Modelling biological macromolecules in solution: 1. The ellipsoid of revolution

Stephen E. Harding
Department of Biochemistry, University of Bristol, Bristol BS8 1TD, UK

and Arthur J. Rowe
Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK
(Received 14 August 1981; revised 18 November 1981)

The problems of determining the axial ratio of biological macromolecules in solution employing the ellipsoid of revolution as a model are discussed and analysed in terms of the sensitivities of the various volume-independent functions available. It is shown that over the whole range of axial ratio only the R *function (Rowe, 1977) is applicable, but the newly derived* Π *and* Λ *functions may have application to macromolecules of axial ratio* >3. *The widely employed* β *function is shown to be entirely unusable in terms of the defined criteria.*

Keywords: Mathematical models; ellipsoid model; volume independent function; sensitivities

Introduction

X-ray crystallography, where applicable, is by far the most accurate method for determining the conformation of macromolecules. Unfortunately, this technique is also the most laborious, and the calculated structures are of a static form of the macromolecule, which may be only an approximation to the dynamic structure in solution. The study of the transport properties of macromolecules in solution (hydrodynamics) yields information directly relevant to the actual structure in solution, but the interpretation of evidence from such techniques poses certain problems which several decades of investigation have not yet overcome.

There are two basic approaches for determining the gross conformation using hydrodynamic techniques. One method is to assume a structure and then to calculate its hydrodynamic properties, for example the intrinsic viscosity, sedimentation coefficient or translational diffusion coefficient, and then to see how much these predicted properties differ from the experimentally determined properties for the unknown structure. The model is then successively changed (refined) until the predicted properties converge to agree with the actual properties. This method has been developed by Bloomfield, Garcia de la Torre and coworkers[1-8]. There is, however, a serious drawback in that the final calculated structure may not be the *only* one that gives these properties.

The alternative approach is to calculate a structure directly from the known hydrodynamic properties. Some general model must, of course, be assumed, but, although the models available from this approach are less precise, the method does not suffer from the uniqueness problem. The most general, and almost universally applied, model is the ellipsoid of revolution, i.e. an ellipsoid with two equal axes[9,10]. In this paper we consider the optimal procedures for a structure calculation based upon this model, and in the second paper of this series[48] show how the restriction to two equal axes may, in appropriate cases, be relaxed.

0141-8130/82/030160-05$03.00

Theory

Volume-independent shape functions

There are several shape functions arising from the various hydrodynamic properties of a macromolecular solution. The most salient are:

(1) the viscosity increment:

$$\nu = \frac{[\eta]}{v_s} \equiv \frac{[\eta] M_r}{N_A V_e} \tag{1}$$

where $[\eta]$ is the intrinsic viscosity (ml g^{-1}), v_s the swollen specific volume (ml g^{-1}), M_r the molecular weight, V_e the volume of a macromolecule (ml) and N_A Avogadro's number.

(2) the translational frictional ratio:

$$\frac{f}{f_0} \equiv \frac{M_r(1-\bar{v}\rho_0)}{N_A 6\pi\eta_0 s}\left(\frac{4\pi}{3V_e}\right)^{1/3} \tag{2}$$

where η_0 is the solvent viscosity, ρ_0 the solvent density and s is the sedimentation coefficient (extrapolated to infinite dilution). Following the convention of Scheraga and Mandelkern[18], f_0 (and θ_0, τ_0 below) refers to a *hydrated* sphere of the same volume.

(3) the reduced molecular covolume:

$$U_{red} = \frac{U}{N_A V_e} \tag{3}$$

where U is the molecular covolume (ml mol^{-1}).

(4) the rotational diffusion ratios:

$$\frac{\theta_i}{\theta_0} = \frac{6\eta_0 V_e}{kT}\theta_i \qquad (i = a, b, c) \tag{4}$$

where a, b, c refer to the semiaxes of the ellipsoid (for a prolate ellipsoid $a > b = c$ and for an oblate $a < b = c$).

