ELSEVIER

Biophysical Chemistry 93 (2001) 129–139

Biophysical Chemistry

www.elsevier.com/locate/bpc

X-Ray and neutron scattering analyses of hydration shells: a molecular interpretation based on sequence predictions and modelling fits

Stephen J. Perkins*

Department of Biochemistry and Molecular Biology, Royal Free and University College Medical School, Rowland Hill Street, London NW3 2PF, UK

Received 26 April 2001; received in revised form 21 June 2001; accepted 8 August 2001

Abstract

Solution scattering is a low resolution diffraction method that provides important structural data on proteins. The ability to model scattering curves by recourse to known crystal structures for proteins under study significantly improves the resolution (and the utility) of the method because of the strict constraints that the crystal structures impose. For these structure determinations, a molecular description of the effect of hydration shells is needed. In calibration studies used for X-ray scattering curve modelling, it has been reproducibly found that a hydration shell is required. In molecular terms, this results from the higher electron density of the hydration shell compared to that of bulk water, which then becomes similar to that of the protein. This is well represented by a level of 0.3 g H_2O/g glycoprotein and a water molecule volume of 0.0245 nm^3. Procedures for the addition of a hydration shell to a sphere model of a protein are described. For neutron scattering fits, it is not necessary to incorporate a hydration shell, as to a good approximation this is not detectable. In molecular terms, this apparent absence of the neutron hydration shell results from the effect of proton exchange on the scattering densities of bulk water and bound water which causes these to be similar but different from that of the protein.

Keywords: X-Ray scattering; Neutron scattering; Analytical ultracentrifugation; Scattering densities; Hydration shell; Sphere modelling

* Tel.: +44-20-7794-0500, ext 4210/4914; fax: +44-20-7794-9645.
E-mail address: s.perkins@rfc.ucl.ac.uk (S.J. Perkins).

0301-4622/01/$ - see front matter
PII: S0301-4622(01)00216-2

1. Introduction

Solution scattering is a diffraction technique used to study overall structures in solution. A sample is irradiated by a collimated, monochromated beam of X-rays or neutrons. The resulting two-dimensional circularly-symmetric diffraction pattern is recorded on a flat area detector system. Radial averaging leads to a one-dimensional scattering curve. Structural and compositional information is obtained at a resolution of approximately 2–4 nm from measurements of the radius of gyration R_G, the intensity $I(0)$ at zero scattering angle, the cross-sectional R_G (R_{XS}) and the distance distribution function (Fig. 1). These parameters provide information on the overall macromolecular dimensions and the molecular weight. The utility of solution scattering has been much improved in recent years by means of a novel strategy in which molecular structures are derived directly from the scattering curves (reviewed in [1–4]). This method starts from known molecular structures for all the subunits within the macromolecule which are used as tight constraints of the scattering data. Four distinct types of constrained modelling studies have been developed, based on either calibration studies, translational–rotational search fits, symmetry-constrained search fits or molecular dynamics-constrained search fits [4]. By these means, the structural precision of scattering modelling analysis can be raised as high as 0.5–1.0 nm.

X-Ray and neutron scattering exhibit very different but complementary properties. X-Ray scattering at synchrotron sources provides high quality curves that are minimally affected by instrumental geometry, as the incident fluxes are sufficiently high to permit the use of ideal pin-hole optics. X-Rays visualise the macromolecule in a high positive solute–solvent contrast. More importantly, X-rays also visualise the macromolecule with a hydration shell surrounding it, and this significantly affects the modelling of the scattering curve. These hydrated structures are the same as those observed when the sedimentation coefficient is measured by analytical ultracentrifugation. In distinction to this X-ray view, neutron scattering is able to visualise macromolecules in positive and negative contrasts by the use of light and heavy water buffers with widely different scattering densities. The range of scattering densities generated by varying the heavy water content is sufficiently wide to encompass the scattering densities of lipids, protein, carbohydrate and DNA/RNA. Neutron contrast variation can therefore be applied, in which one com-

