Solvation Interactions of Proteins in Solution [and Discussion]

F. Franks, M. C. R. Symons and R. J. P. Williams

Phil. Trans. R. Soc. Lond. B 1977 **278**, 89-96
doi: 10.1098/rstb.1977.0032

Phil. Trans. R. Soc. Lond. B. **278**, 89–96 (1977) [89]
Printed in Great Britain

Solvation interactions of proteins in solution

BY F. FRANKS
Biosciences Division, Unilever Research Laboratory Colworth/Welwyn,
Colworth House, Sharnbrook, Bedford

Organizer's note

This lecture was to have been presented by Professor Rufus Lumry but complications resulting from a recent operation prevented his presence at the meeting, nor was he able to prepare a full-scale manuscript. He suggested that Professor Franks should prepare and deliver the lecture in his stead and this course was adopted. Professor Lumry did, however, supply Professor Franks with detailed notes on several recent developments which he felt should be dealt with in the presentation and these are also included in the following extended summary of the lecture delivered by Professor Franks and they are indicated by † at the beginning of the relevant sections.

Several recent reviews describe the interactions of proteins with the solvent medium, the rôle played by the solvent in maintaining conformational stability, and the manner in which changes in the solvent can influence conformational equilibria (Cooke & Kuntz 1974; Kuntz & Kauzmann 1974; Eagland 1975; Berendsen 1975; Franks & Eagland 1975).

Proteins can be studied at different levels of resolution. At one extreme, described by Dr Finney in the first paper of this meeting, the protein is defined by a set of atomic coordinates and any 'solvent effect' is reduced to the examination of discrete water molecules, and their position and orientations with respect to given atomic groupings in the protein.

At the other – low resolution – extreme the protein is usually regarded as an impenetrable body, roughly spherical in its native state, the dimensions of which are governed, at least partly, by its interactions with the surrounding solvent medium. Since native or folded states are inseparable from an aqueous solvent medium and water is characterized by a highly developed intermolecular order, it is concluded that the viability of native states must be associated with this peculiar 'water structure'. This in turn implies that the conformational stability is not solely the result of protein–water, i.e. nearest neighbour, interactions, but is influenced by the more distant water–water interactions. This is shown in diagrammatic form in figure 1, where the A shell includes all those solvent molecules which are nearest neighbours to protein side chains or backbone atoms. Region C represents the unperturbed arrangement – largely tetrahedral – of water molecules and region B results from the incompatibility of the hydrogen bonding regimes in this tetrahedral arrangement and in the primary hydration sphere A.

The following questions relating to the effects of the aqueous solvent in maintaining or perturbing folded states are currently receiving attention:

(1) What is the significance of the A shell? Presumably the motions of water molecules in this primary hydration sphere are determined largely by those of the protein itself or of specific side chain groups.

(2) What is the significance of the B shell? Presumably experimentally accessible properties are determined mainly by water–water correlations which differ from those in the C shell.

Clearly B shell effects are more subtle than A shell effects, but it is here that the influences of temperature, pressure and solvent composition on protein stability operate.

(3) How can minor changes (at least in physico-chemical terms) in the solvent composition promote major upheavals in the protein conformation? On the other hand it is also relevant to ask whether unfolding processes are in fact major upheavals involving the majority of residues. A good case can be made to support the view that a native (N)→denatured (D) state transition may only affect relatively few amino-acid residues.

