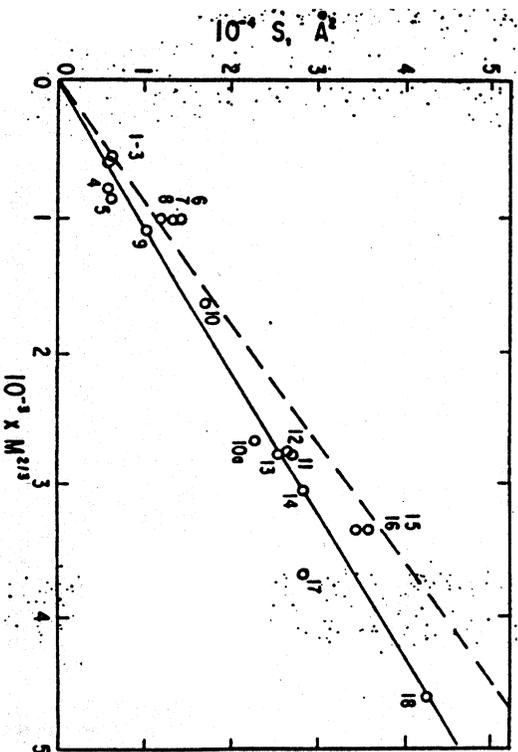


Fig. 4. Plot of surface area from small-angle X-ray scattering vs $1/3$ power of molecular weight for the 19 proteins in Table 1. Solid line (SAXS): linear least-squares fit, with slope 9.49 ± 0.25 . Dashed line (X-ray diffraction): accessible surface area computed from three-dimensional X-ray structure [cf. Eq. (5a)], with slope 11.12. (Adapted from Fig. 4 of Ref. 21.)

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., $A_2 = 9.49 M^{2/3}$ and $V = 1.964 M$), along with the spherical assumption used for X-ray crystallographic data [i.e., $R_0 = (3/5)^{1/2}(3V/4\pi)^{1/3}$], 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 correct since it can be seen from Table I 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_0 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.¹⁸ Whether the observed increase in volume is due to binding of solvent components or to the breathing of the macromolecule cannot be resolved without extensive additional studies. These would include sedimentation in $H_2^{17}O$ and $H_2^{18}O$ for increased solvent density, and small-angle neutron scattering using $H_2^{17}O$ to avoid the increased hydrophobic interactions shown to occur in D_2O .

Further Applications: Calculation of Solution Geometric Parameters

Calculations from Sedimentation Coefficients without SAXS Data (Table III)

In the absence of access to an SAXS instrument, structural parameters of monomeric proteins in solution can be computed with reasonable accuracy from their sedimentation coefficients, anhydrous molecular weights, and partial specific volumes with the aid of the data shown in Figs. 3 and 4. In these figures surface areas and volumes are shown as functions of molecular weight to the $\frac{1}{2}$ power and to the first power, respectively. Although the statistics for the linear least-square fits to the data appear to be good, the direct use of these functions to evaluate geometric parameters from individual experiments is inadvisable because it can lead to large errors.

It has been pointed out²¹ that the increase in hydrodynamic volume (which is equivalent to the SAXS volume) over the partial specific volume, \bar{v}_2 , at moderate salt concentrations is attributable to the hydration of the protein, and that monomeric proteins of molecular weight less than 100,000 have hydration values close to an average of 0.280 g H₂O/g protein, while oligomeric proteins of molecular weight greater than 100,000 have values close to an average of 0.444 g H₂O/g protein. (Recalculated from emended data of Table I of Ref. 21 for Nos. 1-6, 8-10 and for Nos. 12, 14, 15, 17, and 18, respectively.) These average values, in conjunction with the anhydrous molecular weight, yield a reasonable estimate of the particle volume from the relationship

$$V = (M/N)(\bar{v}_2 + A_1\bar{v}_1) \quad (7)$$

where M is the anhydrous molecular weight, N is Avogadro's number, \bar{v}_2 is the partial specific volume of the particle, A_1 is the hydration value, and \bar{v}_1 , the partial specific volume of bulk water, here may be taken as 1 g/ml.

