ARTICLE

Helmut Durchschlag · Peter Zipper

Modeling the hydration of proteins: prediction of structural and hydrodynamic parameters from X-ray diffraction and scattering data

Abstract The implications of protein-water interactions are of importance for understanding the solution behavior of proteins and for analyzing the fine structure of proteins in aqueous solution. Starting from the atomic coordinates, by bead modeling the scattering and hydrodynamic properties of proteins can be predicted reliably (Debye modeling, program HYDRO). By advanced modeling techniques the hydration can be taken into account appropriately: by some kind of rescaling procedures, by modeling a water shell, by iterative comparisons to experimental scattering curves (ab initio modeling) or by special hydration algorithms. In the latter case, the surface topography of proteins is visualized in terms of dot surface points, and the normal vectors to these points are used to construct starting points for placing water molecules in definite positions on the protein envelope. Bead modeling may then be used for shaping the individual atomic or amino acid residues and also for individual water molecules. Among the tuning parameters, the choice of the scaling factor for amino acid hydration and of the molecular volume of bound water turned out to be crucial. The number and position of bound water molecules created by our hydration modeling program HYDCRYST were compared with those derived from X-ray crystallography, and the capability to predict hydration, structural and hydrodynamic parameters (hydrated volume, radius of gyration, translational diffusion and sedimentation coefficients) was compared with the findings generated by the water-shell approach CRYSOL. If the atomic coordinates are unknown, ab initio modeling approaches based on experimental scattering curves can provide model structures for hydrodynamic predictions.

Keywords Hydrodynamics · Modeling · Protein hydration · Small-angle X-ray scattering · X-ray diffraction

Abbreviations AA: amino acid · CS: citrate synthase · EM: electron microscopy · GA: genetic algorithm · IM: initial model · MB: multibody · MS: malate synthase · OE: oblate ellipsoid of revolution · PDB: Protein Data Bank · PE: prolate ellipsoid of revolution · RdM: reduced model · SAXS: small-angle X-ray scattering · WB: whole body

H. Durchschlag (✉)
Institute of Biophysics and Physical Biochemistry,
University of Regensburg, Universitätsstrasse 31,
93040 Regensburg, Germany
E-mail: helmut.durchschlag@biologie.uni-regensburg.de
Tel.: +49-941-9433041
Fax: +49-941-9432813

P. Zipper
Physical Chemistry, Institute of Chemistry,
University of Graz, Heinrichstrasse 28, 8010 Graz, Austria

Introduction

Proteins are the functional forms of polypeptides, built up from 20 different amino acids (AAs) linked by covalent bonds. The biological activity of a protein depends on its three-dimensional (3D) structure, which in turn is determined by its AA sequence, and manifold interactions with other molecules including solvent components (Creighton 1993; Van Holde et al. 1998).

Among the naturally occurring proteins, the water-soluble globular proteins are important representatives; albumins and enzymes belong to this group. Proteins in solution are neither rigid nor motionless, since the bonds in the backbone and the AA side chains allow considerable flexibility; these motions are of fundamental importance for their functional role, e.g. for enzymatic activity. Hydrophilic proteins are surrounded by some kind of hydration layer. Since all living systems exist in largely aqueous media, there is utmost interest in the structural organization of water molecules within and around proteins. Water may be regarded as the "lubricant of life", easing necessary hydrogen bonding

changes: it stabilizes biological structures and allows motions of its constituents.

The solubility of a protein is determined by the composition and distribution of polar and nonpolar AAs on its surface and the surrounding milieu (pH, presence of additives, ionic strength, temperature, pressure, etc.). Polar protein side chains interact more with water molecules than nonpolar chains (Thanki et al. 1988). The extent of water binding is highly influenced by the protein surface topography; ordered waters are three times more likely to be found in solvent-accessible surface grooves than elsewhere on the protein surface; even deep surface grooves typically have the width of a water molecule (Kuhn et al. 1992). Obviously, such crevices represent local energy minima. In crystals, however, the ordered waters, except those in crevices, seem essentially caused by the geometry of the crystal lattice (Levitt and Park 1993). The internal water sites are well conserved in different crystal modifications (Blake et al. 1983), but only part of the surface waters are suggested to be conserved in different crystal forms (Zhang and Matthews 1994). Though, at present, most biophysicists accept the existence of well-ordered water molecules, the hydration problem is still a matter of controversy. With all techniques the identification of bound waters is somewhat subjective, in particular when they are located at the protein surface (e.g., correct choice of residence times and electron density thresholds). In this context it has also to be mentioned that a few authors even deny the requirement of a layer of solvating water to explain the hydrodynamic properties of proteins; for example, the dynamics simulations by Smith and van Gunsteren (1994) invoke the absence of any water of hydration associated with the proteins.

Several types of locally organized water molecules have been described (Kuntz and Kauzmann 1974; Hopfinger 1977; Rashin et al. 1986; Rupley and Careri 1991; Teeter 1991; Kakalis and Kumosinski 1992; Westhof 1993; Gregory 1995; Schoenborn et al. 1995; Finney 1996; Harding 2001a, 2001b; Rowe 2001), in addition to the water molecules of bulk water. Apart from a few "tightly bound" water molecules (e.g., located within the cavity of an active site of an enzyme or in the interior of a protein), a magnitude of partially localized ("bound") water molecules has been identified by several physicochemical techniques, including spectroscopic, thermodynamic, hydrodynamic, X-ray and neutron scattering and diffraction methods. These preferentially bound water molecules are positioned in the first layer of water on the protein surface, they have properties different from those of the bulk water, they are more ordered, less mobile and have, on the average, a higher density than bulk water (Perkins 1986, 2001; Creighton 1993; Gerstein et al. 1995; Gerstein and Chothia 1996; Murphy et al. 1998; Svergun et al. 1998; Ebel et al. 2000; Durchschlag and Zipper 2001; Chaplin 2002; Henchman and McCammon 2002; Merzel and Smith 2002; Smith et al. 2002). In solution, all bound water molecules exchange rapidly with bulk water, with residence times differing over a wide range. As follows from high-resolution NMR studies (Otting et al. 1991; Wüthrich et al. 1992, 1996; Brunne et al. 1993; Otting and Liepinsh 1995; Denisov and Halle 1996), the residence times of waters in the protein interior are in the range of about 10^{-2}-10^{-8} s, while the hydration of the protein surface is characterized by waters in the subnanosecond time scale. By contrast, the motions of the water molecules of the bulk phase (approx. 1 ps) are about one or two orders of magnitude faster (Rupley and Careri 1991; Pal et al. 2002). The transition state between bound and bulk water ("transition water") is identical to a secondary layer of water, which is, however, difficult to identify since the properties of the transition water are only marginally different from those of bulk water (perturbations decay exponentially with the distance from the surface).

A comparison of the results of various physicochemical and modeling techniques has shown that a hydration of 0.35 g water per gram of protein may be considered to be a reasonable default (average) value for the hydration of an average globular protein, corresponding to approximately two water molecules per AA residue (Durchschlag and Zipper 2001). Though it is not a priori clear if different techniques measure the same water entity, a critical comparison of the results obtained from different solution techniques rather points to differences in the assumptions underlying the interpretation of hydration than to real differences. In contrast to earlier interpretations of hydrodynamic data, which indicated rather high levels of hydration (0.53-0.54 g/g: Kuntz and Kauzmann 1974; Squire and Himmel 1979), present studies (including electrostatic and X-ray solution scattering and diffraction data, in addition to hydrodynamic approaches) rather seem to point to a unified picture of protein hydration (Zhou 1995, 2001; Durchschlag and Zipper 2001). However, only a part of the hydration waters seems to be observed by X-ray crystallography: between one half and two ordered water molecules per AA residue are found in protein crystal structures, the observed number strongly dependent on the resolution and state of refinement of the particular crystal structure (Blake et al. 1983; Levitt and Park 1993; Schoenborn et al. 1995; Carugo and Bordo 1999). In the context of the X-ray detectable water molecules, several important questions should be addressed. (1) Are the positions of individual hydration water molecules the same in protein crystals and aqueous solution, i.e. are they assigned to the same AA residues? (2) Is the entire surface of proteins covered with preferentially bound waters? (3) Why is there a discrepancy in the amount of hydration observed by X-ray crystallography and the solution techniques? (4) Is there a distinction in the properties of hydration waters seen by crystallography and solution techniques, respectively?

Modeling of protein structures in aqueous solution for scattering and hydrodynamic purposes involves modeling of both the contributions of AA residues and

water molecules. In the context of hydration contributions, we have to focus our attention mainly on the preferentially bound water molecules, since (because of their extraordinary features) only these contribute to the observable scattering and hydrodynamic behavior. This may be accomplished by quite different approaches:

1. Whole-body (WB) approaches (e.g., Kumosinski and Pessen 1982; Harding 1989, 1995; Durchschlag and Zipper 1997a, 1997b, 1999, 2002b; Winzor et al. 2001).
2. Multibody (MB) approaches: "bead modeling" in the classical sense (e.g., Durchschlag 1975; García de la Torre and Bloomfield 1981; García de la Torre 1989; García de la Torre et al. 1997; Carrasco and García de la Torre 1999; Byron 2000; Behlke 2001; Durchschlag and Zipper 2002b).
3. Use of atomic coordinates obtained from X-ray or neutron crystallography and high-resolution NMR spectroscopy (e.g., Müller 1983, 1991; Durchschlag et al. 1991, 1996; Müller and Schrauber 1992; Beavil et al. 1995; Byron 1997, 2000; Spotorno et al. 1997; Zipper and Durchschlag 1997, 1998, 2000, 2002a, 2002b; Perkins et al. 1998; García de la Torre et al. 2000; García de la Torre 2001; Perkins 2001).
4. Use of density distributions obtained from 3D reconstructions from cryo-electron microscopy (e.g., Zipper and Durchschlag 2000, 2003; García de la Torre et al. 2001; Zipper et al. 2002).
5. Ab initio modeling approaches for the prediction of low-resolution 3D structures from scattering profiles (e.g., Chacón et al. 1998, 2000; Svergun 1999, 2000; Byron and Gilbert 2000; Walther et al. 2000; Zipper and Durchschlag 2003).

The present paper is concerned with different aspects of how hydration contributions are considered in the different approaches applied, with special reference to solution scattering and hydrodynamics. For our calculations we chose proteins of different size [monomeric lysozyme, dimeric citrate synthase (CS), trimeric malate synthase (MS)], all of them well characterized from the physicochemical point of view. In the case of lysozyme and CS, the crystallographic structures are known to high resolution.

For modeling individual, bound water molecules, the atomic coordinates in connection with the surface calculation algorithm SIMS (Vorobjev and Hermans 1997) and our hydration algorithm HYDCRYST (Durchschlag and Zipper 2002a) have been applied. Since the structural organization of water bound to lysozyme is described in some detail by high-resolution X-ray diffraction studies (Blake et al. 1983; Kodandapani et al. 1990; Niimura et al. 1997; Bon et al. 1999), this enzyme was chosen for a comparison with our approach of modeling individual water molecules. Moreover, this enzyme has already been the subject of several molecular dynamics simulations (e.g., Brooks and Karplus 1989; Smith and van Gunsteren 1994). CS was used to compare the results of our hydration strategy with the water-shell approach CRYSOL (Svergun et al. 1995), frequently used for the prediction of the SAXS behavior of proteins from atomic coordinates. Further, we used the ab initio modeling approach GASBOR (Svergun et al. 2001) to predict the solution structure and hydrodynamic properties for an enzyme (MS) whose precise crystal or EM structure is unknown to date. For the prediction of hydrodynamic parameters, in all cases modifications of the program HYDRO (García de la Torre et al. 1994) have been used.

Materials and methods

Sources of data

Atomic coordinates of proteins were obtained from the Protein Data Bank (PDB) (Berman et al. 2000) and molar masses from the AA sequence data as stored in the SWISS-PROT data bank (Bairoch and Apweiler 2000). Scattering and hydrodynamic parameters were extracted from the relevant literature. In particular, the following properties were used: atomic coordinates, scattering intensity $I(h)$, molar mass M, partial specific volume $\bar{v}$, radius of gyration R_G, hydrated volume V, amount of hydration δ_1, sedimentation coefficient s, and translational diffusion coefficient D.

Lysozyme from hen egg-white: atomic coordinates (Diamond 1974): PDB ID: 2LYZ; SWISS-PROT entry: P00698, monomer, $M = 14.313$ kg/mol; $\bar{v} = 0.702$ cm^3/g (Lee and Timasheff 1974); $R_G = 1.43$ nm and $V = 24.2$ nm^3 (Pessen et al. 1971); $\delta_1 = 0.317$ g/g (Pessen and Kumosinski 1985); $R_G = 1.38$ nm (Stuhrmann and Fuess 1976); $R_G = 1.43$ nm (Krigbaum and Kügler 1970); $R_G = 1.52 \pm 0.02$ nm and $D = 11.8 \times 10^{-7}$ cm^2/s (Luzzati et al. 1961); $R_G = 1.54 \pm 0.02$ nm (Svergun et al. 1998); $R_G = 1.55 \pm 0.05$ nm (Ducruix et al. 1996); $s = (1.87 \pm 0.02) \times 10^{-13}$ s and $D = (10.4 \pm 0.1) \times 10^{-7}$ cm^2/s (Colvin 1952); $s = 1.91 \times 10^{-13}$ s and $D = 11.2 \times 10^{-7}$ cm^2/s (Sophianopoulos et al. 1962); $D \approx 11.5 \times 10^{-7}$ cm^2/s (Dubin et al. 1967; Nicoli and Benedek 1976); $D = (10.6 \pm 0.2) \times 10^{-7}$ cm^2/s (Foord et al. 1970); $D = (10.6 \pm 0.1) \times 10^{-7}$ cm^2/s (Dubin et al. 1971, 1973); $D = 10.7 \times 10^{-7}$ cm^2/s (Barel et al. 1972); $D = 10.3$-11.5×10^{-7} cm^2/s (Mikol et al. 1990); $D = (10.2 \pm 0.4) \times 10^{-7}$ cm^2/s (Eberstein et al. 1994); $D = 10.5$-11.6×10^{-7} cm^2/s (Grigsby et al. 2000).

CS from pig heart: atomic coordinates (Remington et al. 1982): PDB ID: 1CTS; SWISS-PROT entry: P00889, homodimer, $M = 97.838$ kg/mol; $I(h)$, $\bar{v} = 0.740$ cm^3/g, $R_G = 2.91 \pm 0.02$ nm, $V = 174.4 \pm 2.0$ nm^3 and $\delta_1 = 0.339 \pm 0.011$ g/g (Durchschlag et al. 1991, 1996); $s = 6.0 \times 10^{-13}$ s and $D = 5.8 \times 10^{-7}$ cm^2/s (Wu and Yang 1970); $s = 6.2 \times 10^{-13}$ s (Singh et al. 1970).

MS from baker's yeast: SWISS-PROT entry: P21826 and P30952, homotrimer, $M = 188.391$ or 188.373 kg/mol; $I(h)$, $M = 187 \pm 3$ kg/mol; $R_G = 3.96 \pm 0.02$ nm, $V = 338 \pm 5$ nm^3 and $\delta_1 = 0.35 \pm 0.02$ g/g (Zipper and Durchschlag 1978); $\bar{v} = 0.745$ cm^3/g, $s = (8.6 \pm 0.1) \times 10^{-13}$ s and $D = (4.4 \pm 0.1) \times 10^{-7}$ cm^2/s (Durchschlag et al. 1996); threefold symmetry (Durchschlag et al. 1978).

Concerning the accuracy of scattering and hydrodynamic data, it has to be mentioned that there is a considerable scatter of experimental results, as follows, for example, from a critical inspection of the data for lysozyme, one of the most frequently investigated proteins. For modeling purposes, therefore, often consensus values are used to fit the parameters (e.g., Allison and Tran 1995).

Apart from experimental deficiencies, the diversity of values is influenced by experimental conditions and evaluation procedures (purity of protein samples, protein concentration, protein aggregation, salt type and concentration, temperature, pH, standardization procedures regarding solvent viscosity and temperature, extrapolation to infinite dilution, etc.). Of course, the absence of

unique data sets (in particular in hydrodynamics) impedes the ultimate assessment of protein hydration and to distinguish unequivocally between different assumptions and specific input parameters (particularly concerning fine tuning parameters of specific hydration models) required for the model calculations. These constraints hold in a similar manner for the prediction of scattering and hydrodynamic parameters, since the changes in protein size caused by hydration are similar in scattering and hydrodynamics.

Modeling approaches for proteins and consideration of hydration contributions

WB approaches

Proteins of simple shape can be modeled as spheres or prolate/oblate ellipsoids of revolution (PE/OE) or other triaxial bodies with unequal axes (cf. Durchschlag and Zipper 2002b). For approaches dealing with scattering and hyrodynamic data, the hydration is taken into account by using either quantities involving the contribution of bound water (e.g. hydrated volume, V, or surface-to-volume ratio, S/V, obtained from SAXS experiments) or by applying qualified assumptions regarding the amount of bound water, δ_1, or some kind of rescaling procedures. In the context of the hydration problem, also the introduction of hydration-independent shape functions has to be addressed (Harding 1995).

MB approaches

In the case of MB approaches, the structure of a protein is approximated by many spheres ("multisphere approaches", "bead modeling"), ranging from a few spherical entities (e.g., corresponding to the number of subunits) to several thousands to model the exact protein shape (cf. Byron 2000; García de la Torre et al. 2000). Thereby even very complex and branched structures may be modeled. Hydration contributions may be considered by rescaling procedures (thereby enlarging the whole molecule or only surface beads) or modeling a water shell.

Modeling starting from the atomic coordinates

Since, nowadays, many 3D structures are available (currently about 20,000 PDB entries), modeling approaches based on the atomic coordinates of proteins seem to be the most realistic manner of acting. This procedure corresponds to a MB approach using plenty of coordinates (cf. Byron 2000; García de la Torre et al. 2000). Problems may arise from the exact values to be used for molecular volume and hydration of the constituents (atoms or AA residues). The molecular volume of a protein may be calculated by summing up the volumes of AA residues (according to Traube or Cohn and Edsall or using similar approaches; see Durchschlag 1986, 2003; Perkins 1986; Durchschlag and Zipper 1997c). Again, the most intriguing question is connected with the consideration of hydration contributions, and may be solved by scaling factors or construction of a water shell surrounding somehow the dry (anhydrous) protein. Moreover, special hydration algorithms for considering the individual water molecules on the protein surface have been suggested recently (Durchschlag and Zipper 2001, 2002a, 2002c; Zipper and Durchschlag 2002a, 2002b). Since these approaches will be used in the following, they will be discussed below in more detail.

Modeling starting from density distributions

Density maps derived from 3D reconstructions from cryo-EM, in combination with bead-modeling strategies, may also be exploited for deriving the protein solution structures (Zipper and Durchschlag 2000, 2003; Zipper et al. 2002). A sophisticated analysis of the voxel density distribution allows us to discriminate between anhydrous and hydrated protein models. By using definite density thresholds, an optimal agreement with experimental solution parameters (V, R_G) can be achieved.

Ab initio modeling based on SAXS profiles

Starting from experimental SAXS curves, without (exact) knowledge of 3D information, modeling may be achieved by classical (trial-and-error) or more advanced (ab initio) shape determinations. Among the advanced methods, in particular the programs DALAI_GA, DAMMIN, SAXS3D and GASBOR have to be mentioned.

Low-resolution models can be obtained (1) from an iterative fitting of scattering curves by the genetic algorithm GA (Chacón et al. 1998, 2000), (2) by the program DAMMIN (Svergun 1999), which uses a multiphase model from densely packed dummy atoms (beads) and employs simulated annealing to find a configuration that fits the data, or (3) by SAXS3D (Walther et al. 2000), a Monte Carlo type reconstruction algorithm, using a "Give'n'Take" algorithm to add and remove beads until an optimum fit to the experimental SAXS profile is reached. With all three programs mentioned (DALAI_GA, DAMMIN, SAXS3D), hydration is considered by fitting the experimental SAXS profile which inherently contains hydration contributions. A comparison of the validity of these programs has been presented elsewhere (Zipper and Durchschlag 2003).

The program GASBOR (Svergun et al. 2001) is similar to the DAMMIN procedure mentioned above (use of simulated annealing), but rather uses a chain-like ensemble of dummy residues to simulate the backbone of the protein structure. Contrary to the above programs, this approach uses dummy water molecules to calculate models of hydrated proteins.

Modeling proteins by considering the contributions of individual water molecules bound to the protein surface

Modern surface calculation programs allow the exact surface topography of proteins to be calculated analytically, e.g. by applying the programs MSRoll (Connolly 1993) or SIMS (Vorobjev and Hermans 1997). Both programs produce a smooth molecular "dot surface" by rolling one or two probe spheres, respectively. In previous studies (Durchschlag and Zipper 2001, 2002a, 2002b, 2002c; Zipper and Durchschlag 2002a, 2002b), we have used both programs for advantageously creating a plethora of dot surface points (N_{dot}) and corresponding normal vectors from the original PDB file. For the calculations in this study, we used the program SIMS, applying different values for the dot density (d_{dot}), a probe radius r_{probe} of 0.145-0.1525 nm (depending on the water volume V_w) and a smoothing probe sphere of $r_{sm} = 0.04$ nm.

The surface points and normal vectors created by an appropriate surface calculation procedure can be used as a reservoir of starting points for hypothetical positions of water molecules on the protein surface. Special hydration algorithms (program HYDCRYST for atomic coordinates, program HYDMODEL for AA coordinates) have now to be applied to select appropriate positions out of this pool of possible points. The number of water molecules assigned to each accessible AA residue was based on hydration numbers derived from NMR experiments on polypeptides (Kuntz 1971) and a minimum distance between adjacent waters ($d_w = 0.29$-0.305 nm). Each selected water molecule was then approximated as a sphere of volume V_w. The molecular volume of bound water was assumed to be smaller than that of the bulk water: 0.0245-0.0284 nm^3, in comparison to 0.0299 nm^3; this corresponds to an increase in the average density of bound waters by ca. 5-20%. No clear-cut decision in favor of a definite water density can be given to date.

In the past we have tested various tuning and fine-tuning parameters, in order to test the surface calculation and hydration approaches, on the one hand, and to modulate water binding in a

wide range, on the other. Above all, scaling procedures had to be applied to achieve different degrees of hydration [e.g., minimum, intermediate (realistic), maximum hydration]. This was achieved, in particular, by introducing a factor f_K, acting directly on the original hydration numbers given by Kuntz (1971), which, of course, do not discern between surface and internal AA residues.

Visualization of models

The resultant anhydrous and hydrated models can be visualized by any 3D molecular graphics program. Our graphics were made with the program RASMOL (Sayle and Milner-White 1995).

Prediction of structural and hydrodynamic properties of hydrated proteins

Prediction of hydration and scattering data

The original hydration numbers of Kuntz (1971) may be used for estimating in good approximation an overall hydration of a given protein, in agreement with the results of various physicochemical investigations (Durchschlag and Zipper 2001). These hydration numbers may also be used to predict a theoretical hydration of accessible AA residues, $N_{w,acc}$, and a maximum hydration, $N_{w,max}$, if all AA residues would be hydrated.

The coordinates of atoms (or coordinates of AA residues derived therefrom) and selected water molecules can be used directly to derive structural parameters such as the radius of gyration, R_G, and hydrated volume, V.

Scattering curves $I(h)$ can also be calculated by MB modeling, using Debye's formula (Glatter and Kratky 1982). For the prediction of SAXS patterns, however, the spheres have to be weighted according to the calculated number of excess electrons (difference between the number of electrons of atoms or AA residues or bound water molecules and the number of electrons in the same volume of bulk water). Of course, structural parameters such as R_G and V can also be derived from the SAXS profiles. If required, $I(h)$ can be converted to pair-distance distribution functions $p(r)$ by Fourier transformation. In principle, $p(r)$ functions can also be calculated directly from atomic coordinates.

For the calculation of the scattering behavior of hydrated protein models we used the following approximations. The protein atoms were assumed as hard spheres; the form factors of these hard spheres were corrected according to Fraser et al. (1978) with respect to the displaced solvent. However, both displaced bulk water molecules ($V_w = 0.0299$ nm^3; electron density: 334 e/nm^3) and bound waters [electron density: 352-408 e/nm^3, depending on the assumed value for V_w (0.0284-0.0245 nm^3)] were treated as Gaussian spheres. Modeling bound water molecules as Gaussian spheres is suggested by their high fluctuation rate.

The program CRYSOL (Svergun et al. 1995) allows us to evaluate solution scattering from atomic coordinates. The program uses multipole expansion of the scattering amplitudes to calculate the spherically averaged scattering profile and takes the hydration into account by assuming a hydration shell (0.3 nm thickness) of constant electron density; crystallographically identified waters are ignored. It can either predict the SAXS curve or fit the experimental curve. The form factors of the dummy solvent atoms used are expressed as described by Fraser et al. (1978). For the envelope function, a quasiuniform grid of angular directions using Fibonacci numbers (Svergun 1994) is evaluated. The SAXS intensity depends on the average displaced volume per atomic group (expressed by the effective atomic radius, r_a) and the contrast of the border layer ($\Delta\rho$). In general, the bound solvent is assumed to be 8% denser than the bulk solvent ($\Delta\rho = 30$ e/nm^3); the maximum contrast of $\Delta\rho$ of about 75 e/nm^3 corresponds to a more pronounced density difference (20%) between bound and bulk water. CRYSOL is able to take information from a GNOM output file (Svergun 1992).

Prediction of hydrodynamic data

Sedimentation coefficients s and translational diffusion coefficients D were calculated by means of the well-established program HYDRO (García de la Torre et al. 1994, 2000), modified for our special purposes. We used a "filling model" strategy to derive both scattering and hydrodynamic quantities. Hydrodynamic models were built from spheres of equal density and more or less unequal size. Though the usage of overlapping spheres is usually discouraged (García de la Torre et al. 2000), it cannot be avoided when creating space-filling models of multiple spheres. Already the representation of the atoms or atomic groups of a crystal structure by spheres corresponding to the volumes of the atoms or groups causes a considerable amount of overlapping (Zipper and Durchschlag 1999). The problem of overlapping non-equal spheres was handled by use of an ad hoc modification in the interaction tensor, as described elsewhere (Zipper and Durchschlag 1997; Carrasco et al. 1999), in order to avoid erroneous results. Usage of overlapping beads does not result in other hydrodynamic computation problems or in the restriction of the number or nature of the parameters to be calculated.

If required for hydrodynamic modeling, for the hydrated models discussed above the initial models (IM) had to be reduced to a manageable number of beads ($N_{beads,IM} \rightarrow N_{beads,RdM}$), preferably by a cubic grid approach (Zipper and Durchschlag 1997). Thereby the given IM (representing the atomic coordinates) was mapped into a cubic grid of cells of given edge length, and finally spheres of appropriate radii were placed at the centers of gravity of mass in all occupied cells.

Our hydrodynamic computations concern translational friction only. Rotational friction and intrinsic viscosity parameters were not considered in our present calculations, in order to avoid the problem of special volume corrections (García de la Torre and Carrasco 1998; Carrasco and García de la Torre 1999; Gmachowski 2001). Moreover, in the case of simple globular particles the values for these quantities are not very sensitive to particle size and shape and therefore of similar magnitude, as may be inferred from a comparison of various values for intrinsic viscosities, essentially ranging between 3 and 4 cm^3/g (Harding 1997; Durchschlag and Zipper 1999).

If one is striving for hydrodynamic parameters only, the internal beads can be removed. The interior residues are not experiencing any contact with the solvent, because hydrodynamic friction occurs at the molecular surface. This can be done effectively by replacing the primary hydrodynamic model by a shell model (García de la Torre et al. 1997, 2000; Carrasco and García de la Torre 1999; García de la Torre 2001); in this approach the hydration problem is solved by assuming for the primary hydrodynamic model an effective hydrodynamic radius of the protein atoms of ca. 0.3 nm instead of the conventional van der Waals value of ca. 0.2 nm.

Since our primary goal was the computation of hydrodynamic properties from both crystal and solution X-ray data, in connection with special site-specific hydration algorithms, we preferred filling models. These models allowed the calculation of both scattering and hydrodynamic properties, without changing the strategy by use of shell models, and allowed application of our site-specific hydration algorithms for the water molecules within and around the proteins.

Results and discussion

Lysozyme

For lysozyme, a great number of PDB structures are available, based on crystallographic data retrieved from different enzyme sources and crystal modifications. We chose a PDB file of hen egg-white lysozyme which contained a major number of water coordinates, in

addition to the coordinates of the protein atoms, to enable a comparison of crystallographically observable water sites with the positions of water molecules created by our hydration modeling algorithm HYDCRYST. The resultant space-filling models are shown in Fig. 1 and details of the underlying models are given in Table 1.

Concerning the results stemming from application of HYDCRYST, Fig. 1 suggests that only part of the surface is covered by water molecules. This is in accord with findings be Lee and Richards (1971), who postulated for lysozyme a considerable percentage (41%) of hydrophobic surface. Obviously, the bound water molecules do not form a uniform hydration shell; instead, they are placed at preferred positions on the protein surface. This is also in accord with the crystallographically detectable waters. HYDCRYST points to the existence of a few water molecules in the protein interior, in full accord with the finding of four internal waters by crystallography (Blake et al. 1983). A close inspection of the positions covered by waters produced by HYDCRYST and observed by crystallography, however, reveals only a moderate coincidence of water sites: only about 29 out of 101 waters listed in the crystallographic data sufficiently overlap

(distance < 0.22 nm; cf. Rashin et al. 1986) with waters produced by HYDCRYST.

Figure 2 illustrates a quantitative comparison of crystal data and HYDCRYST for one selected example ($V_w = 0.0257$ m^3, $f_K = 1.25$) out of 35 analyzed models, with respect to the number of hydrated AA residues and total amount of bound water molecules. Certain AAs, such as R (Arg), N (Asn), D (Asp), E (Glu), K (Lys) and S (Ser), are preferably hydrated, as observed with both approaches. It is not surprising that most representatives of this group are charged AAs which are known to bind a lot of water. Since X-ray crystallography only observes a minor number of bound waters, the total amount of bound waters suggested by HYDCRYST must exceed that found in the crystals.

The example shown in the figure represents one of the best HYDCRYST fits concerning the hydrodynamic data presented in Table 1, if compared to the experimental results for s and D. Unfortunately, the comparison of predicted and experimental values is impaired by the considerable uncertainty in the experimental D values; at present, use of a consensus value of ca. 11×10^{-7} cm^2/s for this quantity is the best choice. Similarly, the comparison of the predicted values for V, δ_1 and $R_{G,SAXS}$ with the experimental ones is satisfactory, particularly when considering the most reliable value for R_G of ca. 1.5 nm. In this context, one has to mention some discrepancy between the values for $R_{G,SAXS}$ and $R_{G,Hydro}$, in particular for high values of V_w. This is obviously caused by different calculation procedures for these two quantities (calculation from excess electrons, radii and coordinates of the beads for $R_{G,SAXS}$, and only bead radii and coordinates in the case of $R_{G,Hydro}$). The model also yields a reasonable value for the hydration: $\delta_1 = 0.345$ g of water per gram of protein. Nearly equally good results are found for f_K values near 1, irrespective of the value chosen for V_w. Clearly, the results for the unhydrated model disagree with the experimental findings. On the other hand, variation of f_K in the HYDCRYST procedure allows us to cover a wide range of possible extents of hydration.

Fig. 1 Different views (**a-d**: successive 90° rotations around the *y*-axis) and the corresponding central slabs (**e-h**) of space-filling models for hydrated lysozyme, showing protein atoms and individual water molecules. The coordinates of the basic atoms of the protein moiety (C, O, N, S) and of O atoms belonging to water molecules were taken from the original PDB file (2LYZ). Alternatively, coordinates of bound water molecules were derived by application of the hydration modeling program HYDCRYST (based on surface calculations by the SIMS program, $d_{dot} = 500$ per nm^2, and applying $V_w = 0.0257$ nm^3 and $f_K = 1.25$). The basic protein atoms are shown in *gray*. Bound waters derived from crystallographic work are visualized in *green* and waters obtained by application of HYDCRYST are displayed in *orange*; waters common to both approaches (overlapping > 20% of their volume) are colored *blue* (derived from crystal data) or *red* (derived from HYDCRYST). Graphics were made with the program RASMOL
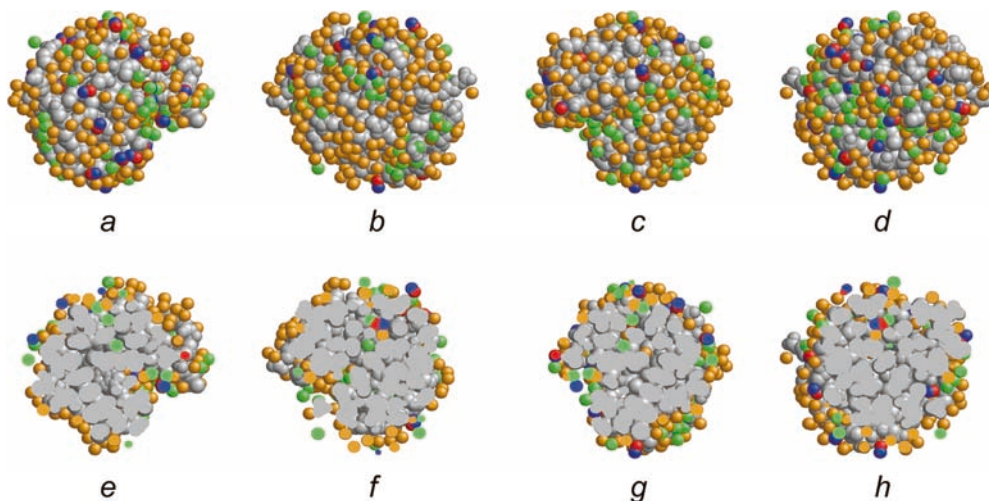
Table 1 Comparison of calculated parameters for hydrated lysozyme models as obtained by HYDCRYST and different input parameters,[a] together with a survey of the experimental values[b]

Input		Output						
V_w (nm³)	f_K	N_w	V (nm³)	δ_1 (g/g)	$R_{G,SAXS}$ (nm)[c]	$R_{G,Hydro}$ (nm)[d]	$s \times 10^{13}$ (s)	$D \times 10^7$ (cm²/s)
Experiment[b]								
			24.2	0.317	1.38-1.55		1.87-1.91	10.2-11.8
Unhydrated model								
		0	16.95	0	1.451	1.386	2.153	12.30
Hydrated models[e]								
0.0245	0.80	199	21.83	0.250	1.492	1.461	1.990	11.36
0.0245	0.90	216	22.24	0.272	1.498	1.470	1.979	11.30
0.0245	1.00	226	22.49	0.284	1.500	1.474	1.968	11.24
0.0245	1.25	277	23.74	0.349	1.513	1.494	1.936	11.05
0.0245	1.50	310	24.55	0.390	1.528	1.516	1.909	10.90
0.0245	1.75	338	25.23	0.425	1.534	1.523	1.900	10.85
0.0245	2.00	351	25.55	0.442	1.540	1.532	1.890	10.79
0.0245	2.50	398	26.70	0.501	1.559	1.559	1.855	10.59
0.0245	3.00	416	27.14	0.524	1.566	1.568	1.845	10.53
0.0257	0.80	198	22.04	0.249	1.484	1.464	1.983	11.32
0.0257	0.90	216	22.50	0.272	1.489	1.473	1.971	11.25
0.0257	1.00	231	22.89	0.291	1.491	1.478	1.961	11.20
0.0257	1.25	274	23.99	0.345	1.505	1.502	1.921	10.97
0.0257	1.50	307	24.84	0.386	1.516	1.521	1.903	10.86
0.0257	1.75	329	25.41	0.414	1.519	1.527	1.893	10.81
0.0257	2.00	339	25.67	0.427	1.524	1.535	1.886	10.77
0.0257	2.50	384	26.82	0.483	1.539	1.561	1.852	10.57
0.0257	3.00	405	27.36	0.510	1.545	1.571	1.841	10.51
0.0269	0.80	196	22.23	0.247	1.477	1.470	1.972	11.26
0.0269	0.90	213	22.68	0.268	1.481	1.480	1.962	11.20
0.0269	1.00	226	23.03	0.284	1.483	1.485	1.954	11.16
0.0269	1.25	271	24.24	0.341	1.491	1.505	1.924	10.98
0.0269	1.50	306	25.18	0.385	1.501	1.526	1.898	10.84
0.0269	2.00	338	26.05	0.425	1.507	1.540	1.878	10.73
0.0269	2.50	381	27.20	0.480	1.520	1.568	1.842	10.52
0.0269	3.00	399	27.69	0.502	1.524	1.577	1.834	10.47
0.0284	0.80	195	22.49	0.245	1.465	1.471	1.966	11.23
0.0284	0.90	210	22.92	0.264	1.467	1.482	1.953	11.15
0.0284	1.00	226	23.37	0.284	1.469	1.490	1.941	11.08
0.0284	1.25	268	24.56	0.337	1.473	1.509	1.915	10.93
0.0284	1.50	298	25.42	0.375	1.479	1.533	1.887	10.78
0.0284	2.00	335	26.47	0.422	1.483	1.550	1.859	10.62
0.0284	2.50	376	27.63	0.473	1.490	1.575	1.836	10.48
0.0284	3.00	392	28.09	0.493	1.492	1.582	1.829	10.44
0.0284	4.00	415	28.74	0.522	1.496	1.597	1.814	10.36

[a]Surface calculations were performed by means of the program SIMS: dot density $d_{dot} = 500$ per nm²; probe radius $r_{probe} = 0.145$, 0.1475, 0.150 or 0.1525 nm (for $V_w = 0.0245$, 0.0257, 0.0269 or 0.0284 nm³, respectively); smoothing probe sphere $r_{sm} = 0.04$ nm
[b]For refs. see text (sources of data). In the case of $R_{G,SAXS}$, experimental values of ca. 1.5 nm seem to be the most realistic ones (cf. Svergun et al. 1998). Since there is some uncertainty in the experimental value for D, for a comparison with predicted values a consensus value of ca. 11×10^{-7} cm²/s should be preferred [$(10.9 \pm 0.5) \times 10^{-7}$ cm²/s would result from an averaging of the

values in the literature, in accord with the consensus value of 11.1×10^{-7} cm²/s used earlier (Kuntz and Kauzmann 1974, Allison and Tran 1995)]
[c]$R_{G,SAXS}$, radius of gyration calculated from bead radii and coordinates and the number of excess electrons
[d]$R_{G,Hydro}$, radius of gyration calculated from bead radii and coordinates
[e]For comparison, the values for maximum hydration based on the values of Kuntz (1971) and for the hydration of accessible AA residues are given: $N_{w,max} = 276$; $N_{w,acc} = 255$-256

Citrate synthase

Table 2 summarizes a variety of hydrated CS models, created by HYDCRYST, and input parameters (V_w, f_K) differing over a wide range. We also calculated several further parameters such as the total volume of bound waters ΣV_w, the number of excess electrons in the bound water $N_{exc.el,w}$, the radius of gyration of the initial model $R_{G,IM}$, which is nearly identical to the radius of gyration

derived from the SAXS curve (data not shown), and the fitting parameter χ (a measure of the goodness of the fit of the experimental or calculated curve, allowing a qualified selection of models under consideration). As in the case of lysozyme, the values found for $R_{G,Hydro}$ generally exceed those of $R_{G,SAXS}$. Some representative CS models are pictured in Fig. 3, which have been obtained by the use of different V_w and f_K values. The figure impressively shows how the extent of hydration
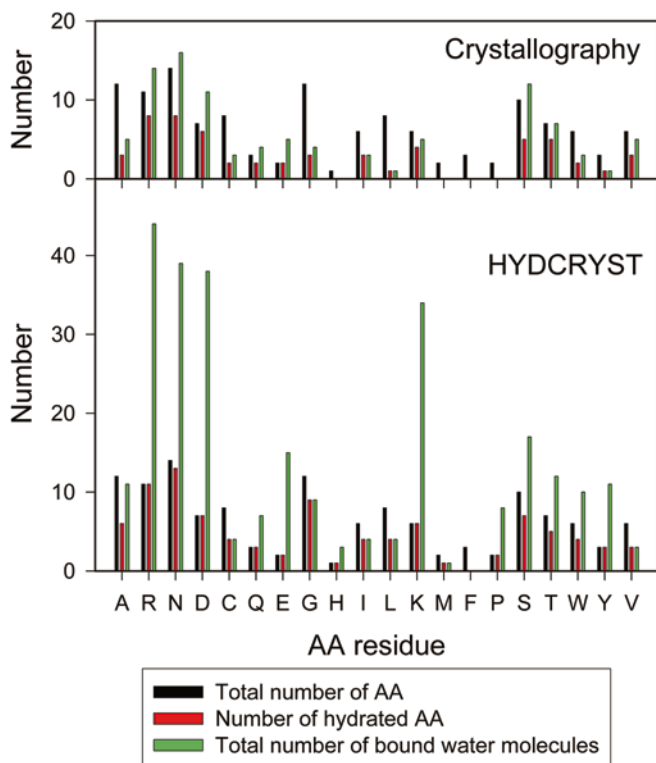
Fig. 2 Histogram showing the total number of AAs of lysozyme, the number of hydrated AAs, together with the total number of bound water molecules. The hydration numbers were retrieved by X-ray crystallography (2LYZ) or HYDCRYST ($V_w = 0.0257$ nm^3, $f_K = 1.25$). AAs are given in alphabetic order of their three-letter code; for short, however, they are symbolized by the one-letter code

can be changed by varying f_K, and to a minor extent also by changing V_w. This, of course, also follows from the values presented in Table 2.

The gallery of scattering curves presented in Fig. 4 (use of different V_w and f_K values) reveals that, in the case of CS, f_K values near 1 obviously fit the data best. Of the water volumes chosen, V_w values of 0.0284 nm^3 (corresponding to a 5% increase in density as compared to bulk water) turn out to be most appropriate. Accordingly, the corresponding parameter predictions given in Table 2 are in fair accord with the experimental data. This may also be taken in a more quantitative way from a comparison of the fitting parameter χ: the best fit of the experimental scattering curve (lowest χ value) is found for $V_w = 0.0284$ nm and $f_K = 1.5$. Though the agreement of V, δ_1 and R_G with experimental values is by no means perfect, the coincidence of the ΣV_w value of 41.8 nm^3 (Table 2) with the volume of 41.49 nm^3 for the water shell derived by CRYSOL (Table 3, footnote f) is surprising. Also the agreement between observed and predicted hydrodynamic parameters (s, D) is satisfactory, irrespective of the value chosen for V_w. Therefore, this model was chosen for illustrating the histogram presented in Fig. 5. Also in the case of CS, the charged amino acids R (Arg), D (Asp), E (Glu) and K (Lys) are the preferred targets for water binding.

Use of the CRYSOL program also leads to a perfect approximation of the scattering behavior, when comparing the experimental scattering profile (Fig. 6). This approach obviously presents an alternative method to simulate hydration in the context of SAXS in a very efficacious way. Application of CRYSOL to the scattering curves of the hydrated models created on the basis of HYDCRYST and listed in Table 2 renders the results summarized in Table 3. For the unhydrated model, CRYSOL yields the contrast $\Delta\rho = 0$, but an appreciable amount of hydration ($\delta_1 = 0.255$ g/g). This discrepancy can be explained by the procedure to calculate the hydration from the ratio of the shell volume to the volume of a single water molecule. Though CRYSOL produces slightly enhanced hydration values (up to 0.29 g/g) with increasing f_K values, the changes are much smaller than implied by our hydrated models. This also follows from a comparison of the numbers of excess electrons of bound water, $N_{exc.el,w}$, with the corresponding numbers in Table 2: the values calculated from $\Delta\rho$ are too small. Use of alternative values for V and r_a did not improve significantly the predictions under consideration. Nevertheless, the scattering curves of the models created by HYDCRYST are fitted perfectly by CRYSOL (not shown).

Malate synthase

GASBOR was chosen as an ab initio modeling program, since this program explicitly models a hydration shell. By contrast, the other ab initio programs (DALAI_GA, DAMMIN, SAXS3D) take hydration into account by merging the hydration contribution with the contributions of the protein building blocks. This may be achieved by some kind of scaling (by modeling more masses at the protein surface compared to the true mass distribution). As input parameters the number of AAs (554 per subunit) and threefold symmetry were used; the trimeric structure of the enzyme was established previously (Durchschlag et al. 1978, 1981; Zipper and Durchschlag 1978). Figure 7 clearly reveals the occurrence of an oblate shape, in agreement with our previous results (Zipper and Durchschlag 1978). The visualization of water molecules, however, discloses that these are only hypothetical water sites needed for the performance of calculations; they represent no physical reality. However, as may be taken from Table 4, the agreement of calculated R_G, s and D values with experimental values is excellent. The predicted values were obtained by equating the volume of a dummy residue of the protein with the mean dry volume of an AA residue. Other ab initio modeling programs gave also a nearly perfect agreement with the observed data (Zipper and Durchschlag 2003), though these programs handle the hydration problem in a way different from GASBOR.

Table 2 Comparison of calculated parameters for hydrated CS models as obtained by HYDCRYST and different input parameters,[a] together with a survey of experimental values[b]

Input		Output											
V_w (nm³)	f_K	$N_{beads,IM}$	N_w	V (nm³)	ΣV_w (nm³)	$N_{exc.el,w}$[c]	δ_1 (g/g)	$R_{G,IM}$ (nm)[d]	χ[e]	$N_{beads,RdM}$	$R_{G,Hydro,RdM}$ (nm)[f]	$s\times10^{13}$ (s)	$D\times10^7$ (cm²/s)
Experiment[b]													
				174.4 ± 2.0			0.339 ± 0.011	2.91 ± 0.02[g]				6.0-6.2	5.8
Unhydrated model													
		6888	0	119.4	0	0	0	2.834	2.79	1034	2.792	6.347	6.076
Hydrated models[h]													
0.0245	0.80	8095	1207	149.0	29.6	2193	0.222	2.911	2.51	1252	2.902	5.993	5.737
0.0245	0.90	8152	1264	150.4	31.0	2297	0.233	2.917	2.73	1265	2.909	5.980	5.726
0.0245	1.00	8234	1346	152.4	33.0	2446	0.248	2.926	3.09	1279	2.921	5.964	5.710
0.0245	1.25	8380	1492	155.9	36.6	2711	0.275	2.938	3.69	1332	2.937	5.928	5.676
0.0245	1.50	8487	1599	158.6	39.2	2905	0.294	2.943	4.11	1337	2.944	5.922	5.670
0.0245	1.75	8571	1683	160.6	41.2	3058	0.310	2.951	4.50	1356	2.954	5.895	5.644
0.0245	2.00	8607	1719	161.5	42.1	3123	0.316	2.953	4.71	1366	2.956	5.885	5.635
0.0257	0.80	8049	1161	149.2	29.8	1644	0.214	2.894	1.80	1261	2.902	6.012	5.756
0.0257	0.90	8111	1223	150.8	31.4	1732	0.225	2.898	1.96	1272	2.909	6.001	5.745
0.0257	1.00	8187	1299	152.8	33.4	1840	0.239	2.906	2.24	1286	2.921	5.972	5.717
0.0257	1.25	8324	1436	156.3	36.9	2034	0.264	2.916	2.71	1324	2.937	5.928	5.676
0.0257	1.50	8428	1540	159.0	39.6	2181	0.283	2.920	3.11	1344	2.944	5.908	5.656
0.0257	1.75	8511	1623	161.1	41.7	2298	0.299	2.926	3.44	1359	2.953	5.888	5.637
0.0257	2.00	8541	1653	161.9	42.5	2341	0.304	2.928	3.58	1368	2.956	5.883	5.632
0.0257	2.50	8641	1753	164.4	45.1	2483	0.323	2.932	4.03	1392	2.963	5.863	5.613
0.0257	3.00	8672	1784	165.2	45.8	2487	0.328	2.934	4.17	1395	2.965	5.858	5.609
0.0269	0.80	8031	1143	150.1	30.7	1161	0.210	2.879	1.30	1261	2.908	6.007	5.751
0.0269	0.90	8094	1206	151.8	32.4	1225	0.222	2.883	1.35	1272	2.916	5.991	5.736
0.0269	1.00	8159	1271	153.6	34.2	1291	0.234	2.889	1.48	1298	2.928	5.948	5.695
0.0269	1.25	8303	1415	157.5	38.1	1437	0.260	2.898	1.80	1327	2.949	5.908	5.656
0.0269	1.50	8397	1509	160.0	40.6	1532	0.278	2.901	2.05	1348	2.955	5.892	5.641
0.0269	1.75	8481	1593	162.2	42.9	1618	0.293	2.906	2.31	1366	2.965	5.868	5.618
0.0269	2.00	8505	1617	162.9	43.5	1642	0.298	2.907	2.39	1371	2.967	5.866	5.616
0.0269	2.50	8587	1699	165.1	45.7	1725	0.313	2.910	2.68	1387	2.974	5.847	5.598
0.0269	3.00	8621	1733	166.0	46.6	1760	0.319	2.911	2.80	1395	2.976	5.840	5.591
0.0284	0.80	8005	1117	151.1	31.7	575	0.206	2.859	1.56	1263	2.916	5.992	5.737
0.0284	0.90	8062	1174	152.7	33.3	604	0.216	2.861	1.48	1271	2.923	5.979	5.725
0.0284	1.00	8122	1234	154.4	35.0	635	0.227	2.863	1.41	1290	2.935	5.955	5.701
0.0284	1.50	8359	1471	161.2	41.8	757	0.271	2.870	1.26	1333	2.958	5.897	5.645
0.0284	2.00	8445	1557	163.6	44.2	801	0.287	2.872	1.26	1347	2.967	5.864	5.614
0.0284	2.50	8526	1638	165.9	46.5	843	0.301	2.873	1.28	1367	2.971	5.849	5.600
0.0284	3.00	8549	1661	166.6	47.2	854	0.306	2.874	1.29	1373	2.973	5.845	5.596
0.0284	4.00	8571	1683	167.2	47.8	866	0.310	2.874	1.30	1374	2.976	5.840	5.591
0.0284	5.00	8579	1691	167.4	48.0	870	0.311	2.875	1.31	1375	2.978	5.839	5.591

[a]Surface calculations were performed by means of the program SIMS: dot density $d_{dot} = 10$ per nm²; probe radius $r_{probe} = 0.145$, 0.1475, 0.150 or 0.1525 nm (for $V_w = 0.0245$, 0.0257, 0.0269 or 0.0284 nm³, respectively); smoothing probe sphere $r_{sm} = 0.04$ nm

[b]Forrefs. see text (sources of data)

[c]Calculated from N_w, V_w and the electron density of the solvent (334 e/nm³)

[d]$R_{G,IM}$, radius of gyration of the initial model calculated from bead radii and coordinates and the number of excess electrons. Similar results may be obtained by a Guinier analysis of calculated scattering curves (SD of R_G values: ±0.001 nm)

[e]χ characterizes the goodness of the fit of the experimental scattering curve, assuming an error band for $I(h)$ of ±3%

[f]$R_{G,Hydro,RdM}$, radius of gyration of the reduced model calculated from bead radii and coordinates

[g]The experimental value for the radius of gyration represents $R_{G,SAXS}$

[h]For comparison, the values for maximum hydration based on the values of Kuntz (1971) and for the hydration of accessible AA residues are given: $N_{w,max} = 2096$; $N_{w,acc} = 1835$-1869, depending on r_{probe}

Conclusions

Water bound to the protein surface or in the protein interior is fundamental to protein folding, stability, function (activity and recognition) and for an understanding of internal water dynamics. A variety of studies of water in and around proteins have made us acutely aware of the problems of this pivotal topic. Recent modeling approaches (Durchschlag and Zipper 2002a) have shown that modeling of individual, bound water molecules is possible, allowing both the scattering and hydrodynamic properties of the hydrated proteins to be calculated. This approach complements the information revealed from solution NMR, crystal diffraction and
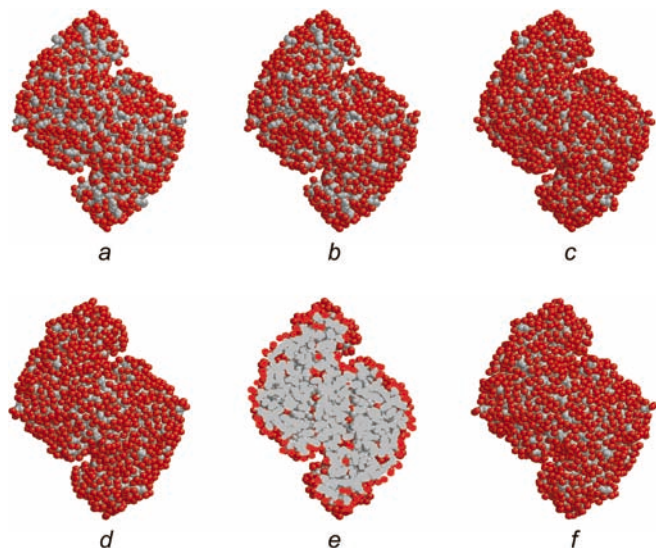
Fig. 3 Selected space-filling models for hydrated CS (**a-d**, **f**) together with an illustrative slab (**e**). Hydration was modeled by means of HYDCRYST, applying different input parameters with respect to water volume (V_w) and scaling factor for the extent of hydration (f_K). The protein atoms are given in *gray* and individual water molecules are highlighted in *red*: (**a**) $V_w = 0.0245$ nm^3, $f_K = 0.8$; (**b**) $V_w = 0.0245$ nm^3, $f_K = 1.0$; (**c**) $V_w = 0.0245$ nm^3, $f_K = 2.0$; (**d**) $V_w = 0.0269$ nm^3, $f_K = 2.0$; (**e**) $V_w = 0.0269$ nm^3, $f_K = 2.0$, central slab; (**f**) $V_w = 0.0284$ nm^3, $f_K = 2.0$

molecular dynamics simulation methods. Though water molecules possess a considerable mobility and have no particular binding site (they may alternate between several equally energetically stable sites), they may be visualized at preferentially occupied sites (Schoenborn et al. 1995). This should be taken into account when discussing conflicting results from different techniques, e.g. when comparing the results detected in crystals, on the one hand, and observed in solution, on the other.

As outlined previously (cf. García de la Torre et al. 2000; Durchschlag and Zipper 2002a, 2002b, 2002c), atomic or other types of coordinates together with bead modeling may be used efficaciously to predict the scattering and hydrodynamic behavior of proteins of simple and complex structure. "Filling-model" strategies have to be preferred, if both scattering and hydrodynamic quantities are to be predicted without changing the computation strategy. Since the behavior of proteins in aqueous solution is influenced both by the molecular characteristics of the protein under investigation and manifold interactions with the solvent water, the hydration problem is still the most crucial problem for examining the solution behavior in detail. This especially holds for the precise interpretation of results stemming from scattering and hydrodynamic studies. As discussed above in detail, hydration contributions can be taken into account by quite different approaches, including rescaling procedures, construction of shell models and

Fig. 4 Normalized scattering curves $I(h)$, of models for unhydrated and hydrated CS, where $h = 4\pi \sin\theta/\lambda$ (2θ = scattering angle, λ = wavelength). The selected models comprise a few representative examples, covering a broad range of water volumes ($V_w = 0.0245$-0.0284 nm^3) and scaling factors ($f_K = 0.8$-2.0) when using the hydration modeling program HYDCRYST. For comparison, the experimental SAXS function (*circles*) is also shown
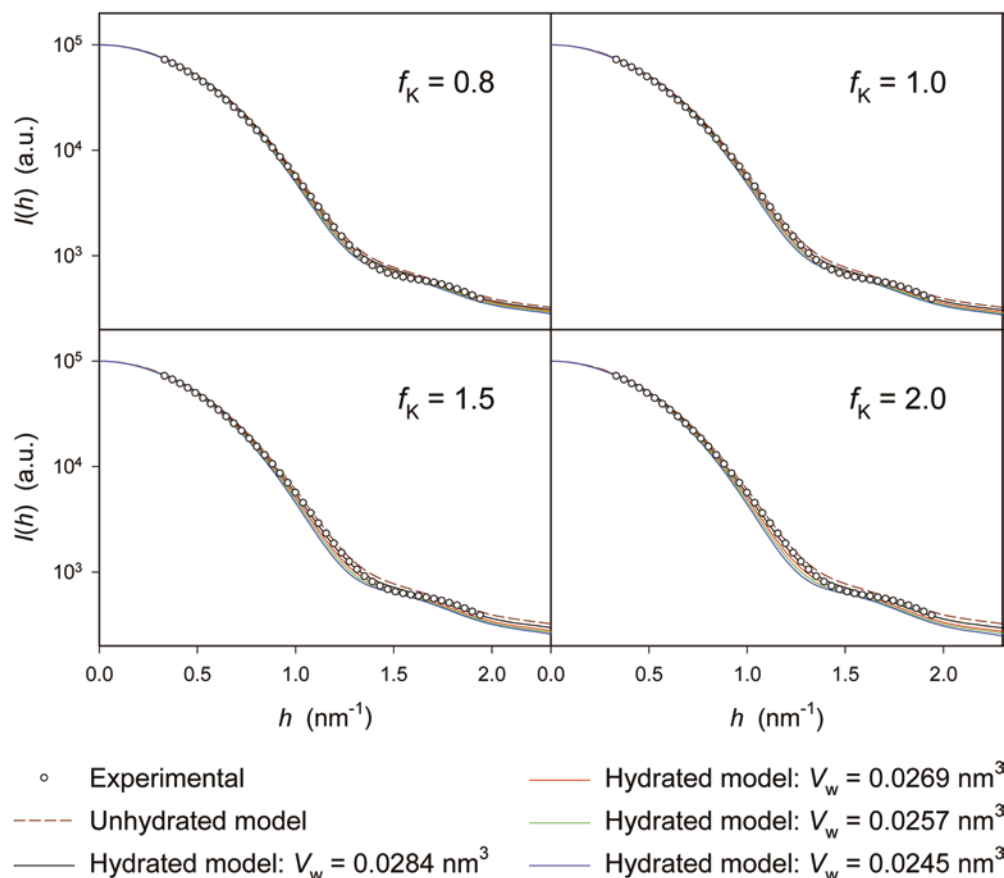
Table 3 Comparison of parameters as obtained by CRYSOLfrom the experimental scattering curve of CS and of calculated curves of hydrated models[a]

Input		Output							
V_w (nm^3)	f_K	$R_{G,e}$ (nm)[b]	V (nm^3)	r_a (nm)	χ[c]	$R_{G,t}$ (nm)[d]	$\Delta\rho$ (e/nm^3)[e]	δ_1 (g/g)	$N_{exc.el,w}$[f]
Experimental SAXS curve									
		2.938 ± 0.086	127.3	0.180	0.840	2.862	3	0.257	124
Unhydrated model									
		2.844 ± 0.039	119.9	0.162	0.355	2.837	0	0.255	0
Hydrated models									
0.0245	0.80	2.928 ± 0.038	119.9	0.168	0.650	2.923	25	0.274	1037
0.0245	0.90	2.933 ± 0.038	119.9	0.170	0.589	2.931	28	0.276	1162
0.0245	1.00	2.943 ± 0.038	119.9	0.172	0.632	2.938	30	0.278	1245
0.0245	1.25	2.956 ± 0.038	120.5	0.164	0.593	2.957	35	0.282	1452
0.0245	1.50	2.962 ± 0.038	120.5	0.164	0.488	2.972	40	0.285	1660
0.0245	1.75	2.971 ± 0.037	121.1	0.158	0.484	2.983	43	0.287	1784
0.0245	2.00	2.973 ± 0.037	121.1	0.156	0.513	2.983	43	0.287	1784
0.0257	0.80	2.909 ± 0.038	119.9	0.166	0.541	2.907	20	0.270	830
0.0257	0.90	2.913 ± 0.038	119.9	0.162	0.544	2.908	20	0.270	830
0.0257	1.00	2.921 ± 0.038	119.9	0.162	0.544	2.916	22	0.272	913
0.0257	1.25	2.932 ± 0.038	120.5	0.156	0.494	2.927	25	0.274	1037
0.0257	1.50	2.937 ± 0.038	120.5	0.160	0.423	2.943	30	0.278	1245
0.0257	1.75	2.944 ± 0.038	120.5	0.160	0.483	2.950	33	0.280	1369
0.0257	2.00	2.946 ± 0.038	121.1	0.150	0.492	2.954	33	0.280	1369
0.0257	2.50	2.951 ± 0.038	121.1	0.148	0.513	2.962	35	0.282	1452
0.0257	3.00	2.953 ± 0.038	121.7	0.140	0.517	2.966	35	0.282	1452
0.0269	0.80	2.893 ± 0.038	119.9	0.164	0.434	2.891	15	0.266	622
0.0269	0.90	2.896 ± 0.038	119.9	0.162	0.395	2.891	15	0.266	622
0.0269	1.00	2.902 ± 0.038	119.9	0.164	0.410	2.899	18	0.268	747
0.0269	1.25	2.912 ± 0.038	119.9	0.164	0.337	2.908	20	0.270	830
0.0269	1.50	2.916 ± 0.038	120.5	0.152	0.283	2.912	20	0.270	830
0.0269	1.75	2.921 ± 0.038	120.5	0.154	0.341	2.920	22	0.272	913
0.0269	2.00	2.922 ± 0.038	120.5	0.152	0.360	2.920	22	0.272	913
0.0269	2.50	2.926 ± 0.038	121.1	0.146	0.321	2.931	25	0.274	1037
0.0269	3.00	2.927 ± 0.038	121.1	0.144	0.309	2.932	25	0.274	1037
0.0284	0.80	2.870 ± 0.039	119.9	0.162	0.350	2.865	7	0.261	290
0.0284	0.90	2.872 ± 0.039	119.9	0.160	0.324	2.865	7	0.261	290
0.0284	1.00	2.875 ± 0.039	119.9	0.158	0.328	2.866	7	0.261	290
0.0284	1.50	2.882 ± 0.039	120.5	0.152	0.272	2.877	10	0.263	415
0.0284	2.00	2.884 ± 0.039	120.5	0.156	0.275	2.885	13	0.264	539
0.0284	2.50	2.886 ± 0.039	120.5	0.154	0.258	2.885	13	0.264	539
0.0284	3.00	2.886 ± 0.039	120.5	0.152	0.261	2.886	13	0.264	539
0.0284	4.00	2.887 ± 0.039	120.5	0.152	0.274	2.886	13	0.264	539
0.0284	5.00	2.888 ± 0.039	120.5	0.152	0.280	2.886	13	0.264	539

[a]Calculations are based on the following parameters: maximum order of harmonics: 20; order of Fibonacci grid: 18. Structural and hydrodynamic parameters of the models are given in Table 2
[b]Radius of gyration determined by Guinier analysis of the input scattering curve, assuming an error band for $I(h)$ of ±3%

[c]χ characterizes the goodness of the fit of the experimental scattering curve or the scattering curve of the respective model
[d]Theoretical radius of gyration
[e]Contrast of the hydration shell
[f]Calculated from $\Delta\rho$ and the assumed volume of the hydration shell (41.49 nm^3)

sophisticated algorithms for determining the protein surface and hydration contributions. In all cases, hydrated protein models yield much more accurate predictions of scattering and hydrodynamic data than the anhydrous model.

Our hydration modeling approach HYDCRYST yields the number of bound water molecules and suggests possible positions for them. As shown in previous studies (Durchschlag and Zipper 2002a, 2002c), the choice of most input parameters is of subordinate importance, except those for scaling the amount of AA hydration (f_K) and molecular water volume (V_w). A comparative analysis of our results, however, indicates that values for f_K between 1 and 2 and V_w of ca. 0.0269 ± 0.0015 nm^3 (corresponding to an increase of water density by 10 ± 5% compared to bulk water) would represent realistic default values for modeling bound water. This can be proven by the finding of reasonable values for protein hydration (δ_1), by best fits to experimental scattering curves (e.g., expressed by low values for the fitting parameter χ) and by correctly predicting both scattering and hydrodynamic parameters (V, R_G, s, D). However, if required, a higher coverage of the protein envelope by waters can be achieved by usage of enhanced f_K values.
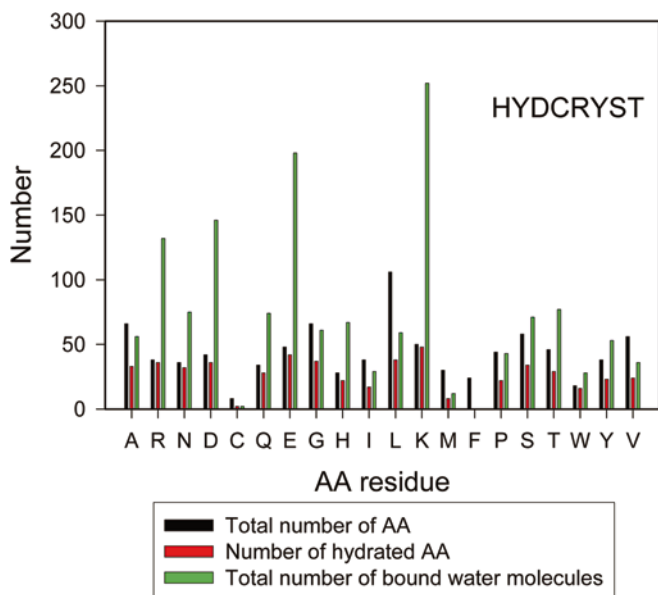
Fig. 5 Histogram showing the total number of AAs of CS, the number of hydrated AAs, together with the total number of bound water molecules. The hydration numbers were derived from application of HYDCRYST ($V_w = 0.0284$ nm^3, $f_K = 1.5$). AAs are given in alphabetical order of their three-letter code; for short, however, they are symbolized by the one-letter code
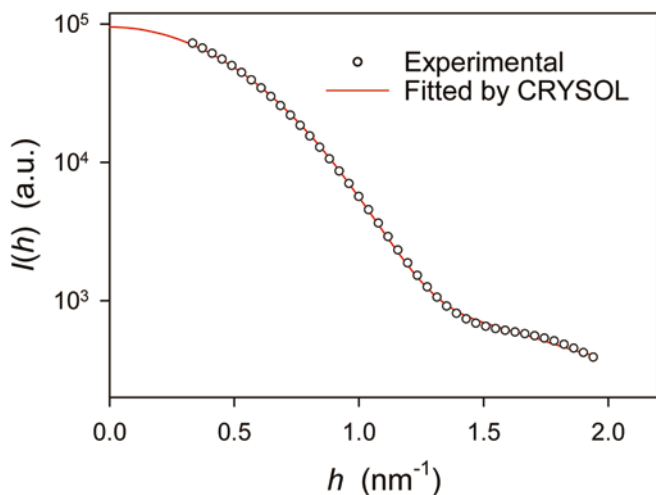

Fig. 6 Normalized scattering curve $I(h)$ of hydrated CS as created by CRYSOL, together with the experimental profile (circles)

Fig. 7 Top, side and bottom views (a-c) of space-filling models for hydrated MS, created by the ab initio modeling program GASBOR after 50 consecutive cycles. The basic protein units (an assembly of 1662 spherical dummy residues) are displayed in gray and a variety (987) of hypothetical water molecules are shown in red. Thereby the protein is surrounded by some kind of hydration layer of 0.3 nm thickness
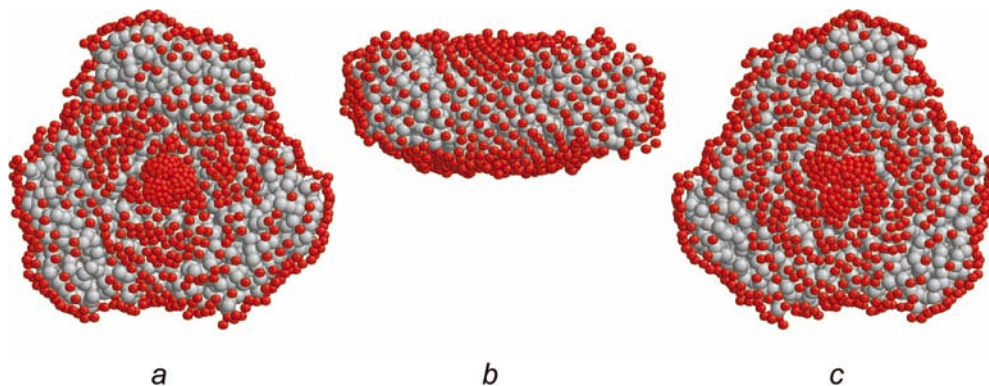
If only the prediction of scattering and/or hydrodynamic parameters is desired, the other approaches also tested, CRYSOL (which also uses atomic coordinates but calculates a water shell), GASBOR (an ab initio modeling approach which uses dummy waters) and, presumably, some others, which consider hydration contributions somehow, are successful. It should be kept in mind, however, that these methods use hydration contributions only as an auxiliary means for performing calculations, and, in general, no statements concerning exact amount of hydration or water placement should be inferred. By contrast, our HYDCRYST approach aims at evolving a biophysically more realistic picture by placing water molecules at sites where, in principle, they could be for a moment, both from geometric and energetic considerations. In this context it should be stressed that the initial values used by us for hydration, the hydration numbers by Kuntz (1971), are very close to the numbers by Hopfinger (1977) which have been derived from thermodynamic considerations (cf. Durchschlag and Zipper 2002a). Consequently, these numbers already include energetic contributions.

The amount of hydration achieved by HYDCRYST for lysozyme and CS (about 0.3 g of water per gram of protein for the above-mentioned default values for f_K and V_w) is within the range of δ_1 values commonly obtained (Durchschlag and Zipper 2001). On the other hand, only part of bound water molecules is usually detected by crystallographic methods. This may also be taken from the histogram in Fig. 2, comparing HYDCRYST data with crystallographically identified waters. Both approaches show a preferred water binding of charged AA residues, though differing in the assessment of hydrated AAs and total amount of water binding. A critical comparison of the water sites monitored by HYDCRYST or crystallography indicates only moderate coincidence. This scenario, however, is by no means surprising, since crystallography only depicts a static picture, and the results of HYDCRYST, admittedly, are influenced by the sequence of steps to be performed when applying the hydration algorithm. This, however, is no matter of concern, since water molecules permanently fluctuate and precise positions on the protein surface make no sense. Neutron diffraction studies (Bon et al. 1999) also show that only part of

Table 4 Comparison of experimental and calculated structural and hydrodynamic parameters, as obtained by various approaches from experimental scattering curves of MS utilizing both slit and cone collimation data[a]

Method	Output			
	V (nm^3)	R_G (nm)	$s \times 10^{13}$ (s)	$D \times 10^7$ (cm^2/s)
Experiment (slit and cone)[b]	338 ± 5	3.96 ± 0.02	8.6 ± 0.1	4.4 ± 0.1
Whole body (OE: V, R_G)[c]			8.95	4.55
GASBOR (slit and cone)	258[d]	3.89	8.65	4.42
GA (slit and cone)[e]	232[f]	4.02	8.63	4.41
	313[g]	4.03	8.55	4.37
DAMMIN (slit and cone)[e]	221[f]	4.03	8.69	4.44
	299[g]	4.03	8.62	4.41
SAXS3D (slit)[e]	248[f]	4.06	8.58	4.38
	334[g]	4.06	8.51	4.35

[a] Zipper and Durchschlag (1978)
[b] For refs. see text
[c] Durchschlag and Zipper (1997a); OE: oblate ellipsoid
[d] Volume consisting of dry protein volume plus 987 water molecules ($V_w = 0.0269$ nm^3)

[e] Zipper and Durchschlag (2003)
[f] Original volume of the model created by the applied program (sum of bead volumes)
[g] Volume of the model after correction for the packing density of spheres

crystallographically localized water molecules sit in definite orientations; they rather form a constantly fluctuating water network covering the protein surface. Recent water simulations also indicate that there is little preference for a water to occupy a particular site; nevertheless, the concept of hydration sites is useful for describing the water structure around proteins (Henchman and McCammon 2002).

At present, only two proteins (lysozyme, CS) have been used for a detailed comparison of crystallographic, SAXS and hydrodynamic data with respect to hydration contributions. Based on these data, it appears that the method applied underestimates slightly both the sedimentation and diffusion coefficients (Tables 1 and 2). Further computations including additional proteins, improved modeling procedures and a fine tuning of input parameters (e.g., probe radius) will presumably show if this is indeed the case. These calculations could also disclose if the original hydration values of Kuntz (1971) have to be modified to avoid an overestimation of hydration; the present results indicate that even values of f_K lower than unity would be required to properly match some data.

Summarizing our present knowledge, the hydration of globular proteins has a significant influence on the interpretation and prediction of scattering and hydrodynamic data in aqueous solution. The correct prediction of the molecular properties of proteins is possible if the hydration is taken into account by rescaling procedures (e.g., by a uniform expansion of whole-body or multibody models), by the assumption of a uniform water shell (e.g. used in the public domain program HYDROPRO; García de la Torre et al. 2000), or by a specific hydration model (exploiting specific, preferential hydration sites) as outlined in this study. If only scattering or hydrodynamic data are to be predicted, all mentioned approaches can be used. All of them allow the prediction of solution properties which deviate only a few percent from the experimental data, provided

hydration contributions are taken into account somehow. The accuracy of all experimental data in the literature is obviously not sufficient to make a clear choice in favor of a definite approach to be used. From the biophysical point of view, however, our specific hydration model is much more realistic. It allows many details of hydration to be discussed in terms of various important aspects, for example aiming at the localization of preferentially bound and internal water molecules, at the elucidation of molecular mechanisms of enzyme activity and stability, at the exact volume occupancy in the context of ligand binding, transport phenomena, and drug design projects, etc. Possible implications in connection with radiation-induced events in proteins have been mentioned recently (Durchschlag et al. 2003). Because of the fluctuation of the water molecules, the assumption of time-averaged positions of bound waters with average properties (e.g. regarding their density) are sufficient for realistic computer simulations of hydrated protein structures.

References

Allison SA, Tran VT (1995) Modeling the electrophoresis of rigid polyions: application to lysozyme. Biophys J 68:2261–2270

Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 28:45–48

Barel AO, Prieels JP, Maes E, Looze Y, Léonis J (1972) Comparative physicochemical studies of human α-lactalbumin and human lysozyme. Biochim Biophys Acta 257:288–296

Beavil AJ, Young RJ, Sutton BJ, Perkins SJ (1995) Bent domain structure of recombinant human IgE-Fc in solution by X-ray and neutron scattering in conjunction with an automated curve fitting procedure. Biochemistry 34:14449–14461

Behlke J (2001) Protein oligomerization requires correct folding of the protomers. Nova Acta Leopoldina Suppl 16:93–95

Berman HM, Westbrook Z, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242

Blake CCF, Pulford WCA, Artymiuk PJ (1983) X-ray studies of water in crystals of lysozyme. J Mol Biol 167:693–723

Bon C, Lehmann MS, Wilkinson C (1999) Quasi-Laue neutron-diffraction study of the water arrangement in crystals of triclinic hen egg-white lysozyme. Acta Crystallogr Sect D 55:978–987

Brooks CL III, Karplus M (1989) Solvent effects on protein motion and protein effects on solvent motion. Dynamics of the active site region of lysozyme. J Mol Biol 208:159–181

Brunne RM, Liepinsh E, Otting G, Wüthrich K, van Gunsteren WF (1993) Hydration of proteins. A comparison of experimental residence times of water molecules solvating the bovine pancreatic trypsin inhibitor with theoretical model calculations. J Mol Biol 231:1040–1048

Byron O (1997) Construction of hydrodynamic bead models from high-resolution X-ray crystallographic or nuclear magnetic resonance data. Biophys J 72:408–415

Byron O (2000) Hydrodynamic bead modeling of biological macromolecules. Methods Enzymol 321:278–304

Byron O, Gilbert RJC (2000) Neutron scattering: good news for biotechnology. Curr Opin Biotechnol 11:72–80

Carrasco B, García de la Torre J (1999) Hydrodynamic properties of rigid particles: comparison of different modeling and computational procedures. Biophys J 75:3044–3057

Carrasco B, García de la Torre J, Zipper P (1999) Calculation of hydrodynamic properties of macromolecular bead models with overlapping spheres. Eur Biophys J 28:510–515

Carugo O, Bordo D (1999) How many water molecules can be detected by protein crystallography? Acta Crystallogr Sect D 55:479–483

Chacón P, Morán F, Díaz JF, Pantos E, Andreu JM (1998) Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. Biophys J 74:2760–2775

Chacón P, Díaz JF, Morán F, Andreu JM (2000) Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. J Mol Biol 299:1289–1302

Chaplin M (2002) Water structure and behavior. http://www.sbu.ac.uk/water/

Colvin JR (1952) The size and shape of lysozyme. Can J Chem 30:831–834

Connolly ML (1993) The molecular surface package. J Mol Graph 11:139–141

Creighton TE (1993) Proteins: structures and molecular properties, 2nd edn. Freeman, New York

Denisov VP, Halle B (1996) Protein hydration dynamics in aqueous solution. Faraday Discuss Chem Soc 103:227–244

Diamond R (1974) Real-space refinement of the structure of hen egg-white lysozyme. J Mol Biol 82:371–391

Dubin SB, Lunacek JH, Benedek GB (1967) Observation of the spectrum of light scattered by solutions of biological macromolecules. Proc Natl Acad Sci USA 57:1164–1171

Dubin SB, Clark NA, Benedek GB (1971) Measurement of the rotational diffusion coefficient of lysozyme by depolarized light scattering: configuration of lysozyme in solution. J Chem Phys 54:5158–5164

Dubin SB, Feher G, Benedek GB (1973) Study of the chemical denaturation of lysozyme by optical mixing spectroscopy. Biochemistry 12:714–720

Ducruix A, Guilloteau JP, Riès-Kautt M, Tardieu A (1996) Protein interactions as seen by solution X-ray scattering prior to crystallogenesis. J Crystal Growth 168:28–39

Durchschlag H (1975) X-ray small-angle studies of the pyruvate dehydrogenase core complex from *Escherichia coli* K-12. II. Subunit structure of the core complex. Biophys Struct Mechanism 1:169–188

Durchschlag H (1986) Specific volumes of biological macromolecules and some other molecules of biological interest: In: Hinz H-J (ed) Thermodynamic data for biochemistry and biotechnology. Springer, Berlin Heidelberg New York, pp 45–128

Durchschlag H (2003) Partial specific volumes and other volumetric properties of proteins and substances related to protein chemistry, In: Hinz H-J (ed) Landolt-Börnstein new series, VII/2A. Springer, Berlin Heidelberg New York (in press)

Durchschlag H, Zipper P (1997a) Prediction of hydrodynamic parameters of biopolymers from small-angle scattering data. J Appl Crystallogr 30:1112–1124

Durchschlag H, Zipper P (1997b) Calculation of hydrodynamic parameters of biopolymers from scattering data using whole-body approaches. Prog Colloid Polym Sci 107:43–57

Durchschlag H, Zipper P (1997c) Calculation of partial specific volumes and other volumetric properties of small molecules and polymers. J Appl Crystallogr 30:803–807

Durchschlag H, Zipper P (1999) Calculation of structural parameters from hydrodynamic data. Prog Colloid Polym Sci 113:87–105

Durchschlag H, Zipper P (2001) Comparative investigations of biopolymer hydration by physicochemical and modeling techniques. Biophys Chem 93:141–157

Durchschlag H, Zipper P (2002a) Modelling of protein hydration. J Phys Condens Matter 14:2439–2452

Durchschlag H, Zipper P (2002b) Correlations between crystallographic, small-angle scattering and hydrodynamic data of biopolymers. Prog Colloid Polym Sci 119:121–130

Durchschlag H, Zipper P (2002c) Modeling of protein hydration with respect to X-ray scattering and hydrodynamics. Prog Colloid Polym Sci 119:131–140

Durchschlag H, Bogner F, Wilhelm D, Jaenicke R, Zipper P, Mayer F (1978) The subunit structure of malate synthase from baker's yeast. Hoppe-Seyler's Z Physiol Chem 359:1077

Durchschlag H, Biedermann G, Eggerer H (1981) Large-scale purification and some properties of malate synthase from baker's yeast. Eur J Biochem 114:255–262

Durchschlag H, Zipper P, Wilfing R, Purr G (1991) Detection of small conformational changes of proteins by small-angle scattering. J Appl Crystallogr 24:822–831

Durchschlag H, Zipper P, Purr G, Jaenicke R (1996) Comparative studies of structural properties and conformational changes of proteins by analytical ultracentrifugation and other techniques. Colloid Polym Sci 274:117–137

Durchschlag H, Hefferle T, Zipper P (2003) Comparative investigations of the effects of X- and UV-irradiation on lysozyme in the absence or presence of additives. Radiat Phys Chem (in press)

Ebel C, Eisenberg H, Ghirlando R (2000) Probing protein-sugar interactions. Biophys J 78:385–393

Eberstein W, Georgalis Y, Saenger W (1994) Molecular interactions in crystallizing lysozyme solutions studied by photon correlation spectroscopy. J Crystal Growth 143:71–78

Finney JL (1996) Overview lecture: hydration processes in biological and macromolecular systems. Faraday Discuss Chem Soc 103:1–18

Foord R, Jakeman E, Oliver CJ, Pike ER, Blagrove RJ, Wood E, Peacocke AR (1970) Determination of diffusion coefficients of haemocyanin at low concentration by intensity fluctuation spectroscopy of scattered laser light. Nature 227:242–245

Fraser RDB, MacRae TP, Suzuki E (1978) An improved method for calculating the contribution of solvent to the X-ray diffraction pattern of biological molecules. J Appl Crystallogr 11:693–694

García de la Torre J (1989) Hydrodynamic properties of macromolecular assemblies. In: Harding SE, Rowe AJ (eds) Dynamic properties of biomolecular assemblies. Royal Society of Chemistry, Cambridge, pp 3–31

García de la Torre J (2001) Hydration from hydrodynamics. General considerations and applications of bead modelling to globular proteins. Biophys Chem 93:159–170

García de la Torre J, Bloomfield VA (1981) Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications. Q Rev Biophys 14:81–139

García de la Torre J, Carrasco B (1998) Intrinsic viscosity and rotational diffusion of bead models for rigid macromolecules and bioparticles. Eur Biophys J 27:549–557

García de la Torre J, Navarro S, López Martínez MC, Díaz FG, López Cascales JJ (1994) HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules. Biophys J 67:530–531

García de la Torre J, Carrasco B, Harding SE (1997) SOLPRO: theory and computer program for the prediction of SOLution PROperties of rigid macromolecules and bioparticles. Eur Biophys J 25:361–372

García de la Torre J, Huertas ML, Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. Biophys J 78:719–730

García de la Torre J, Llorca O, Carrascosa JL, Valpuesta JM (2001) HYDROMIC: prediction of hydrodynamic properties of rigid macromolecular structures obtained from electron microscopy images. Eur Biophys J 30:457–462

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

Gerstein M, Tsai J, Levitt M (1995) The volume of atoms on the protein surface: calculated from simulation, using Voronoi polyhedra. J Mol Biol 249:955–966

Glatter O, Kratky O (eds) (1982) Small angle X-ray scattering. Academic Press, London

Gmachowski L (2001) Intrinsic viscosity of bead models for macromolecules and bioparticles. Eur Biophys J 30:453–456

Gregory RB (ed) (1995) Protein-solvent interactions. Dekker, New York

Grigsby JJ, Blanch HW, Prausnitz JM (2000) Diffusivities of lysozyme in aqueous $MgCl_2$ solutions from dynamic light-scattering data: effect of protein and salt concentrations. J Phys Chem B 104:3645–3650

Harding SE (1989) Modelling the gross conformation of assemblies using hydrodynamics: the whole body approach. In: Harding SE, Rowe AJ (eds) Dynamic properties of biomolecular assemblies. Royal Society of Chemistry, Cambridge, pp 32–56

Harding SE (1995) On the hydrodynamic analysis of macromolecular conformation. Biophys Chem 55:69–93

Harding SE (1997) The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution. Prog Biophys Mol Biol 68:207–262

Harding SE (2001a) The hydration problem in solution biophysics: an introduction. Biophys Chem 93:87–91

Harding SE (ed) (2001b) Special issue: the hydration problem in solution biophysics. Biophys Chem 93:87–246

Henchman RH, McCammon JA (2002) Extracting hydration sites around proteins from explicit water simulations. J Comput Chem 23:861–869

Hopfinger AJ (1977) Intermolecular interactions and biomolecular organization. Wiley, New York

Kakalis LT, Kumosinski TF (1992) The dynamics of water in protein solutions: the field dispersion of deuterium NMR longitudinal relaxation. Biophys Chem 43:39–49

Kodandapani R, Suresh CG, Vijayan M (1990) Crystal structure of low humidity tetragonal lysozyme at 2.1-Å resolution. Variability in hydration shell and its structural consequences. J Biol Chem 265:16126–16131

Krigbaum WR, Kügler FR (1970) Molecular conformation of egg-white lysozyme and bovine α-lactalbumin in solution. Biochemistry 9:1216–1223

Kuhn LA, Siani MA, Pique ME, Fisher CL, Getzoff ED, Tainer JA (1992) The interdependence of protein surface topography and bound water molecules revealed by surface accessibility and fractal density measures. J Mol Biol 228:13–22

Kumosinski TF, Pessen H (1982) Estimation of sedimentation coefficients of globular proteins: an application of small-angle X-ray scattering. Arch Biochem Biophys 219:89–100

Kuntz ID (1971) Hydration of macromolecules. III. Hydration of polypeptides. J Am Chem Soc 93:514–516

Kuntz ID Jr, Kauzmann W (1974) Hydration of proteins and polypeptides. Adv Protein Chem 28:239–345

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

Lee JC, Timasheff SN (1974) Partial specific volumes and interactions with solvent components of proteins in guanidine hydrochloride. Biochemistry 13:257–265

Levitt M, Park PH (1993) Water: now you see it, now you don't. Structure 1:223–226

Luzzati V, Witz J, Nicolaieff A (1961) Détermination de la masse et des dimensions des protéines en solution par la diffusion centrale des rayons X mesurée à l'échelle absolue: exemple du lysozyme. J Mol Biol 3:367–378

Merzel F, Smith JC (2002) SASSIM: a method for calculating small-angle X-ray and neutron scattering and the associated molecular envelope from explicit-atom models of solvated proteins. Acta Crystallogr Sect D 58:242–249

Mikol V, Hirsch E, Giegé R (1990) Diagnostic of precipitant for biomacromolecule crystallization by quasi-elastic light-scattering. J Mol Biol 213:187–195

Müller JJ (1983) Calculation of scattering curves for macromolecules in solution and comparison with results of methods using effective atomic scattering factors. J Appl Crystallogr 16:74–82

Müller JJ (1991) Prediction of the rotational diffusion behavior of biopolymers on the basis of their solution or crystal structure. Biopolymers 31:149–160

Müller JJ, Schrauber H (1992) The inertia-equivalent ellipsoid: a link between atomic structure and low-resolution models of small globular proteins determined by small-angle X-ray scattering. J Appl Crystallogr 25:181–191

Murphy LR, Matubayasi N, Payne VA, Levy RM (1998) Protein hydration and unfolding: insights from experimental partial specific volumes and unfolded protein models. Fold Des 3:105–118

Nicoli DF, Benedek GB (1976) Study of thermal denaturation of lysozyme and other globular proteins by light-scattering spectroscopy. Biopolymers 15:2421–2437

Niimura N, Minezaki Y, Nonaka T, Castagna J-C, Cipriani F, Høghøj P, Lehmann MS, Wilkinson C (1997) Neutron Laue diffractometry with an imaging plate provides an effective data collection regime for neutron protein crystallography. Nat Struct Biol 4:909–914

Otting G, Liepinsh E (1995) Protein hydration viewed by high-resolution NMR spectroscopy: implications for magnetic resonance image contrast. Acc Chem Res 28:171–177

Otting G, Liepinsh E, Wüthrich K (1991) Protein hydration in aqueous solution. Science 254:974–980

Pal SK, Peon J, Zewail AH (2002) Biological water at the protein interface: dynamical solvation probed directly with femtosecond resolution. Proc Natl Acad Sci USA 99:1763–1768

Perkins SJ (1986) 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 157:169–180

Perkins SJ (2001) X-ray and neutron scattering analyses of hydration shells: a molecular interpretation based on sequence predictions and modelling fits. Biophys Chem 93:129–139

Perkins SJ, Ashton AW, Boehm MK, Chamberlain D (1998) Molecular structures from low angle X-ray and neutron scattering studies. Int J Biol Macromol 22:1–16

Pessen H, Kumosinski TF (1985) Measurements of protein hydration by various techniques. Methods Enzymol 117:219–255

Pessen H, Kumosinski TF, Timasheff SN (1971) The use of small-angle X-ray scattering to determine protein conformation. J Agric Food Chem 19:698–702

Rashin AA, Iofin M, Honig B (1986) Internal cavities and buried waters in globular proteins. Biochemistry 25:3619–3625

502

Remington S, Wiegand G, Huber R (1982) Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution. J Mol Biol 158:111–152

Rowe AJ (2001) Probing hydration and the stability of protein solutions: a colloid science approach. Biophys Chem 93:93–101

Rupley JA, Careri G (1991) Protein hydration and function. Adv Protein Chem 41:37–172

Sayle RA, Milner-White EJ (1995) RASMOL: biomolecular graphics for all. Trends Biochem Sci 20:374–376

Schoenborn BP, Garcia A, Knott R (1995) Hydration in protein crystallography. Prog Biophys Mol Biol 64:105–119

Singh M, Brooks GC, Srere P (1970) Subunit structure and chemical characteristics of pig heart citrate synthase. J Biol Chem 245:4636–4640

Smith JC, Merzel F, Verma CS, Fischer S (2002) Protein hydration water: Structure and thermodynamics. J Mol Liquids 101:27–33

Smith PE, van Gunsteren WF (1994) Translational and rotational diffusion of proteins. J Mol Biol 236:629–636

Sophianopoulos AJ, Rhodes CK, Holcomb DN, Van Holde KE (1962) Physical studies of lysozyme. I. Characterization. J Biol Chem 237:1107–1112

Spotorno B, Piccinini L, Tassara G, Ruggiero C, Nardini M, Molina F, Rocco M (1997) BEAMS (BEAds Modelling System): a set of computer programs for the generation, the visualization and the computation of the hydrodynamic and conformational properties of bead models of proteins. Eur Biophys J 25:373-384; erratum 26:417

Squire PG, Himmel ME (1979) Hydrodynamics and protein hydration. Arch Biochem Biophys 196:165–177

Stuhrmann HB, Fuess H (1976) A neutron small-angle scattering study of hen egg-white lysozyme. Acta Crystallogr Sect A 32:67–74

Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr 25:495–503

Svergun DI (1994) Solution scattering from biopolymers: advanced contrast-variation data analysis. Acta Crystallogr Sect A 50:391–402

Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J 76:2879–2886

Svergun DI (2000) Advanced solution scattering data analysis methods and their applications. J Appl Crystallogr 33:530–534

Svergun D, Barberato C, Koch MHJ (1995) CRYSOL: a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J Appl Crystallogr 28:768–773

Svergun DI, Richard S, Koch MHJ, Sayers Z, Kuprin S, Zaccai G (1998) Protein hydration in solution: experimental observation by x-ray and neutron scattering. Proc Natl Acad Sci USA 95:2267–2272

Svergun DI, Petoukhov MV, Koch MHJ (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys J 80:2946–2953

Teeter MM (1991) Water-protein interactions: theory and experiment. Annu Rev Biophys Biophys Chem 20:577–600

Thanki N, Thornton JM, Goodfellow JM (1988) Distributions of water around amino acid residues in proteins. J Mol Biol 202:637–657

Van Holde KE, Johnson WC, Ho PS (1998) Principles of physical biochemistry. Prentice Hall, Upper Saddle River, NJ

Vorobjev YN, Hermans J (1997) SIMS: computation of a smooth invariant molecular surface. Biophys J 73:722–732

Walther D, Cohen FE, Doniach S (2000) Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray solution scattering data for biomolecules. J Appl Crystallogr 33:350–363

Westhof E (ed) (1993) Water and biological macromolecules. Macmillan, London

Winzor DJ, Carrington LE, Harding SE (2001) Analysis of the thermodynamic non-ideality in terms of protein solvation. Biophys Chem 93:231–240

Wu J-Y, Yang JT (1970) Physicochemical characterization of citrate synthase and its subunits. J Biol Chem 245:212–218

Wüthrich K, Otting G, Liepinsh E (1992) Protein hydration in aqueous solution. Faraday Discuss Chem Soc 93:35–45

Wüthrich K, Billeter M, Güntert P, Luginbühl P, Riek R, Wider G (1996) NMR studies of the hydration of biological macromolecules. Faraday Discuss Chem Soc 103:245–253

Zhang X-J, Matthews BW (1994) Conservation of solvent-binding sites in 10 crystal forms of T4 lysozyme. Protein Sci 3:1031–1039

Zhou H-X (1995) Calculation of translational friction and intrinsic viscosity. II. Application to globular proteins. Biophys J 69:2298–2303

Zhou H-X (2001) A unified picture of protein hydration: prediction of hydrodynamic properties from known structures. Biophys Chem 93:171–179

Zipper P, Durchschlag H (1978) Small-angle X-ray scattering on malate synthase from baker's yeast. The native substrate-free enzyme and enzyme-substrate complexes. Eur J Biochem 87:85–99

Zipper P, Durchschlag H (1997) Calculation of hydrodynamic parameters of proteins from crystallographic data using multibody approaches. Prog Colloid Polym Sci 107:58–71

Zipper P, Durchschlag H (1998) Recent advances in the calculation of hydrodynamic parameters from crystallographic data by multibody approaches. Biochem Soc Trans 26:726–731

Zipper P, Durchschlag H (1999) Prediction of hydrodynamic parameters from 3D structures. Prog Colloid Polym Sci 113:106–113

Zipper P, Durchschlag H (2000) Prediction of hydrodynamic and small-angle scattering parameters from crystal and electron microscopic structures. J Appl Crystallogr 33:788–792

Zipper P, Durchschlag H (2002a) Prediction of structural and hydrodynamic parameters of hydrated proteins by computer modeling based on the results from high-resolution techniques. Physica A 304:283–293

Zipper P, Durchschlag H (2002b) Modeling of complex protein structures. Physica A 314:613–622

Zipper P, Durchschlag H (2003) Modeling of protein solution structures. J Appl Crystallogr (in press)

Zipper P, Krebs A, Durchschlag H (2002) Prediction of hydrodynamic parameters of Lumbricus terrestris hemoglobin from small-angle X-ray and electron microscopic structures. Prog Colloid Polym Sci 119:141–148