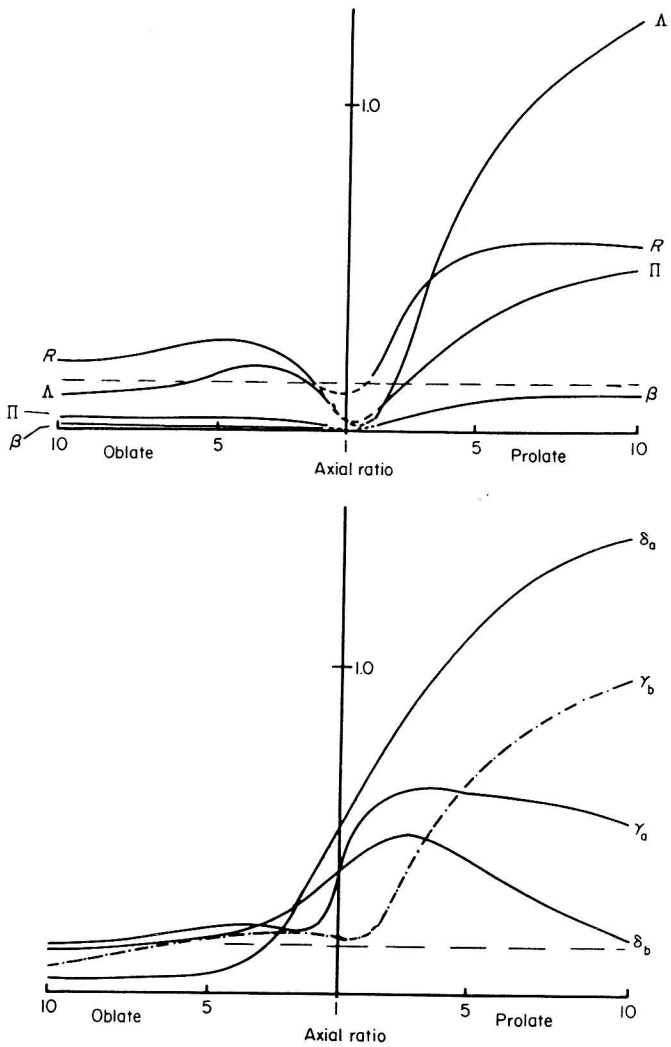

Figure 1 Rate of change (1st derivative) of the relative error in various volume-independent hydrodynamic functions (defined in the text) with respect to the relative imprecision in the axial ratio of the assumed ellipsoid of revolution. Numerical differentiation has been performed on values of the various functions computed at intervals of 1 in axial ratio, by taking the analytical derivative of the local coefficients of a sliding strip least-squares quadratic fit. The horizontal broken line in each case indicates the minimum value which this derivative must have if an axial ratio precise to $\pm 20\%$ is to be retrieved from the measured function, the latter being assumed to be precise to $\pm 3\%$. Rates for the ψ and Ψ functions are not plotted, in the interests of clarity. They are close to the baseline (i.e. the function is very insensitive) for all values of axial ratio

(5) the dielectric dispersion relaxation time ratios:

$$\frac{\tau_i}{\tau_0}=\frac{kT}{3\eta_0 V_e}\tau_i \qquad (i=a,\ b,\ c) \tag{5}$$

where, for ellipsoids of revolution

$$\tau_a=\frac{1}{2\theta_b} \qquad \tau_b=\tau_c=\frac{1}{\theta_a+\theta_b} \tag{6}$$

(6) the harmonic mean rotational relaxation time ratio:

$$\frac{\tau_h}{\tau_0}=\frac{3}{(\tau_0/\tau_a)+(2\tau_0/\tau_b)}\equiv\frac{kT}{3\eta_0 V_e}\tau_h \tag{7}$$

Explicit relations for ν[11,12], f/f_0[13], U_{red}[14], θ_i/θ_0[15,10], τ_i/τ_0[15,16] (and hence τ_h/τ_0)[17] in terms of axial ratio for ellipsoids of revolution have been given.

The determination of all these functions, however, requires a knowledge of the molecular volume $V_e=(M_r v_s/N_A)$. V_e can be eliminated by various combinations of equations (1)–(7) to yield shape functions which are volume-independent:

$$\beta=\frac{N_A^{1/3}}{(16\,200\pi^2)^{1/3}}\frac{\nu^{1/3}}{f/f_0}\equiv\frac{N_A s[\eta]^{1/3}\eta_0}{M_r^{2/3}(1-\bar{\nu}\rho_0)100^{1/3}} \tag{8}$$