For the surface area calculation for proteins of molecular weight less than 100,000 one can use the solution value from SAXS, $S = 9.49 \pm 0.25 M^{2/3}$. However, on examination of Fig. 4 it is evident that for some of these proteins the surface area might come closer to the accessible surface area value, 11.12 $M^{2/3}$. This discrepancy must be taken into account when shape factors are calculated from sedimentation values. For this reason this analysis is divided into separate procedures: one for globular monomeric proteins of molecular weight less than 100,000 (Table IIIA, where both the accessible surface area and the solution surface area must be included in the considerations), and one for oligomeric proteins of molecular weight greater than 100,000 (where only the solution surface area can be used).

For Monomers. Since monomeric proteins of molecular weight less than 100,000 generally can be represented by prolate ellipsoids of revolution with an average smooth-surface axial ratio of 1.8 (see Table I), the criterion can be adopted that any calculated smooth-surface axial ratio greatly (e.g., more than 25%) in excess of this value is *prima facie* not reasonable and therefore not to be accepted. Assuming the appropriate average of 0.280 g/g for the extent of hydration, one can estimate the hydrated volume V from \bar{v}_2 and M by Eq. (7). From the $S_{20,w}^0$, a frictional coefficient (f/f_0) can be estimated by use of Eqs. (1a) and (1b), and from this a rugose axial ratio $(a/b)_{2,k,p}$ by Eq. (2a). Substitution of this ratio for p in Eq. (3b) gives $r_z = R_G S/V$. S can be estimated either by Eq. (5a) from the accessible surface area ($S = 11.12 M^{2/3}$) or from the solution area ($S = 9.49 M^{2/3}$), and corresponding values of R_G can be calculated from Eq. (3b). Table IIIA shows the results of such calculations performed on proteins 1-10 of Table I (with the exception of apo-RBP at pH 3.0), $R_{k,p}$ and R_a being the radii of gyration thus obtained from solution surface areas (this chapter) and from accessible surface areas, respectively. To determine which radius of gyration is more nearly correct, smooth-surface axial ratios $[(a/b)_{1,k,p}]$, derived from solution surface area, and $(a/b)_{2,k,p}$, derived from accessible surface area, are evaluated from V and each R_G by Eq. (3a). Since values of $3V/(4\pi\bar{r}_G^3)$ greater than 2.15 are geometrically meaningless for ellipsoids of revolution (noted under the respective axial ratios in Table III as NO, for "not obtainable"), one chooses the R_G value for which $(a/b)_1$ is not impossible or, if both are possible, the lower value. If both $(a/b)_{1,k,p}$ and $(a/b)_{2,k,p}$ have been designated as NO, it is to be assumed that $(a/b)_{2,k,p}$ does not differ from $(a/b)_{1,k,p}$ because the molecule has a relatively smooth surface and is best represented by a radius of gyration R_{sm} , calculated from the value of $3V/(4\pi\bar{r}_G^3)$, obtained in this instance from the rugose axial ratio $p = (a/b)_{2,k,p}$ by Eq. (3a). This is the case with chymotrypsinogen and α -chymotrypsin, which both have essentially the same $(a/b)_1$ and $(a/b)_2$ values (Table I). The results of this procedure can be verified in Table III, where the calculated radii of gyration in italics are in good agreement with the experimental values from SAXS. (An exception is pepsin, whose surface area from SAXS is also much larger than that from accessible surface.) It is seen also that calculated rugose-surface $[(a/b)_{2,k,p}]$ as well as smooth-surface $[(a/b)_{1,k,p}]$ or $(a/b)_{2,k,p}$ values are in reasonable agreement with axial ratios determined experimentally by SAXS.

For Oligomers. For oligomeric proteins of molecular weight greater than 100,000, radii of gyration and smooth-surface axial ratios can be calculated as for monomers. Here, however, only the solution surface area relationship, $S = 9.49 M^{2/3}$, can be used. Whether the particle is a

