The UltraScan SOlution MOdeler (US-SOMO) Software Suite and its application to the modeling of integrin $\alpha_{IIb}\beta_3$.

Brookes E[1], Demeler B[1] and Rocco M[2]

[1]Dept. of Biochemistry, UTHSCSA [2]Biopolimeri e Proteomica, Istituto Nazionale per la Ricerca sul Cancro

Abstract

The functions of biological macromolecules are highly correlated with their structures, which can often be determined at atomic resolution using X-ray crystallography or nuclear magnetic resonance (NMR) methods. However, these are not problem-free operations. For instance, suitable crystals might be difficult or impossible to obtain, and packing constraints might induce conformational changes, while current NMR methods still have size limitations. Lower-resolution solution parameters such as translational and rotational diffusion coefficients, and the sedimentation coefficient, are instead relatively easily measured. Computing these parameters from atomic-level resolution structures can then offer additional insights, like allowing the investigator to verify the correspondence between crystal and solution structures, or, in multi-resolution modeling, to discriminate between alternative models. The UltraScan SOlution MOdeler (US-SOMO) suite presented in this work provides a comprehensive, flexible interface for accurately computing solution parameters from atomic-level structures of biomacromolecules through bead-modeling and cubic grid approaches.

Introduction

Computing hydrodynamic parameters, such as the translational diffusion coefficient D_t, the sedimentation coefficient s, the rotational relaxation time τ_h, and the intrinsic viscosity $[\eta]$ from atomic-level structures is not a straightforward task. Bead modeling, where a macromolecule is represented by an array of non-overlapping spheres of variable diameter, is the most widely used technique to compute the hydrodynamics, and various methods have been implemented in public-domain computer programs. To overcome the limitations of existing software, we have developed the freely available UltraScan SOlution MOdeler (US-SOMO) suite. US-SOMO provides a flexible interface with a user-friendly GUI (see Figure 1) for accurately computing a multitude of solution parameters from 3D structures using two bead-modeling approaches.

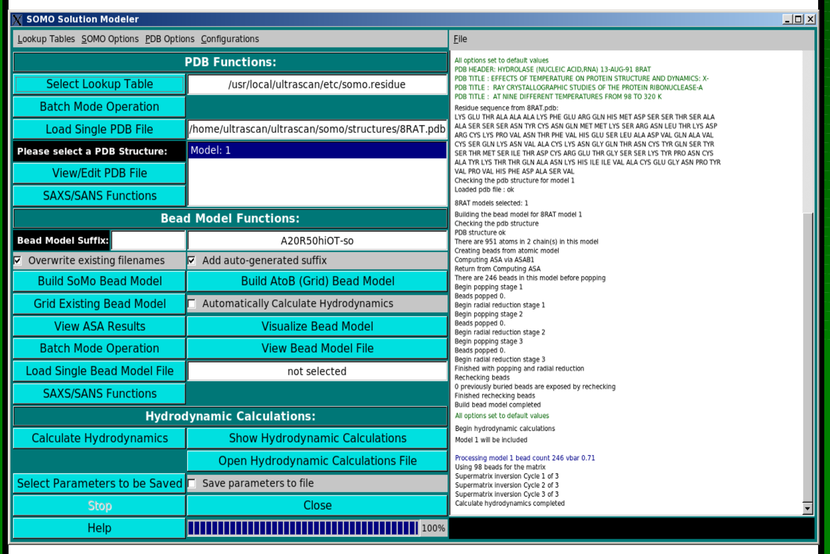

Figure 1: *The US-SOMO main GUI panel, with the commands module on the left side and the progress window on the right side (shown is the processing of the RNase A 8RAT.PDB structure).*

US-SOMO layout and properties

As shown in Figure 1, the main US-SOMO GUI has three sections, dealing with PDB uploading/parsing and related operations, bead model(s) generation/editing/visualization, and hydrodynamic computations. US-SOMO operates by comparing atom and residue entries in the PDB file with those stored in lookup tables, and then properly assigning radii, masses and theoretical hydration values. The lookup tables are editable from a pull-down menu, and include a hybridizations table [TTCG99], an atom table (where atomic groups are defined), a SAXS coefficients table, and a residue table, where the global properties of each residue (e.g. the partial specific volume) and the rules governing their conversion to one or more beads are also defined. Structures and models are visualized with RasMol. Approximate methods to deal with structures with incomplete residues or with non-coded residues are provided, accessible from the PDB-options pull-down menu. Several fine-control panels can be opened from the SOMO options pull-down menu.

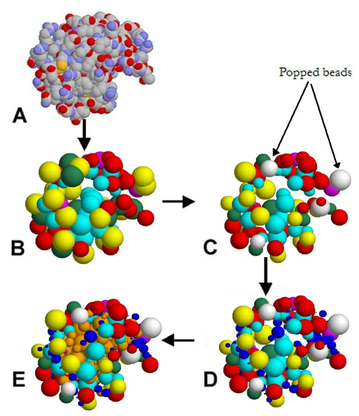

Figure 2: *The SoMo direct correspondence method applied to the RNase A 8RAT.PDB structure.*

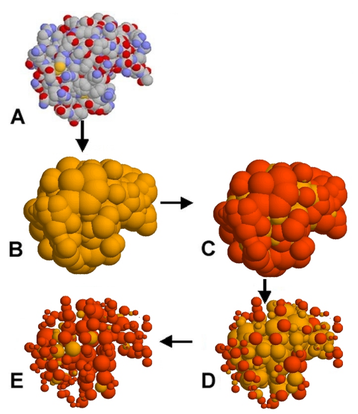

Figure 3: *The A2B grid (5 Å) method applied to the RNase A 8RAT.PDB structure.*

US-SOMO bead generation routines

Two methods are available, a direct correspondence (SoMo) method (Figure 2) and a grid (A2B) approach (Figure 3). In the SoMo method (Figure 2), an ASA routine identifies exposed and buried side- and main-chain segments in the original structure (A); beads representing the exposed side-chains (cyan, red, yellow & green beads) are then generated (volume = sum of the atoms volumes + theoretical H2O volume) and placed according to specific rules (B); overlaps between beads are removed by fusion above a certain threshold (white beads) and then hierarchically or synchronously proportionally reducing the beads' radii, while moving their centers outwardly by the same amount (C; this preserves the original surface); the same procedure is then used, without the outward translation, for the main-chain exposed beads (D, blue beads); finally, buried residues beads (orange) are generated, placed and their overlaps removed (E).

US-SOMO A2B bead generation

The A2B grid method method (Figure 3) proceeds similarly to the direct correspondence method: all beads are placed according to the grid size (B); surface exposed beads (red) are identified (C); surface beads overlaps are removed (D); buried beads (orange) overlaps are removed and the ensemble is re-screened for exposed/buried beads (E).

Performance of US-SOMO

Table 1 lists the comparison between experimental and computed D_t, τ_h, and $[\eta]$ for the SoMo method using the hierarchical or the synchronous overlap removal procedures for a series of proteins ranging from ~14,000 to ~149,000 Daltons in molecular weight. Buried beads are not used in the computations, significantly reducing the computer load. Note the excellent agreement for D_t (usually < 2%) across the entire range, and the good performance for τ_h and $[\eta]$ (usually within 5%). Similar results are obtained using the A2B approach (not shown). US-SOMO is fully described and validated in two publications [BDRR09],[BDR10].

Table 1. Comparison between experimental (Exp.) and computed (Comp.) values of $D^0_{(20,w)}$, $\tau^0_{h(20,w)}$ and $[\eta]$ for several test proteins using the two different overlap removal procedures (HI, hierarchical; SY, synchronous) within the direct correspondence SoMo bead modeling method (ASA cutoff 20 Å; peptide bond rule on; fusion threshold 70%; outward translation on; ASA re-check threshold 50%).

Protein	PDB	MW	$D^0_{(20,w)}$, % diff. from exp.			$\tau^0_{h(20,w)}$, % diff. from exp.			$[\eta]$, % diff. from exp.		
			F			ns			$cm3 \cdot g^{-1}$		
			Exp.	Comp., HI	Comp., SY	Exp.	Comp., HI	Comp., SY	Exp.	Comp., HI	Comp., SY
RNase A	8RAT	13 682	11.6 ± 0.3	11.8 (+1.7)	11.8 (+1.7)	8.05 ± 0.51	7.79 (−3.2)	7.74 (−3.9)	3.30 ± 0.04	3.24 (−1.82)	3.22 (−2.42)
α-Lactalbumin	1A4V[a]	15 793	10.9	10.9 (0)	10.9 (0)	10.3 ± 2.7	10.25 (−0.5)	10.14 (−1.6)	na[b]	3.63 (nd[c])	3.59 (nd)
Myoglobin (CO)	1DWR	17 521	10.7	10.9 (+1.9)	10.9 (+1.9)	10 ± 1	10.0 (0)	9.88 (−1.2)	na	3.24 (nd)	3.20 (nd)
Chymotrypsinogen A	2CGA	25 666	9.5	9.64 (+1.5)	9.64 (+1.5)	na	13.8 (nd)	13.7 (nd)	3.21	3.13 (−2.5)	3.10 (−3.4)
β-Lactoglobulin	1BEB	36 608	7.85 ± 0.08	7.92 (+0.9)	7.92 (+0.9)	23.2	25.1 (+8.2)	24.8 (+7.1)	na	3.92 (nd)	3.88 (nd)
Ovalbumin	1OVA	43 157	7.73 ± 0.04	7.78 (+0.7)	7.78 (+0.7)	20.9	26.2 (+25.4)	25.9 (+24.1)	4.0 ± 0.5	3.48 (−13.0)	3.44 (−14.0)
Hemoglobin CO	1HCO	64 557	6.9	6.98 (+1.2)	6.98 (+1.2)	na	35.6 (nd)	35.4 (nd)	na	3.28 (nd)	3.25 (nd)
Hemoglobin oxi	1GZX	64 573	7.21	6.99 (−3.1)	6.99 (−3.1)	35.4	34.9 (−1.1)	34.7 (−1.8)	3.16 ± 0.2	3.18 (+0.6)	3.16 (0.0)
Citrate synthase	1CTS	97 838	5.8	5.86 (+1.0)	5.86 (+1.0)	na	59.5 (nd)	59.1 (nd)	3.95	3.51 (−11.1)	3.49 (−11.7)
G3PD apo	2GD1	143 540	5.0	5.08 (+1.6)	5.08 (+1.6)	na	90.2 (nd)	89.6 (nd)	na	3.59 (nd)	3.56 (nd)
G3PD holo	1GD1	146 431	5.3	5.10 (−3.8)	5.10 (−3.8)	na	88.9 (nd)	88.6 (nd)	3.45	3.48 (+0.9)	3.46 (+0.3)
Lactate dehydrogenase	5LDH	148 636	5.06 ± 0.15	5.14 (+1.6)	5.14 (+1.6)	na	89.1 (nd)	88.3 (nd)	na	3.46 (nd)	3.43 (nd)

[a]Modified with the addition of a $Nag_4Man_3Gal_2$ biantennary carbohydrate chain attached to Asn45.[11] [b]na: not available; [c]nd: not done.

Future

As we continue to improve US-SOMO, we are developing the ability to simulate SAXS and SANS data from atomic structures. We will also implement Brownian Dynamics and Discrete Molecular Dynamic methods to compute hydrodynamic parameters for structures presenting local or larger scale flexibility issues.

References

[TTCG99] Tsai, J.,Taylor, R., Chothia, C, and Gerstein, M. The Packing Density in Proteins: Standard Radii and Volumes, JMB ***1999*** *290:253-66.*

[BDRR09] Brookes, E., Demeler, B., Rosano, C., and Rocco, M. The implementation of SOMO (SOlution MOdeler) in the UltraScan analytical ultracentrifugaion data analysis suite: enhanced capabilities allow the reliable hydronamic modeling of virtually any kind of biomacromolecule. Eur Biophys J: ***2009****-02-21. Epub ahead of print. PMID: 19234696*

[BDR10] Brookes, E., Demeler, B., and Rocco, M. Developments int the US-SOMO Bead Modeling Suite: New Features in the Direct Residue-to-Bead Method, Improved Grid Routines, and Influence of Accessible Surface Area Screening. Macromolecular Bioscience: ***2010****-05-17. Epub ahead of print.*