(Ref 18)

$$\psi=\frac{U_{red}}{162\pi^2}\left(\frac{f_0}{f}\right)^3\equiv\frac{U\eta^3 N_A^2 s^3}{M_r^3(1-\bar{\nu}\rho_0)^3} \quad \text{(Ref 14)} \tag{9}$$

$$\Pi=\frac{U_{red}}{\nu}\equiv\frac{U}{[\eta]M_r} \quad \text{(Refs 10, 49)} \tag{10}$$

$$\Psi=\left(\frac{\tau_0}{\tau_h}\right)^{1/3}\left(\frac{f}{f_0}\right)\equiv\left(\frac{4\pi\eta_0}{kT}\right)^{1/3}\frac{M_r(1-\bar{\nu}\rho_0)}{6\pi\eta_0 N_A s}\left(\frac{1}{\tau_h}\right)^{1/3} \tag{11}$$

(Refs 17, 19)

$$\Lambda=\left(\frac{\tau_0}{\tau_h}\right)\nu\equiv\frac{3\eta_0[\eta]M_r}{N_A kT\tau_h} \quad \text{(Ref 20)} \tag{12}$$

$$R=\frac{2}{\nu}\left[1+\left(\frac{f}{f_0}\right)^3\right]\equiv\frac{k_s}{[\eta]} \quad \text{(Ref 21)} \tag{13}$$

where k_s is the sedimentation concentration regression coefficient given by:

$$s_c=s(1-k_s c)\simeq s(1+k_s c)^{-1} \tag{14}$$

and where s_c and s are the sedimentation coefficients at concentration c and infinite dilution, respectively,

$$\delta_i=\frac{\theta_i}{\theta_0}\nu\equiv\frac{6\eta_0\theta_i[\eta]M_r}{N_A kT} \tag{15}$$

(Refs 18, 19, 10)

$$\gamma_i=\left(\frac{f}{f_0}\right)^3\frac{\tau_0}{\tau_i}\equiv\frac{1}{54\pi^2 N_A^3 kT}\cdot\frac{M_r^3(1-\bar{\nu}\rho_0)}{s^3\eta_0^2\tau_i} \tag{16}$$

(Ref 19)

Results and discussion

Many of the volume-independent functions given in equations (8)–(16) above are extremely insensitive to axial ratio and sensitive to experimental error. Nevertheless, many workers have applied them, notably the β function, without adequate regard for the inaccuracies involved in their use. It is therefore both important and interesting to compare quantitatively their sensitivities to axial ratio and to experimental error.

Sensitivity

In *Figure 1* we plot the fractional change in the function ($\beta, R\ldots$) arising from a given fractional change in the axial

Table 1 Use of the R function to predict the conformation of various macromolecules in solution in terms of an ellipsoid of revolution model

Protein	k_s (ml g^{-1})	k_η (ml g^{-1})	$[\eta]$ (ml g^{-1})	R	Axial ratio	Model-dependent ($\bar{v}_s/\bar{v}$)	Model-independent ($\bar{v}_s/\bar{v}$)	Conclusion
Apoferritin[24]	8	12	5.16	1.55	1.45[a,b]	2.6[a,b]	1.5	Approx. spherical: agrees with X-ray crystallography and electron microscopy[39]
BSA[24]	5.5	7.7	2.75	2.0	—	—	1.4	Not a hydrodynamic ellipsoid (c.f. $\beta<2.1$)[46]
Fibrinogen[25]	7	14	7.8	0.9	6.3[a]	1.1[a]	2.0	Prolate ellipsoid ~6:1. Agrees with electron microscopy[26]
Ovalbumin[40]	5.45[c]	6.6	3.49	1.56	1.5[a,b]	1.5[a,b]	1.2	Approx. spherical
C-protein[27]	11	15.4	12.6	0.87	26.0[b], 6.65[a]	0.9[b], 2.12[a]	1.4	Oblate ellipsoid ~25:1
Myosin[28,29]	85	92	234	0.38	30[a]	4.3[a]	1.1	Not hydrodynamic ellipsoids of revolution
Synthetic A-filaments[30]	160.8	366	176	0.9	19.5[a]	16[a]	2.3	
Collagen sonicates[31]								
$M_r=352\,000$	308	880	1252	0.246	80[a]	2.28[a]	2.85	Prolate ~80:1
$M_r=330\,000$	291	756	1078	0.270	64[a]	2.85[a]	2.60	Prolate ~65:1
$M_r=273\,000$	241	564	639	0.377	30[a]	6.12[a]	2.34	Not hydrodynamic ellipsoids of revolution
$M_r=227\,000$	193	428	400	0.483	18[a]	9.13[a]	2.22	

[a] Prolate ellipsoid; [b] oblate ellipsoid; [c] corrected to solution density[21]

ratio. A large change in the function denotes high sensitivity, and it is clear that the various functions differ widely, both as a function of axial ratio and among themselves. To evaluate the practical use of the functions we define the following criteria:

(1) Estimates of axial ratio of worse than $\pm 20\%$ precision are of little or no interest.