TABLE III
STRUCTURAL CALCULATIONS FROM $s_{20,w}^0$

Protein	$R_{K-P}, \text{\AA}^a$ $R_{as}, \text{\AA}^b$	$R_{sm}, \text{\AA}^c$	$R_{exp}, \text{\AA}^d$	$(a/b)_{2,K-P}^e$	$(a/b)_{2,exp}^d$	$(a/b)_{1,K-P}^e$ $(a/b)_{1,as}^b$	$(a/b)_{1,exp}^d$
A. Globular proteins, $M < 100,000$							
1. Ribonuclease	18.6 15.8		14.8	3.93	3.69	3.24 2.24	1.87
2. Lysozyme	15.0 12.7		14.3	2.90	2.90	1.83 NO ^e	
3. α -Lactalbumin	13.8 11.7		14.5	2.66	2.81	1.11 NO	1.43
4. α -Chymotrypsin	13.2 11.2	17.2	18.0	1.72	2.02	NO NO	2.0
5. Chymotrypsinogen A	14.5 12.3	18.6	18.1	1.94	2.12	NO NO	2.0
6. Pepsin	27.8 23.6		20.5	4.25	4.76	3.81 2.74	2.0
8. RBP, holo and apo at pH 7.0 ^f	24.7 20.9		19.8	3.78	3.62	3.14 2.14	1.76
9. β -Lactoglobulin dimer	23.8 20.2		21.6	3.31	2.93	2.61 1.69	2.13
10. Bovine serum albumin ^g	35.2 29.8		30.6	4.27	3.88	3.88 2.81	2.90
B. Globular proteins, $M > 100,000^h$							
10a. Lactate dehydrogenase, M ₂	29.6 33.1	33.8	34.7	0.514	0.409	NO -0.650	0.409 0.347
11. β -Lactoglobulin A octamer	33.1 33.8		34.4	0.438	0.255	-0.650 0.536	0.347 0.636
12. Glyceraldehyde-3-phosphate dehydrogenase, apo	33.8 42.4		32.1	0.406	0.389	0.536 0.280 ⁱ	0.636 0.363
14. Malate synthase	42.4 49.2	41.0	39.6	0.314	0.363	0.280 ⁱ 0.195 ^j	0.363 0.321
15. Pyruvate kinase, apo	49.2 40.4	44.4	43.5	0.270	0.298	0.195 ^j 1.72	0.321 1.91
17. Catalase	40.4 57.6		39.8	2.39	2.24	1.72 3.37	1.91 1.98
18. Glutamate dehydrogenase	57.6		47.0	3.43	2.30	3.37	1.98
C. Spherical viruses							
Protein	$(a/b)^k$	$V_{K-P}, \text{\AA}^3$	$V_{exp}, \text{\AA}^3$	$R_{K-P}, \text{\AA}$	$R_{exp}, \text{\AA}$		
19. Turnip yellow mosaic virus	1.0	10.91×10^6	11.49×10^6	106.5	108		
20. Southern bean mosaic virus	1.0	15.30×10^6	12.25×10^6	119	111		

^a Subscript "K-P" refers to calculations as described under Applications.

^b Subscript "as" and corresponding figures refer to calculations based on accessible surface.

^c Subscript "sm" refers to calculations based on smooth-surface model.

^d Subscript "exp" refers to experimental values from Table I.

^e NO = "not obtainable," because corresponding $r_1 > 2.15$ (see discussion under Monomers).

^f The two forms have been found to have the same value of $s_{20,w}^0$ at pH 7.0 (see Table 14). The data under No. 7 of Table I refer to the apo form at pH 3.0, where the riboflavin is released.

^g Hydration of 0.280 assumed since BSA is monomeric.

^h For proteins in Section C of this table, which are all oligomers, an average hydration is taken as 0.444 (see section on Calculations from Sedimentation Coefficients).

ⁱ Native, as opposed to fructose diphosphate liganded.

^j Value not usable because less than $(a/b)_2$ (see discussion under Oligomers).

^k From electron microscopy.

prolate or oblate ellipsoid of revolution or a cylinder cannot be determined without SAXS experimental data, and therefore the radius of gyration $R_{k,p}$ and smooth-surface axial ratio $(a/b)_{1,k,p}$ must be estimated for both prolate and oblate ellipsoids of revolution. The volume again can be calculated from the anhydrous molecular weight, now assuming the appropriate average hydration of 0.444 g H₂O/g protein. As before, from the sedimentation coefficient, with a knowledge of the molecular weight, volume, and partial specific volume, the frictional coefficient and rugose axial ratio can be obtained by Eqs. (1) and (2). The solution surface areas derived in this chapter can be used in the expressions relating surface areas, volume, and radius of gyration to rugose axial ratio, Eqs. (3b) or (3c), to calculate the radius of gyration for either prolate or oblate models. Smooth-surface axial ratios can then be calculated from Eq. (3a). If the parameter r_1 is greater than 2.15 (geometrically meaningless) the rugose-surface axial ratio is assumed equal to the smooth-surface axial ratio, and the radius of gyration R_{sm} is now calculated from $(a/b)_{2,k,p}$ values with the expression $3V/(4\pi R_0^2)$ [Eq. (3a)]. (This is the case for No. 10a.)