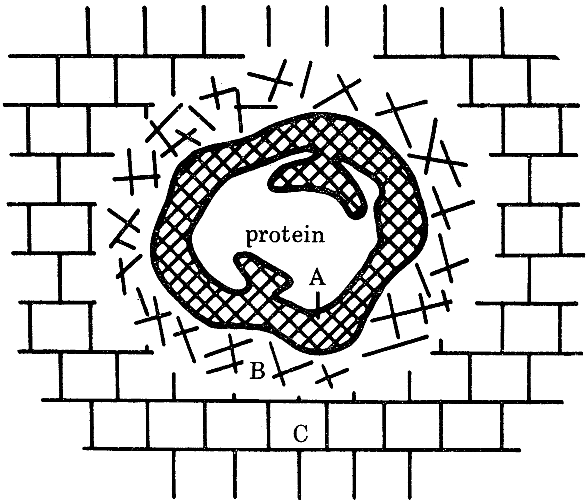

Figure 1. Diagrammatic representation of the protein environment in solution. The protein can be regarded as a hydrodynamic sphere with a primary hydration shell A in which the molecular motions are largely determined by those of the polar protein sites. C is the unperturbed water 'structure' and region B arises from the spatial and orientational mismatch between regions A and C. (From Franks & Eagland 1975.)

(4) By what mechanisms do solvent molecules enter the interior of a folded protein? If they can form a continuous proton exchange path (Grotthus) to the A shell, then what is the practical meaning of the 'interior' of the protein?

(5) Many experimental techniques are currently employed to measure 'protein hydration' and solvent mediated effects on conformational stability. It needs to be established whether a given method monitors A shell or B shell water only, or the combined effects of the two regions. Thus X-ray diffraction methods in their present state of refinement are limited to the study of locations of some A shell and internal water molecules, but for instance hydrodynamic measurements might well be affected by the B shell. The recent review by Kuntz & Krauzmann (1974) discusses the significance of such measurements in some detail.

Classification of hydration interactions

The operative interactions can be classified along the lines discussed by Franks in the second paper of this meeting, namely:

(1) Electrostatic: ion–ion and ion–dipole; these are limited to the A shell.

(2) Dipole–dipole and hydrogen bonding: these involve peptide groups and polar side chains (e.g. histidine, glutamic acid, lysine, tyrosine, serine) and are likely to be very sensitive to the spatial and orientational distribution of the hydrogen bonding sites. Also involved will be water–water interactions in the B shell.

(3) Hydrophobic effects: these probably involve hydrophobic cage formation around exposed apolar residues (leucine, isoleucine, alanine, valine). This type of interaction is weak and long range, entropic in origin and increases in magnitude with rising temperature.

The *solvation* free energies (as distinct from intramolecular contributions) of the N and D states are likely to be the resultant of the individual effects as classified above, and $\Delta G_{N\rightarrow D}$ reflects net change in the solvation effects only. These can be brought about by changes in temperature, pressure or solvent composition.

Experimental investigations of A shell hydration

Apart from X-ray diffraction, the most informative methods are high resolution proton magnetic resonance and differential scanning calorimetry. Both these techniques have been applied to monitor the 'unfreezable water', i.e. the amount of water per unit weight of protein which does not exhibit the usual freezing behaviour near 0 °C.

The former method was used extensively by Kuntz and his colleagues (1971) to study the hydration of amino acids and hence to obtain an estimate of the hydration of proteins. Where both N and D states were examined the degree of hydration was the same (within experimental error) and this suggests that the method monitors the properties of A shell water.

Table 1. Hydration of amino acid residues at −35 °C

hydration mol H_2O/mol residue	amino acid
0	Phe
1	Gly, Ile, Leu, Val, Cys, Met (Ala = 1.5)
2	Trp, Ser, Thr, Gln, Asn, Asp (pH 4), Glu (pH 4)
3	Pro, Arg (pH 10), Arg^+, Tyr
4	Hyp, His^+, [Lys^+, Lys (pH 10) = 4.5]
6	Asp^-
7.5	Glu^-, Tyr^- (pH 12)

(From Kuntz 1971; and also printed with permission in Franks & Eagland 1975.)

Table 1 shows the amino acid hydration numbers assigned by Kuntz. The data indicate that ionic residues are more heavily hydrated than their un-ionized counterparts, that hydrophobic hydration cannot be detected (which is not surprising, since the measurements were performed at −35 °C), and that proline is comparatively heavily hydrated, which may explain the high degree of solvent exposure of proline residues in globular proteins. There are several unresolved problems relating to this particular experiment approach for the quantification of protein hydration (Falk, Poole & Goymour 1970), its main virtue being that it works!

