# Chapter 10
# Calculation of Hydrodynamic Parameters: US-SOMO

**Emre Brookes and Mattia Rocco**

**Abstract** Although it now offers a number of complementary applications for data processing and multiresolution modeling of (bio)macromolecules in solution, the UltraScan SOlution MOdeler (US-SOMO) suite of programs was initially developed for the computation of their solution properties starting from atomic-resolution structures and their comparison with experimental data. In this chapter, we will give a brief overview of the basic principles behind the hydrodynamic parameter computational methods available in US-SOMO and outline their operation.

**Keywords** Diffusion coefficient • Sedimentation coefficient • Intrinsic viscosity • Bead modeling • Multiresolution modeling • Solution structure

## 10.1 Introduction

US-SOMO (http://somo.uthscsa.edu/) was initially started as a graphical user interface (GUI) within the analytical ultracentrifugation data analysis program UltraScan (Demeler 2005) for the SOlution MOdeler (*SoMo*) method developed by the Rocco and Byron labs (Rai et al. 2005). The previously published *AtoB* grid method (Byron 1997) was also available in US-SOMO from its initial release (Brookes et al. 2010a). Both methods were developed for the computation of the hydrodynamic parameters starting from high-resolution structures of (bio)macromolecules using different bead modeling procedures, trying to avoid some of the drawbacks present in other approaches. Since then, it has grown to include other methods such as *Zeno* (http://www.stevens.edu/zeno/; Kang et al. 2004) and *BEST* (http://esmeralda.

E. Brookes
Department of Biochemistry, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA
e-mail: emre@biochem.uthscsa.edu

M. Rocco (✉)
Biopolimeri e Proteomica, IRCCS AOU San Martino-IST, Istituto Nazionale per la Ricerca sul Cancro, Genova, Italy
e-mail: mattia.rocco@hsanmartino.it

sfsu.edu; Aragon 2004), the former directly implemented, the latter operating on a supercompute cluster through a dedicated interface (see also Chap. 12). Small-angle scattering (SAS) data analysis and simulation modules have been subsequently added (Brookes et al. 2010b, 2013a), and discrete molecular dynamics (DMD) procedures (Ding and Dokholyan 2006; Dokholyan et al. 1998) for the expansion of conformational space when dealing with flexibility issues have been implemented, again operating on a supercompute cluster. While the overarching goal of the US-SOMO suite is to provide a full toolbox for the multiresolution modeling of (bio)macromolecules, in this chapter we will deal only with the features relating to hydrodynamic computation. Recent literature summarizing the other US-SOMO capabilities is available (Brookes et al. 2012; Rocco and Brookes 2014).

## 10.2 Operational Principles of the Bead Modeling Methods

Two bead modeling approaches are available in US-SOMO, *SoMo* (Rai et al. 2005) and *AtoB* (Byron 1997), with the computations carried out, in their original implementation, by solving a system of $n$ linear equations with $3n$ unknowns using the coefficients "supermatrix" inversion (*SMI*) procedure (Brookes et al. 2010a; García de la Torre and Bloomfield 1981; Spotorno et al. 1997). The *SoMo* method is based on a direct correspondence between the structural elements of a (bio)macromolecule and the beads used to represent it. Appropriately positioned beads of different radii are used, but since in the *SMI* procedure the hydrodynamic parameters are computed using the Rotne-Prager-Yamakawa hydrodynamic interaction tensor as modified by García de la Torre and Bloomfield (1981), valid for assemblies of variable-sized beads only if they do not overlap (see below), overlaps must be removed after the initial set of beads is defined. For proteins, a distinction is made between side- and main-chain segments, each represented with a bead as the default option. Two alternatives are available for representing the latter: the main chain of each $n$th residue $(N-CA-C-CO)_n$ or the peptide bond between the $n$th and $(n\text{th} + 1)$ residues $(CA-C-CO)_n N_{(n+1)}$. The second is the default option, because it reduces the chances of overlaps between the main- and side-chain beads. The initial spatial location of each bead is chosen according to the nature of the segment it represents. For the main chain (peptide bond) and for the hydrophobic and nonpolar side chains, the bead is placed at the center of mass of the atoms involved, while for polar and charged side chains, the bead is located toward the end of the side chain. Similar rules are employed for the sugar units in carbohydrates and for the sugars and bases forming the nucleotides in RNA/DNA. Prosthetic groups are likewise treated. The anhydrous volume of each bead is defined by the sum of the anhydrous atomic volumes of the atoms it represents, taken from literature analyses of crystallographic data (Perkins 1986; Tsai et al. 1999; Nadassy et al. 2001; Voss and Gerstein 2005). Alternatively, volumes can be calculated from structural models using dedicated software (*e.g.*, the 3 V Contact Volume Calculator: http://3vee.molmovdb.org/volumeCalc.php; Voss and Gerstein 2010).

The *AtoB* method relies instead on a cubic grid approach to "assign" atoms to a particular bead (Byron 1997). The initial volume of each bead is then simply calculated by summing up all the assigned atom volumes, and the position of each bead is defined either at the center of mass of its constituent atoms or the center of the cubelet. The overlap removal problem also applies to the *AtoB* method. The resolution of the model is controlled by the chosen size of the grid spacing.

A key, common aspect of the two bead modeling methods available in US-SOMO is a more realistic treatment of the water of hydration, in contrast to the uniform expansion of the model or to the addition of a uniform shell on the model surface as utilized by other approaches (*e.g.*, see Chaps. 11 and 12).

In the *SoMo* direct correspondence method, a number of water molecules are assigned to each bead, based on the theoretical, statistical hydration values determined by Kuntz and Kauzmann (1974) for each residue using NMR freezing. The volume of these water molecules is taken to be different from that of bulk water molecules, on the basis of crystallographic studies (Gerstein and Chothia 1996). Although this representation of the hydration effect as "bound" water molecules is not correct in principle, it turns out that it compensates quite well for the real physical effects involving changes in local viscosity and density at the protein/water interface (Halle and Davidovic 2003) (see also Rocco et al. (2012) and Chap. 12). This procedure can in general be applied to other types of biomacromolecules (*e.g.,* nucleic acids, carbohydrates, etc.). Operationally, the volume of the theoretically bound water molecules is then added to each corresponding bead, thereby accounting for the local variation in hydration. This approach is also implemented in the revised *AtoB* grid method available in US-SOMO, where water molecules are assigned to atoms within residues. Currently (May 2016), the waters/atoms assignment is provided for amino acid and carbohydrate residues only, but experienced users can define their own values for other residues modifying the *somo.residue* lookup table (see below).

Another innovation in both the *SoMo* and *AtoB* methods is a prescreening of the (bio)macromolecule to identify buried and exposed patches. This information is then associated with atoms/residues-representing beads, which are then labeled as being either buried or exposed. A further distinction is also made between exposed main- and side-chain segments. This information is utilized to greatly reduce the number of beads that are subsequently included in hydrodynamic computations using the *SMI* procedure, because only the beads that contribute to the surface frictional interaction with solvent are then considered.