Another constraint arises from the relationship between r_1 and r_2 , when it is remembered that the point in utilizing r_2 is that it, in contrast to r_1 , takes into account the excess surface area due to rugosity by translating it, for hydrodynamic purposes, into an axial ratio that is enhanced in the sense of indicating increased anisotropy.¹⁴ Therefore, $(a/b)_2$ for a prolate ellipsoid can never be smaller than $(a/b)_1$, and vice versa for an oblate ellipsoid. If the calculated values violate one of these constraints (as they do for Nos. 14 and 15), they are not usable (so indicated in the table), nor are the related $R_{k,p}$ and R_{sm} . Instead, R_{sm} calculated as under Monomers, is utilized.

Table IIIB shows the results of such calculations performed on Nos. 10a-18 of Table I, using only their shape, determined by SAXS, as listed in that table. Dogfish lactate dehydrogenase, β -lactoglobulin A octamer, and bakers' yeast glyceraldehyde-3-phosphate dehydrogenase have values of $3V/(4\pi R_0^2)$ greater than 2.15, based on the calculated $R_{k,p}$. Hence, the smooth-surface axial ratio $(a/b)_{1,k,p}$ must be assumed equal to the rugose-surface axial ratio, from which the italicized R_G is calculated. These derived radii of gyration and axial ratios, shown in Table IIIB, are in fair agreement with the experimental R_G and (a/b) from SAXS in all cases except glutamate dehydrogenase. There, neither the calculated smooth-surface and rugose-surface axial ratios nor the R_G agree with the experimentally derived SAXS values. The sedimentation coefficient predicted from SAXS in Table I for glutamate dehydrogenase, however, is quite close to the experimental value. A value of hydration larger than 0.444 for this protein²¹ is a likely cause for the disagreement between

calculated and experimental values of the parameters. It is of interest that this is the only apoprotein in the data set of Tables I and III for which such a large disagreement exists.

For Ligand-Induced Shape Changes. Three proteins listed in Table I (not counting the two virus particles, Nos. 19 and 20) have not been used in Table III. They are apo-RBP at pH 3.0 (No. 7), the holo form of glyceraldehyde-3-phosphate dehydrogenase (No. 13), and the liganded form of pyruvate kinase (No. 16). These proteins in each case have undergone some structural (either conformational or configurational) change from their respective native form, as may be seen from the structural parameters listed in Table I. Three types of structural change are conceivable: (1) a change in volume, with essentially constant axial ratio, as in the holo form of glyceraldehyde-3-phosphate dehydrogenase; (2) a change in axial ratio, with constant volume, as in the holo (liganded) form of pyruvate kinase, and (3) changes in both volume and axial ratio, as in the apo form of RBP at pH 3.0.

To determine from sedimentation velocity data in the absence of SAXS information which of these three categories would be appropriate for an unknown protein undergoing a structural rearrangement, some other accessory information is required. For this purpose, changes in the preferential hydration of the particle could be measured under both conditions (for example by density gradient ultracentrifugation in high salt solutions, by pycnometry, or by another suitable method), bearing in mind that changes in such a quantity can usually be determined with higher precision than the absolute value of the quantity itself. Assuming unchanged salt binding, changes in preferential hydration are, by the definition of this quantity, equal to changes in total hydration²⁶; the latter, in turn, are reflected in changes in hydrated particle volume by Eq. (7). Thus, in the case of glyceraldehyde-3-phosphate dehydrogenase, Sloan and Velick²⁷ found a decrease in preferential hydration of 0.075 g/g concurrent with the binding of the coenzyme nicotinamide adenine dinucleotide (NAD), as determined from sedimentation velocity and relative viscosity data. (The ready availability since then of advanced density-measuring instrumentation^{28,29} has considerably simplified the precise determination of preferential hydrations.²⁶) From Eq. (7), with $b_2 = 0.737$ ml/g and $A_1 = 0.444$ g/g (the average value), a decrease in $(b_2 + A_1)$ by 0.075 corresponds to a decrease in V of 6.35%. This compares with

* J. C. Lee, K. Gakko, and S. N. Timasheff, this series, Vol. 61, p. 26.