(2) Functions can be calculated from experimental data to a precision of better than $\pm 3\%$, but seldom to *much* better precision (see below). In terms of these criteria — admittedly slightly subjective ones — we see from *Figure 1* that many of the defined functions are unusable, especially at low axial ratio. The R and δ_a functions are the most sensitive functions for the whole range but the newly defined Π and Λ functions may have application for particles of small (but not very small) asymmetry, and the rotational functions (δ_b, γ_a and γ_b) for prolate ellipsoids.

In addition to its purely mathematical sensitivity to shape variation, a shape function must be judged with respect to its insensitivity to experimental error. It is readily seen from equations (8), (9), (11) and (16) that the β, ψ,Ψ, γ_a and γ_b functions require a relatively large number of measurements to be made, and many terms appearing in these equations are either squared or cubed. On the other hand, the Λ function [equation (12)] requires knowledge of the harmonic mean rotational relaxation time, τ_h, the measurement of which suffers from problems of internal rotation of the chromophore and segmental rotation of the macromolecule, a good example being fibrinogen[20,22]. In order to determine δ_b or γ_a, a knowledge of the rotational diffusion coefficient θ_b (or, alternatively, τ_a, see equation (6)) is required. According to Benoit[42], for ellipsoids of revolution there will be one electric birefringence relaxation time τ_n related to θ_b by[19] $\tau_n = 1/(6\theta_b)$. Using this technique, θ_b has been measured to a precision of $\pm 1.5\%$ for haemoglobin and Squire[19], determining the corresponding value of γ_a, has found that the axial ratio of haemoglobin lies within the range 1.2 (prolate) to 2.6 (oblate), corresponding to an error in γ_a of $\sim \pm 5\%$.

The evaluation of δ_a and γ_b is more complicated since they require θ_a and τ_b, respectively, and hence, from equation (6), the resolution of two dielectric relaxation times. Resolution is almost impossible without constraints in the analysis of the dispersion curve, but Moser *et al.*[43,44] have developed a technique whereby θ_b (or equivalently τ_a) is first obtained by electric birefringence decay as above and then used as a constraint in the analysis of the dielectric dispersion curve to obtain τ_b (and hence θ_a). In their programs, Moser *et al.* have added the further constraint that the corresponding γ_a and γ_b functions give the same axial ratio, and for bovine serum albumin they obtain a prolate ellipsoid of axial ratio 3.0, compared with 'unconstrained' values from γ_a and γ_b, respectively, of 3.5 and 5.1. The ratio of the two dielectric relaxation times[45] is also a volume-independent function of axial ratio; an axial ratio of 3.0 is again obtained for BSA[44]. There remains some doubt, however, from other measurements as to whether BSA is ellipsoidal at all (see *Table 1* and Ref 46).

Although these algorithms apparently produce resolution of the two-term dielectric dispersion curve for an axial ratio ~3, it is not known whether adequate resolution is possible for more symmetric particles, i.e. where the relaxation times will be closer. Another difficulty[19] is that the method is limited to solutions of low conductivity so that macromolecules in physiological conditions cannot be examined, measurements being restricted to those pH values and ionic strengths at which proteins have minimum solubility. Finally, the relaxation times are concentration dependent and have to be extrapolated to infinite dilution[47].

The degree of experimental uncertainty associated with a calculation of values for all the various functions can be

Table 2 Crystallographic dimensions of some globular proteins

Protein	Dimensions (Å)	Reference
Carboxypeptidase	50 × 42 × 38	32
Myoglobin	43 × 35 × 23	33
Cytochrome *c*	25 × 25 × 35	34
Lysozyme	45 × 30 × 30	35
Ribonuclease	38 × 28 × 22	36
Pre-albumin	70 × 55 × 50	37
Haemoglobin	64 × 55 × 50	38

estimated from values assigned to the errors in the molecular parameters from which these functions are calculated. We assign these as follows:

$s_{20,w}$ 0.2%	θ_a 2%
M_r 1.0%	θ_b 1.5%
$[\eta](1+0.5^*)\%$	τ_a 1.5%
$\bar{v}$ 0.5*%	τ_b 2.5%
τ_h 2.0%	
k_s $(1.0+0.5^*)\%$	
c 0.5%	
$U(1.5+0.5^*)\%$	

The asterisked quantities reflect the contribution of uncertainty in the concentration measurement which being a systematic error will cancel in certain cases. From these assigned errors, using normal statistical procedures we may estimate the resulting uncertainty in the derived functions as:

β 1.7%	δ_a 2.7%
ψ 4.4%	δ_b 2.3%
Π 2.1%	γ_a 2.6%
Ψ 2.0%	γ_b 3.3%
Λ 2.7%	
R 1.4%	

The error in the final estimated parameters is ~2% in most cases, although the R function is a little better than this and the ψ function significantly worse. The purely mathematical sensitivity of the various functions discussed above and illustrated in *Figure 1* is therefore the dominating factor in deciding which function is the most readily usable.

The R function requires knowledge only of the sedimentation regression coefficient k_s and the intrinsic viscosity $[\eta]$ which can both be determined accurately by fitting data to a new universal equation for transport at all solute concentrations up to the critical packing fraction[2,3,10]. This insensitivity to experimental error strengthens the conclusion from consideration of the sensitivity to axial ratio that the R function is by far the most applicable to protein systems. Also, any systematic errors in solute concentration cancel in the ratio $k_s/[\eta]$.

Criterion for the goodness of fit of a chosen model

Although, using procedures defined above, a hydrodynamically equivalent prolate or oblate ellipsoid of revolution can be fitted with reasonable precision to a protein structure in solution it could still be different from either ellipsoid, i.e. the choice of an ellipsoid of revolution in the first instance may be a poor approximation to the true structure.

A useful criterion for determining whether the true structure resembles an ellipsoid of revolution — or any model for which data are available — is a comparison of the model-dependent with the model-independent estimate of the swollen molecular volume, V_e, swollen specific volume, v_s, or the 'swelling' ratio, $v_s/\bar{v}$, for the protein.

The model-dependent estimate can be found by back-substitution of the axial ratio for the ellipsoid of revolution, found from the R function, into equation (1) for the viscosity increment, from which $v_s/\bar{v}$ can be found. An estimate that is *independent* of any assumed model can be found from the ratio of the viscosity regression coefficient, k_η, to the sedimentation regression coefficient[21], i.e.:

$$\frac{v_s}{\bar{v}} = \frac{k_\eta}{k_s}$$

Such a comparison is given for several proteins in *Table 1*.

Conclusions

It is evident from *Table 1* that the application of an ellipsoid of revolution model to many protein systems in solution is not valid. One exception is fibrinogen, for which the result predicted by the hydrodynamic data appears to agree quite accurately with that from electron microscopy.

The most likely protein system to which an ellipsoid of revolution would be a valid model is that of globular proteins, whose shapes, as their name implies, are reasonably regular. A perusal of the shapes and dimensions of globular proteins predicted by X-ray crystallography illustrates that in some cases an ellipsoid of revolution model could well be valid, e.g. lysozyme and cytochrome *c* (*Table 2*; see also Table 1 of Squire and Himmel[41]). In many cases, however, such as carboxypeptidase and myoglobin, the distinction as to whether the protein is better modelled either by a prolate or oblate ellipsoid may be arbitrary and, indeed, impossible in some cases.

It would be a significant step forward, therefore, if the restriction of two equal axes on the ellipsoid were removed to allow use of the more general 'triaxial' ellipsoid. However, due either to the lack of the necessary theoretical relationships linking the axial dimensions of the ellipsoid with experimental parameters, or, even if they are available, due to the lack of the necessary experimental precision, numerical inversion procedures or data analysis techniques, this model has not yet been available. A very recent study has shown, however, that this restriction can now be removed[48].