The steps required to generate bead models in the *SoMo* and *AtoB* methods are illustrated in Fig. 10.1. In *SoMo*, the accessible surface area (ASA) is first determined, assigning each main and side chain as being either buried or exposed (the colors used refer to the nature of the placed residues; see Fig. 10.1 legend). The beads corresponding to exposed side chains are subsequently placed (A → B). The overlaps between these beads are then removed (B → C), first fusing together beads that overlap by more than a preset threshold and then proportionally reducing the bead radii either hierarchically (the couple with the largest overlap first and then the others) or synchronously (the radii of all overlapping beads are reduced
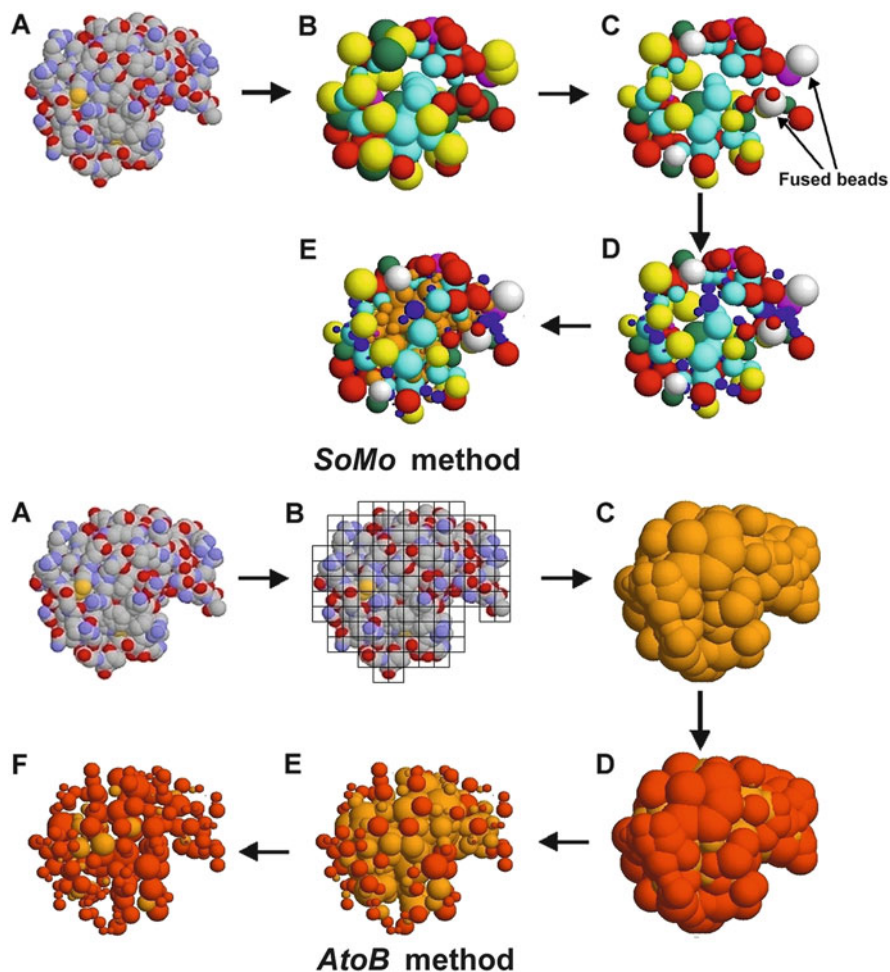
**Fig. 10.1** Schematic representation of the *SoMo* (*top*) and *AtoB* (*bottom*) bead model generation methods. A test protein atomic structure is shown in space-filling mode in step A in both procedures. The monodimensional grid visualized in step B of the *AtoB* method panel is in practice a three-dimensional "cage." The steps in both procedures (A → B, etc.) are described in detail in the text. The color coding in the *SoMo* B → E steps is *blue*, main chain; *cyan*, hydrophobic; *magenta*, nonpolar; *red,* polar; *yellow*, basic; *green*, acidic; *white*, fused beads; and *orange*, buried beads. The color coding in the *AtoB* steps D → F is *orange*, buried beads, and *red*, exposed beads

by a percentage of their original radius, and the procedure is repeated until no overlaps remain). An important procedure is implemented in this step to preserve the original surface as much as possible: while their radii are reduced, the bead centers are moved outwardly along a line connecting them to the center of mass of the (bio)macromolecule by an equal amount ("outward translation," OT). In the subsequent step (C → D), the main-chain exposed beads (blue) are placed and their

overlaps removed using one of the procedures described above but without the OT (this choice is made because the peptide bond segments do not usually protrude from the protein surface as most of the exposed side chains do). In the last step (D → E), buried residue beads (orange) are placed and their overlaps removed, again without OT. An ASA screen is then performed again on the final bead model, because some beads might have changed their exposed/buried status during the overlap removal procedure.

In *AtoB*, a cubic grid of a selected spacing is first placed on the original structure and atoms are "assigned" to cubes (A → B). In a single step, all beads are generated summing up the hydrated volumes of the atoms in each cube, and beads are placed according to the centering method chosen (B → C). An ASA screen is then performed (C → D; orange, buried; red, exposed). Overlaps are subsequently removed first in the exposed subset (D → E) using preferentially the synchronous procedure (default; the hierarchical procedure is also available), with OT. The same procedure is then applied to the buried subset, without OT, and the entire set is rescreened for ASA, resulting in many more beads becoming exposed (E → F).

Very recently, on the basis of an extensive investigation of the performance of the main available methods/programs used to compute the hydrodynamic parameters starting from atomic-resolution structures, it was found that the best results could be obtained by utilizing *SoMo*-type bead models *without* removing the overlaps between them and using the *Zeno* method for the computations (Rocco and Byron 2015). This approach is now directly available within US-SOMO (see Sect. 10.3).

## 10.3 The US-SOMO Main GUI Interface and Option Settings

### 10.3.1 PDB Function Area

The first button ("Select Lookup Table") allows the user to change the main reference file containing all the information necessary to properly recognize each residue and the atoms within it (the automatically uploaded default file is shown in the corresponding field) (see Fig. 10.2). The main lookup table and the other tables necessary for its construction can be edited from the top bar pull-down menu ("Lookup Tables"). The proper coding of each residue is a fundamental step in hydrodynamic bead model generation in US-SOMO (as well as for SAS computations, not dealt with here), and the tables contain the atomic radii, hydration numbers, SAXS/SANS coefficients, and the atoms to bead conversion/bead positioning rules.

Although advanced editors are available within US-SOMO (see Figs. 10.3 and 10.4), coding for atoms/residues and assignment to beads are not simple operations, as they entail knowledge of several physicochemical properties. The hybridization state of each non-H atom (see Tsai et al. 1999) and its related properties (*i.e.*,
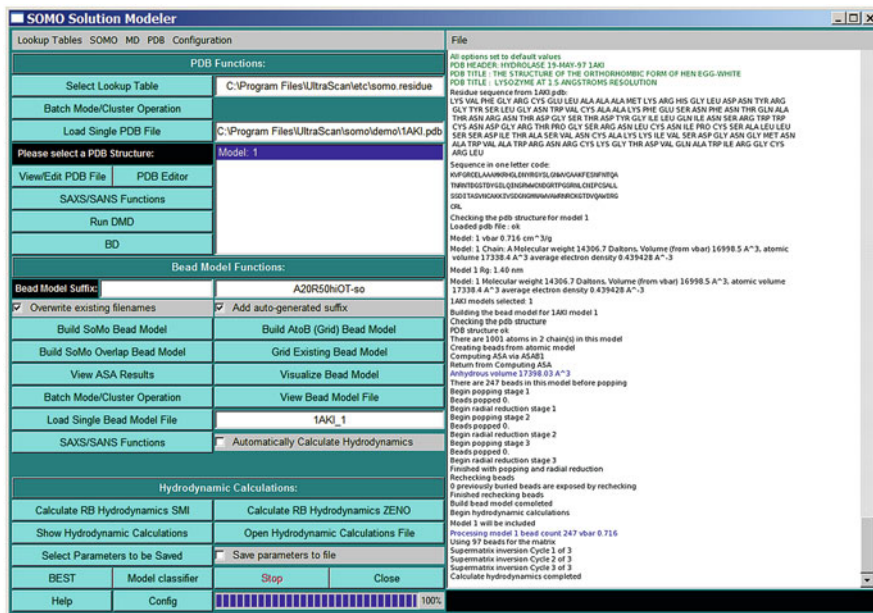
**Fig. 10.2** The US-SOMO main panel GUI. The *left side* of the window is divided into three subpanels: PDB Functions, Bead Model Functions, and Hydrodynamic Calculations. The *right-side* panel reports on structure loading/verification, modeling, and calculation progress (Shown with a reduced font size are the steps in the processing of the 1AKI.pdb RNase A structure)



**Fig. 10.3** The "Add/Edit Hybridization Lookup Table" (*left panel*) and "Add/Edit Atom Lookup Table" (*right panel*) modules of US-SOMO
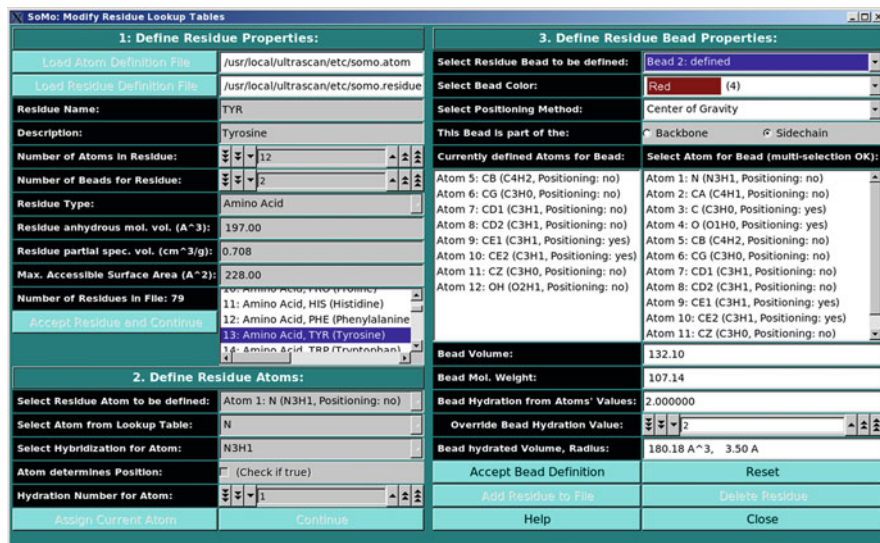
**Fig. 10.4** The "Add/Edit Residue Lookup Table" module of US-SOMO