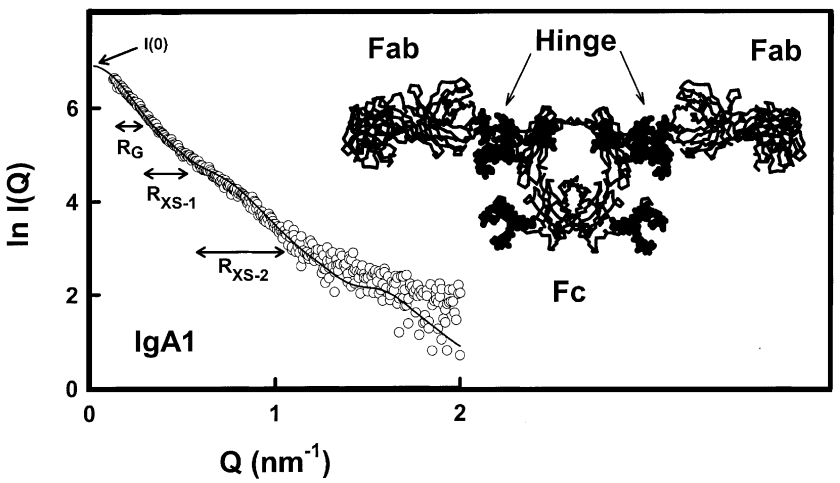

Fig. 1. X-Ray scattering curve fit for a hydrated model of the immunoglobulin molecule IgA1. The α-chain trace of the IgA1 model is shown, with the positions of carbohydrate shown in bold. The two Fab and the Fc fragments are orientated in a T-shape structure joined by an extended hinge. The $I(0)$ value and the Q ranges used to calculate the R_G and two R_{XS} values are indicated as arrowed. Adapted from Boehm et al. [7].

ponent is matched out in its scattering density by the same scattering density from an appropriate choice of heavy water buffer to make it invisible (i.e. $I(0) = 0$). Neutron scattering is also characterised by the absence of radiation damage effects sometimes encountered with X-ray synchrotron radiation, and also by the ease of determining molecular weights. Most importantly, to a good approximation, the hydration shell is not visible in neutron scattering. This affects the observed neutron scattering curve and in turn simplifies the application of structural modelling strategies.

The modelling of a scattering curve requires both a known molecular structure for the macromolecule as well as one for the hydration shell surrounding it. This hydration shell is sufficiently large to be detectable in the scattering experiment, and the evidence relevant to this is reviewed first. As scattering experiments are based on the observation of the difference in electron or nuclear scattering density between the macromolecule and solvent, the observation of hydration shells is next discussed in terms of the calculation of partial specific volumes and scattering densities from known compositions. This provides a molecular description for the analysis of the hydration shell, this being the primary viewpoint of this article. Finally, methods are described for the addition of a hydration shell to the structure whose scattering curve is to be modelled, together with the outcome of X-ray and neutron calibration tests that validate these hydration methods.

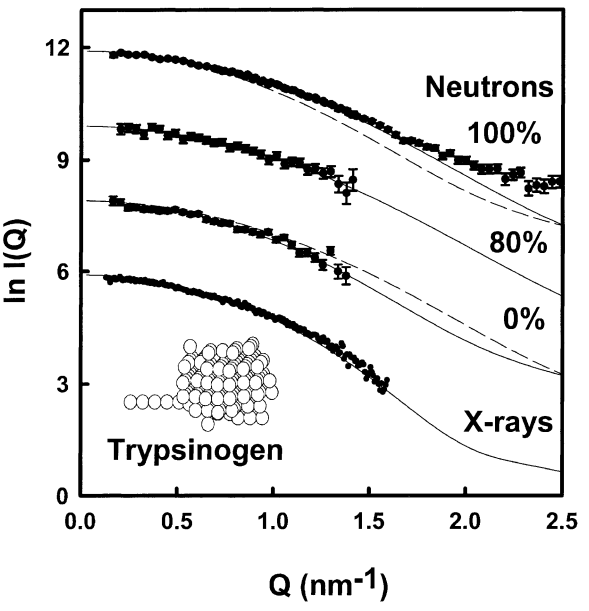

Fig. 2. Calibration studies of scattering curves based on the crystal structure of bovine trypsinogen. The neutron curve fits in three contrasts were achieved using an unhydrated structure and a two-density sphere model. The contrast dependence of the neutron curves is shown by the dashed lines showing the 0% ^2H_2O curve fit in comparison with the 100% ^2H_2O data, and the 100% ^2H_2O curve fit in comparison with the 0% ^2H_2O data. The X-ray curve fit was achieved using a hydrated structure with 0.3 g H_2O/g protein and a single density model. Adapted from Perkins et al. [12].

2. Observation of hydration shells by scattering data analysis

The usual assumption of 0.3 g H_2O/g protein leads to a hydration shell whose volume is approximately 33% of the protein volume. This is therefore detectable by a low resolution method such as scattering. A hydration shell corresponds approximately to the mass of a monolayer of water molecules surrounding the protein (see below). To place this volume change in perspective, it is useful to note that a probe of radius 0.14 nm (diameter 0.28 nm) is often used to analyse solvent-accessible sidechains at protein surfaces [5], and that the mean distance between water molecules in bulk water is 0.31 nm [6]. This means that the monolayer of water molecules surrounding a protein will increase each of the triaxial dimensions of the protein by approximately 0.6 nm.

