3

Anhydrous and Hydrated Protein Models Derived from High-Resolution and Low-Resolution Techniques

Helmut Durchschlag[1] and Peter Zipper[2]
[1]*Institute of Biophysics and Physical Biochemistry, University of Regensburg, Regensburg,*
[2]*Physical Chemistry, Institute of Chemistry, University of Graz, Graz,*
[1]*Germany,*
[2]*Austria*

1. Introduction

High-resolution crystallography provides information on the precise 3D structure of proteins or other (bio)macromolecules, including ligands and several water molecules. In the case of small proteins, NMR techniques may be equally purposive. Reconstructions from cryo-electron microscopy may also supply precise anhydrous models. The most intimate structural details of proteins can be visualized from application of these high-resolution techniques (Creighton, 2010a, 2010b; Serdyuk et al., 2007).

In general, however, only a small fraction of preferentially bound water molecules is identified by crystallographic techniques, owing to insufficient resolution and hydrogen bond network building. Discrete waters should have hydrogen bonded contact(s) to other solvent molecules or to protein; poorly placed waters tend to drift away during refinement (Rupp, 2010). Further deficiencies/errors in crystallographic work are due to missing parts (amino acid (AA) residues) and the occurrence of various radiation damages (Ravelli & Garman, 2006).

By contrast, the data from low-resolution solution techniques, such as small-angle X-ray scattering (SAXS) and hydrodynamics (analytical ultracentrifugation, viscometry etc.), inherently contain hydration and yield hydrated protein models (Durchschlag et al., 2007; Zipper & Durchschlag, 2010b). For many reasons, knowledge of hydration is essential for understanding the behaviour of proteins in solution and manifold interactions.

Water is important for structure, stability, dynamics, and function of native proteins; it is also involved in guiding protein folding, and, consequently, needs to be involved in protein structure predictions and modelling of folding pathways (Levy & Onuchic, 2006; Papoian et al., 2004). Minimizing the number of hydrophobic side-chains exposed to aqueous solvent is a major driving force behind protein structure formation. In a typical protein, a tightly packed core contains more than 80 % of the non-polar side chains and water molecules are

generally excluded from protein interiors. There is a clear correlation between the hydrophobicities of AA residues and their tendency to occur in the interior (Creighton, 2010b). Water soluble proteins have all ionized groups on the surface, exposed to the solvent. Polar groups tend to be paired in hydrogen bonds. Fixed, preferentially bound water molecules occur in positions where they can build hydrogen bonds to polar groups. Overall, to some extent both bound and free waters are of relevance for the structure of proteins and their formation. Only the existence of water leads to the distinction between hydrophobic and hydrophilic parts. Water is the 'lubricant of life'.

Biophysically relevant hydration details are mandatorily required in certain cases: understanding protein interactions as precondition for flexibility, dynamics and functionality (Creighton, 2010a, 2010b; Serdyuk et al., 2007); very precise comparisons of high-resolution 3D structures and molecular parameters with data from solution techniques (Durchschlag et al., 2007; Durchschlag & Zipper, 2008; Zipper & Durchschlag, 2010b); gaining insight into radiation damage events (Durchschlag et al., 2003; Durchschlag & Zipper, 2007); construction of tailor-made nano-compounds in context with drug-design projects and the development of functionalized surfaces and polymers by mimicking proteins (Durchschlag & Zipper, 2008); improvement/check of crystallographic/NMR data with regard to hydration waters (Durchschlag & Zipper, 2003).

Experimental determinations of volume, surface, and hydration properties of proteins by SAXS or other low-resolution techniques turned out to be rather inaccurate, whereas the hydration numbers for individual AA residues, derived from NMR spectroscopy or thermodynamic considerations, seem to be rather precise (Durchschlag & Zipper, 2001, 2008). For different pH values, of course, different hydration numbers for the AA residues have been found.

Unlike experimental data, calculations of volume and surface properties of simple and complex proteins proved to yield reliable results, if based on the properties of the molecular constituents and the coordinates (Durchschlag & Zipper, 2005, 2008). Modern analytical procedures and programs yield the anhydrous volume (van der Waals volume), summing up the contributions of the constituents. For analyzing the surface area of proteins, rolling-ball mechanisms were most effective, characterizing either an anhydrous or a hydrated protein. Prediction of the values for hydrated volume and surface requires assumptions/estimations/findings for the amount of water bound to the protein, e.g. application of a shell model, blowing up (surface) AA residues, and use of individual waters. For the precise prediction of structural and hydrodynamic parameters (volume, surface, radius of gyration, sedimentation and diffusion coefficients, intrinsic viscosity etc.) realistic assumptions or estimations are sufficient, whereas visualization of biophysically realistic hydrated protein models obviously necessitates more sophisticated, advanced modelling techniques.

A critical inspection of anhydrous and hydrated protein models obtained by crystallography with models derived from quite different experimental techniques and calculation approaches allows a scrutinized comparison of the models under analysis. Among a variety of problems, the amount of hydration and the position of the individual water molecules turned out to be the most crucial points. To meet this challenge, a variety of techniques and approaches were examined and both models and molecular parameters were analyzed: (i)

Conventional and *ab initio* modelling approaches signify satisfactory agreement between crystal- and SAXS-based protein models, provided hydration contributions and other precautions are taken into account (Durchschlag et al., 2007). (ii) Recourse to crystallographic or model data also allows scattering and hydrodynamic modelling; in the case of multibead structures novel modelling refinements (e.g., efficient bead reductions) have to be adopted (Zipper et al., 2005; Zipper & Durchschlag, 2007, 2010a, 2010b). (iii) The creation of hydrated models from cryo-electron microscopy data necessitates qualified assumptions regarding hydration, e.g. in terms of voxel densities (Zipper & Durchschlag, 2002b). (iv) Combining the exact surface topography (molecular dot surface; derived from atomic or amino acid coordinates of proteins or appropriate models) and our recent hydration algorithms (program HYDCRYST) allows the prediction of individual water molecules preferentially bound to certain amino acid residues (Durchschlag & Zipper, 2001, 2002a, 2002b, 2003, 2004, 2005, 2006, 2008; Zipper & Durchschlag, 2002a, 2002b).

2. Methods

2.1 Coarse-grained models

Sphere (S), prolate and oblate ellipsoids of revolution (PE, OE) of different axial ratios, and hollow spheres (HS) of different hollowness may serve as approximations for simple anhydrous protein structures. Selected structures were modelled by assemblies of equal-sized densely packed beads (Zipper & Durchschlag, 2010a). By variation of the bead radius (r_b), the number of beads (N_b), i.e. the extent of the reduction process, can be varied systematically. The coordinates of the beads were generated by an in-house program.

2.2 Proteins: Recourse to experimental results and data banks

SAXS experiments yield scattering intensity profiles, $I(h)$, and pair-distance distribution functions, $p(r)$, in addition to a variety of molecular parameters such as radius of gyration, R_G, hydrated volume, V, and maximum particle diameter, d_{max} (Glatter & Kratky, 1982). Shape reconstructions may be obtained by conventional modelling or advanced *ab initio* modelling approaches, preferably by programs based on simulated annealing or on a genetic algorithm (Zipper et al., 2005). Hydrodynamic properties such as sedimentation coefficient, s, translational diffusion coefficient, D, and intrinsic viscosity, $[\eta]$, are determined by analytical ultracentrifugation and viscometry, respectively. Atomic coordinates and masses of proteins were obtained from the PDB and SWISS-PROT data banks (Berman et al., 2000; Boeckmann et al., 2003). For hydration predictions by HYDCRYST, only the coordinates of the protein were taken, i.e. in this case all crystallographically found waters were discarded.

2.3 Model construction, data reduction, and prediction of structures and molecular parameters

Size, shape and properties of simple and complex proteins can be calculated by bead modelling procedures: in context with solution techniques, but also in the case of crystallographic and EM data (Zipper et al., 2005; Byron, 2008). Pilot tests applying whole-body approaches yield rough estimates of scattering and hydrodynamic molecular

parameters and allow estimates of partial specific volume, molecular volume and overall hydration (Durchschlag & Zipper, 2005). Among the more advanced approaches applied, Debye modelling (Glatter & Kratky, 1982) and the program CRYSOL (Svergun et al., 1995) may be used for the calculation of scattering profiles from (atomic) coordinates. For hydrodynamic modelling, the HYDRO program suite (consisting of several programs and many modern adaptations; e.g., programs HYDRO++ and HYDROPRO) (García de la Torre et al., 1994, 2000, 2007, 2010; Ortega et al., 2011a, 2011b), and ZENO (Kang et al., 2004; Mansfield & Douglas, 2008) should be emphasized. Filling-model strategies (instead of shell models usually applied for hydrodynamic modelling) have to be used for analyzing both scattering and hydrodynamic quantities.

In a broad range of circumstances, models composed of a multitude of beads have to be handled, in particular when resorting to (atomic) coordinates of huge macromolecules or if the biophysically relevant fine structure of hydrated models is required. Various programs, e.g. AtoB (Byron, 1997), SOMO (Brookes et al., 2010a, 2010b; Rai et al., 2005), PDB2AT, PDB2AM, and MAP2GRID (Zipper & Durchschlag, 2007) have been developed to transform crystallographic/NMR data information to bead models (unreduced initial models based on atoms, atomic groups, or amino acid residues, and reduced models at certain reduction levels). The hydration program HYDCRYST allows the efficient prediction of individual water molecules preferentially bound to proteins provided the accessible surface area has been calculated previously by a surface calculation program. Several complementary tools may be required, e.g., for conversion between data formats, calculation of scattering functions, or for a correct visualization of reduced and/or hydrated structures by RASMOL (Sayle & Milner-White, 1995).

2.4 Calculation of volumes and surfaces

Volume and surface properties of protein molecules can be calculated using crystallographic data sets as those deposited in the PDB. The molecular volumes thus obtained are either anhydrous volumes or, because of scarce waters, poorly hydrated volumes (Durchschlag & Zipper, 2008). Surface characteristics of proteins are obtained by using analytical surface calculation programs such as SIMS based on the rolling-ball strategy (Vorobjev & Hermans, 1997), yielding molecular surface and solvent-accessible surface areas (i.e. anhydrous and hydrated surfaces), in addition to a smooth 'molecular dot surface' required for values for the solvent-excluded volume and appliance of advanced hydration modelling strategies. The program SIMS may be applied either to the atomic coordinates or newly-created coordinates (e.g., gravity centres of AA residues) (Durchschlag & Zipper, 2008).

2.5 Calculation of hydration

The water molecules bound preferentially to proteins have properties different from those of the bulk water (higher order, lower mobility, higher density) (Durchschlag & Zipper, 2003). For roughly estimating the overall hydration (Durchschlag & Zipper, 2004, 2008), the hydration numbers found for the individual AA residues (Kuntz, 1971) may be used.