²⁷ D. L. Sloan and S. F. Velick, *J. Biol. Chem.* **248**, 5419 (1973).

²⁸ D. W. Kupke, in "Physical Principles and Techniques of Protein Chemistry" (S. L. Timasheff, ed.), Part C, p. 1. Academic Press, New York, 1973.

²⁹ O. Kratky, H. Leopold, and H. Stabinger, this series, Vol. 27, p. 98.

a corresponding volume decrease of 5.37% reported by Durchschlag *et al.*³⁰ for this same enzyme, although under slightly different conditions (at pH 8.5, 40°²⁷ vs pH 7.4, 25°²⁶). This was accompanied by a 0.4 Å decrease in R_G , with almost no change in axial ratio. As can be seen from Eqs. (1a) and (1b) and Table I after allowing for the change in molecular weight, the amount of volume contraction or decrease in hydration can account to within about 1% for the change in sedimentation coefficient upon binding of NAD. This treatment, therefore, may be useful in determining the contribution of volume contraction (or decrease in hydration), as well as of the change in axial ratio, to changes in sedimentation coefficients for proteins under altered environmental conditions.

If, as in the case of binding of pyruvate kinase, the hydrated volume remains constant (indicated, for instance, by unchanged preferential hydration), any change in sedimentation coefficient must be attributed to a change in axial ratio alone. (Malate synthase, No. 14, behaves similarly on binding to substrate or to an analog,³¹ but it does not furnish a good quantitative example because changes here are exceedingly subtle and literature data do not appear to be sufficiently explicit to permit the present kind of analysis.) For pyruvate kinase, evaluating V and S as indicated before, the increase in $s_{20,w}^0$ from 8.70 to 8.81 corresponds to a decrease in f/f_0 from 1.148 to 1.139 and an increase in axial ratio from 0.270 to 0.281. The corresponding dimensionless ratio r_2 ($\equiv R_G S/V$) would be required to go from 4.11 to 4.00 (r_2 being appropriate because axial ratios derived from frictional coefficients relate to the rugose surface). Assuming the surface area remains essentially constant, and with the volume also constant, the change in sedimentation coefficient therefore translates into an unambiguous surface change in radius of gyration as well. (The assumption of approximate surface constancy for proteins in excess of M_r 100,000, mentioned earlier, is verified by inspection of the pertinent data of Table I, from which it may be seen that changes in S are of the order of only 3% for both pyruvate kinase, $M \approx 190,000$, and glyceraldehyde-3-phosphate dehydrogenase, $M \approx 140,000$.) It follows that R_G should change in proportion, i.e., at a ratio of 4.11/4.00 = 1.028. This prediction, made without SAXS, agrees well with the SAXS radii of gyration of Table I, according to which this ratio is 43.5/42.25 = 1.024. To estimate R_G values individually, following the procedure described under Oligomers gives, from the above $(a/b)_{2,k-p}$ of 0.0270 and 0.281: R_{k-p} , 49.2 and 47.7; r_1 , 0.749 and 0.825; $(a/b)_{1,k-p}$, 0.194 and 0.216. Inasmuch as for oblate ellipsoids $(a/b)_1$ may not be smaller than $(a/b)_2$ (see remarks under Oligo-

mers), these values of R_{k-p} are not usable. The use of 0.270 and 0.281 for p in Eq. (3a), however, gives r_1 values of 1.012 and 1.049, from which R_{cm} of 44.4 and 44.1 are obtained as best estimates, in error by less than 4% when compared to the SAXS values of 43.5 and 42.5, respectively.

By contrast, for the apo form of riboflavin-binding protein, $M \approx 35,000$, in going from pH 7 to pH 3 the change in S (see Table I) amounts to a decrease of over 12%. This is so substantial that only the product of R_G and S can be evaluated from the rugose-surface axial ratio and the estimated hydrated volume, but not R_G by itself. SAXS measurements would need to be made to find the separate contributions of radius of gyration and surface area.