Table 2 is a summary of protein hydration data compiled by Kuntz & Kauzmann (1974). The significance of these results has also recently been discussed by Franks & Eagland (1975).

† From general energetic considerations it would seem reasonable that most of the A shell water is held in the vicinity of NH_3^+ and COO^- groups by electrostrictive forces. However, diffraction studies have not revealed well developed counter-ion structures. On the other hand there is uncertainty about earlier claims that counter-ions are excluded from the A shell of cytochrome c. In apparent contradiction, proton exchange studies (see below) indicate that OH′ ions readily enter all parts of lysozyme, carbonic anhydrase and chymotrypsin. The actual

molecular nature of the exchanging species is uncertain – in the gas phase the hydrated hydroxyl ion is $OH'.2H_2O$. There are therefore several unresolved questions surrounding the nature and function of the A shell.

TABLE 2. A COMPARISON OF PROTEIN HYDRATION DATA OBTAINED BY DIFFERENT TECHNIQUES

		hydration mol H_2O/mol protein				
protein	mol. mass	n.m.r. (-35 °C)	scanning calorimetry (non-freezable water)	hydro-dynamic	isopiestic ($a_w = 0.92$)	structural X-ray
ribonuclease	13700	—	—	434	266	—
lysozyme	14300	270	238	365	199	—
myoglobin	18000	240	—	460	320	110
β-lactoglobulin	36700	—	1121	918	612	—
ovalbumin	45000	825	—	375	750	—
b.s.a.	68000	1511	1208 1851	1624	1208	—
haemoglobin	64500	1504	1147	1792	1326	—
subtilisn	27300	—	—	—	—	273
chymotrypsinogen	25700	485	—	543	414	14, 100‡

‡ This value is for the 'internal' localized water.
(Taken from Franks & Eagland 1975.)

THE SECONDARY HYDRATION SPHERE – THE B SHELL

The existence and importance of the B shell can be inferred from the incompatibility of the spatial and orientational configurations of water molecules in the A shell with those in bulk unperturbed water. Also there is plenty of evidence supporting the importance of hydrophobic hydration and hydrophobic interaction, and such processes will of necessity involve the participation of B shell water molecules.

There are also subtle but very specific medium effects on conformational stability, e.g. the lyotropic ionic series which appears to operate primarily via specificities in ion hydration. These ionic processes extend to quite specific buffer effects, so that pH alone is not a truly diagnostic property. Thus at the same pH, sulphate and phosphate buffer systems affect the thermal unfolding of ribonuclease in different ways (Ginsburg & Carrol 1965).

The effects of ions and non-electrolytes on the N$\longrightarrow$D transition are popularly treated in terms of the binding of the third component to the protein (component 2) at the expense of water binding (component 1). In thermodynamic terms binding isotherms must be based on one of two assumptions: either binding is assumed to be stoichiometric, in which case a series of equilibrium constants is fitted to experimentally determined numbers of molecules or ions bound, or the assumption is made that sites exist with different binding affinities, in which case independent site binding constants are fitted to the experimental data. Neither of the above treatments provides information about solvent binding or its displacement by other ligands. Although it is usually quite easy to fit experimental data by a binding isotherm, the thermodynamic case is strengthened if independent, *non-thermodynamic* evidence for binding is available.

The *absolute* amount of species 3 bound to the protein is given by

$$A_3 = (\partial m_3/\partial m_2)_{T,\mu_1,\mu_3}+m_3 A_1,$$

where concentrations are expressed in moles of ligand per mole of protein. A_1 is the *absolute* amount of water bound and μ_i are the chemical potentials. Further, by the Gibbs–Duhem equation

$$A_1 = (\partial m_1/\partial m_2)_{T,\,\mu_1}$$
$$= -\frac{m_1}{m_3}\left(\frac{\partial m_3}{\partial m_2}\right)_{T,\,\mu_3}.$$

The uncertainties in A_1 are well demonstrated by the various estimates in table 2.

In the literature dealing with ligand binding to proteins definitions of binding are often vague, e.g. do ligand molecules (or ions) replace water molecules and, if so, how many, and how are binding sites defined.