Knowledge of the exact anhydrous surface topography in terms of dot surface points, as obtained by SIMS, however, enables usage of advanced hydration algorithms. The normal vectors at these anhydrous dot surface points allow the creation of a huge amount of

hypothetical points for potential positions of water molecules; these waters are located at the normal distance above the anhydrous protein surface, i.e. at a distance usually corresponding to the water radius (probe radius). For the selection of preferential positions for bound water molecules, a new version of our in-house program HYDCRYST (now including also essentials of the related program HYDMODEL) is applied (creation of models based on the initial atomic coordinates of the protein or reduced models derived from the coordinates of AA residues) (Durchschlag & Zipper, 2001, 2002a, 2003). The selected waters, assigned to the accessible AA residues, are then attached to the dry protein models. Overall, the method is based on geometrical and energetic constraints, owing to the placing of definite hydration numbers to definite AA residues. The extent of hydration can be modulated step by step by several input parameters, in particular by variation of the dot density and a scaling factor acting on the hydration numbers used. Thereby different extents of water binding (minimum, medium, and maximum hydration) may be simulated.

2.6 Calculation of SAXS functions and of structural and hydrodynamic parameters

Radius of gyration, R_G, and the pair-distance distribution function, $p(r)$, of the protein models were calculated from the radii and coordinates of the constituent beads and the bead volume as statistical weight. The $p(r)$ function gives the relative number of distances of two points inside a particle as a function of the distance r. Calculation approaches and the prediction of structural and hydrodynamic parameters have been described in detail previously (Zipper et al., 2005).

3. Results

In the first instance, whole-body models of different shape were selected, to test the applicability of the reduction and calculation approaches applied. Based on these findings, anhydrous models for many proteins were constructed and their predicted structural and hydrodynamic properties were compared to experimental data. Because of obvious deficiencies of the anhydrous (dry) protein models, sophisticated hydration strategies had to be developed in the following, to account for the hydration contributions of the protein structures in solution. Application of our advanced hydration algorithms to various proteins revealed the effectiveness of the approaches applied.

3.1 Coarse-grained models

Models of quite different shape were selected, to prove the effectiveness of the approaches under consideration. Fig. 1 depicts a few illustrative examples: a sphere (S), a hollow sphere (HS), a prolate ellipsoid (PE) and an oblate ellipsoid (OE), composed of a multitude of beads, together with a set of the reduced models. As may be expected, the distance distribution functions, $p(r)$, of the different model structures vary considerably (Fig. 2), with respect to their form, position of maximum, and particle diameter. However, the profiles of unreduced and reduced models coincide nearly completely, proving the applicability of the model reduction process applied. As may be taken from the structural and hydrodynamic parameters presented in Table 1, the comparison of whole-body and multibead models indicates good accordance (data for R_G, D (and therefore also for s), and $[\eta]$). As shown previously (Zipper & Durchschlag, 2010a), choice of $[\eta]_{RVC}$ represents a reasonable approximation for the value of the intrinsic viscosity.

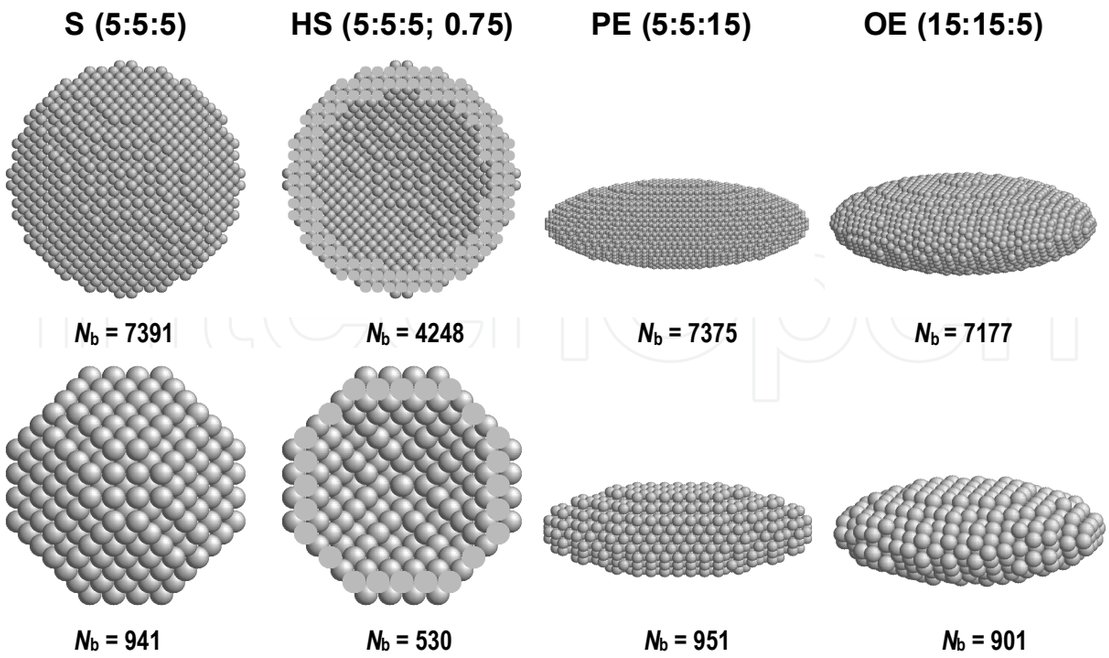


Fig. 1. Selected space-filling multibead models for sphere (S), hollow sphere (HS) with r_i/r_o = 0.75, and prolate and oblate ellipsoids of revolution (PE, OE). Each model is represented by a number of beads (N_b), depending on size, hollowness and axial ratio of the model under consideration. By increasing the bead radius of the model, N_b can be reduced considerably.

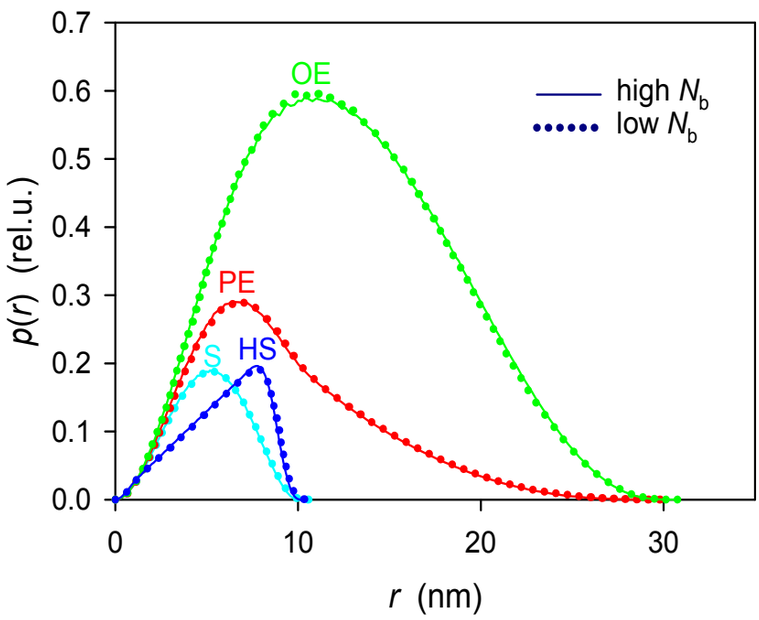


Fig. 2. Distance distribution functions $p(r)$ of the multibead models (S, HS, PE, OE) shown in Fig. 1. The profiles represent initial and reduced numbers of beads (high and low N_b); their integral values are proportional to the overall model volumes.

Shape	*a*:*b*:*c* (nm)	N_b	r_b (nm)	$V \times 10^{-3}$ (nm^3)	R_G (nm)	$D \times 10^7$ (cm^2/s)	$[\eta]_{NVC}$ (cm^3/g)	[(cr
S	**5:5:5, WB**			0.5236	3.873	4.286	2.639 [b]	
	–, MB	10777 – 65	0.20 – 1.00	0.357 ± 0.027	3.90 ± 0.10	4.46 ± 0.35 [c]	2.58 ± 0.22 [c]	4.38
		7391 – 65	0.225 – 1.00			4.32 ± 0.15 [d]	2.58 ± 0.23 [d]	4.38
		5577 – 65	0.25 – 1.00					
HS	**5:5:5, WB, r_i/r_o=0.75**			0.3027	4.449	4.286	4.564 [e]	
	–, MB	6300 - 104	0.20 – 0.80	0.212 ± 0.009	4.49 ± 0.04	4.28 ± 0.04 [d]	4.58 ± 0.13 [d]	6.43
PE	**5:5:15, WB**			1.5708	7.416	2.671	3.889 [f]	
	–, MB	9395 – 233	0.30 – 1.00	1.065 ± 0.030	7.42 ± 0.09	2.69 ± 0.03 [d]	3.82 ± 0.12 [d]	5.61
		5949 – 233	0.35 – 1.00					
OE	**15:15:5, WB**			4.7124	9.747	1.865	3.620 [g]	
	–, MB	10007 – 757	0.425 – 1.00	3.213 ± 0.040	9.77 ± 0.05	1.87 ± 0.01 [d]	3.60 ± 0.06 [d]	5.40
		7177 – 757	0.475 – 1.00					

S: sphere, HS: hollow sphere, PE: prolate ellipsoid of revolution, OE: oblate ellipsoid of revolution, WB: whole-b
of beads, r_b: radius of beads, V: total volume, R_G: radius of gyration, D: translational diffusion coefficient, $[\eta]$: int
correction, FVC: full volume correction, RVC: reduced volume correction, AVC: adjusted volume correction, M:

[a] All values were computed by HYDRO++9beta.
[b] Based on M = 298.8 kg/mol.
[c] Obtained by executing HYDRO in single-precision mode.
[d] Obtained by executing HYDRO in double-precision mode.
[e] Based on M = 172.7 kg/mol.
[f] Based on M = 896.3 kg/mol.
[g] Based on M = 2688.8 kg/mol.

Table 1. Comparison of structural and hydrodynamic parameters of whole-body and multi-bead m

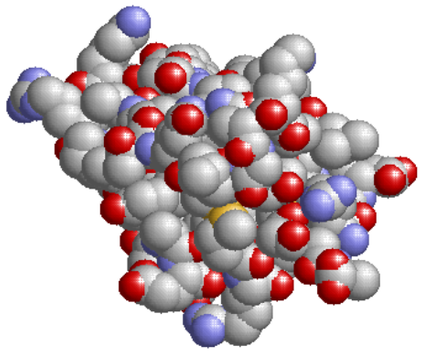
3.2 Anhydrous and hydrated protein models

Bovine pancreatic trypsin inhibitor (BPTI), a nonconjugated protein, was chosen as a small model protein, to demonstrate different types of anhydrous models and descriptions of reduction procedures to be applied in scattering and hydrodynamic modelling.