References

1 Bloomfield, V. A., Dalton, W. O. and Van Holde, K. E. *Biopolymers* 1967, **5**, 135

2 Garcia de la Torre, J. and Bloomfield, V. A. *Biopolymers* 1977, **16**, 1747

3 Garcia de la Torre, J. and Bloomfield, V. A. *Biopolymers* 1977, **16**, 1765

4 Garcia de la Torre, J. and Bloomfield, V. A. *Biopolymers* 1977, **16**, 1779

5 Garcia, de la Torre, J. and Bloomfield, V. A. *Biopolymers* 1978, **17**, 1605
6 Wilson, R. W. and Bloomfield, V. A. *Biopolymers* 1979, **18**, 1205
7 Wilson, R. W. and Bloomfield, V. A. *Biopolymers* 1979, **18**, 1543
8 Garcia Bernal, J. M. and Garcia de la Torre, J. *Biopolymers* 1980, **19**, 751
9 Tanford, C. 'Physical Chemistry of Macromolecules' J. Wiley, New York, 1961
10 Harding, S. E. *PhD Thesis* Leicester University (1980)
11 Simha, R. *J. Phys. Chem.* 1940, **44**, 25
12 Saito, N. *J. Phys. Soc. (Jpn)* 1951, **6**, 297
13 Perrin, F. *J. Phys. Radium* 1936, **7**, 1
14 Jeffrey, P. D., Nichol, L. W., Turner, D. R. and Winzor, D. J. *J. Phys. Chem.* 1977, **81**, 776
15 Perrin, F. *J. Phys. Radium* 1934, **5**, 497
16 Koenig, S. H. *Biopolymers* 1975, **14**, 2421
17 Squire, P. G. *Biochim. Biophys. Acta* 1970, **221**, 425
18 Scheraga, H. A. and Mandelkern, L. *J. Am. Chem. Soc.* 1953, **79**, 179
19 Squire, P. G. *Electro-Opt. Ser.* 1978, **2**, 565
20 Harding, S. E. *Biochem. J.* 1980, **189**, 359
21 Rowe, A. J. *Biopolymers* 1977, **16**, 2595
22 Johnson, P. and Mihalyi, E. *Biochim. Biophys. Acta* 1965, **102**, 476
23 Rowe, A. J. manuscript in preparation
24 Rowe, A. J. and Pancholi, A. unpublished work
25 Rowe, A. J. and Mihalyi, E. unpublished work
26 Fowler, A. E. and Erickson, H. P. *J. Mol. Biol.* 1979, **134**, 241
27 Offer, G., Moos, C. and Starr, R. *J. Mol. Biol.* 1973, **74**, 653
28 Emes, C. H. *PhD Thesis* Leicester University (1977)
29 Emes, C. H. and Rowe, A. J. *Biochim. Biophys. Acta* 1978, **537**, 110
30 Emes, C. H. and Rowe, A. J. *Biochim. Biophys. Acta* 1978, **537**, 125
31 from Nishihara, T. and Doty, P. *Proc. Natl. Acad. Sci. USA* 1958, **44**, 411
32 Lipscomb, W. *Proc. Robert A. Welch Found. Conf. Chem. Res.* 1971, **15**, 134
33 Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, W. G., Wycoff, H. and Phillips, D. C. *Nature* 1958, **181**, 662
34 Dickerson, R. and Geiss, I. 'The Structure and Action of Proteins', Benjamin, California, 1969
35 Blake, C. C. F., Loenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. and Sarma, V. R. *Nature* 1965, **206**, 757
36 Kartha, G., Bellow, J. and Harker, D. *Nature* 1967, **213**, 862
37 Blake, C. C. F., Geisow, M. J. and Oatley, S. J. *J. Mol. Biol.* 1978, **121**, 339
38 Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. and North, A. C. T. *Nature* 1960, **185**, 416
39 Harrison, P. M. *J. Mol. Biol.* 1959, **1**, 69
40 from Holt, J. C. *PhD. Thesis* London University (1970); see also Harding, S. E. *Int. J. Biol. Macromol.* 1981, **3**, 398
41 Squire, P. G. and Himmel *Arch. Biochem. Biophys.* 1979, **196**, 165
42 Benoit, H. *Ann. Physik* 1951, **6**, 561
43 Moser, P., Squire, P. G. and O'Konski, C. T. *J. Phys. Chem.* 1966, **70**, 744
44 Squire, P. G., Moser, P. and O'Konski, C. T. *Biochemistry* 1968, **7**, 4261
45 Wyman, J. and Ingalls, E. N. *J. Biol. Chem.* 1943, **147**, 297
46 McCammon, J. A., Deutch, J. M. and Bloomfield, V. A. *Biopolymers* 1975, **14**, 2479
47 Riddiford, C. L. and Jennings, B. *Biopolymers* 1967, **5**, 757
48 Harding, S. E. and Rowe, A. J. *Biopolymers* in press
49 Harding, S. E. *Int. J. Biol. Macromol.* 1981, **3**, 340