The possible inconsistencies inherent in interpretations of spectroscopic data in terms of binding and conformational transitions are analysed in Franks & Eagland (1975) for the system water–lysozyme–2-chloroethanol (Timasheff 1970).

The general uncertainties surrounding medium effects can be strikingly demonstrated by a survey of the recent literature on the chaotropic action of urea on protein conformational stability. The application of techniques as diverse as u.v. absorption, pH titration, circular dichroism, optical rotatory dispersion, ultrafiltration, calorimetry, dilatometry, n.m.r., and disulphide assay to the urea induced N⟶D transition of globular proteins has led to widely different interpretations as to the mode of action of the urea (Franks & Eagland 1975).

On the basis of the available evidence and particularly in the light of our knowledge regarding the effects of urea on water, we consider it likely that urea acts via the B shell (or indeed the C shell) by robbing water of those properties which render it unique in maintaining native structures.

Convincing evidence in favour of the involvement of the B shell in protein stabilization/destabilization processes has been produced by Brandts & Hunt (1967) in their detailed study of the system water–ribonuclease–ethanol.

The simplified model of the N⟶D transition as an exposure of apolar residues to the solvent is also due for a re-examination. Zipp & Kauzmann (1973) have shown quite conclusively that the pressure denaturation of metmyoglobin is not consistent with the classical model of the hydrophobic interaction, and neither are various estimates of the compressibilities of native proteins.

† THE NATURE OF THE PROTEIN INTERIOR

On the question of the nature of the 'inside' of the protein, proton exchange measurements with lysozyme show that basically two processes can be distinguished – one of which has a very low energy of activation, $E^{\ddagger} \approx 9$ kcal (38 kJ) mol^{-1}, and that *N*-acetylglucosamine binding affects *all* proton exchange rates, leading to an increased $E^{\ddagger}$. The exchange catalysts, H^+, H_2O or most probably OH', are presumably able to penetrate to all proton sites and assume the necessary geometry for exchange, and this occurs with an expenditure of enthalpy of $\ngtr 9$ kcal mol^{-1}.

The 'porosity' of the N state is also demonstrated by the tryptophan fluorescence quenching produced by acrylamide. Thus ribonuclease has a single tryptophan residue, buried in the interior; its fluorescence is quenched by acrylamide, and the process has an $E^{\ddagger}$ of 9 kcal mol^{-1}. The rate of diffusion of acrylamide inside the protein is only 30 times slower than in aqueous solution (Eftink & Ghiron 1975). Similar experiments with oxygen as quencher suggests that this molecule is retarded only tenfold by the protein.

The observations that small molecules readily penetrate proteins suggest the existence of a pool of free volume, possibly divided into small cells, giving rise to defect structures, holes, which can move in a cooperative manner to provide channels and occlusion pathways for water. The postulate of this free volume implies sub-optimal bonding in the interior of the protein. This is in agreement with studies on model peptides which show that 25 % of the NH groups in the random coil (D state) do not hydrogen bond to the solvent. The free energy of the random coil in water is therefore higher than had previously been assumed, and since $\Delta G_{N \rightarrow D}$ is known, this means that the free energy of the folded protein is also higher. In fact by optimizing NH–CO placements, the stability of proteins could be vastly increased (e.g. proteins in thermophilic bacteria can have transition temperatures of > 90°). High folded state stability is not a common property of most proteins, so that the thesis of incomplete bonding, leading to mobile defects in the interior is consistent with experimental results. These mobile defects are shown in diagrammatic form in figure 2. We thus have a family of conformers of nearly the same free energy, allowing the redistribution of small amounts of free volume to facilitate the migration of water and possibly other small molecules or ions. Various dynamic measurements suggest conformer life times ranging from 10^{-9} to 10^{-2} s.