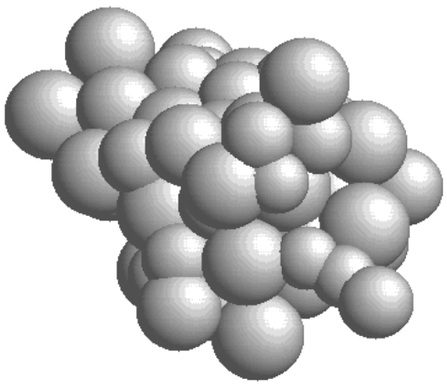
Fig. 3 compares the initial anhydrous model, the unreduced BPTI model based on atomic coordinates (A), with several approaches for reduced models (B-E). The most straightforward types of reduced models are based on AA residue coordinates, either on the entire residue or on parts of it (such as main chain and side chain, a concept adopted from SOMO). Procedures, such as SOMO and AtoB, are used in the program UltraScan II (Brookes et al., 2010a, 2010b; Demeler, 2005). An inspection of the $p(r)$ functions (Fig. 4) clearly proves that the straightforward procedures result in perfect agreement with the initial model, while the applied versions of SOMO, and in particular of AtoB, lead to remarkable discrepancies. This is partly due to the fast, but unprecise modus operandi for calculation of $p(r)$ in the present versions of UltraScan II (the errors caused by neglecting the bead radii increase with decreasing bead numbers). However, part of the disagreement of the $p(r)$ functions of SOMO and AtoB models with those of the initial and straightforward reduction procedures seems to be caused by the models themselves. The models created by SOMO and AtoB tend to be rather artificial and far from space-filling (Fig. 3) because of avoidance of overlapping beads. Our straightforward approaches, on the other hand, do not prevent the occurrence of bead overlaps. Even overlapping unequal-sized beads can be used for hydrodynamic modelling, either by applying a special expression for the interaction tensor in HYDRO (Zipper & Durchschlag, 1997, 1999) or by using ZENO. However, in principle, all types of anhydrous model constructions may be used as starting points for the creation of hydrated models. For obvious reasons, however, in the following we will restrict our considerations to our straightforward in-house approaches for the model reduction process.

The 30 S ribosomal subunit of *Thermus thermophilus* was used as a representative of a conjugated protein. It is a nucleoprotein composed of protein and ribonucleic acid moieties. As may be taken from Fig. 5, also in this case modelling may be achieved on the level of atomic coordinates (unreduced model: 51792 beads) and by applying AA and nucleotide residue coordinates (reduced model: 3917 beads) as well. Different approaches for the reduction process (running mean and grid reductions) and varying input parameters were applied, to test the applicability of the reduction procedures. The overall impression of the images obtained for all the models is satisfactory (Fig. 6). The comparison of the $p(r)$ functions, however, reveals that use of the cubic grid approach and equal-sized beads on lattice points (CLE) may fail (Fig. 7). And the same conclusion can be drawn from a meticulous inspection of the molecular parameters listed in Table 2. Again, the CLE variant rather leads to erroneous predictions of structural and hydrodynamic parameters.

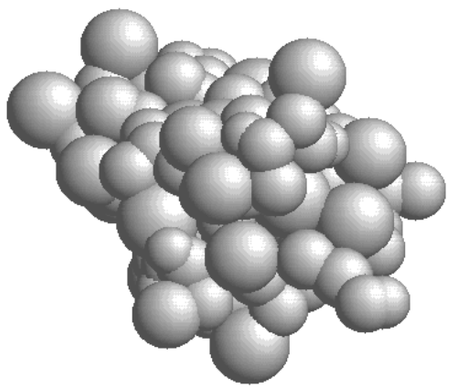
Fig. 8 illustrates the steps of the hydration approach HYDRCRYST, again applied to the model protein BPTI. Starting from the anhydrous protein model (A), a myriad of dot surface points is created (B). The normal vectors at these points are used for the localization of potential water points (C). Selecting water molecules on the protein surface by HYDCRYST yields the water molecules to be tracked on the protein surface, eventually establishing a realistic hydrated protein model, revealing, however, no complete water shell (D). In comparison to this calculated model, the structure based on crystal data demonstrates only a few bound water molecules (E).



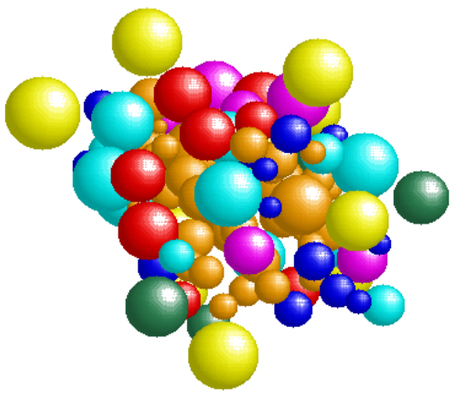
(A) N_b = 454, V = 7.781 nm^3, R_G = 1.103 nm



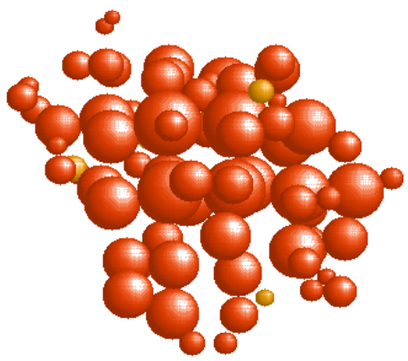
(B) N_b = 58, V = 7.781 nm^3, R_G = 1.106 nm



(C) N_b = 110, V = 7.781 nm^3, R_G = 1.107 nm



(D) N_b = 110, V = 3.666 nm^3, R_G = 1.220 nm



(E) N_b = 84, V = 2.358 nm^3, R_G = 1.129 nm

Fig. 3. Different types of anhydrous models for bovine pancreatic trypsin inhibitor (4PTI), created by different approaches: (A) Unreduced model for the protein based on atomic coordinates; the basic atoms are given in CPK colors, (B) Reduced model based on whole AA residues, (C) Reduced model based on AA residues split into a main-chain and a side-chain moiety, (D) SOMO model created by UltraScan II, (E) AtoB model created by UltraScan II. The given values for V and R_G were calculated from the coordinates and radii of beads.

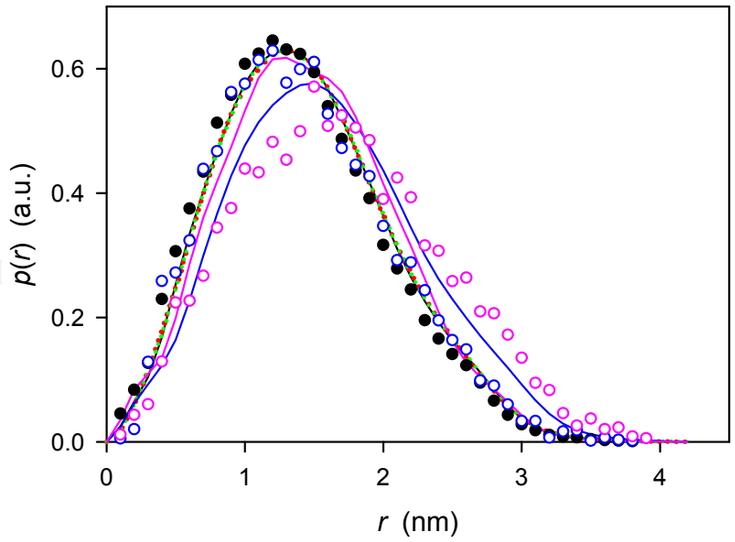


Fig. 4. Distance distribution functions $p(r)$ of the different types of anhydrous models for bovine pancreatic trypsin inhibitor (4PTI) shown in Fig. 3. The black line and the dotted red and green curves are the profiles obtained for the unreduced model and the models reduced to whole (red) or split (green) AA residues; the blue and pink lines represent the profiles calculated for the SOMO (blue) and AtoB (pink) models created by UltraScan II. These profiles were calculated by means of in-house programs from the coordinates and radii of the beads. The circles illustrate the $p(r)$ profiles of the unreduced model (black) and of the SOMO (blue) and AtoB (pink) model as provided directly by UltraScan II; these profiles were calculated from the coordinates of the beads only.

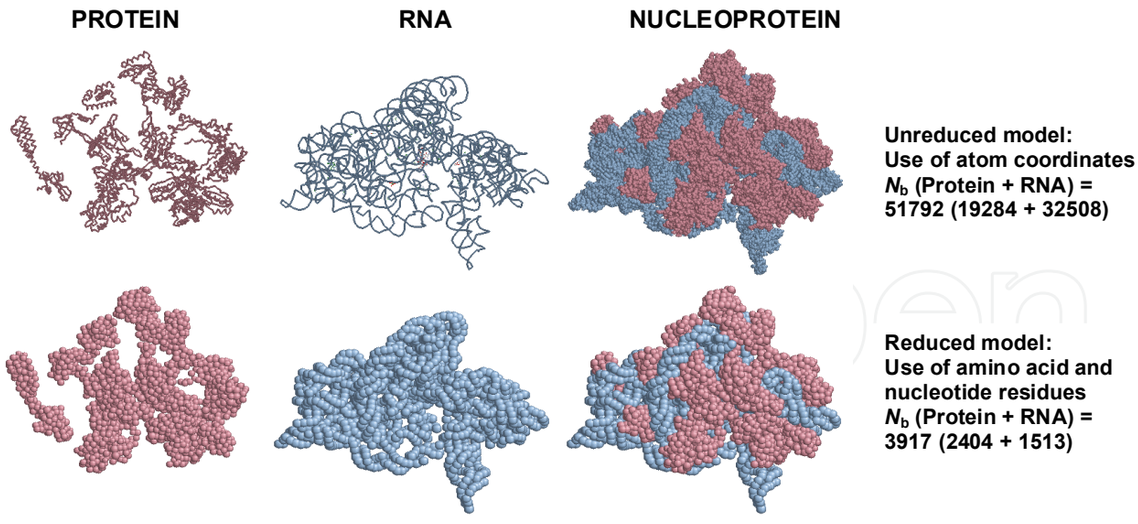


Fig. 5. Anhydrous models for the *Thermus thermophilus* 30 S ribosomal subunit (1FJG), a nucleoprotein composed of protein and RNA. The principle constituents (protein moiety in bluetint, and nucleic acid moiety in pinktint) are shown in backbone and space-filling formats, while the nucleoprotein is shown in space-filling format only. The images in the upper row represent the models based on atomic coordinates, while the lower row signifies the models based on the coordinates of AA and nucleotide residues.