Calculations from Difference Sedimentation and SAXS Data and Electron Microscopy

The surface areas of globular proteins may be calculated from sedimentation coefficients in conjunction with SAXS volumes and R_G values. Since the Soule-Porod plots from which SAXS values for S/V are derived tend to be imprecise (requiring, preferably the use of a symmetrically scanning apparatus to determine the true z_c angle), whereas R_G and V can be much more readily determined precisely, surface areas are easier to calculate through use of Eqs. (1)-(3). Small changes in protein surface areas induced by biological processes or environmental conditions can thus be detected by means of difference sedimentation analysis in conjunction with R_G and V values from SAXS. For the most accurate results, the molecular weight used in Eq. (1) should be obtained from sequence data or, at least, from sedimentation equilibrium, rather than from SAXS.

Section C of Table III shows the predicted radius of gyration and volume from sedimentation coefficients of two virus particles, on the basis of spherical shape as revealed by electron microscopy. These calculated values of the radius of gyration are in fair agreement with the ones determined by SAXS. Hence, it is reasonable to determine the shape of a large particle with smooth surface by electron microscopy, assuming that the fixation technique used has not distorted the sample significantly. The radius of gyration and volume can then be determined accurately from the sedimentation coefficient and the axial ratio from electron microscopy.

Note Added

Since the preparation of this manuscript, the authors have been made aware of results of work not previously available to them which bears directly on the subject of this chapter (Prof. G. Damaschun, East-Berlin, personal communication). Parameters of 10 additional globular proteins

³⁰ H. Durchschlag, G. Puchwein, O. Kratky, I. Schuster, and K. Kirschner, *Eur. J. Biochem.* **19**, 9 (1971).

³¹ D. Zipper and H. Durchschlag, *Eur. J. Biochem.* **87**, 85 (1978).

and 2 small RNA molecules have been examined with respect to the relationships of Ref. 21 and, although differences between predicted and experimental sedimentation coefficients were in several instances larger than those above, the findings in general represent excellent confirmation of the semiempirical procedure described here.³² The reader may further find it of interest that Damaschun and co-workers have developed methods also of calculating hydrodynamic parameters from atomic coordinates or from SAXS many-body models^{33,34} and of estimating the thickness of solvation layers from combined SAXS and quasi-elastic light scattering.³⁵

Acknowledgment

The authors have benefited greatly from discussions with H. M. Farrell, Jr. and from the valuable advice of J. C. Lee.

- ³¹ J. J. Müller, H. Damaschun, G. Damaschun, K. Gast, P. Pletz, and D. Zirwer, *Strud. Biophys.* **102**, 171 (1984).
³² J. J. Müller, O. Glatter, D. Zirwer, and G. Damaschun, *Strud. Biophys.* **93**, 39 (1983).
³³ J. J. Müller, D. Zirwer, G. Damaschun, H. Welfe, K. Gast, and P. Pletz, *Strud. Biophys.* **96**, 103 (1983).
³⁴ K. Gast, D. Zirwer, P. Pletz, J. J. Müller, G. Damaschun, and H. Welfe, in "Physical Optics of Dynamic Phenomena and Processes in Macromolecular Systems," de Gruyter, Berlin (in press).

[12] The Use of Covolume in the Estimation of Protein Axial Ratios

By LAWRENCE W. NICHOL and DONALD J. WINZOR

For several decades attempts have been made to assess the overall geometry of protein molecules in solution by visualizing them as ellipsoids of revolution.¹⁻⁵ With the convention that a denotes the length of the semimajor axis of the ellipse and b that of the semiminor axis, rotation of the ellipse about these axes results, respectively, in prolate and oblate ellipsoids both with axial ratio $a/b \geq 1$, the limiting case where a equals b being a sphere. The ultimate aim in this context is to view the hydrated

- ¹ T. Svedberg and K. O. Pedersen, "The Ultracentrifuge," Oxford Univ. Press, London and New York, 1940.
² J. L. Oncley, *Ann. N.Y. Acad. Sci.* **41**, 121 (1941).
³ H. K. Schachman, "Ultracentrifugation in Biochemistry," Academic Press, New York, 1959.
⁴ J. T. Yang, *Adv. Protein Chem.* **16**, 323 (1961).
⁵ C. Tanford, "Physical Chemistry of Macromolecules," Wiley, New York, 1961.