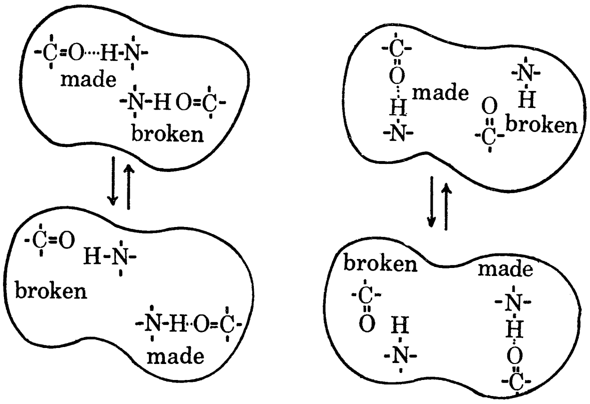

FIGURE 2. Peptide hydrogen bonding in proteins. Poor bonding situations migrate by disruption at good bonding situations in other regions of the protein. The mobile defect system is highly cooperative, so that the pairwise migration process shown is an oversimplification.

We can further postulate that specific inhibitors reduce the potential energy and total volume of the mobile defect system, therefore $E^{\ddagger}$ for proton exchange increases, but uniformly so for all protons. The existence of available volume is also suggested by the internal binding of xenon, HgI_3', and neo-pentane.

Transitions within the conformer family may be subject to high activation barriers which should be affected by the ionization state of the protein, solvent composition and any chemical modifications.

In order further to study internal rearrangements which take place upon the binding of small molecules, Frauenfelder and his collaborators (1975) studied the CO binding to myoglobin by progressively cooling the protein solution, and changing the solvent, as required, to keep the system liquid. They found a barrier between the exterior of the protein and some intermediate pocket which is very sensitive to solvent conditions, surface modification of the protein and external environmental effects in general. The second barrier is between this intermediate pocket and the hole, containing in its surface the haem iron. The second barrier is insensitive to environmental factors but very sensitive to modifications in amino-acid composition and chemical modifications of the protein.

Detailed studies show a broad distribution of conformers separated by energy barriers of about 1.5 kcal (6.3 kJ) $\mathrm{mol^{-1}}$. Rough calculations on rates of interconversion suggest 10^3–10^9 $\mathrm{s^{-1}}$.

Perhaps it is not strictly in order to regard single water molecules or OH′ ions diffusing through the protein as part of the solvent, but at the present time it is not known whether there is not a continuous path between 'outside' water and the interior, with a life time long enough to allow the various proton exchange processes to take place. It must be borne in mind that water diffuses rapidly across barriers which should really be impermeable, e.g. phospholipid bilayer systems. The diffusion rates are certainly not simply related to the size of the H_2O molecule. The indications certainly are that protein hydration does not stop at the protein/water interface as shown diagrammatically in figure 1 but that both water and other small molecule species readily diffuse into the protein to interact with sites which should normally be well buried in the interior.

REFERENCES (Franks)

Berendsen, H. J. C. 1975 *Water – a comprehensive treatise* (ed. F. Franks), vol. 5, ch. 6, pp. 293–330. New York: Plenum Press.

Brandts, J. F. & Hunt, L. 1967 The thermodynamics of protein denaturation. III. The denaturation of ribonuclease in water and in aqueous urea and aqueous ethanol mixtures. *J. Am. chem. Soc.* **89**, 4826–4838.

Cooke, R. & Kuntz, I. D. 1974 Properties of water in biological systems. *A. Rev. Biophys. Bioeng.* **3**, 95–126.

Eagland, D. 1975 *Water – a comprehensive treatise* (ed. F. Franks), vol. 4, ch. 5, pp. 305–518. New York: Plenum Press.

Eftink, M. R. & Ghiron, C. A. 1975 Dynamics of protein matrix revealed by fluorescence quenching. *Proc. Natn. Acad. Sci., U.S.A.* **72**, 3290–3294.

Falk, M., Poole, A. G. & Goymour, C. G. 1970 Infrared study of the state of water in the hydration shell of DNA. *Can. J. Chem.* **48**, 1536–1542.

Franks, F. & Eagland, D. 1975 The role of solvent interactions in protein conformation. *Crit. Rev. Biochem.* **3**, 165–219.

Frauenfelder, 1975 *Biochemistry* **14**, 5355–5373.