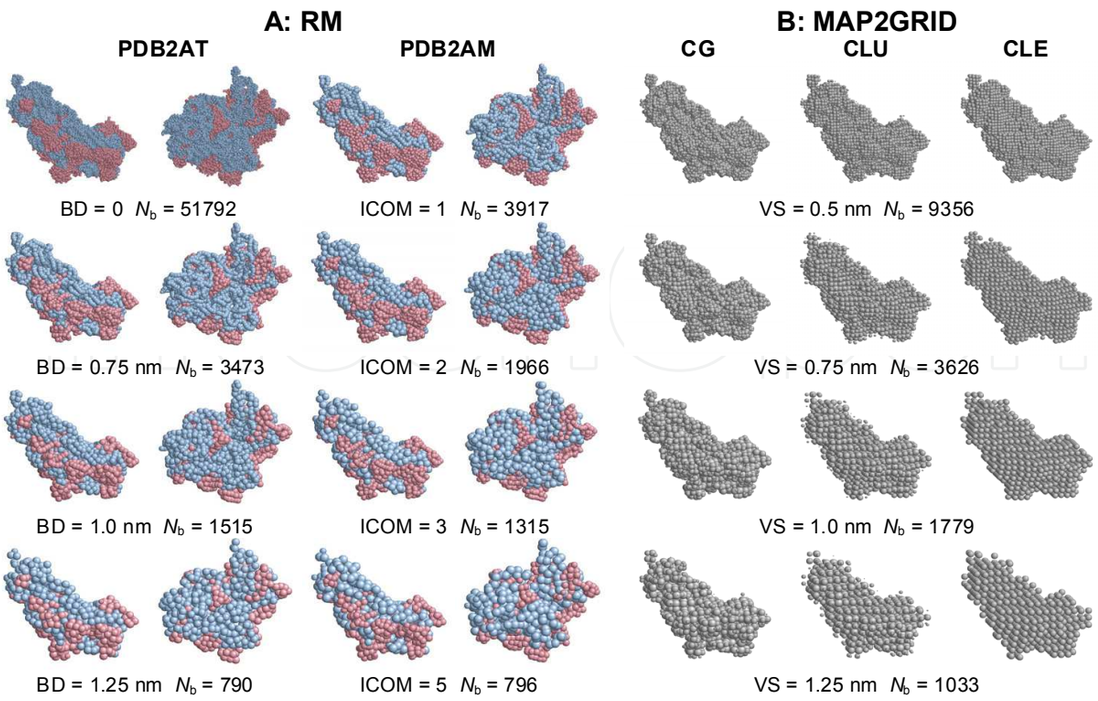


Fig. 6. Selected bead models for the *Thermus thermophilus* 30 S ribosomal subunit (1FJG), obtained by various approaches. (A) Running mean (RM) reductions to the crystal structure, applying different bead diameters BD (program PDB2AT) or compression indices ICOM (program PDB2AM) when transforming the original crystallographic information to various kinds of bead models. The models obtained with BD = 0 or ICOM = 1 correspond to the original crystal structure in atomic coordinates or residue coordinates, respectively. (B) Grid reductions (program MAP2GRID), mapping a given structure into a 3D cubic grid (C) of chosen voxel size (VS) and placing the unequal-sized (U) or equal-sized (E) beads at local centres of gravity (G) or on lattice points (L). Note that CG implicitly means CGU since the G approach is only applicable together with unequal-sized beads.

Model Method/mode	Parameter	N_b	V [a] (nm^3)	R_G [a] (nm)	χ [b]	$D \times 10^7$ [c] (cm^2/s)	$s \times 10^{13}$ [d] (s)	$[\eta]$ [e] (cm^3/g)
Crystal structure,	Initial	51792	800.49	6.652				
anhydrous			485.1	6.652		2.70	31.1	4.20
PDB2AT	BD (nm)							
	0.5	10484	800.49	6.650	0.0044			
			571.5	6.654		2.70	31.1	4.20
	1.0	1515	800.49	6.647	0.0079			
			692.3	6.650		2.71	31.3	4.15
	1.5	465	800.49	6.626	0.0374			
			676.5	6.633		2.75	31.7	3.99
PDB2AM	ICOM							
	1	3917	800.62	6.651	0.0050			
			664.5	6.661		2.70	31.1	4.21

	3	1315	800.62	6.646	0.0080			
			699.7	6.650		2.71	31.2	4.15
	5	796	800.62	6.638	0.0189			
			691.9	6.647		2.72	31.4	4.10
MAP2GRID	VS (nm)							
CG	0.5	9356	800.49	6.651	0.0032			
			664.0	6.649		2.69	31.0	4.24
	1.0	1779	800.49	6.650	0.0097			
			728.6	6.643		2.69	31.0	4.25
	1.5	678	800.49	6.640	0.0222			
			751.9	6.635		2.70	31.1	4.21
CLU	0.5	9356	800.49	6.656	0.0111			
			730.5	6.668		2.64	30.4	4.49
	1.0	1779	800.49	6.682	0.0388			
			776.9	6.685		2.58	29.8	4.77
	1.5	678	800.49	6.717	0.0767			
			790.4	6.721		2.52	29.1	5.16
CLE	0.5	9356	800.49	6.658	0.0753			
			779.7	6.744		2.60	30.0	4.71
	1.0	1779	800.49	6.679	0.286			
			799.3	7.013		2.50	28.8	5.27
	1.5	678	800.49	6.711	0.584			
			800.7	7.451		2.38	27.4	6.09

N_b: number of beads, V: total volume, R_G: radius of gyration, χ: goodness of the fit, D: translational diffusion coefficient, s: sedimentation coefficient, $[\eta]$: intrinsic viscosity; PDB2AT: program generating running-mean models by merging as many atoms in sequential order as fit into a bead of given diameter BD, PDB2AM: program generating running-mean models by merging a given number (compression index ICOM) of AA and nucleotide residues in sequential order to one bead, MAP2GRID: program generating bead models by mapping a given model onto a cubic or hexagonal grid of given edge length (voxel size) VS; CG, CLU, CLE: cubic grid models composed of unequal-sized beads placed on local centres of gravity (CG) or on lattice points (CLU) or of equal-sized beads placed on lattice points (CLE).

[a] The values given in the first line were obtained from the reduction program, the values given in the second line were obtained by the ZENO approach. Discrepancies in V are mainly due to the overlap of beads; the discrepancies in R_G for the models reduced with MAP2GRID (calculation mode CLE) result from the neglect of the special weights of the beads of these models by the ZENO approach.

[b] The values were obtained by comparing the calculated $p(r)$ function of the reduced models with the $p(r)$ function of the initial model. For computing the $p(r)$ functions weighting by volume was assumed throughout.

[c] The values are accurate to about $\pm 0.03 \times 10^{-7}$ cm^2/s.

[d] The values are accurate to about $\pm 0.3 \times 10^{-13}$ s.

[e] The values are accurate usually to about ± 0.06 cm^3/g, except for the models reduced by MAP2GRID where broader error bands are encountered (for mode CLU up to ± 0.08 cm^3/g, for mode CLE up to ± 0.09 cm^3/g).

Table 2. Comparison of structural and hydrodynamic parameters of anhydrous bead models for *Thermus thermophilus* 30 S ribosomal subunit (1FJG), generated by various reduction approaches.

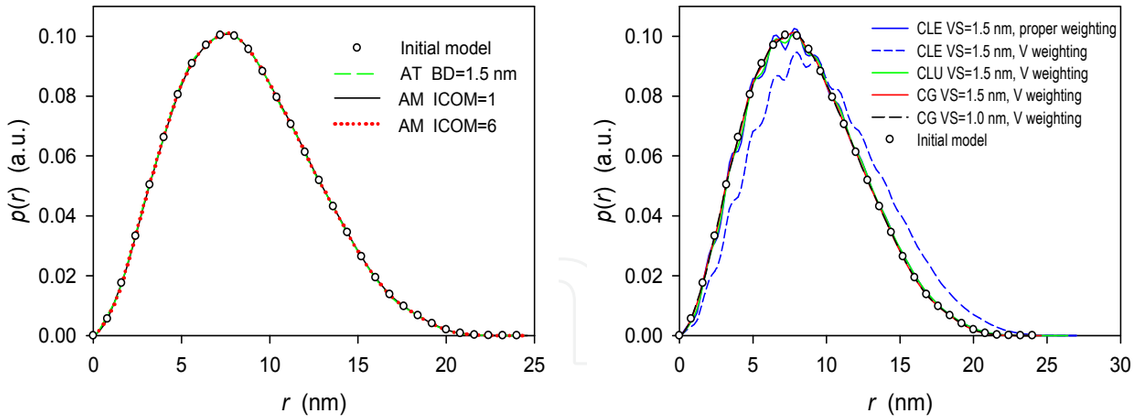


Fig. 7. Comparison of distance distribution functions $p(r)$ of selected unreduced and reduced bead models for the *Thermus thermophilus* 30 S ribosomal subunit (1FJG) as explained and shown in Fig. 6. The term 'V weighting' means that in the calculation of $p(r)$ the beads are weighted according to their volume. This kind of weighting is correct for grid models of type CG and CLU, but obviously not for type CLE .

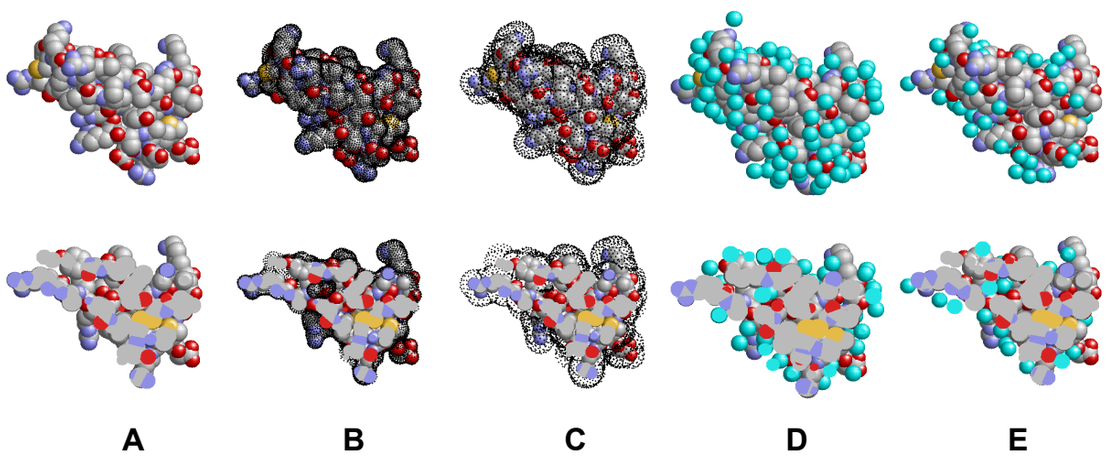


Fig. 8. Space-filling models for anhydrous and hydrated bovine pancreatic trypsin inhibitor (4PTI), together with central slabs: (A) Model for the anhydrous protein based on atomic coordinates; the basic atoms are given in CPK colors. (B) The anhydrous model and dot surface points (in black) created for the anhydrous contour. (C) The anhydrous model and surface points created by HYDCRYST for the contour of potential water points (in black) located at a certain distance from the initial surface points. (D) Model for the hydrated protein as obtained by HYDCRYST; bound waters are displayed in cyan. (E) Model for the hydrated protein as obtained by crystallography; waters are shown in cyan.

Some characteristics of BPTI and the model under discussion are outlined in Table 3, together with the properties of some other selected proteins discussed later on. In the following, our hydration approaches were applied to several proteins differing in size and complexity.