In typical experimental scattering data acquisitions, data are recorded using scattering cameras located at the high-flux X-ray and neutron sources at the SRS Daresbury synchrotron, ISIS Rutherford-Appleton pulsed neutron source, or the Institut Laue Langevin neutron reactor [3]. Purified protein samples are well-dialysed against their buffers before the experiment. The scattering from both the sample and buffers are measured in alternation. After normalisation, data reduction involves the subtraction of the buffer run

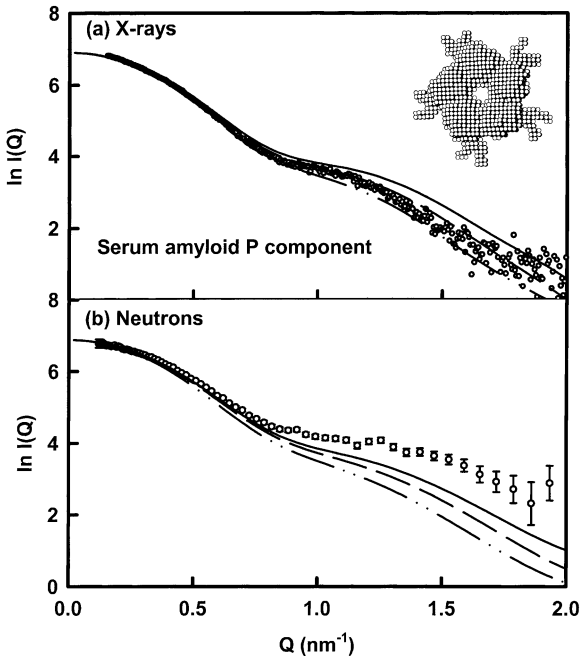

Fig. 3. Calibration studies of scattering curves based on the crystal structure of pentameric serum amyloid P component. Five oligosaccharide chains were added at the periphery of the structure, the sphere model for which is shown as an inset. Adapted from Ashton et al. [13]. (a) The X-ray curve (open circles) is shown in comparison with the calculated curve based on a hydration of 0.3 g H_2O/g glycoprotein (dashed line) and with two other calculated curves based on hydrations of 0.0 g H_2O/g glycoprotein (solid line) and 0.6 g H_2O/g glycoprotein (dashed–dotted line). (b) The neutron curve measured using 100% ^2H_2O buffers on instrument LOQ (open circles) is shown in comparison with the calculated curve based on a hydration of 0.0 g H_2O/g glycoprotein (solid line) and with two other calculated curves based on hydrations of 0.3 g H_2O/g glycoprotein (dashed line) and 0.6 g H_2O/g glycoprotein (dashed–dotted line). The neutron curve deviation at large Q is attributed to a uniform background from incoherent scattering due to protons in the sample.

(which includes any scattering arising from the cell and its holder) from the sample run. This gives the scattering curve intensities $I(Q)$ as a function of Q, the scattering vector ($Q = 4\pi \sin \theta/\lambda$; 2θ = scattering angle; λ = wavelength), as illustrated in Figs. 1–3.

Scattering data analysis reveals evidence for a hydration shell as follows:

1. Guinier analyses of ln $I(Q)$ vs. Q^2 at low Q values (in the Q range between approx. 0.1 and 0.5 nm^{-1}, depending on the macromolecular dimensions) will result in a radius of gyration (R_G) and the intensity of the forward scattering at zero Q (the $I(0)$ value) (Fig. 1). Since the overall macromolecular dimensions are monitored by the R_G value, and these dimensions can be large, the R_G parameter on its own is not sensitive to the presence or absence of a surface layer of water molecules of thickness 0.3 nm.
2. If the macromolecule is elongated, cross-sectional Guinier analyses of ln $[I(Q)Q]$ vs. Q^2 plots at intermediate Q values (between Q values of approx. 0.3 and 1.0 nm^{-1} for antibody molecules [7]) will give the radius of gyration of the cross-section R_{XS} (Fig. 1). In this Q range, the dimensions that are probed by the scattering vectors are shorter, and more comparable with those of the hydration shell. The cross-sectional Guinier analyses show that the X-ray R_{XS} values are often larger than those of the neutron R_{XS} values measured in ^2H_2O buffers. One explanation of this difference is that it corresponds to the effect of measuring the protein in high positive and negative solute–solvent contrasts. This explanation would, however, require very large scattering density inhomogeneities within the protein to become observable. The more plausible alternative explanation is that the difference reflects a hydration shell surrounding the protein that is seen by X-rays but not by neutron scattering.
3. The Fourier transformation of the full $I(Q)$ curve between Q of 0.05 and 2 nm^{-1} will yield the distant distribution function $P(r)$. This corresponds to the summation of all the distances r between all the atoms within the macromolecule, weighted by their scattering densities. The $P(r)$ curve decreases to zero at the maximum dimension L of the macromolecule. As the maximum dimension of a hydrated macromolecule will be increased by 0.6 nm compared to one in which the hydration shell is not visible, the hydration shell is in principle detectable by comparing the X-ray and neutron $P(r)$ curves if the shell is not

visible by neutron scattering. In practice, as the errors in $P(r)$ are higher at large r values and L is generally determined to approximately? 1 nm, this small predicted difference in L is not measurable.

4. The dependence of the neutron $I(0)$ values on the ratio of light and heavy water present in the buffer provides the clearest evidence that neutron scattering views a protein as an unhydrated macromolecule [6]. An explanation of this requires the corresponding analysis of partial specific volumes, which is considered in the next section.
5. Absolute molecular weight calculations of proteins can be made from neutron $I(0)/c$ values if they are normalised using the protein concentration c and referenced to the isotropic incoherent scattering of a water standard [8]. These result in the sequence-derived molecular weight (i.e. not that including the molecular weight of the hydration shell), indicating that it is the unhydrated protein that is studied by neutrons. Relative X-ray $I(0)$ values can be measured for molecular weight calculations, and will correspond to hydrated molecular weights. Since the standard itself is hydrated, it is not possible to identify a hydrated molecular weight as such from this comparison.

3. Calculation of partial specific volumes from sequences

Depending on how the hydration shell is incorporated in these, five different methods to calculate the partial specific volume $\bar{v}$ of a protein yields different values. To evaluate this, the amino acid residue volumes in nm^3 were derived from each approach, then these were used to predict the protein $\bar{v}$ values for comparison with a set of 12 experimentally-measured $\bar{v}$ values [6]. The null difference in predicted and experimental $\bar{v}$ values showed that the Cohn and Edsall values based on densitometric and molar group volumes gave fortuitously good agreement with experiment. The Zamyatnin volumes based on densitometric work with the amino acids systematically underestimated the experimental $\bar{v}$ by 0.01–0.02 ml/g. The Richards volume based on molar group summations and the Chothia and Perkins volume based on protein and amino acid crystal structures, respectively, all overestimated the $\bar{v}$ values by 0.04–0.05 ml/g. Since the error of the experimental $\bar{v}$ values is 0.01 ml/g, the differences between the densitometric and crystallographic approaches are significant. This analysis of these different approaches resulted in a consensus set of $\bar{v}$ values for the 20 amino acid residues for use in predicting $\bar{v}$ values for proteins and glycoproteins.

The hydration shell accounts for the above differences in $\bar{v}$ values once it is noted that the volume behaviour of water molecules is altered in different environments. From the density of water at 4°C, the volume of a free water molecule is calculated to be 0.0299 nm^3. The volume properties of water are well known. For example, on freezing, the mean water molecule volume in the I_h and I_c forms of ice is increased to 0.0323 nm^3. Alterations in the hydrogen bonding arrangements between water molecules thus have a large effect on their volume. Hence, in bulk liquid water or at the protein surface, the water molecule volume will be determined by its hydrogen bonding. Here, it is interesting that the mean water molecule volume in hydrated amino acid crystal structures was calculated to be 0.0245 nm^3 [6]. This is less than that of a free water molecule. An independent calculation of this volume based on 393 water molecules sandwiched between the protein surface and a second layer of water molecules also gave 0.0245 nm^3 [9]. So the ordered positioning of water molecules at the protein surface is seen to reduce its volume. If the volume of a water molecule is 0.0245 nm^3 in a typical hydration shell, while that of bulk water is 0.0299 nm^3, the variation in $\bar{v}$ values in Table 1 can be explained. The densitometry $\bar{v}$ values for amino acids are too low when used to calculate protein $\bar{v}$ values because the free amino acids bind water molecules that are not present in the folded protein, while the crystallography $\bar{v}$ values are too high when used to calculate protein $\bar{v}$ values

Table 1
Partial specific volumes of proteins calculated by five methods

Authors	Method	Difference between observed and calculated $\bar{v}$ values (ml/g)[a]
Cohn and Edsall, 1943	Densitometry/ molar groups	0.00
Zamyatnin, 1972, 1984	Densitometry	−0.01, −0.02
Richards, 1974	Molar groups	0.04
Chothia, 1975	Protein crystal sturctures	0.04
Perkins, 1986	Small crystal structures	0.05
Perkins, 1986	Consensus	0.00

[a]The full calculations are given in Table 1 of Perkins [6].

because they are determined independently of bound water molecules.

Hydration shells are directly observable in protein crystal structures that are refined at a high enough resolution. The water molecules at their surfaces are in well-defined 'electrostricted' hydrogen-bonded positions, unlike the loose ones that are typical of bulk water. The second column of Table 2 summarises the total water molecules observed in the first hydration shell immediately adjacent to the protein surface and beyond the first shell. The total corresponds to hydration levels of 0.16–0.37 g H_2O/g protein. The total of bound water molecules required to correct the protein $\bar{v}$ value calculated from crystallographic volumes to the consensus $\bar{v}$ value is shown in the third column. This total is derived from the water molecule volume change from 0.0299 to 0.0245 nm^3, and resulted in hydration levels of 0.20–0.28 g H_2O/g protein. The two estimates are comparable with each other. This indicates that an electrostricted, denser monolayer of water molecules surrounding a protein corresponds to the widely-assumed hydration level of 0.3 g H_2O/g protein, i.e. this concept provides a molecular view of hydration.

4. Calculation of nuclear and electron scattering densities

The lack of effect of the hydration shell in neutron scattering experiments is best observed when scattering densities are determined from matchpoint calculations. The observed value of $\sqrt{I(0)}$ depends on the scattering density of the solvent, which varies from -56×10^{-6} nm^{-2} for H_2O to 640×10^{-6} nm^{-2} for ^2H_2O. A graph of the normalised $I(0)$ values against the volume fraction of heavy water is linear. The $\sqrt{I(0)}$ value disappears at the matchpoint when the scattering density of the heavy water buffer becomes the same as the protein with its hydration shell, typically at approximately 40% ^2H_2O. This means that the experimental neutron matchpoint value for a protein can be compared with the predicted matchpoint calculated from the five $\bar{v}$ approaches of Table 1 [6]. This prediction corresponds to the

Table 2
Observed and predicted contents of water molecules in crystal structures

Protein	Observed water (crystal)[a]	Predicted water[b]
BPTI	63 (43)	76–93
Rubredoxin	123 (80)	68–83
Erythrocrurin	– (111)	171–223
Actinidin	272 (163)	289–326
Penicillopepsin	319 (264)	440–528
Lysozyme (human)	140 (95)	178–188
Lysozyme (tortoise)	128 (90)	162–195

[a]Unbracketted: total of water molecules observed in the crystal structure. Bracketted: water molecules in the first hydration shell.

[b]These are the total number of water molecules of volume 0.0245 nm^3 that correspond to a hydration of 0.20–0.28 g H_2O/g protein.

Table 3
Neutron scattering matchpoints for 11 proteins calculated by three methods[a]

Authors	Method	Difference between observed and calculated matchpoint (% ^2H_2O)
Zamyatnin, 1972, 1984	Densitometry	2.7 ± 1.0
Chothia, 1975; Perkins, 1986	Crystal structures	-0.2 ± 0.7
Perkins, 1986	Consensus $\bar{v}$	1.9 ± 1.0

[a] The full calculations are given in Table 5 of Perkins [6].

total of neutron scattering lengths for all the atoms in the protein divided by the protein volume. As the scattering lengths of ^1H and ^2H nuclei are very different, the prediction needs an estimate of the degree of non-exchange of the exchangeable protons at NH and OH groups that are buried within the protein core. This exchange is easily defined by ^1H-NMR, being low at approximately 10% of all the mainchain exchangeable protons. Consequently Table 3 shows that the predicted matchpoint based on the crystal structure $\bar{v}$ values shows good agreement with the experimental matchpoints, but those based on the densitometric and the consensus $\bar{v}$ values deviate from them. This shows that the hydration shell known to be present from partial specific volume and crystallographic studies has no effect on matchpoint predictions, i.e. it is invisible.

The apparent absence of the hydration shell in neutron scattering can be explained by calculations of its scattering density [10]. The neutron experiment observes the difference in scattering density between the protein and its solvent. For bulk H_2O, the neutron scattering density is given by the summation of scattering lengths divided by its volume to give -56×10^{-6} nm^{-2} (Table 4). The equivalent value for the hydration shell is -69×10^{-6} nm^{-2}, and that for a protein is 19×10^{-6} nm^{-2}. So the scattering density of the hydration shell differs by -23% from that of bulk water, while that of the protein differs by $+134\%$ in the opposite direction. It should be noted that this difference of $+134\%$ is an average for the amino acids in the protein, given that their scattering densities vary between 8×10^{-6} nm^{-2} (Ile) to 31×10^{-6} nm^{-2} (Asp), i.e. a difference of $+114$ to $+155\%$ relative to that of bulk water. Two consequences follow from these calculations. Firstly, it is not possible for the neutron scattering density of the hydration shell to even approximate that of the protein, so the shell will make a negligible contribution to the neutron scattering

Table 4
Scattering densities of hydration shells

Technique	Component	Scattering density[a]	Percentage difference from free solvent
Neutron scattering in H_2O	Free H_2O	-56×10^{-6} nm^{-2}	–
	Bound H_2O	-69×10^{-6} nm^{-2}	-23%
	Protein	19×10^{-6} nm^{-2}	$+134\%$
Neutron scattering in ^2H_2O	Free ^2H_2O	640×10^{-6} nm^{-2}	–
	Bound ^2H_2O	782×10^{-6} nm^{-2}	$+22\%$
	Protein	32×10^{-6} nm^{-2}	-95%
X-Ray scattering in H_2O	Free water	334 e nm^{-3}	–
	Bound water	408 e nm^{-3}	$+22\%$
	Protein	425 e nm^{-3}	$+27\%$

[a] The data used to calculate values for this Table are taken from Perkins [10].

properties of the protein. Secondly, since the 23% difference in scattering density between bulk and hydration water is similar to the 31% variation in scattering densities within the common 20 amino acids found within the protein, it is not possible to distinguish between two different types of water in the same way that it is not possible to distinguish between the 20 amino acids. The net result is that, to a good approximation, the hydration shell surrounding a protein is not detectable in a neutron experiment with H_2O buffer.

The apparent absence of the hydration shell in neutron scattering experiments performed in ^2H_2O buffers is likewise explained by scattering density calculations, in which the effect of full $^1H-^2H$ exchange in the hydration shell also makes this undetectable [10]. The scattering density of bulk heavy water is 640×10^{-6} nm^{-2}, and that of the deuterated hydration shell is +22% different from this at 782×10^{-6} nm^{-2} (Table 4). The scattering density of the protein has now increased to 32×10^{-6} nm^{-2} for reason of $^1H-^2H$ exchange of its exchangeable protons. This is now −95% different, again in the opposite direction from that of bulk water, and varies in the range between −92 and −98% if the individual amino acids are considered. Again, the large contrast difference between the protein and heavy water components means that, to a good approximation, the hydration shell is also not detectable.

The calculation of electron densities is different in that the hydration shell of a protein is shown to be detectable by X-ray scattering [10]. The electron density of bulk water is 334 e nm^{-3}. This is increased to 408 e nm^{-3} in the protein hydration shell, a +22% increase compared to bulk water (Table 4). The hydration shell electron density is now close to that of a protein which is 425 e nm^{-3}, or +27% higher than that of bulk water. The electron density of the individual amino acids varies in the range between +10% and +51% of that of bulk water. These calculations show that the electron density of the hydration shell is indistinguishable from those of the 20 common amino acids, and it will therefore contribute to the protein scattering curve and be detectable.

5. Methods to add a hydration shell to a protein structure

The constrained automated modelling of the X-ray and neutron scattering curve $I(Q)$ in terms of known atomic structures [4] uses small spheres of uniform density to represent the known protein crystal or NMR subunit structures. This is a versatile and direct approach that allows the irregularities in the protein surface to be represented in the simulation. Three other major curve modelling methods which start from either ellipsoids, spherical harmonics or genetic algorithms do not have this advantage. In constrained modelling, the hydration shell is not present in the known structure and has to be added (see below). Initially, the total volume of the spheres must be the same as that of the dry protein, i.e. the volume as seen in the crystal structure. This is achieved by a grid transformation, in which the atomic coordinates for the protein of interest are placed within a three-dimensional grid of cubes. A sphere of equal volume to the cube is placed at the centre of each cube if a user-specified cutoff for the minimum number of atoms contained within a cube is satisfied. A cube of side length of 0.5 or 0.6 nm in combination with a cutoff of four atoms often produces sphere models within 2% of the protein dry volume calculated from its composition. If residues are missing from the crystal structure, they are either replaced in their approximate positions by standard molecular modelling approaches, or the cube side and cutoff are adjusted so that the resulting spheres correspond to the full dry volume. For neutron work, the scattering curve is directly calculated from these spheres using Debye's Law adapted to small spheres [1,2]. The reason for the use of small spheres is that their form factor in the Debye equation becomes almost invariant in the Q range used for the simulation of $I(Q)$.

Two methods have been used to incorporate the hydration shell detected by X-ray scattering. In the first one, for compact proteins such as trypsinogen (Fig. 2), a linear expansion is applied to the radius of the spheres and their coordinates

using a factor of approximately 1.09–1.10 [11,12], in order that the total volume of the spheres becomes the sum of the dry protein volume and the volume of the hydration shell. The latter is calculated assuming a hydration of 0.3 g H_2O/g glycoprotein and a water volume of 0.0245 nm^3. This procedure is, however, unsatisfactory if the protein has a void space at its centre, as this void becomes larger on expansion, while hydration is expected to make this smaller. Examples include the centre of the Fc fragment of antibodies (Fig. 1) and the pentameric discoid structure of serum amyloid P component (Fig. 3). In the second method, the proteins are hydrated using an improved approach HYPRO, in which hydration was achieved by the addition of spheres to the protein surface [7,13]. The algorithm is based on the addition of 26 cubes around each of the cubes in the dry model. Their size of 0.5–0.6 nm is comparable to the thickness of the hydration shell. In principal, this addition is adjustable to add any number of between 1 and 26 cubes per cube in order to achieve the desired hydration level. The resulting model contains a large number of duplicate cubes. This redundancy is eliminated by a reapplication of the original grid transformation in which the origin of the grid is now displaced by 0.1 nm. This transformation is also adjustable by varying the cutoff for the number of duplicate cubes required in order to optimise the desired hydrated volume. If higher hydration levels are needed, HYPRO can be reapplied to the hydrated model, recycling as often as required.

6. Tests of hydration shell modelling by scattering and sedimentation analyses

A number of calibration studies support the premise that the hydration shell is undetectable by neutron scattering but is detectable by X-ray scattering [11–13]. The crystal structures for α-chymotrypsin, β-trypsin, α_1-antitrypsin and pentameric serum amyloid P component are suitable for these as they correspond to protein structures that are rigid and well-defined in solution and span a wide molecular weight range. These provide a suitable basis for comparison with scattering data. At the low molecular weight end, Fig. 2 illustrates the outcome of these tests for bovine trypsinogen (molecular weight 23 900) [12]. In the case of the neutron data, a clear contrast dependence was observed, and this was allowed for by the use of two-density models. No hydration shell was required for any of the fits in three contrasts. In distinction to this, an excellent X-ray curve fit for a single-density hydrated model of trypsinogen was obtained, in which the hydration was set as 0.3 g H_2O/g protein. At the high molecular weight end, internal density fluctuations become less important as they are averaged out over a much larger structure. Only single-density models were therefore needed to fit the scattering curves. Fig. 3 illustrates the X-ray and neutron curve fits with hydrated and unhydrated structures for the serum amyloid P component (molecular weight 127 000) using these single-density models [13]. All the curves in Fig. 3 are normalised to an $I(0)$ value of 1000 for ease of comparison. As the hydration is increased, the protein becomes larger, meaning that the predicted scattering curve decreases more rapidly at larger Q (cf. the discussion of R_{XS} values in Section 2(ii)). The X-ray fit in Fig. 3a was best achieved with a hydration of 0.3 g H_2O/g glycoprotein over the full Q range (dashed line). Even though a contribution from a flat background of incoherent scattering is evident at large Q in Fig. 3b, and uncorrected, the neutron curve at low Q is seen to be best fitted to an unhydrated glycoprotein structure (solid line).

The hydrated small sphere models used for X-ray scattering simulations can also be used for the calculation of frictional coefficients. This means that ultracentrifugation as well as scattering data can be used to refine a given sphere model for a protein. Hydration shells influence the calculation of the sedimentation coefficient in two ways. Firstly, it affects the value of the protein $\bar{v}$ value required to convert the frictional coefficient to the sedimentation coefficient (see above). Secondly, it affects the expansion of the protein structure in order to allow for a monolayer of water molecules of thickness 0.3 nm. Variations in the hydration level of 0.0 or 0.6 g H_2O/g

protein in a study of the serine proteases showed that the sedimentation coefficient is altered by ± 0.2 S [12].

7. Conclusions and future perspectives

The proposal that water in a hydration shell is denser than bulk water is supported by several approaches including densitometry and scattering. The main conclusions from the scattering data and their modelling can be summarised as follows:

1. A hydration shell of 0.3 g H_2O/g protein that is denser than bulk water explains the experimentally-observed partial specific volumes of proteins when this is combined with the crystallographic volume of the protein.
2. The hydration shell is not detectable in experimentally-observed neutron matchpoint determinations of proteins. This is explained by the calculation of similar neutron scattering densities of bulk and hydration water assuming full $^1H-^2H$ exchange, both of which are similar, but significantly different from that of the protein.
3. In X-ray scattering, calculations show that the hydration shell has a similar electron density to that of the protein, and is detectable by X-ray scattering because both values are different from that of bulk water.
4. It is necessary to include the hydration shell for the modelling of X-ray scattering curve fits based on crystal structures, but not so in the analogous neutron scattering curve fits.

Scattering analyses of hydration shells by others generally conclude that the hydration shell is visible in X-ray scattering. For example, in the analysis of Svergun and Zaccai [14], the key experimental observations are the same as in our calibration studies [11–13], namely that the X-ray scattering curve for a given protein corresponds to a physically larger object than that from the neutron scattering curve. This agrees with the observation of a hydration shell by X-ray scattering that is not seen by neutron scattering. Interestingly a multi-density curve fit approach was used with both the X-ray and neutron data. This showed that a hydration shell that is denser than that of bulk water could be identified from both the X-ray and neutron data, even though as expected the weighting accorded to the hydration shell is noticeably weaker for the neutron data compared to the X-ray data.