Ginsburg, A. & Carrol, W. R. 1965 Some specific ion effect on the conformation and thermal stability of ribonuclease. *Biochemistry* **4**, 2159–2174.

Kuntz, I. D. 1971 Hydration of macromolecules. III. Hydration of polypeptides. *J. Am. chem. Soc.* **93**, 514–516.

Kuntz, I. D. & Kauzmann, W. 1974 Hydration of proteins and polypeptides. *Adv. Protein Chem.* **28**, 239–345.

Timasheff, S. N. 1970 Protein-solvent interactions and protein conformation. *Accounts Chem. Res.* **3**, 62–68.

Zipp, A. & Kauzmann, W. 1973 Pressure denaturation of metmyoglobin. *Biochemistry* **12**, 4217–4228.

Discussion

M. C. R. SYMONS. (*Department of Chemistry, Leicester University*). With reference to the freezing experiments and the KF effect described by Professor Franks, it may be significant that there are often relatively high concentrations of salts and other small molecules in aqueous protein solutions. In the absence of protein, on cooling, a narrow proton resonance characteristic of concentrated solutions would be obtained, strongly shifted from the normal water resonance. If, as I presume to be the case, only one resonance is obtained from such protein solutions, one must conclude that the electrolyte has become concentrated in the water associated with the protein. This means that this 'mobile' water is readily accessible to ions and small molecules during the freezing process.

R. J. P. Williams, F.R.S. (*Wadham College, Oxford OX1 3PN*). The study of proteins by high resolution n.m.r. sheds a rather different light on the properties of proteins and their possible hydration spheres than that shed by other methods. The advantage of the high resolution n.m.r. method is that it permits an examination of a protein at the level of the individual atoms. Therefore the motions of very small parts of a protein can be discerned. The work of Campbell, Dobson, Moore and myself on lysozyme and cytochrome c shows that the following types of motion can be observed apart from the overall rotation and translation of the protein.

(*a*) A general breathing which is common to the whole protein and which has a time constant of 10^{-9} s.

(*b*) Fast rotations of methyl groups, time constant 10^{-12} s.

(*c*) Fast 'waving' motions of surface groups such as lysines, time constant 10^{-10} s.

(*d*) Fast movement of groups on and off metal ions revealed by the rates of spin state equilibria, time constant $\leqslant 10^{-4}$ s.

(*e*) Fast flipping of tyrosine, phenylalanine, alanine, leucine and possibly several other amino-acid side chains, time constant $\leqslant 10^{-4}$ s but with a wide spectrum of rates.

(*f*) Slow conformational flipping of small amplitude of some tryptophans, time constant $\sim 10^{-2}$ s.

(*g*) Ill-defined movement of sections of a protein which may involve small portions of the backbone.

These different and complex motions would indicate that the study of overall relaxation processes, of hydrogen/deuterium exchange, and of the thermodynamic properties of proteins would be very complicated sums. In the case of hydrogen/deuterium exchange our ability to study proteins in any mixture of D_2O/H_2O has allowed us to follow the rates of exchange of assigned specific protons in the NH region. While we have not done a full analysis of the activation energy of exchange the very different rate constants for exchange of different groups indicate that a large variation in the heats and entropies of activation of these processes is to be expected. Thus while we share Lumry's views that there are many fluctuation processes inside proteins the nature of these fluctuations may well be very specific properties of individual proteins. Such fluctuations could be as important in the control of the properties of a protein as is the 'structure' seen by X-rays. These fluctuations are affected by substrate binding or chemical modification.

Turning from the inside to the outside of the protein the rapid movements of surface residues would suggest that there would be mixing of water layers with a time constant 10^{-10} s. We would suspect that these motions would be affected somewhat by ion-pairing between with charged amino-acid side chains and counter-ions of salts. It is not obvious that the surface detail of a protein crystallized from high salt will be similar to that in a low salt medium in solution. Finally some proteins such as chromagranin A have very open, largely unfolded, structures and the residual structure could well be under the control of salt and water activity. All intermediate situations are to be expected.