Carbonmonoxide myoglobin has been used as an example of a liganded protein (ligands haem and carbonmonoxide CMO). The hydration approach is demonstrated both for the

unreduced model based on atomic coordinates and the reduced model utilizing AA residue coordinates (Fig. 9). Evidently, the approach works in both cases.

Protein	PDB entry	M (kg/mol)	N_{AA}	Type of coordinates (AT or AM)	N_b (anhydrous model)	N_w (HYDCRYST)	N_b (hydrated model)	N_w (crystal)
BPTI	4PTI	6.5 [a]	58	AT	454	125	579	60
Apoferritin	2W0O	464.7 [b]	4080	AM	4080	6519	10599	5352
Aquaporin 1	1J4N	26.2 [c]	249	AT	1852	345	2197	114
Aquaporin Z	2O9D	47.3 [d]	464	AT	3356	577	3933	96

Symbols and abbreviations: M: molar mass; N_{AA}: number of AA residues; N_b: number of beads used for modelling; N_w: number of water molecules (of radius r_w = 0.145 nm in the case of HYDCRYST); AT: atomic coordinates; AM: AA residue coordinates.

[a] Composed of one chain.

[b] Contains 24 identical chains.

[c] Asymmetric unit containing one chain.

[d] Asymmetric unit containing two nearly identical chains.

Table 3. Properties of selected proteins and models used for hydration calculations.

A representative of large proteins, apoferritin, is shown in Fig. 10, highlighting some further aspects of particular interest. Water molecules are bound both to the outer and inner surface of this hollow protein; in addition, a few waters seem to occur also in some channels between interior and outside. Again, the number of waters predicted by HYDCRYST exceeds the number found by crystallography (Table 3).

A comparison of structural and hydrodynamic parameters of anhydrous and hydrated models for apoferritin (Table 4) reveals that the properties of the hydrated protein (V, R_G, D, s, $[\eta]$) considerably deviate from those of the anhydrous one. It is quite obvious that a critical comparison of crystallographic data with the findings of solution techniques requires strict consideration of hydration contributions.

There is compelling evidence that water molecules can also be visualized in typical water channels of membrane proteins. The aquaporins shown in Fig. 11 are illustrative examples showing that the existence of a water channel as predicted by HYDCRYST is in full agreement with the crystallographic waters found in this case (Table 3). Some pilot tests also showed that the width of channels can be determined by variation of the probe radius.

Quantifying the results of anhydrous and hydrated proteins in terms of distance distribution functions $p(r)$ again demonstrates that the properties of hydrated models differ significantly from anhydrous models. For the proteins selected (BPTI, apoferritin, aquaporin 1, aquaporin Z), this is shown in Fig. 12. The profiles for the proteins plus the HYDCRYST waters deviate significantly from those of the anhydrous proteins. This effect is most pronounced in the case of small proteins (such as BPTI), whereas, due to the size ratio between protein and waters, the difference is rather insignificant with large molecules (such as apoferritin). For a given protein, owing to the scarce numbers of waters in the crystal data set, generally the functions obtained on the basis of the proteins plus crystal waters rather resemble the anhydrous proteins.

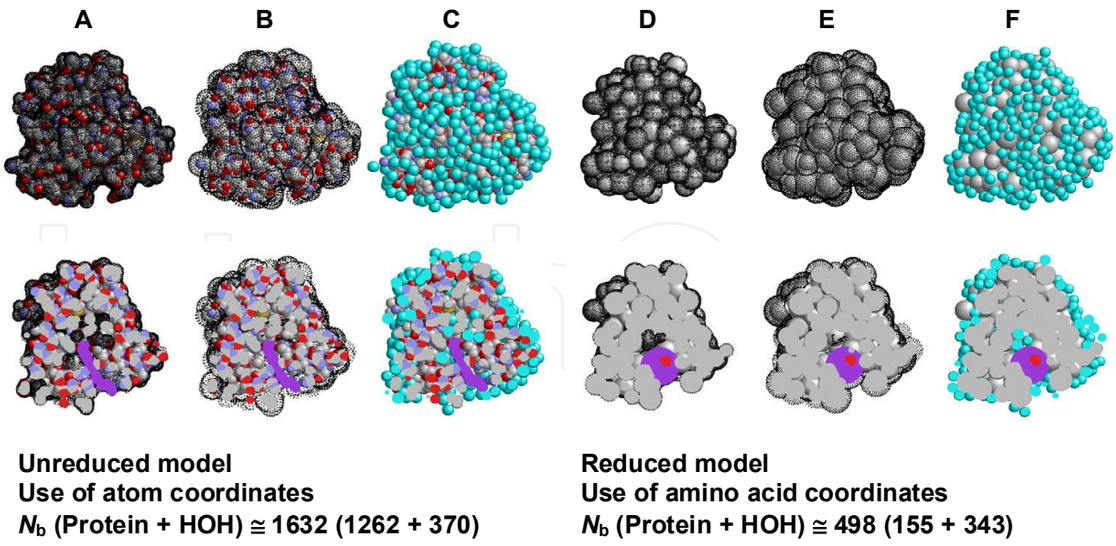


Fig. 9. Space-filling models for anhydrous and hydrated carbonmonoxide myglobin (1VXC), together with central slabs: (A-C) Unreduced model for the anhydrous protein based on atomic coordinates plus dot surface points (A) or surface points created by HYDCRYST (B) or for the hydrated protein as obtained by HYDCRYST (C). The basic atoms of the protein and the ligand CMO are given in CPK colors and the ligand haem in purple; dot surface points and potential water points are shown in black, and the bound waters are displayed in cyan. (D-E) Reduced model for the anhydrous protein based on AA residue coordinates plus dot surface points (D) or surface points created by HYDCRYST (E) or for the hydrated protein as obtained by HYDCRYST (F). AA residues are shown in gray, and the ligands haem and CMO in purple and red, respectively.

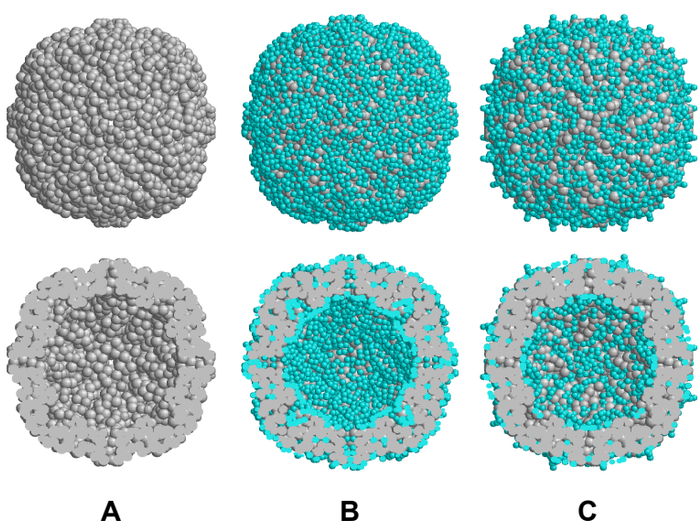


Fig. 10. Space-filling model for anhydrous and hydrated apoferritin (2W0O), together with illustrative central slabs. (A) Model for the anhydrous protein based on AA residue coordinates; AA residues are given in gray. (B) Model for the hydrated protein as obtained by HYDCRYST; bound waters are displayed in cyan. (C) Model for the hydrated protein as obtained by crystallography; waters are shown in cyan.

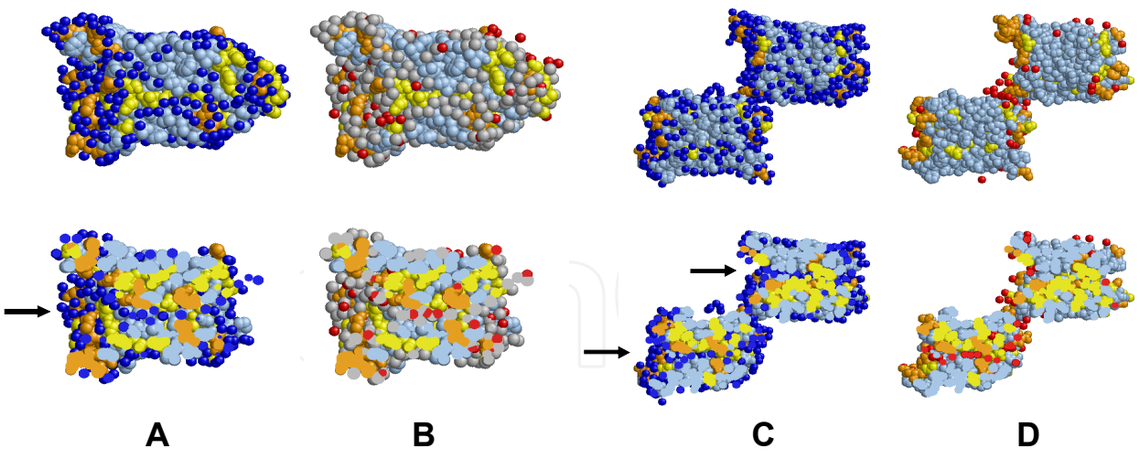


Fig. 11. Space-filling models and slabs for the asymmetric units of aquaporin 1 (1J4N; A, B) and aquaporin Z (2O9D; C, D), together with illustrative central slabs. Models for the anhydrous proteins are based on atomic coordinates; the atoms of hydrophobic AA residues are drawn in bluetint, those of polar residues in yellow and those of charged residues in orange. Models for the hydrated proteins refer to proteins plus waters predicted by HYDCRYST (A, C) or found by crystallography (B, D); bound waters are displayed in blue (A,C) and red (B,D), respectively. Water channels are indicated by arrows.

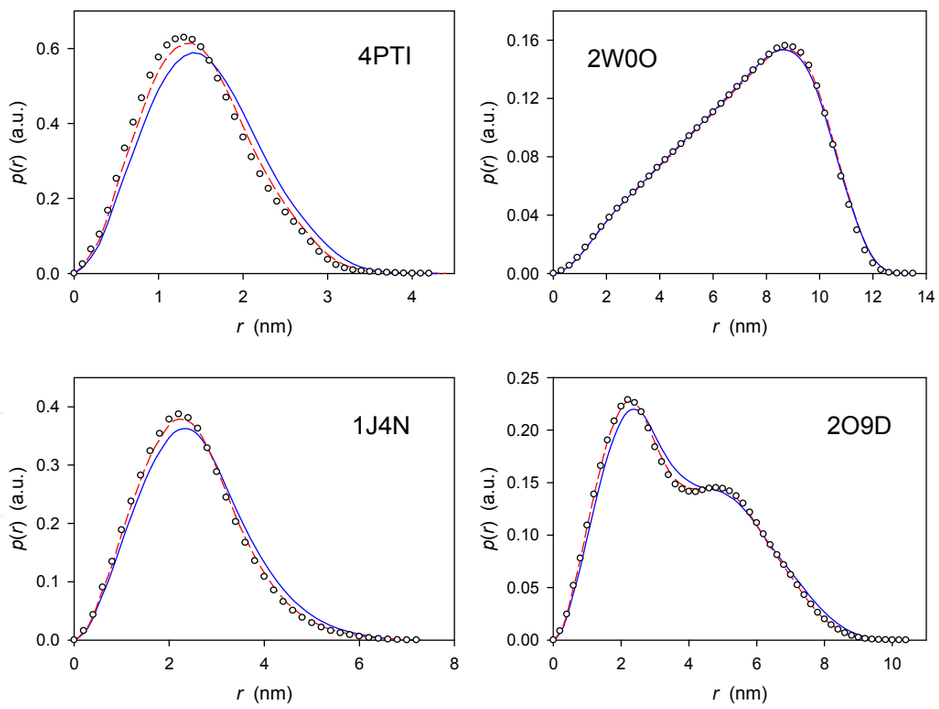


Fig. 12. Comparison of distance distribution functions $p(r)$ of space-filling multibead models for the selected anhydrous and hydrated proteins: BPTI (4PTI), apoferritin (2W0O), aquaporin 1 (1J4N), aquaporin Z (2O9D). ○: anhydrous model; ——: hydrated model: protein plus waters predicted by HYDCRYST; – – –: hydrated model: protein plus waters found by crystallography.