molecular weight including the H atoms attached to it, radius, etc.) are defined in a
first table (default, *somo.hybrid*, currently containing 42 entries; see Fig. 10.3, left
panel). Since in PDB files each type of atom (*e.g.*, C, O, N) can have many different
"names" (*e.g.*, C1, OG, N3), a second table is built where the atom names present
in the PDB entries are linked to the proper hybridization and associated parameters
(default, *somo.atom*, currently containing 629 entries; see Fig. 10.3, right panel).
Both tables are connected to a third basic table containing the SAXS coefficients
(default, *somo.saxs_atoms*; editor not shown). Finally, the residues making up a
(bio)macromolecule are stored in the main lookup table (default, *somo.residue*, cur-
rently containing 122 residues, including all standard and some nonstandard amino
acids, ribo- and deoxyribonucleosides/nucleotides, and carbohydrates, plus some
lipids, detergents, and various prosthetic groups). In Fig. 10.4, the editor module
for the residue lookup table is shown. A detailed description of these procedures
is provided in the US-SOMO help files, accessible by pressing the "Help" button
located at the bottom of each GUI module (see also the Supplementary Information
of Brookes et al. (2010a, b)). It is important to emphasize that for reliable results,
all atoms/residues present in the sample for which experimental data are collected
must also be present in the structural model used for the hydrodynamic parameter
computations. However, to avoid the complicated task of encoding new residues,
skipping noncoded atoms/residues or approximate methods to represent them are
provided, the latter being now the default option. A warning message will appear
if noncoded atoms/residues are found in a structure, and the user can proceed with
the approximate method or chose a different option. Both are controlled from the
"PDB" pull-down menu in the top bar. If skipping is chosen (not recommended),

the user is warned about the risks of underestimating the molecular weight (*mw*) of the model and of miscalculating its partial specific volume (*psv*; both parameters are needed to compute the sedimentation coefficient from the computed translational frictional coefficient of the bead model, and the *mw* is needed for the computation of the intrinsic viscosity [η]). If the correct *mw* and *psv* are available, the user can enter them in the appropriate US-SOMO modules (see below). This of course would not take into account the lack of friction due to the skipped residue(s). The approximate "automatic bead builder" instead at least partially compensates for it and will roughly define a single "side-chain" bead for each noncoded residue. This procedure is based on an "average" volume for each atom (with an "average" *mw* and hydration number), from which a global volume (and *mw*) is calculated. An "average" radius for each atom is also provided for the ASA routines. The bead is then placed at the center of mass of all the atoms within the noncoded residue, and an "average" *psv* is assigned to it. All these "average" values can be modified in the Miscellaneous Options panel (see below), allowing the user to tune them to the type of noncoded residue (*e.g.*, amino acid, sugar, nucleotide, etc.). As with the "skip" option, if available "true" *mw* and *psv* values should be entered anyway in the appropriate US-SOMO modules. Likewise, in the more common case of when incomplete (but coded) residues are present in the PDB file, the default option is to use an approximate method to generate and place a bead. In this case, since the residue is encoded, *mw* and *psv* are computed as for complete structures. If the missing atom(s) are not marked in the *somo.residue* table as needed to position the bead, the approximation will lead to a "normal" bead, indistinguishable from what would be obtained for a complete residue. Otherwise, the level of the approximation will depend on the number and position of the missing atom(s). As long as there is even a single atom belonging to a coded residue, a bead representing it can be generated. Again, a warning message pops-up if incomplete residues are found in a structure, and the user can proceed or chose another option, like stopping or skipping the whole residue (not recommended), by selecting it in the "PDB" pull-down menu. Of course, there is no cure in US-SOMO for totally missing residues: the users are urged to complete their structures using external methods (*e.g.*, ROBETTA, http://robetta.bakerlab.org/ (Kim et al. 2004); I-TASSER, http://zhanglab.ccmb.med.umich.edu/I-TASSER/ (Roy et al. 2010); MODELLER, https://salilab.org/modeller/ (Eswar et al. 2006)). Missing atoms within coded protein residues can be added by WHATIF (http://swift.cmbi.ru.nl/servers/html/index.html; Vriend 1990).

PDB files can be individually loaded using the "Load Single PDB file" button or in batch mode (the latter will open a new window with advanced functions; see Sect. 10.5). When NMR-style files are opened, all models present are listed in the field provided, and either individual or multiple/all models can then be selected for further operations. Each structure is automatically visualized upon loading using RasMol (http://www.bernstein-plus-sons.com/software/rasmol/; Sayle and Milner-White 1995).

PDB files can be viewed in text mode and manually edited by pressing the "View/Edit PDB Files" button. Alternatively, an advanced PDB editor is also available, including cut/splice capabilities and the possibility to extract individual models from NMR-style files or to create NMR-style files from single models. These two functions are particularly useful as a complement to the DMD utility (Ding and Dokholyan 2006; Dokholyan et al. 1998) (accessed by pressing the "Run DMD" button, see Sect. 10.5), *e.g.*, to splice generated multiple conformations of a connecting segment between two static domains. The DMD utility will not be dealt with in detail in this chapter.

SAXS/SANS functions allowing computations directly on the atomic structure can also be accessed from this area (not dealt with here; see Brookes et al. (2012, 2013a); Rocco and Brookes (2014)). A Brownian Dynamics (BD) module (in preparation) will be also available in the future for the hydrodynamic parameter computation for flexible/partially disordered structures.

The Miscellaneous Options menu (Fig. 10.5, left-side panel), in addition to the "Average Parameters for Automatic Bead Builder" settings, contains the *psv* ("vbar") controls. The *psv* can be automatically computed from the composi-
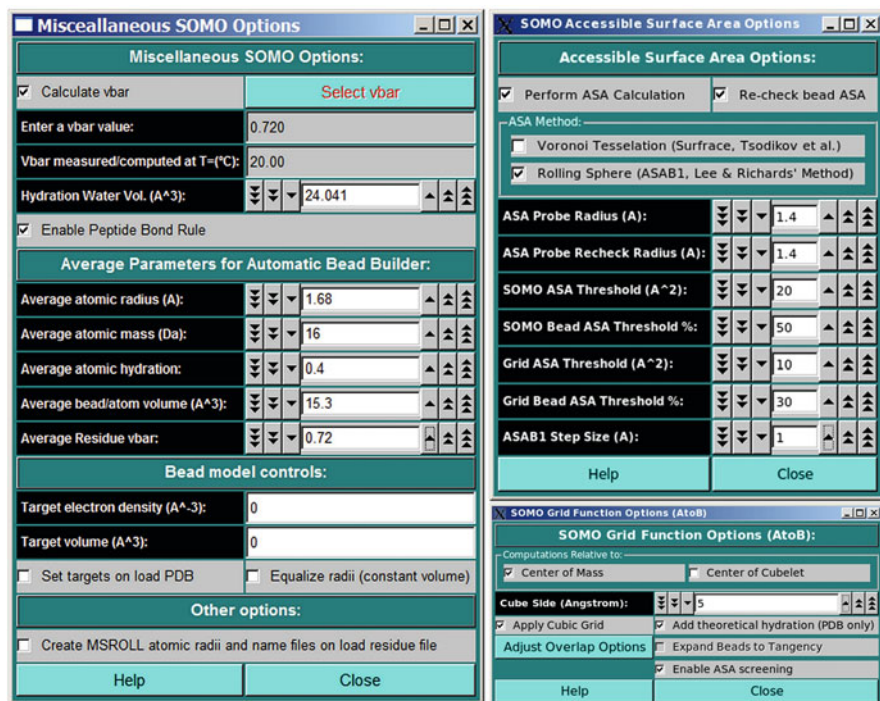


**Fig. 10.5** The US-SOMO "Miscellaneous Options" (left-side panel), "Accessible Surface Area Options" (*top right-side panel*), and "Grid Functions Options" (AtoB) (*bottom right-side panel*) modules

tion using the matching between the residues in the PDB file and those in the *somo.residue* lookup table ("Calculate vbar" checkbox selected), it can be uploaded from a database by pressing the "Select vbar" button, or it can be manually entered in the "Enter a vbar value" field. In the latter case, the "vbar measured/computed at T = (°C)" field should also be updated (default, 20 °C). Thus, if the temperature entered is different from 20 °C, the program can then recalculate a proper *psv* at the standard 20 °C T to which all hydrodynamic parameters are standardized by default (see also the Hydrodynamic Computations Options module).

Another important entry in this module is the volume assigned to the hydration waters, controlled in the "Hydration Water Vol (A^3)" field (default, 24.041 $\text{Å}^3$; Gerstein and Chothia 1996). This is the volume that will be added to the sum of the anhydrous atom volumes for each water molecule assigned to a bead.

The "Enable Peptide Bond Rule" checkbox controls if this rule is used by the *SoMo* method. With it, the peptide bond segment is used for the main-chain beads of a protein structure. These beads are thus positioned at the center of gravity of the $(CA\text{-}C\text{-}O)_n$-$(N)_{(n+1)}$ atoms, except when PRO is the $(n+1)$ residue. In this case, the peptide bond bead is positioned at the center of gravity of the $(CA\text{-}C\text{-}O)_n$ atoms. Additional rules control the generation of the OXT bead and of the first N atom at the beginning of each protein chain. All these rules are controlled by "special" residues in the *somo.residue* table. To gain total control over the positioning, volumes and masses of every bead, the "Enable Peptide Bond Rule" checkbox should be deselected (default, selected, but if breaks are found in a chain, it is disabled). The "Bead Model Controls" (for SAS work) and "Other options" (relating to BEST operations, see Sect. 10.6) sections will not be dealt with here.

### 10.3.2 BEAD Model Function Area

In this section, new bead models can be generated from selected PDB structures according to one of the three methods available, *SoMo* (without overlaps), *AtoB* (also without overlaps), and *SoMo* with overlaps (see Fig. 10.2). The various menus with the options and settings in the bead generation routines are accessible from the "SOMO" pull-down menu in the top bar.

The ASA options are controlled by the "Accessible Surface Area Options" module (Fig. 10.5, top right-side panel). By default, the "Perform ASA Calculation" and "Re-check bead ASA" checkboxes are selected, allowing the assignment of each bead in the final model to either an exposed or buried status. The hydrodynamic computations with the *SMI* procedure can then be carried out on the exposed beads subset only, greatly reducing the computational load (see Sect. 10.2). The default method is the Lee and Richards (1971) rolling sphere algorithm ("ASAB1"), but a Voronoi tessellation method ("Surfracer"; Tsodikov et al. 2002) is also available. The ASA probe radii can be independently set for the original structure and for the resulting bead model (default, both 1.4 Å). The "SOMO ASA threshold (A^2)" and the "Grid ASA threshold (A^2)" fields control the levels above which a main or

side chain will be considered exposed to the solvent in the *SoMo* method and above which primary beads will be considered exposed to the solvent in the *AtoB* (Grid) method, respectively (defaults, 20 and 10 Å$^2$, respectively). The "SOMO bead ASA threshold (%)" and "Grid Bead ASA Threshold (%)" fields set the minimum % of the surface of a bead that must be accessible to reclassify as exposed a bead previously considered to be buried in the *SoMo* and *AtoB* methods, respectively (defaults, 50 % and 30 %, respectively). Finally, the "ASAB1 step size (A)" field defines the increment between the 2D slices, to be integrated, in which the structure (or the model) is subdivided in the rolling sphere method (default, 1 Å).

The options for the *AtoB* grid method can be seen in Fig. 10.5, bottom right-side panel. The positioning method can be either the center of mass of the atoms assigned to each bead or to the center of the cubelet. The grid size can be set here (default, 5 Å). "Apply Cubic Grid" allows the grid procedure to be executed (default, active). It could be deselected to allow the use of the Grid module for overlap removal of a previously loaded bead model. The "Add theoretical hydration (PDB only)" checkbox will enable the addition of the theoretically bound water molecules volume to those of the atoms assigned to a bead. The "Adjust Overlap Options" button will open the *AtoB* overlap reduction options module (see below and Fig. 10.6). Finally, the "Enable ASA screening" checkbox will allow the user to select/deselect that routine (default, selected). The other checkbox controls a function still under development.

The overlap reduction routines have several options that can be accessed from two dedicated modules, one for the *SoMo* and the other for the *AtoB* methods (see Fig. 10.6). A common "overlap cutoff" field, which determines the level of precision in computing the overlaps between beads (default, 0.001 in the model units) is present at the top. Each module then has three different sections, dealing with the overlaps between exposed side-chain beads only, between main- and side-chain beads, and between buried beads for the *SoMo* method, while the distinctions are made between exposed grid beads and buried grid beads for the *AtoB* method. For the latter, in case no ASA screen is selected, there is a specific panel for the overlap reduction settings. All the options visible in Fig. 10.6 are common in the three
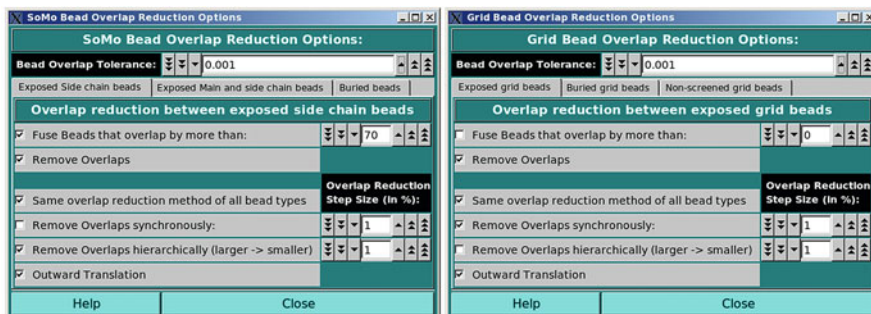


**Fig. 10.6** The *SoMo* (*left side*) and *AtoB* (*right side*) overlap reduction options modules

sections, except for the outward translation which is present only in the exposed side chains and exposed grid beads sections (see the US-SOMO Help pages for a complete description of all the features available in these modules).

The transformation process from an atomic-level structure to a bead model is activated by pressing either the "Build SoMo Bead Model," the "Build AtoB (Grid) Bead Model," or the "Build SoMo Overlap Bead Model" buttons. Bead models thus generated are automatically saved in a file, whose name is the PDB filename with "_1" added and the extension ".bead_model." Filenames are by default customized by adding a suffix containing a coding of the method used and its settings (which can be turned off by deselecting the "Add auto-generated suffix" checkbox) and additionally by entering a user-selected suffix in the "Bead Model Suffix." The auto suffix will have the "-so," "-a2b," or "-so_ovlp" extensions if the bead model was generated with the *SoMo*, *AtoB*, or *SoMo* with overlaps methods, respectively, and will contain a series of "codes" for the ASA parameter settings and the bead model generation options (see the US-SOMO main Help pages for a complete description of this feature). If the resulting filename is already present in the operating directory, a pop-up menu will offer several choices, including overwriting. This step can be automatically bypassed by selecting the "Overwrite existing filenames" checkbox. Options are available to adjust the bead model(s) file format by selecting the "Bead Model Output Options" from the "SOMO" pull-down menu (not shown). If both the "Overwrite existing filenames" and the "Automatically Calculate Hydrodynamics" checkboxes are selected, the program will complete the full process of generating a bead model and computing its hydrodynamic parameters unattended. This is especially useful when relatively large structures are examined. By default, the *SMI* procedure will be called if *SoMo* or *AtoB* models without overlaps are generated, while *Zeno* will be used if *SoMo* models with overlaps are produced. If the "Automatically Calculate Hydrodynamics" checkbox is not selected (default option), at the end of the model building phase the progress bar will be at 100 %, and the bead model(s) can be visualized with RasMol by clicking on "Visualize Bead Model" (recommended, comparing the original structure with the bead model could reveal previously unforeseen problems). A warning: if a NMR-style file has been uploaded and several/all models selected for bead model generation, pressing "Visualize Bead Model" will open a RasMol window for each one!

The "Grid Existing Bead Model" function allows reduction of the resolution of a previously generated bead model by applying a grid procedure. This button is not available until a PDB file has been processed with any of the bead modeling primary options (see above) or until a previously generated bead model file has been loaded (see below). If this operation is launched, the "-a2bg" suffix is automatically added to the filename of the new bead model.

The results of ASA screening of the original PDB file are written in a text-format file, which can be opened by pressing the "View ASA Results" button. Likewise, a bead model file can be opened in text mode by pressing the "View Bead Model File" button. Bead models previously generated by US-SOMO or coming from other sources like DAMMIN/DAMMIF (Franke and Svergun 2009) can be further

processed here by either uploading a single model ("Load Single Bead Model File") or using the "Batch Mode/Cluster Operation" (see Sect. 10.5).

The "SAXS/SANS Functions" button will open the SAS module allowing operations on the current bead model (not dealt with in this chapter).

### 10.3.3 Hydrodynamic Calculations Area

The options setting for the two hydrodynamic calculation methods using bead models offered in US-SOMO (see Fig. 10.2) are shown in Fig. 10.7. In the left side of Fig. 10.7, the options for the default García de la Torre-Bloomfield *SMI* inversion method (Rai et al. 2005; Brookes et al. 2010a; García de la Torre and Bloomfield 1981; Spotorno et al. 1997) are shown. By default, all calculations are performed for structures (bead models) whose dimensions are in Å, and in standard conditions, *i.e.*, in water at 20 °C. The top part of the hydrodynamic calculations options module lists these values. Users wishing to compute hydrodynamics under different conditions, or using bead models on another scale, can change the required parameters here.
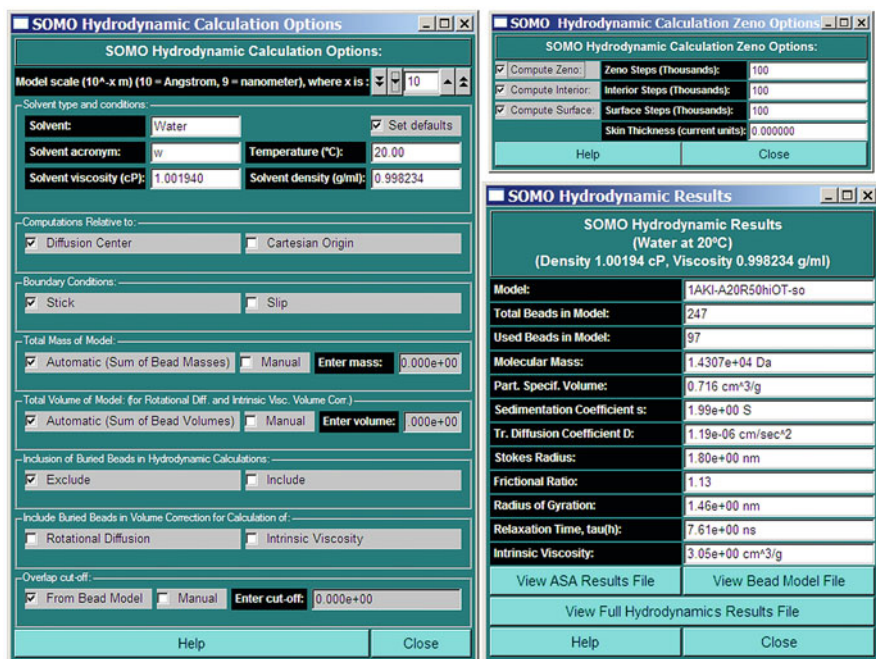


**Fig. 10.7** The "Hydrodynamic Calculations Options" modules and "Hydrodynamic Results" panel. *Left side*, options for the standard *SMI* method. *Right side*, *top*, options for the alternative *Zeno* method. *Right side*, *bottom*, "Hydrodynamic Results" pop-up panel

These definitions also apply to the *Zeno* method. By default, all *SMI* computations are carried out relative to the diffusion center of the model, and under the stick boundary conditions (García de la Torre and Bloomfield 1981), but the alternative Cartesian origin, and slip boundary conditions are respectively available. The total mass and total volume of the model (both necessary for the computations of [η], the latter also for the so-called volume correction for the rotational diffusion and [η]; see Spotorno et al. (1997) and references therein) are by default automatically computed from the beads' values. Users can, however, override either of these values by selecting the "Manual" checkbox and entering appropriate values. Entering a manual mass value is especially important when the bead model derives from an incomplete structure and/or including noncoded residues (see the PDB Functions area). By default, the beads labeled as being buried are excluded from the *SMI* hydrodynamic computations, but this can be overridden by selecting the "Include" checkbox. In such a case, it becomes possible to include or exclude (default) the buried beads from the "volume correction" computations for either or both the rotational diffusion and [η]. Finally, the "overlap cutoff," *i.e.,* the level of precision in checking the bead overlaps (see Fig. 10.6), can be set to manual with a different value, to allow for greater overlap tolerance when processing beads generated by other programs (*e.g.*, DAMMIN/DAMMIF).

The *Zeno* computational method involves enclosing an arbitrarily shaped probe object within a sphere and launching random walks from this sphere. The probing trajectories either hit or return to the launch surface ('loss'), whereupon the trajectory is either terminated or reinitiated (Kang et al. 2004). A summary of the ideas behind *Zeno* is given in its dedicated Help page in the US-SOMO manual. In the *Zeno* options module shown in Fig. 10.7, top right side, the first checkbox allows selection of the *Zeno* computation. This will launch a Monte Carlo numerical path integration that generates a large number of random walks in the space outside the body. Sums taken over these random walks yield the electrostatic capacity, the polarizability tensor, the intrinsic conductivity, and, most relevant here, the hydrodynamic radius $R_h$, the translational diffusion and frictional coefficients $D_t$ and $f_t$, the intrinsic viscosity [η], and the hydrodynamic volume $V_h$. The main option of interest here is the number of steps in the "Zeno Steps (Thousands)" field (default, 1000), which controls the accuracy of the calculations at the cost of increasing computational time. The reader is referred to the *Zeno* Help page within US-SOMO for more information on this and the other operations available by selecting the other two checkboxes, as well on the "skin thickness" field.

Once one or more bead model(s) have been generated, or a single existing bead model has been uploaded, the hydrodynamic calculations are started by pressing either "Calculate RB Hydrodynamics SMI" or "Calculate RB Hydrodynamics ZENO" (the "RB" stands for "Rigid Body," meaning that the computations are in the rigid body frame approximation). In a recent examination of the performance of the main available hydrodynamic computations methods/programs (Rocco and

Byron 2015), it was found that while only a slight improvement in accuracy was observed when *SoMo* models without overlaps were processed with *Zeno* in respect to the *SMI* procedure, a significant improvement was present when the primary models with overlaps were employed. Therefore, US-SOMO now offers directly both procedures, but since the *SMI* cannot be used when overlaps are present, the "Calculate RB Hydrodynamics SMI" button will not be available when the "Build SoMo Overlap Bead Model" has been used to generate the model(s). Once the calculations are completed (bottom progress bar at 100 % and "Calculate Hydrodynamics Complete" appears in the progress window), a subset of the results can be visualized by pressing "Show Hydrodynamic Calculations." In the SOMO Hydrodynamic Results pop-up panel (see Fig. 10.7, bottom right side), the conditions under which the calculations were performed are stated first (default, $H_2O$ @ 20 °C). There a series of the most commonly used parameters are reported, among which are the sedimentation coefficient $s$, $D_t$, $R_h$, the frictional ratio $f/f_0$, the radius of gyration $R_g$, the harmonic mean of the relaxation times $\tau_h$, and $[\eta]$ (the $\tau_h$ field will not be populated if *Zeno* is used). The full list of all the parameters entered/computed is saved in a text-format file that can be opened by pressing the "View Full Hydrodynamic Results File" button in the hydrodynamic results pop-up panel or the "Open Hydrodynamic Calculations File" button in the main panel.

The "Select Parameters to be Saved" button will open another window (see Fig. 10.8) where the user can interactively select among all conditions, results, and parameter values available, to be saved in a comma-separated variable (csv) file for further manipulations either with external spreadsheet programs or by the US-SOMO Model Classifier (see Sect. 10.4). Selecting the "Save parameters to a file" checkbox will enable this feature. "BEST" will open another pop-up window where results from the *BEST* hydrodynamic computation program as implemented within US-SOMO can be analyzed (see Sect. 10.6). "Stop" will halt any operation.
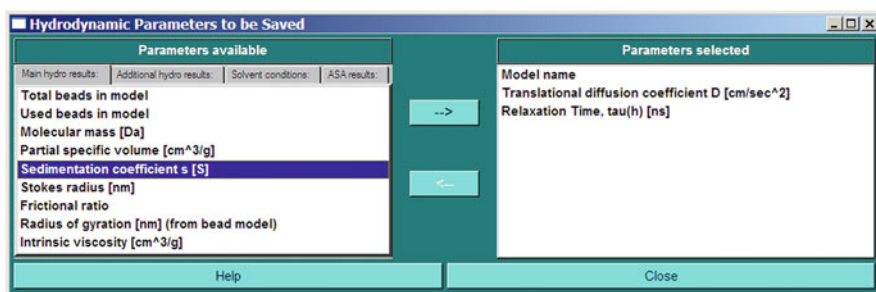


**Fig. 10.8** The "Select Parameters to be Saved" module

## 10.4　The US-SOMO Model Classifier Module

This module presents a tool for selecting a best matching model among a series of models, by comparing their calculated hydrodynamic parameters with user-provided experimental values. Several ranking methods are available in case more than one experimental parameter is known.

In Fig. 10.9, top, the GUI of the Model Classifier is shown. First, the experimental parameters to be used are entered. The selectable parameters are the sedimentation coefficient $s$ [S], the diffusion coefficient $D_t$ [cm$^2$/s], the Stokes' radius $R_h$ [nm], the frictional ratio $f/f_0$, the radius of gyration $R_g$ (nm), the harmonic mean of the



**Fig. 10.9** *Top*, the US-SOMO "Model Classifier" interface; shown are the settings and a run using 16 NMR-derived models of RNaseA (2AAS.pdb) whose hydrodynamic parameters were computed and compared with experimental values (Taken from (Brookes et al. 2010a)). *Bottom*, the results of the run are shown through the dedicated "Model Classifier" viewer

relaxation times $\tau_h$ [ns], and the intrinsic viscosity [$\eta$] [cm$^3$/g]. The methods used to sort the computed results against the experimental values are set next using several alternative criteria, listed under the "Sort results" group. In the "Using percentage difference" subgroup, they can be ranked by % absolute difference or by the weighted sum of % absolute differences. The first is the simplest procedure, ranking the parameters in a descending order (*i.e.*, 1 = most relevant) in the "Rank" field. The second ranks over multiple parameters without specifically assigning a numerical rank to each parameter. This is accomplished by computing a weighted sum of absolute differences of every included parameter. The user-defined weights do not have to add up to 1, and experimental data with higher confidence should be assigned higher weights.

Alternatively, in the "Equivalence class controls" subgroup, the results can be sorted by equivalence class rank. Equivalence classes partition a range of values. A value that falls into a specific equivalence class is equivalent to all other values within the equivalence class. The range runs from the "Minimum model value" to the "Maximum model value" and is composed of "Number of partitions" equivalence classes. The equivalence class that contains the experimental value is given a distance of zero. Equivalence classes next to the one containing the experimental value are given a distance of 1 and so on. Adding up the distances of each of the selected variables gives the equivalence class rank.

The last three columns under the "Add columns to results" label allow the addition of the experimental values and an additional % difference field to the Model Classifier results (the absolute differences are reported by default if the first ranking method is chosen). The current parameters and the criteria used for the sorting can be saved in a file (extension *.smp) by clicking on the "Save Parameters" button, while "Reset Parameters" will clear all fields. Previously saved parameters can be reloaded by clicking on the "Load Parameters" button.

In the bottom part, the parameters calculated for the models are uploaded. They should be in *.csv files, most easily generated using the "Save parameters to a file" checkbox (and the "Select Parameters to be Saved" module) in the Hydrodynamic Calculations section of the main US-SOMO window (see above). Only the parameters present in the *.csv files, identified through their headers, will then be available in the "Select to enable variable comparison" column. Pressing the "Load" button will open the file system dialog and allow import of the required *.csv files into the left-side window. Files can then be selected by clicking on each filename, which will transfer them to the right-side window, or by pressing the "Select All" button. "Remove" will remove files from the list. "Merge" will join the selected files from the list into one csv file. "Set min/max" will set the "Minimum model value" and "Maximum model value" from the values found in the selected files. The files listed in the right-side window can be then selected for processing by individually clicking on them or by pressing the "Select All" button.

Once files have been selected, the Model Classifier can be launched by pressing the "Process" button, and the progress window at the far right will be updated. At the end, pressing "View" will open a window with all the selected columns, as shown in

the bottom part of Fig. 10.9. By pressing the "Save" button, a file system dialog will open to allow the results to be saved in a new *.csv file, which can then be opened with a standard spreadsheet.

## 10.5   The US-SOMO Batch Mode/Cluster Operation Module

This module (Fig. 10.10, top) was conceived to allow the unattended processing of multiple files for both hydrodynamic, DMD, and SAS calculations. Since some of these operations can be performed only on a remote supercompute cluster, access to
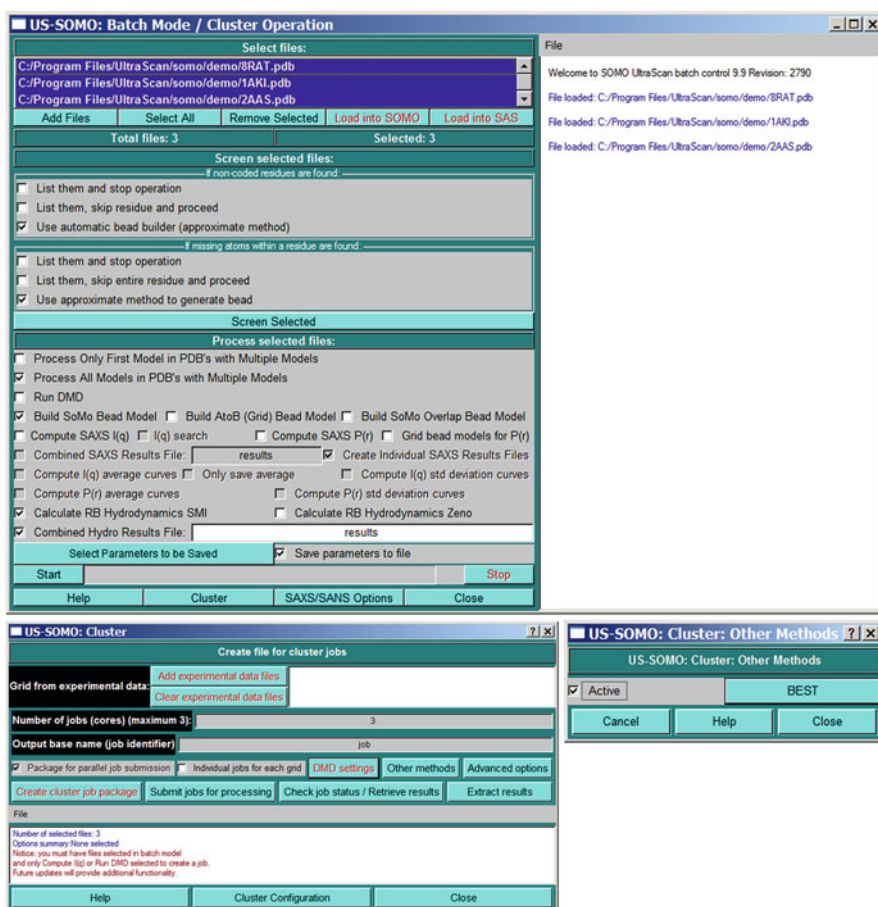


**Fig. 10.10** *Top*, the US-SOMO "Batch Mode/Cluster Operation" module. *Bottom*, *left side*, the US-SOMO "Cluster" module GUI. *Bottom*, *right side*, the pop-up "Cluster: Other Methods" pane

the Cluster interface ("Cluster" button; Fig. 10.10, bottom left) is provided within this module.

Operations begin by loading file(s) using the "Add Files" button in the "Select files" section and then either selecting a subset by clicking on individual filenames or all files with the "Select All" button. Files can be removed from the list with the "Remove Selected" button. The "Load into SOMO" and "Load into SAS" buttons become available only if a single file is selected and will just transfer it to either the main US-SOMO or to the SAS modules, respectively. Both PDB and bead model files can be uploaded and selected in this module, but some operations like the "Run DMD" can be performed only on atomic-level structures.

In the "Screen selected files" section, the user can control the level of tolerance for both noncoded residues (first three checkboxes) and for incomplete residues (second three checkboxes; see Sect. 10.3 for a complete discussion of these features). Pressing "Screen Selected" will then verify if the selected files comply with the US-SOMO requirements for processing. This is a relatively quick step and is highly recommended before launching a batch mode operation, since it will be performed anyway when each single file is processed, but the operations will be halted if noncomplying files are then found. A prescreen will allow users to correct the situation and permit fully unattended operations thereafter.

Operations are chosen in the "Process selected files" section. The first two checkboxes control if just the first model or all the models are to be processed when NMR-style files are uploaded. The "Run DMD" checkbox will allow a DMD run to be performed on chosen PDB file(s) (not dealt with in this chapter). The three bead modeling methods available within US-SOMO can be alternatively chosen by selecting either the "Build SoMo Bead Model," the "Build AtoB (Grid) Bead Model," or the "Build SoMo Overlap Bead Model" checkboxes. Next follows a series of checkboxes related to SAS operations, which will be not described here (see the Batch Mode/Cluster Operation Help page for a detailed description of these options). The "Calculate RB Hydrodynamics SMI" or the alternative "Calculate RB Hydrodynamics Zeno" checkboxes (the latter automatically selected if the "Build SoMo Overlap Bead Model" method is checked) allow the hydrodynamic calculations to be performed for bead models, either already present in the uploaded files or after generation from uploaded PDB structures. The "Combined Hydro Results File" checkbox allows saving the hydrodynamic parameter computation results performed on all bead models in a single file, with the averages of all parameters, instead of separate files for each model. A filename for the single results file must be provided in the dedicated space. Otherwise, each file will be named using the general US-SOMO rules and the prefixes present in the main program panel. As with single file operation, selected parameters can be chosen and saved in a *.csv file by accessing the "Select Parameters to be Saved" module (see Fig. 10.8) and selecting the "Save parameters to file" checkbox. The operations are launched by pressing the "Start" button and can be aborted at any stage by pressing the "Stop" button. After launching, the various operations will be reported in the right-side progress window, and the progress bar will become active.

With the exception of DMD, all the other options listed in the Batch Mode/Cluster Operation module can be carried out locally. However, some can be computationally intensive and might require supercomputing in order to be efficiently carried out. For this reason, a cluster interface has been developed, accessible by pressing the "Cluster" button. A complete description of this module is, however, beyond the scope of this chapter, and only a general overview and the *BEST* application (see Sect. 10.6) will be described. See Brookes et al. (2012, 2013b) for more information on cluster usage.

In Fig. 10.10, the main GUI of the Cluster module is shown (bottom left), together with the "Other Methods" pop-up panel that is launched from the "Other methods" button (bottom right). The top part of the module ("Grid from experimental data") deals with SAXS settings not described here. The "Number of jobs (cores) (maximum #)" is adjusted to the number of independent structures considered when the "Package for parallel job submission" checkbox is selected; it can be changed but the value should not be above the maximum # indicated. The "DMD settings" and "Advanced options" buttons will open the DMD settings and the SAXS advanced settings panels, respectively; they will not be discussed in this chapter. Currently, *BEST* is the only option available under "Other Methods."

Once the options have been set, the cluster submission procedure begins with pressing the "Create cluster job package" button. The package is then submitted to the cluster by pressing "Submit jobs for processing," which will open a cluster dialog panel (not shown) where jobs can be seen, clusters can be selected, the status of the operation(s) monitored, and from where the results can be retrieved. The cluster dialog panel can be accessed at any time by pressing the "Check job status/Retrieve results" button. Once the packaged results have been transferred back to the local machine, full datasets can be extracted by pressing the "Extract results" button. All these steps are described at length in the cluster Help section, and cluster access can be defined and configured through a dedicated panel accessed by pressing the "Cluster Configuration" button (not shown).

## 10.6 The US-SOMO BEST Interfaces

*BEST* is a software package for the computation of the hydrodynamic properties of (bio)macromolecules that relies on the direct evaluation of the frictional forces acting on surface elements (Aragon 2004) [see also Chap. 12]. *BEST* is made available under US-SOMO as an alternative method to the bead modeling methods we offer for the computation of the hydrodynamic parameters starting from a high-resolution structure. In principle, *BEST* can produce more accurate values with respect to the bead modeling procedures, especially with regard to the rotational diffusion and the intrinsic viscosity, since no "volume correction" is needed. However, some issues such as the proper consideration of the hydration (a recently done comparison between the various hydrodynamic methods (Rocco and Byron 2015) has evidenced that the current *BEST* implementation slightly

underestimates, $\sim -3$ %, the translational frictional properties of proteins), and the requirement to extrapolate values to zero triangle size (see below) need to be considered with care. To this end, we also provide an interface to visualize and statistically analyze the *BEST* results (Fig. 10.11, bottom). Moreover, *BEST* is very computationally intensive, and when many structures are analyzed, for instance when dealing with conformational variability/flexibility, bead modeling can offer a more practical alternative. Due to its requirements, *BEST* is offered only on a supercompute cluster within US-SOMO. To perform the calculations, in *BEST* the smooth atomic surface of the structure needs to be transformed into an ensemble
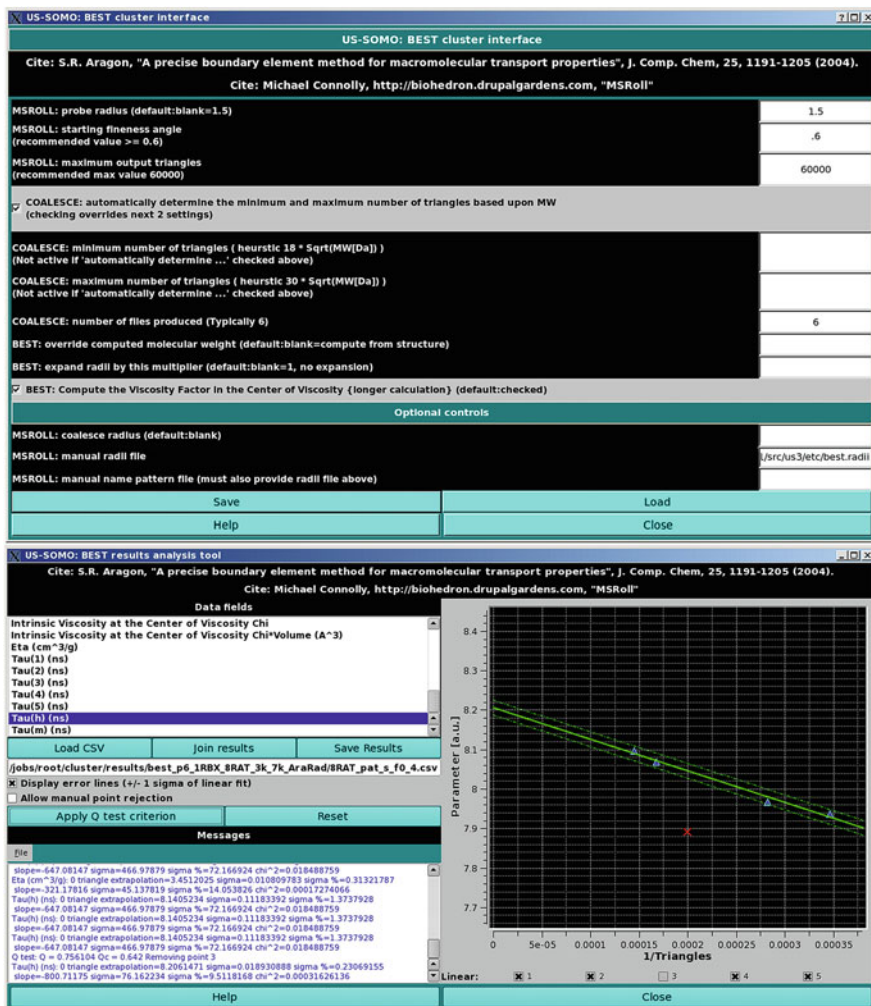


**Fig. 10.11** *Top*, the US-SOMO "*BEST* cluster interface" module. *Bottom*, the "*BEST* results analysis tools" module

of triangular elements, allowing the correct evaluation of the surface resistance integrals (Aragon 2004). First, the external program *MSROLL* (Connolly 1993) is used by *BEST* to generate an initial high-resolution triangulated surface of the structure under examination. Then, the *BEST* module *COALESCE* will produce a series of triangulated structures with different resolutions from the initial *MSROLL*-generated surface. The hydrodynamic properties calculated for each triangulated structure are then extrapolated to zero triangle size. The full principles and detailed operation of *BEST* can be found in Aragon (2004) and in Chap. 12, and here we will just describe the various tools and settings we provide.

In the cluster "Other Options" panel (see Sect. 10.5), pressing "BEST" will open the *BEST* settings interface panel. As shown in Fig. 10.11, top, the settings interface allows the user to set the *MSROLL* (Connolly 1993) probe radius (default 1.5 Å), finesse angle, and maximum number of triangles. The options for the *BEST* module *COALESCE* are set next. The first checkbox allows the automatic determination of the optimal maximum and minimum number of triangles based on a heuristic approach involving the structure's molecular weight (see Chap. 12). If this checkbox is left unchecked, these values can be manually entered in the next two fields. The number of files generated, used then for the extrapolation to zero triangle size, is entered in the following field (4 is the minimum suggested value, but 6 can allow for a better checking of the extrapolation). The last two fields in this part of the module allow the user to enter a molecular weight different from that calculated by *BEST* from the structure and to expand (or shrink) the atom radii used by *MSROLL* to compute the surface, which are optimized in *BEST* to take into account a uniform layer of hydration (see Chap. 12). By default, the atomic radii internally used by *BEST* are selected (available in a file called *best.radii*), but any other properly formatted radii file can be uploaded in the "Optional controls" section (a *MSROLL*-formatted radii file can automatically be generated from the values present in the *somo.residue* entries by selecting the "Create MSROLL atomic radii and names files on load residue file" checkbox in the Miscellaneous SOMO Options module; see Fig. 10.5, left panel). Finally, the "BEST: Compute the Viscosity Factor in the Center of Viscosity (longer calculation)" checkbox if unselected will speed up the calculations but at the cost of accuracy (default, checked).

The US-SOMO *BEST* implementation includes assembling all calculated parameters for each model in a csv file. Following retrieval from the cluster and extraction, the *BEST* results can be uploaded in the "BEST results analysis tool" (Fig. 10.11, bottom), accessible by pressing "BEST" from the main panel (see Fig. 10.2).

Upon loading using the "Load CSV" button, the "Data fields" panel will list all the parameters computed by *BEST*, selectable by clicking on each one. The data associated with the selected parameter are plotted *vs.* 1/(number of triangles) in the right-side graphics window, together with a linear regression line and a series of checkboxes corresponding to each data point (see Fig. 10.11, bottom). The "Join results" button allows merging of the separate csv files that the US-SOMO *BEST* implementation will generate, for instance, from NMR-style files, into a single csv file with the data for every parameter grouped together. In this way, averaging can then be easily performed using an external spreadsheet program.

Since *BEST* requires an extrapolation procedure to produce the final value for each parameter, the US-SOMO implementation provides automatic (recommended) or manual (non-recommended) ways to reject outliers from the regression. First, by selecting the "Display error lines (+/− 1 sigma of linear fit)" checkbox, two dotted lines corresponding to ± 1 standard deviation (SD) will be traced along the regression line. By pressing the "Allow Q test criterion" button, Dixon's *Q*-test (Dixon 1951) is performed and will reject a single outlier if the outlier's computed *Q* value is greater than the critical *Q* value set at a 90 % confidence level. In the example shown in Fig. 10.11, bottom, the third point (marked with a red "X") has been rejected, and the ± 1 SD lines are retraced after its exclusion (signaled also in the checkboxes below the graph). If more than one point visually appears to be problematic in the regression, it is suggested to rerun the computations including more points. "Reset" will re-include all points in the linear regression. The updated regression data are shown each time in the "Messages" window. All parameters within a single csv file can be independently analyzed. At the end, a new csv file containing all the updated extrapolated values can be saved by pressing the "Save Results" button.

## 10.7   Conclusions

US-SOMO has now grown into a hub harboring different methods useful in multiresolution modeling. In this chapter, we have dealt only with the hydrodynamic methods which are directly linked to the parameters that AUC can provide. A verification of the accuracy with which the *SoMo* and *AtoB* methods can reproduce experimentally determined hydrodynamic parameters had been presented before (Brookes et al. 2010a, b), with generally more accurate results than alternative bead modeling methodologies. The cost paid is that US-SOMO requires a detailed coding of each residue in order to appropriately convert it into bead(s), somewhat limiting its direct application. However, approximate methods dealing with noncoded residues are provided. In addition, recently two other hydrodynamic computation methods have been implemented within US-SOMO, *Zeno*, which can operate on arbitrarily shaped models (Kang et al. 2004), and *BEST*, using the alternative boundary elements methodology (Aragon 2004) (for the latter, see Chap. 12). A full comparison between all the hydrodynamics methods currently available in US-SOMO, and with *HYDROPRO* (see Chap. 11) using a well-defined set of proteins with carefully verified literature translational diffusion and sedimentation experimental parameters, has been very recently carried out (Rocco and Byron 2015). The results evidenced a slight overestimation on average of $D_t$ and $s$ by the *SoMo* approach ($\sim + 2$ %) and a slightly larger underestimation of the same parameters by *BEST* and *HYDROPRO* ($\sim -3$ % and $\sim -4$ %, respectively). The best results with the standard implementations were obtained using the US-SOMO *AtoB* with a 5 Å grid size ($\sim + 1$ %). However, a combination of the *SoMo* bead model generation method, without overlap removal, and the *Zeno* computational

tool produced even better results (~0 %) (Rocco and Byron 2015). For this reason, this new combination has been already implemented within US-SOMO. With the future release of a much faster *Zeno* code (J. Douglas, NIST, Gaithersburg, MD, USA, personal communication), this approach could become the method of choice in US-SOMO for the computation of translational frictional properties and [η]. If the computation of rotational diffusion is sought, *BEST* could represent a viable alternative, since it is based on a correct hydrodynamic treatment, even if it is quite computationally intensive, requires an extrapolation to zero plate size, and treats hydration as a uniform layer.

# References

Aragon SR (2004) A precise boundary element method for macromolecular transport properties. J Comput Chem 25:1191–1205

Brookes E, Demeler B, Rosano C, Rocco M (2010a) 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 biomacro-molecule. Eur Biophys J 39:423–435

Brookes E, Demeler B, Rocco M (2010b) Developments in 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. Macromol Biosci 10:746–753

Brookes E, Singh R, Pierce M et al (2012) UltraScan Solution Modeler: integrated hydrodynamic parameter and small angle scattering computation and fitting tools. In: Proceedings of the 1st conference of the extreme science and engineering discovery environment: bridging from the eXtreme to the campus and beyond. ACM, New York. Article #42. doi:10.1145/2335755.2335839

Brookes E, Pérez J, Cardinali B et al (2013a) Fibrinogen species as resolved by HPLC-SAXS data processing within the UltraScan SOlution MOdeler (US SOMO) enhanced SAS module. J Appl Cryst 46:1823–1833

Brookes E, Singh R, Pierce M et al (2013b) US-SOMO cluster methods: year one perspective. XSEDE'13. In: Proceedings of the conference on extreme science and engineering discovery environment: gateway to discovery, ACM, New York. doi:10.1145/24847622484815

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

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

Demeler B (2005) UltraScan, a comprehensive data analysis software package for analytical ultracentrifugation experiments. In: Scott DJ, Harding SE, Rowe AJ (eds) Modern analytical ultracentrifugation: techniques and methods. Royal Society of Chemistry, Cambridge, pp 210–229

Ding F, Dokholyan NV (2006) Emergence of protein fold families through rational design. PLoS Comput Biol 2:e85

Dixon WJ (1951) Ratios involving extreme values. Ann Math Stat 22:68–78

Dokholyan NV, Buldyrev SV, Stanley HE, Shaknovich EI (1998) Discrete molecular dynamics studies of the folding of a protein-like model. Fold Des 3:577–587

Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, Pieper U, Sali A (2006) Comparative protein structure modeling With MODELLER. Curr Protoc Bioinf Supplement 15:5.6.1–5.6.30, John Wiley & Sons, Inc.

Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J Appl Cryst 42:342–346

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

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

Halle B, Davidovic M (2003) Biomolecular hydration: from water dynamics to hydrodynamics. Proc Natl Acad Sci U S A 100:12135–12140

Kang EH, Mansfield ML, Douglas JF (2004) Numerical path integration technique for the calculation of transport properties of proteins. Phys Rev E 69:031918/1-11

Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 32(S2):W526–W531 (WEB SERVER ISSUE)

Kuntz ID, Kauzmann W (1974) Hydration of proteins and polypeptides. In: Anfinsen CB, Edsall JT, Richards FM (eds) Advances in protein chemistry, vol 28. Academic, New York, pp 239–345

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

Nadassy K, Tomas-Oliveira I, Alberts I et al (2001) Standard atomic volumes in double-stranded DNA and packing in protein-DNA interfaces. Nucleic Acids Res 29:3362–3376

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

Rai N, Nöllmann M, Spotorno B et al (2005) SOMO (SOlution MOdeler): differences between X-ray and NMR-derived bead models suggest a role for side chain flexibility in protein hydrodynamics. Structure 13:723–734

Rocco M, Brookes E (2014) Dynamical aspects of biomacromolecular multi-resolution modelling using the UltraScan Solution Modeler (US-SOMO) suite. In: Howard JA, Sparkles HA (eds) The future of dynamical structural science. NATO science for peace and security series A: chemistry and biology, Springer-Verlag, Berlin/Heidelberg. Chapter 13, pp 185–194. doi:10.1007/978-94-017-8550-1_13

Rocco M, Byron O (2015) Computing translational diffusion and sedimentation coefficients: an evaluation of experimental data and programs. Eur Biophys J 44:417–431

Rocco M, Brookes EH, Byron O (2012) US-SOMO: methods for the construction and hydration of macromolecular hydrodynamic models. In: Roberts GCK (ed) Encyclopedia of biophysics. Springer, Berlin/Heidelberg. doi:10.1007/978-3-642-16712-6

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5:725–738

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

Spotorno B, Piccinini L, Tassara G et al (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)

Tsai J, Taylor R, Chothia C, Gerstein M (1999) The packing density in proteins: standard radii and volumes. J Mol Biol 290:253–266

Tsodikov OV, Record MT Jr, Sergeev YV (2002) Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. J Comput Chem 23:600–609

Voss NR, Gerstein M (2005) Calculation of standard atomic volumes for RNA and comparison with proteins: RNA is packed more tightly. J Mol Biol 346:477–492

Voss NR, Gerstein M (2010) 3V: cavity, channel and cleft volume calculator and extractor. Nucleic Acids Res 38:W555–W562

Vriend G (1990) WHAT IF: a molecular modeling and drug design program. J Mol Graph 8:52–56