There are several debates in relation to hydration shells. Different estimates of the level of excess density of the hydration shell have been proposed, where an increase of between 3 and 20% was determined by scattering in Svergun et al. [14] and by others, while we have estimated an increase of 22–23% (Table 4). Our approach is based on the molecular consequences of a single monolayer of water molecules at 0.3 g H_2O/g glycoprotein, which has either the same scattering density as that of the glycoprotein (X-rays) or is not (neutrons). While this outcome is most useful for the automation of our constrained modelling strategies [7,13], this only constitutes a good approximation. Different views have been proposed for the distribution of the hydration shell. Thus, it is possible that bound water in apposition to surface charged groups may differ in density from water that is apposed to surface non-polar groups. It is also probable that further layers of more loosely-bound water molecules may be attached to the primary hydration shell. While more sophisticated modelling algorithms may enable these factors to be considered, it is not clear whether they will have a significant affect on the precision of the modelling fits.

Acknowledgements

The Wellcome Trust, the Biotechnology and Biological Sciences Research Council, and the Clement Wheeler-Bennett Trust are thanked for grant support for the many studies used to derive the results mentioned here.

References

[1] S.J. Perkins, A.W. Ashton, M.K. Boehm, D. Chamberlain, Molecular structures from low angle X-ray and neutron scattering studies, Int. J. Biol. Macromol. 22 (1998) 1–16.

[2] S.J. Perkins, C.G. Ullman, N.C. Brissett, D. Chamberlain, M.K. Boehm, Analogy and solution scattering modelling: new structural strategies for the multidomain proteins of complement, cartilage and the immunoglobulin superfamily, Immunol. Rev. 163 (1998) 237–250.

[3] S.J. Perkins, Protein–ligand interactions: A practical approach in: B. Chowdhry, S.E. Harding (Eds.), High-Flux X-Ray and Neutron Scattering Studies, 1, Oxford University Press, Oxford, 2000, pp. 223–262.

[4] S.J. Perkins, Applications of highly constrained molecular modelling scattering curve fits to biologically important proteins, Fibre Diffraction Rev. (2001) (in press).

[5] B. Lee, F.M. Richards, The interpretation of protein structures: estimation of static accessibility, J. Mol. Biol. 55 (1971) 379–400.

[6] S.J. Perkins, Protein volumes and hydration effects: the calculation of partial specific volumes, neutron scattering matchpoints and 280 nm absorption coefficients for proteins and glycoproteins from amino acid sequences, Eur. J. Biochem. 157 (1986) 169–180.

[7] M.K. Boehm, J.M. Woof, M.A. Kerr, S.J. Perkins, The Fab and Fc fragments of IgA1 exhibit a different arrangement from that in IgG: a study by X-ray and neutron solution scattering and homology modelling, J. Mol. Biol. 286 (1999) 1421–1447.

[8] B. Jacrot, G. Zaccai, Determination of molecular weight by neutron scattering, Biopolymers 20 (1981) 2413–2426.

[9] M. Gerstein, C. Chothia, Packing at the protein–water interface, Proc. Natl. Acad. Sci. USA 93 (1996) 10167–10172.

[10] S.J. Perkins, New comprehensive biochemistry in: A. Neuberger, L.L.M. Deenen (Eds.), X-Ray and Neutron Solution Scattering, 11B, Elsevier, Amsterdam, 1988, pp. 143–265.

[11] K.F. Smith, R.A. Harrison, S.J. Perkins, Structural comparisons of the native and reaction centre cleaved forms of α_1-antitrypsin by neutron and X-ray solution scattering, Biochem. J. 267 (1990) 203–212.

[12] S.J. Perkins, K.F. Smith, J.M. Kilpatrick, J.E. Volanakis, R.B. Sim, Modelling of the serine protease fold by X-ray and neutron scattering and sedimentation analyses: its occurrence in factor D of the complement system, Biochem. J. 295 (1993) 87–99.

[13] A.W. Ashton, M.K. Boehm, J.R. Gallimore, M.B. Pepys, S.J. Perkins, Pentameric and decameric structures in solution of the serum amyloid P component by X-ray and neutron scattering and molecular modelling analyses, J. Mol. Biol. 272 (1997) 408–422.

[14] D.I. Svergun, S. Richards, M.H.J. Koch, Z. Sayers, S. Kuprin, G. Zaccai, Protein hydration in solution: Experimental observation by X-ray and neutron scattering, Proc. Natl. Acad. Sci. USA 95 (1998) 2267–2272.