Model Method/mode	f_K	N_w	N_b	V (nm³)	R_G (nm)	$D \times 10^{7}$ [a] (cm²/s)	$s \times 10^{13}$ [a] (s)	$[\eta]$ [b] (cm³/g)
Crystal structure, anhydrous			32736 [c]	564.0	5.305	3.45 [d]	18.2 [d]	3.21 [d]
Anhydrous RM models:								
AT, BD = 1.0 nm			1056	564.0	5.289	3.52	18.64	3.23
AM, ICOM = 1			4080	564.0	5.292	3.49	18.48	3.25
						3.45 [d]	18.2 [d]	3.21 [d]
AM, ICOM = 10			408	564.0	5.282	3.54	18.72	3.26
Anhydrous CG models:								
AT, VS = 1.0 nm			1082	564.0	5.296	3.48	18.38	3.36
AM, VS = 1.0 nm			854	564.0	5.303	3.51	18.56	3.29
Hydrated models:								
AM, SC0 (low hydration)	1.0	6519	10599	723.0	5.323	3.33	17.64	3.71
						3.31 [d]	17.5 [d]	3.62 [d]
AM, SC0, CG, VS = 1.0 nm			1235	723.0	5.324	3.40	17.96	3.63
AM, EC0	1.0	[6519]	4080	723.0	5.318	3.42	18.09	3.48
						3.37 [d]	17.8 [d]	3.44 [d]
AM, SC9 (high hydration)	4.0	8347	12427 [c]	767.6	5.365	3.29 [d]	17.4 [d]	3.72 [d]
AM, SC9, CG, VS = 1.0 nm			1262	767.6	5.360	3.37	17.83	3.72
AM, EC9	4.0	[8347]	4080	767.6	5.361	3.39	17.92	3.59
						3.34 [d]	17.6 [d]	3.55 [d]

f_K: hydration factor acting on the hydration numbers, N_w: number of water molecules, N_b: number of beads, V: total volume, R_G: radius of gyration, D: translational diffusion coefficient, s: sedimentation coefficient, $[\eta]$: intrinsic viscosity; AT: models based on atoms or atomic groups, AM: models based on AA residues, BD: bead diameter, ICOM: compression index, VS: voxel size, SC: hydration is expressed by attachment of N_w discrete beads to the anhydrous model, EC: hydration is expressed by appropriately increased dimensions of the solvent-accessible beads of the anhydrous model, CG: models obtained by mapping a given model onto a cubic grid (edge length VS) and placing the resulting unequal-sized beads on local centres of gravity.

[a] If not stated otherwise the data were obtained by means of program HYDRO.

[b] If not stated otherwise the data represent $[\eta]_{RVC}$ obtained by means of program HYDRO applying the approach of reduced volume correction.

[c] Too many beads for prediction of hydrodynamic parameters by means of program HYDRO.

[d] The data were obtained by means of program ZENO; the limits of error typically amount to ± 0.04 x 10^{-7} cm²/s for D, ± 0.2 x 10^{-13} s for s, and ± 0.05 cm³/g for $[\eta]$.

Table 4. Comparison of structural and hydrodynamic parameters of anhydrous and hydrated models for apoferritin (2W0O).

The multidrug resistance transporter Sav1866 is another example of a membrane protein. In this case crystallography was able to identify only very few water molecules, while our calculative approach HYDCRYST reveals many waters (Fig. 13). In context of membrane proteins, it has to be mentioned, however, that the preferential water molecules identified only indicate that at these positions individual waters could exist in principle, provided that they have contact with water. On the other hand, the model also shows that waters are preferentially bound to charged and polar residues, while hydrophobic residues are avoided.

Finally, modelling of the giant protein *Lumbricus terrestris* haemoglobin calls for tough measures. Fig. 14 demonstrates SAXS-based experimental scattering intensities, $I(h)$, and distance distribution functions, $p(r)$, of the native HBL complex and of its dodecameric subunit, respectively. The two profiles render information of the molecules under investigation in the reciprocal space and real space, respectively. The two profiles also show impressively the inverse relation between particle size (real space) and the decay of scattering intensity (reciprocal space) by reflecting the different size of the subunit and the complex.

Modelling of the anhydrous and hydrated HBL complexes requires consideration of several measures and precautions regarding model reduction and careful attention of AA residues missing in the crystal structure. Additional, appropriately located beads may serve as substitutes for the missing residues. Figs. 15 and 16 demonstrate that all structural features are retained through the reduction process, and the possibility to generate hydrated objects in both cases.

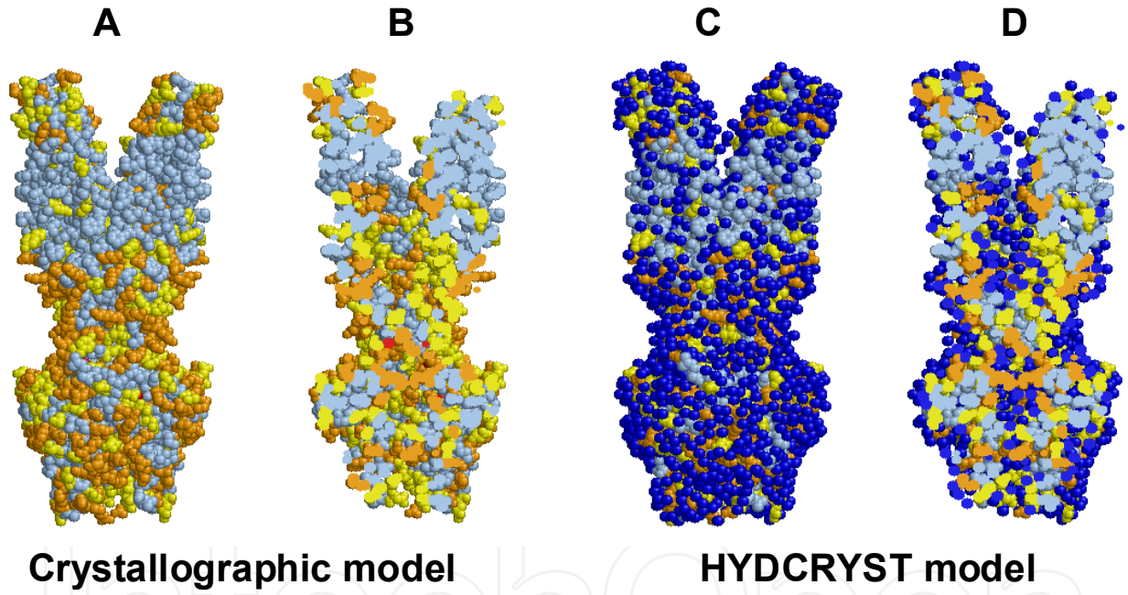


Fig. 13. Space-filling models and slabs for the hydrated multidrug resistance transporter Sav1866 (2ONJ), based on crystallographic data (A, B) and HYDCRYST modelling procedures (C, D). Groups of special AA residues are highlighted in bluetint (hydrophobic residues), yellow (polar residues), or orange (charged residues). Waters localized by crystallography are shown in red, and waters identified by HYDCRYST are displayed in blue.

Advanced 3D reconstructions of the HBL complex, obtained from cryoelectron microscopy (Krebs et al., 1998), yield a voxel density distribution that can be interpreted in terms of anhydrous and hydrated protein volumes (Fig. 17). Thus, the EM-based data allow construction of anhydrous protein models and hydrated models as well (Fig. 18). The resemblance between the crystallography-based and EM-based protein models is striking.

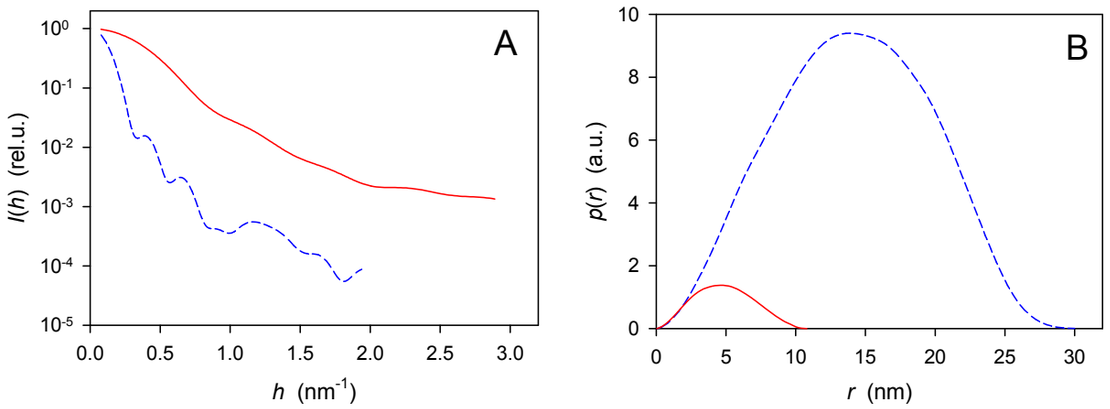


Fig. 14. Experimental scattering profiles of the HBL complex of *Lumbricus terrestris* haemoglobin (blue, dashed lines) and its dodecameric subunit (red, solid lines). (A) Scattering intensities, $I(h)$, normalized to $I(0)=1$; (B) distance distribution functions, $p(r)$, with areas under the $p(r)$ functions proportional to the particle masses.

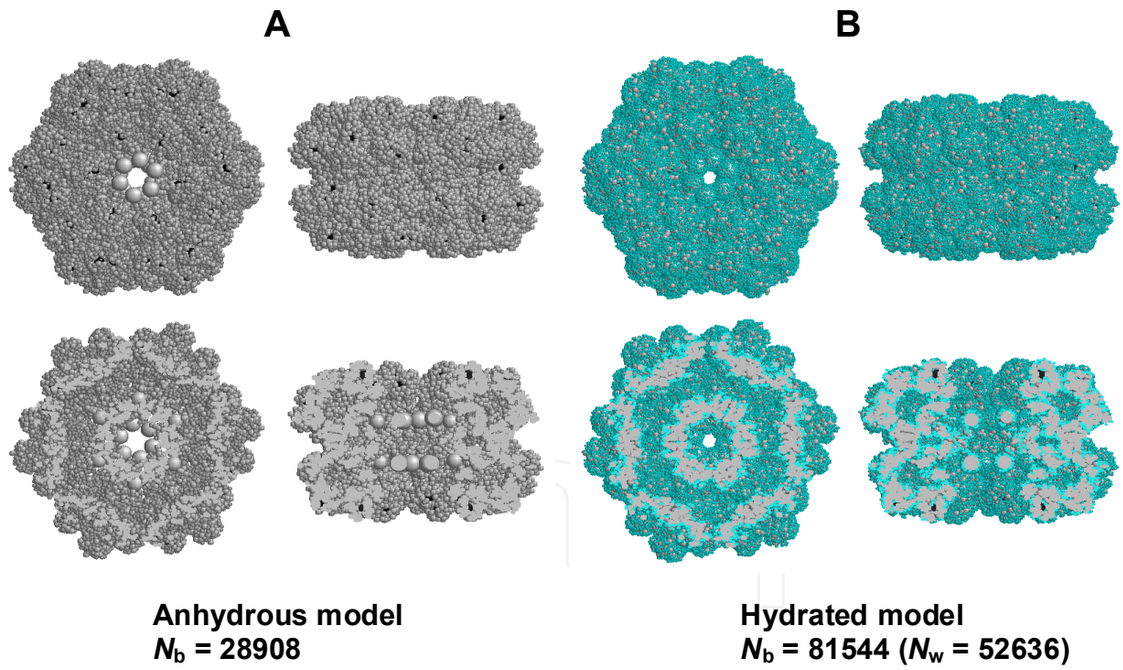


Fig. 15. Space-filling models for anhydrous and hydrated HBL complexes of *Lumbricus terrestris* haemoglobin (2GTL), together with illustrative central slabs. (A) Model for the anhydrous protein based on AA and substitute residue coordinates. AA residues are displayed in gray and haem groups in black; 24 additional large beads (in gray) are substitutes for the AA residues missing in the crystal structure at the N and C termini of the linker chains. (B) Model for the hydrated protein as obtained by HYDMODEL; bound waters are displayed in cyan.

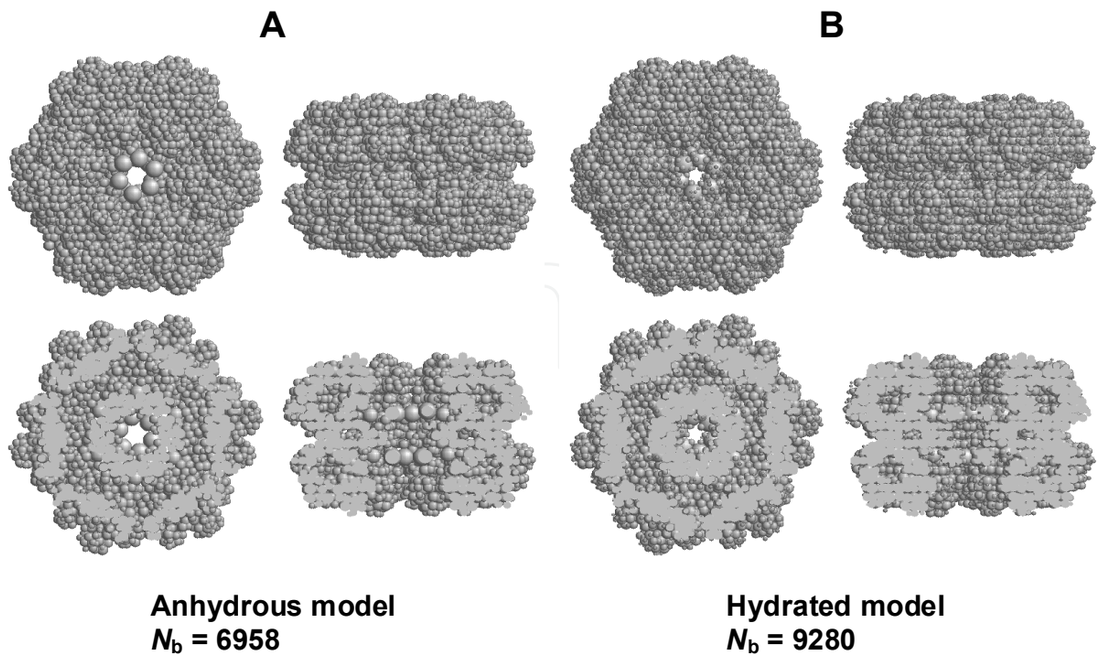


Fig. 16. Space-filling reduced models for anhydrous and hydrated HBL complexes of *Lumbricus terrestris* haemoglobin (2GTL), together with illustrative central slabs. The models were generated by means of MAP2GRID, by mapping the anhydrous and hydrated models shown in Fig. 15 onto hexagonal grids of edge length 1.05 nm and placing the resulting unequal-sized beads on local gravity centres. (A) Model for the anhydrous protein. (B) Model for the hydrated protein.

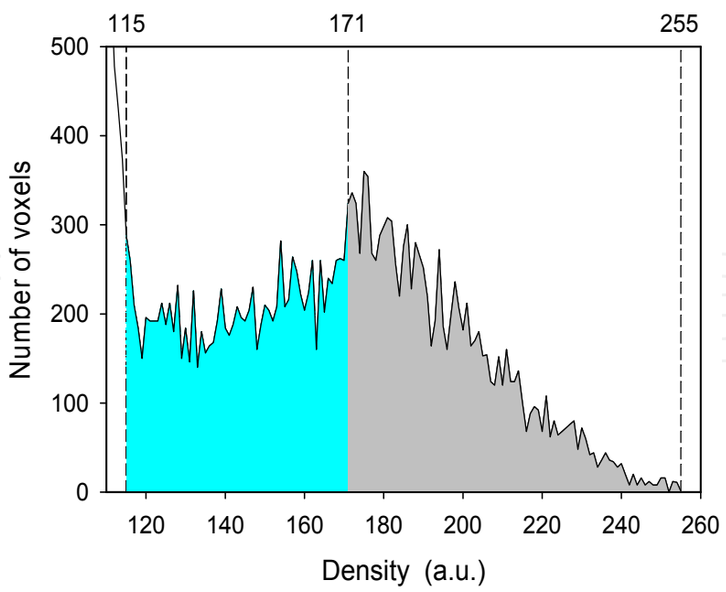


Fig. 17. Voxel density distribution of a 3D reconstruction of the HBL complex of *Lumbricus terrestris* haemoglobin as obtained from cryoelectron microscopy. The density peak between 115 and 255 is caused by the protein; the threshold at 171 is compatible with the anhydrous protein volume, while lower thresholds correspond to hydrated volumes.

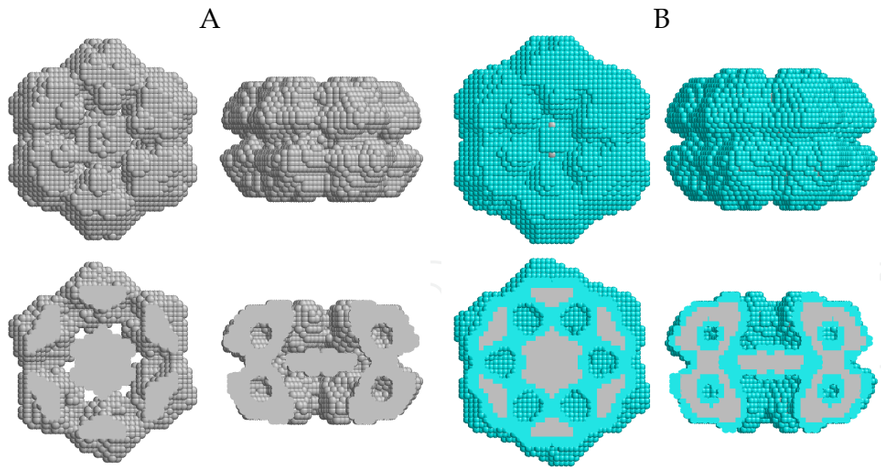


Fig. 18. Top and side views of EM models for the HBL complex of *Lumbricus terrestris* haemoglobin. The different colors represent the anhydrous and hydrated models (shown in gray and cyan, respectively) as given by the thresholds and color codes in Fig. 17.

4. Discussion

Several classes of coarse-grained models and biologically relevant proteins of different size and complexity have been used, to test the applied approaches concerning bead reduction and hydration algorithms, and to compare protein structures and molecular parameters obtained by quite different high-resolution and low-resolution techniques in the crystal, solid state or solution.

With all models and proteins, the bead reduction steps applied were successful, provided the reduction process was not too excessive and grid approaches with beads on lattice points are avoided. Reductions by a factor of 10 provided parameter predictions identical to the initial models, while reductions exceeding this factor can lead to slight deviations.

In the case of proteins, usage of precise anhydrous 3D models (derived from atomic or amino acid coordinates or appropriate models) along with computation of the exact surface topography (molecular dot surface) and our recent hydration approaches (program HYDCRYST) allow the prediction of discrete water molecules preferentially bound to particular residues. In this context, various approaches and procedural methods were tested: sequence of assignment to accessible residues, atomic vs. amino acid coordinates, original vs. coarse-grained models, fine-tuning of input parameters, variation of channel characteristics (e.g., width), rugosity effects. A critical comparison of the water sites on the surface, in active centres, ligand binding sites, interior, crevices, channels, contact areas etc. proves far-reaching identity of crystallographic data, if available, and our predictions. Examples presented include proteins ranging from simple to complex, multisubunit, liganded proteins and water-channels in membrane proteins (e.g., aquaporins) as well. Our hydration algorithms allow the prediction of the number and position of discrete water molecules, even in those cases where no or scant crystallographic waters or water channels have been identified. Our approaches may be used in the future as useful tools for improving crystal data.

The good agreement of the results found for hydrated models by crystallography, SAXS, and other techniques offers the possibility to complement different techniques and to predict details such as the localization of potential water sites - even in those cases where no

or insufficient amounts of waters, water clusters or water channels have been identified by crystallography. Variation of input parameters (such as probe radius) should also allow the width and type of channels (other than water channels) to be established. Visualization of protein sites of special concern (charged, hydrophilic and hydrophobic residues and patches, radiosensitive groups, active centres of enzymes, ligand binding sites, docking sites and contact areas) together with individual waters provides the basis for a much deeper understanding of the mechanisms of biological action and effective biotechnological application.

Considering quite different proteins, it can be stated that the majority of waters is bound to polar AA residues, located primarily outside. Some hydrophobic patches on the protein surface prevent the formation of a perfect, uniform water shell. The hydrated models also reveal that some water molecules can be found in the protein interior (e.g. in crevices, clefts, cavities or between subunits).

Both structural and hydrodynamic parameters and scattering profiles proved good agreement between observed and predicted quantities. In conclusion, about two water molecules were found per AA residue, corresponding to about 0.35 g of water per g of protein. Anhydrous and hydrated models differ substantially in their volume-to-mass ratios (1.2 vs. 1.6-1.7). Detailed data on various proteins may be found in previous papers (Durchschlag et al., 2007; Durchschlag & Zipper, 2001, 2002a, 2002b, 2003, 2004, 2005, 2006, 2008; Zipper & Durchschlag, 2002a, 2002b, 2010b).

5. Acknowledgement

The authors are much obliged to several scientists for use of their computer programs: to Y.N. Vorobjev for SIMS, B. Demeler for UltraScan, to J. García de la Torre for the HYDRO suite, to M.L. Mansfield for ZENO, and R.A. Sayle for RASMOL.

6. References

Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N. & Bourne, P.E. (2000). The protein data bank. *Nucleic Acids Res*, 28, 235-242

Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M.-C.; Estreicher, A.; Gasteiger, E.; Martin, M.J.; Michoud, K.; O'Donovan, C.; Phan, I.; Pilbout, S. & Schneider, M. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res*, 31, 365-370

Brookes, E.; Demeler, B. & Rocco, M. (2010a). Developments in the US-SOMO bead modeling suite: New Features in the direct residue-to-bead method, improved grid routines, and influence of accessibile surface area screening. *Macromol Biosci*, 10, 746-753

Brookes, E.; Demeler, B.; Rosano, C. & Rocco, M. (2010b). The implementation of SOMO (SOlution MOdeller) in the UltraScan analytical ultracentrifugation data analysis suite: enhanced capabilities allow the reliable hydrodynamic modeling of virtually any kind of biomacromolecule. *Eur Biophys J*, 39, 423-435

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. (2008). Hydrodynamic modeling: the solution conformation of macromolecules and their complexes. *Methods Cell Biol*, 84, 327-373

Creighton, T.E. (2010a). *The Physical and Chemical Basis of Molecular Biology*. Helvetian Press, ISBN 978-0-9564781-0-8, www.HelvetianPress.com

Creighton, T.E. (2010b). *The Biophysical Chemistry of Nucleic Acids & Proteins.* Helvetian Press, ISBN 978-0-9564781-1-5, www.HelvetianPress.com

Demeler, B. (2005). UltraScan – A comprehensive data analysis software package for analytical ultracentrifugation experiments. In: *Analytical Ultracentrifugation: Techniques and Methods.* Scott, D.J., Harding, S.E. & Rowe, A.J. (Eds.), pp. 210-230, Royal Society of Chemistry, ISBN 0-85404-547-3, Cambridge, UK

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*, 67, 479-486

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). Modeling of protein hydration with respect to X-ray scattering and hydrodynamics. *Prog Colloid Polymer Sci*, 119, 131-140

Durchschlag, H. & Zipper, P. (2003). Modeling the hydration of proteins: prediction of structural and hydrodynamic parameters from X-ray diffraction and scattering data. *Eur Biophys J*, 32, 487-502

Durchschlag, H. & Zipper, P. (2004). Modeling the hydration of proteins at different pH values. *Prog Colloid Polymer Sci*, 127, 98-112

Durchschlag, H. & Zipper, P. (2005). Calculation of volume, surface, and hydration properties of biopolymers. In: *Analytical Ultracentrifugation: Techniques and Methods.* Scott, D.J., Harding, S.E. & Rowe, A.J. (Eds.), pp. 389-431, Royal Society of Chemistry, ISBN 0-85404-547-3, Cambridge, UK

Durchschlag, H. & Zipper, P. (2006). Tracking water molecules on protein surfaces. *Bussei Kenkyu*, 87, 68-69

Durchschlag, H. & Zipper, P. (2007). X-ray-based structural models for the *in situ* X-irradiation of a sulfur-containing enzyme. *Radiat Phys Chem*, 76, 1295-1301

Durchschlag, H. & Zipper, P. (2008). Volume, surface and hydration properties of proteins. *Prog Colloid Polymer Sci*, 134, 19-29

Durchschlag, H.; Zipper, P. & Krebs, A. (2007). A comparison of protein models obtained by small-angle X-ray scattering and crystallography. *J Appl Cryst*, 40, 1123-1134

García de la Torre, J.; Amorós, D. & Ortega, A. (2010). Intrinsic viscosity of bead models for macromolecules and nanoparticles. *Eur Biophys J*, 39, 381-388

García de la Torre, J.; del Rio Echenique, G. & Ortega, A. (2007). Improved calculation of rotational diffusion and intrinsic viscosity of bead models for macromolecules and nanoparticles. *J Phys Chem B*, 111, 955-961

García de la Torre, J.; Huertas, M.L. & 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.; Navarro, S.; López Martínez, M.C., Díaz, F.G. & López Cascales, J. (1994) HYDRO: A computer program for the prediction of hydrodynamic properties of macromolecules. *Biophys J*, 67, 530-531

Glatter, O. & Kratky, O. (Eds.). (1982). *Small Angle X-ray Scattering*. Academic Press, London, ISBN 0-12-286280-5

Kang, E.-H.; Mansfield, M.L. & Douglas, J.F. (2004). Numerical path integration technique for the calculation of transport properties of proteins. *Phys Rev E*, 69, 031918

Krebs, A.; Lamy, J.; Vinogradov, S.N. & Zipper, P. (1998). *Lumbricus terrestris* hemoglobin: a comparison of small-angle X-ray scattering and cryoelectron microscopy data. *Biopolymers*, 45, 289-298

Kuntz, I.D. (1971). Hydration of macromolecules. III. Hydration of polypeptides. *J Am Chem Soc*, 93, 514-516

Levy, Y. & Onuchic, J.N. (2006). Water mediation in protein folding and molecular recognition. *Annu Rev Biophys Biomol Struct*, 35, 389-415

Mansfield, M.L & Douglas, J.F. (2008). Improved path integration method for estimating the intrinsic viscosity of arbitrarily shaped particles. *Phys Rev E*, 78, 046712

Ortega, A.; Amorós, D. & García de la Torre, J. (2011a). Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys J*, 101, 892-898

Ortega, A.; Amorós, D. & García de la Torre, J. (2011b). Global fit and structure optimization of flexible and rigid macromolecules and nanoparticles from analytical ultracentrifugation and other dilute solution properties. *Methods*, 54, 115-123

Papoian, G.A.; Ulander, J.; Eastwood, M.P., Luthey-Schulten, Z. & Wolynes, P.G. (2004). Water in protein structure prediction. *Proc Natl Acad Sci USA*, 101, 3352-3357

Rai, N.; Nöllmann, M.; Spotorno, B.; Tassara, G.; Byron, O. & Rocco, M. (2005). *SOMO* (*SO*lution *MO*deler): Differences between X-ray and NMR-derived bead models suggest a role for side chain flexibility in protein hydrodynamics. *Structure*, 13, 723-734

Ravelli, R.B.G. & Garman, E.F. (2006). Radiation damage in macromolecular cryo-crystallography. *Curr Opin Struct Biol,* 16, 624-629

Rupp, B. (2010). *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. Garland Science, ISBN 978-0-8153-4081-2, New York

Sayle, R.A. & Milner-White, E.J. (1995). RASMOL: biomolecular graphics for all. *Trends Biochem Sci*, 20, 374-376

Serdyuk, I.N.; Zaccai, N.R. & Zaccai, J. (2007). *Methods in Molecular Biophysics: Structure, Dynamics, Function*. Cambridge University Press, ISBN-13 978-0-521-81524-6, Cambridge, UK

Svergun, D.; Barberato C. & Koch, M.H.J. (1995). *CRYSOL* – a program to evaluate X-ray solution sattering of biological macromolecules from atomic coordinates. *J Appl Cryst*, 28, 768-773

Vorobjev, Y.N. & Hermans, J. (1997). SIMS: computation of a smooth invariant molecular surface. *Biophys J*, 73, 722-732

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. (1999). Prediction of hydrodynamic parameters from 3D structures. *Prog Colloid Polym Sci*, 113, 106-113

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

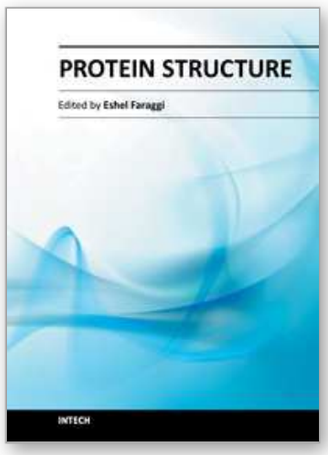
Zipper, P. & Durchschlag, H. (2002b). Modeling of complex protein structures. *Physica A*, 314, 613-622

Zipper, P. & Durchschlag, H. (2007). Modeling complex biological macromolecules: reduction of multibead models. *J Biol Phys*, 33, 523-539

Zipper, P. & Durchschlag, H. (2010a). Hydrodynamic multibead modeling: problems, pitfalls and solutions. 1. Ellipsoid models. *Eur Biophys J*, 39, 437-447

Zipper, P. & Durchschlag, H. (2010b). Hydrodynamic multibead modeling: problems, pitfalls and solutions. 2. Proteins. *Eur Biophys J*, 39, 481-495

Zipper, P.; Durchschlag, H. & Krebs, A. (2005). Modelling of biopolymers. In: *Analytical Ultracentrifugation: Techniques and Methods*. Scott, D.J., Harding, S.E. & Rowe, A.J. (Eds.), pp. 320-371, Royal Society of Chemistry, ISBN 0-85404-547-3, Cambridge, UK



Protein Structure
Edited by Dr. Eshel Faraggi

ISBN 978-953-51-0555-8
Hard cover, 396 pages
Publisher InTech
Published online 20, April, 2012
Published in print edition April, 2012

Since the dawn of recorded history, and probably even before, men and women have been grasping at the mechanisms by which they themselves exist. Only relatively recently, did this grasp yield anything of substance, and only within the last several decades did the proteins play a pivotal role in this existence. In this expose on the topic of protein structure some of the current issues in this scientific field are discussed. The aim is that a non-expert can gain some appreciation for the intricacies involved, and in the current state of affairs. The expert meanwhile, we hope, can gain a deeper understanding of the topic.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Helmut Durchschlag and Peter Zipper (2012). Anhydrous and Hydrated Protein Models Derived from High-Resolution and Low-Resolution Techniques, Protein Structure, Dr. Eshel Faraggi (Ed.), ISBN: 978-953-51-0555-8, InTech, Available from: http://www.intechopen.com/books/protein-structure/anhydrous-and-hydrated-protein-models-derived-from-high-resolution-and-low-resolution-techniques

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen