US-SOMO (SOlution MOdeler): versatile and reliable hydrodynamic and SAS modeling of biomacromolecules within the UltraScan AUC data analysis software

Emre Brookes and Borries Demeler
Department of Biochemistry, The University of Texas Health Science Center at San Antonio, TX, USA

Mattia Rocco
Biopolymers and Proteomics, AOU-IRCCS San Martino-IST, National Institute for Research on Cancer, Genova, Italy

In collaboration also with:

Olwyn Byron, Glasgow University, Scotland, UK
Patrice Vachette, CNRS-Université Paris-Sud, Orsay, FR
Javier Pérez, SWING, Synchrotron SOLEIL, St. Aubin, FR.

Hydrodynamics.
Stokes-Einstein's law

The translational frictional coefficient of f a sphere of radius r in a solvent of viscosity η is:

$$f = 6\pi\eta r$$

For a body of unknown shape having a translational frictional coefficient f, we can define the Stokes radius R_s as that of a sphere having the same f

$$R_s = r = f/6\pi\eta$$

Translational diffusion coefficient D_t:

$$D_t = k_b T/f$$

where k_b = Boltzmann's constant

T = solution temperature (°K)

Sedimentation coefficient s:

$$s = (M/N_0)(1 - v_2\rho)/f$$

where M = mass

N_0 = Avogadro's number

v_2 = partial specific volume

ρ = solvent density

Early applications of hydrodynamics: representing proteins as simple geometrical objects (spheres, ellipsoids of revolution)

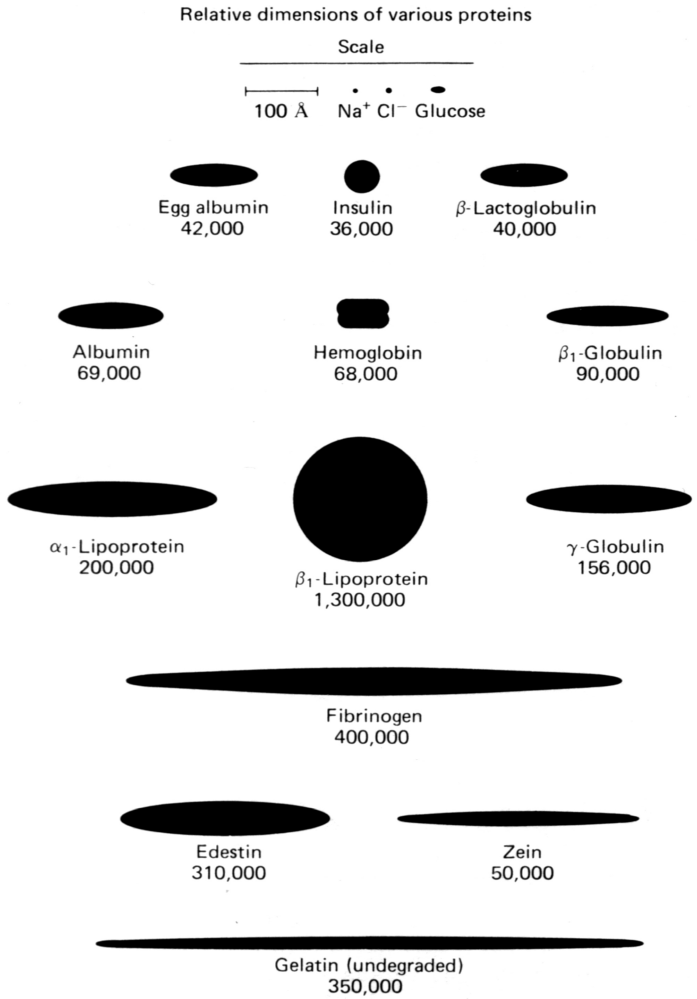
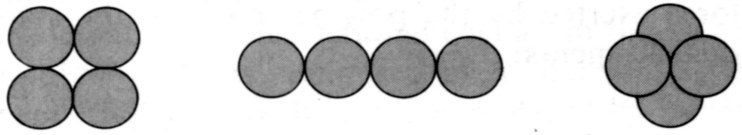
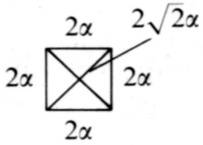
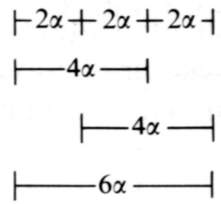
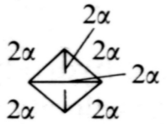
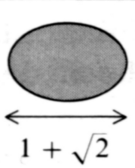
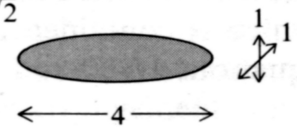


FIGURE 7-25. Dimensions of the ellipsoid of revolution that best account for the hydrodynamic properties (viscosity and frictional coefficient) of various protein molecules. [After W. J. Moore, *Physical Chemistry*, Prentice-Hall, Englewood Cliffs, N. J., 1972.)

Early applications of hydrodynamics: representing proteins as simple geometrical objects (spheres, ellipsoids of revolution)

Subunit arrangement			
Distances (each counted twice in Eqn. 10-33)			
$f/f_{monomer}$	1.70	1.92	1.60
f relative to tetrahedron	1.06	1.20	1.00
Ellipsoid model			 (Sphere)
f/f_{sph} from Eqn. 5-12	1.07	1.18	1.00

Cantor & Schimmel, Biophysical Chemistry Part II: Techniques for the study of biological strcuture and function

The "hydration" issue

Cantor and Schimmel - Biophysical Chemistry Part II, 1980

Table 12-2
Hydrations of biopolymers computed by using shapes known from x-ray diffraction or electron microscopy

Sample	Known axial ratio§	Hydration (δ_1 in g/g) based on		
		Viscosity	Diffusion	Sedimentation
Bushy stunt virus	1.0	0.65	0.71	0.71
Carboxypeptidase	1.25	——	0.30	0.69
Cytochrome *c*	1.48	——	0.18	0.24
Hemoglobin	1.3	0.62	0.52	0.75
Lysozyme	1.5	0.34	0.52	0.52
Myoglobin	1.76	0.44	0.50	0.42
Tobacco mosaic virus	18	0.32	0.1–0.7	0.26

§Axial ratios are for prolate ellipsoids, except for cytochrome *c*, which is oblate.

SOURCE: After I. D. Kuntz, Jr., and W. Kauzmann, in *Advances in Protein Chemistry*, vol. 28, ed. C. B. Anfinsen, J. T. Edsall, and F. M. Richards (New York: Academic Press, 1974), p. 239.

Advances in Protein Chemistry 28:239-345, 1974

HYDRATION OF PROTEINS AND POLYPEPTIDES

By I. D. KUNTZ, JR. and W. KAUZMANN

Department of Pharmaceutical Chemistry, University of California, San Francisco, California, and Department of Chemistry, Princeton University, Princeton, New Jersey

Hydration from NMR freezing experiments

TABLE XXII
Proposed Amino Acid Hydrations Based on Nuclear Magnetic Resonance Studies of Polypeptides[a]

Amino acid residues[b]	Hydration[c]
Ionic	
Asp^-	6
Glu^-	7
Tyr^-	7
Arg^+	3
His^+	4[d]
Lys^+	4
Polar	
Asn	2
Gln	2[e]
Pro	3
Ser, Thr	2[e]
Trp	2[e]
Asp	2
Glu	2
Tyr	(3)
Arg	3
Lys	4
Nonpolar	
Ala	1
Gly	1
Phe	(0)
Val	1
Ile, Leu, Met	1[e]

[a] After Kuntz (1971a).
[b] Standard three-letter code.
[c] Moles of water per mole of amino acid.
[d] As Lys^+.
[e] Assumed values based on one water molecule per amide plus one water molecule per side-chain polar group.

TABLE XXIII
Prediction of Protein Hydration from Composition and Polypeptide Results[a]

Protein, native	Hydration (g H_2O/g protein)	
	Calculated[b]	Observed[c]
Lysozyme	0.36	0.34
Myoglobin	0.45	0.42
Chymotrypsinogen	0.39	0.34
Chymotrypsin	0.36	0.33
Ovalbumin	0.37	0.33
Bovine serum albumin (BSA)	0.45	0.40
Hemoglobin (denatured)	0.42	0.42
BSA + urea	0.45	0.44
BSA, pH 3	0.32[d]	0.30

[a] After Kuntz (1971a); see Table XXII.
[b] Calculation assumes that *all* residues are fully hydrated. This is perhaps reasonable for the denatured proteins but leads to a small positive error unless allowance is made for "buried" groups. This correction was done for lysozyme, yielding a calculated value of 0.335.
[c] NMR freezing experiments.
[d] Calculation assumes that all carboxyl groups are uncharged at pH 3.

Reconciling hydration dynamics with hydrodynamics

PNAS 100:12135-12140, 2003

Biomolecular hydration: From water dynamics to hydrodynamics

Bertil Halle* and Monika Davidovic

Department of Biophysical Chemistry, Lund University, Box 124, SE-22100 Lund, Sweden

1-The static picture of biomolecular hydration is fundamentally inconsistent with magnetic relaxation dispersion experiments and molecular dynamics simulations, which both reveal a highly dynamic interface where rotation and exchange of nearly all water molecules are several orders of magnitude faster than biomolecular diffusion.

2-Waters near the biomolecular surface have a different density, and alter the local viscosity. It turns out that considering a number of “tightly bound”, static water molecules compensate well for this local viscosity effect, otherwise very hard to be directly taken into account.

Bead modeling methods: from an idea of V. Bloomfield, further developed by D.C. Teller, to HYDROPRO (J. García de la Torre)

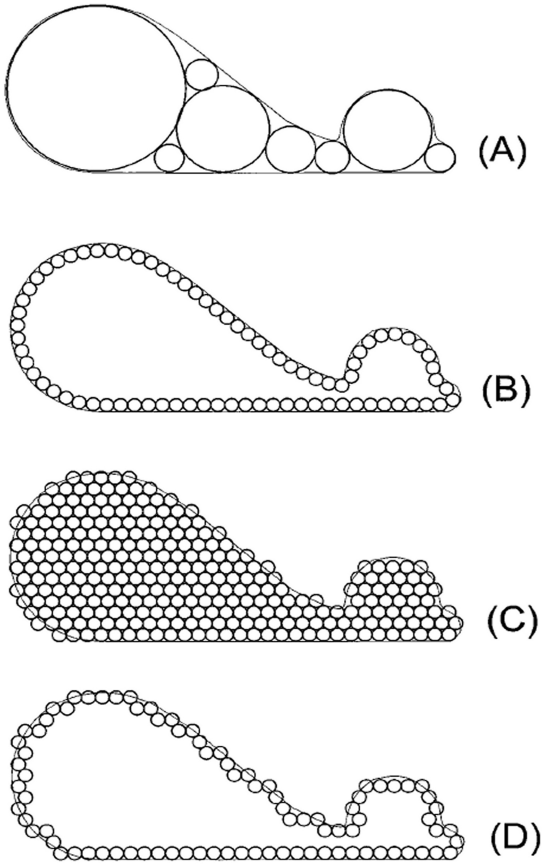


FIGURE 1 Two-dimensional analogies of the various model types. (*A*) A bead model (in strict sense). (*B*) Shell model. (*C*) Filling model. (*D*) Rough-shell model.

Carrasco & García de la Torre
Bioph. J. 75, 3044-3057, 1999

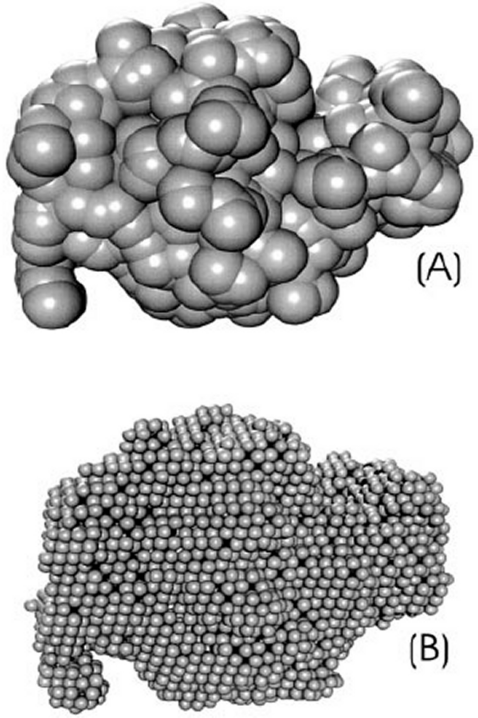


FIGURE 2 (*A*) A bead-per-atom (BPA) model of lysozyme, which we take as the primary hydrodynamic particle (PHP) that represents this protein. The atomic element radius (AER) is $a = 3$ Å. (*B*) A shell model (SHE), derived from the PHP, used for hydrodynamic calculations. The radius of the small beads in this case is $\sigma = 0.8$ Å.

García de la Torre, Huertas & Carrasco
Bioph. J. 78, 719-730, 2000

Low Reynolds number hydrodynamics

$$\begin{pmatrix} \mathbf{F} \\ \mathbf{T}_O \end{pmatrix} = \begin{pmatrix} \Xi_t & \Xi_{O,c}^T \\ \Xi_{O,c} & \Xi_{O,r} \end{pmatrix} \begin{pmatrix} \mathbf{u}_O \\ \boldsymbol{\omega} \end{pmatrix}$$

$$\mathbf{F}_i \, (6\pi \, \eta_0 \, \sigma_i)^{-1} + \sum_{j=1}^{N}{}' \, \mathbf{T}_{ij} \cdot \mathbf{F}_j = (\mathbf{u}_i - \mathbf{v}_i^0)$$

$$\mathbf{T}_{ij} = (8\pi \, \eta_0 \, \mathbf{R}_{ij})^{-1} \left[\mathbf{I} + \frac{\mathbf{R}_{ij} \mathbf{R}_{ij}}{\mathbf{R}_{ij}^2} + \frac{(\sigma_i^2 + \sigma_j^2)}{\mathbf{R}_{ij}^2} \left(\tfrac{1}{3} \mathbf{I} - \frac{\mathbf{R}_{ij} \mathbf{R}_{ij}}{\mathbf{R}_{ij}^2} \right) \right]$$

$$\sum_{j=1}^{N} \mathbf{B}_{ij} \cdot \mathbf{F}_j = (\mathbf{u}_i - \mathbf{v}_i^0) \qquad \mathbf{B}_{ij} = \delta_{ij} \, \frac{\mathbf{I}}{6\pi \, \eta_0 \, \sigma_i} + (1 - \delta_{ij}) \, \mathbf{T}_{ij} \qquad C = \mathcal{B}^{-1}$$

$$\Xi_t = \sum_i \sum_j \mathbf{C}_{ij} \qquad \Xi_{O,c} = \sum_i \sum_j \mathbf{U}_i \cdot \mathbf{C}_{ij} \qquad \Xi_{O,r} = -\sum_i \sum_j \mathbf{U}_i \cdot \mathbf{C}_{ij} \cdot \mathbf{U}_j + 6 \, \eta_0 \, \mathrm{V} \, \mathbf{I}$$

TO COMPUTE THE PARAMETERS THAT CAN BE MEASURED EXPERIMENTALLY, A COMPROMISE MUST BE REACHED BETWEEN A GOOD REPRESENTATION OF THE SURFACE OF THE PROTEIN AND A LOW NUMBER OF FRICTIONAL ELEMENTS (BEADS).

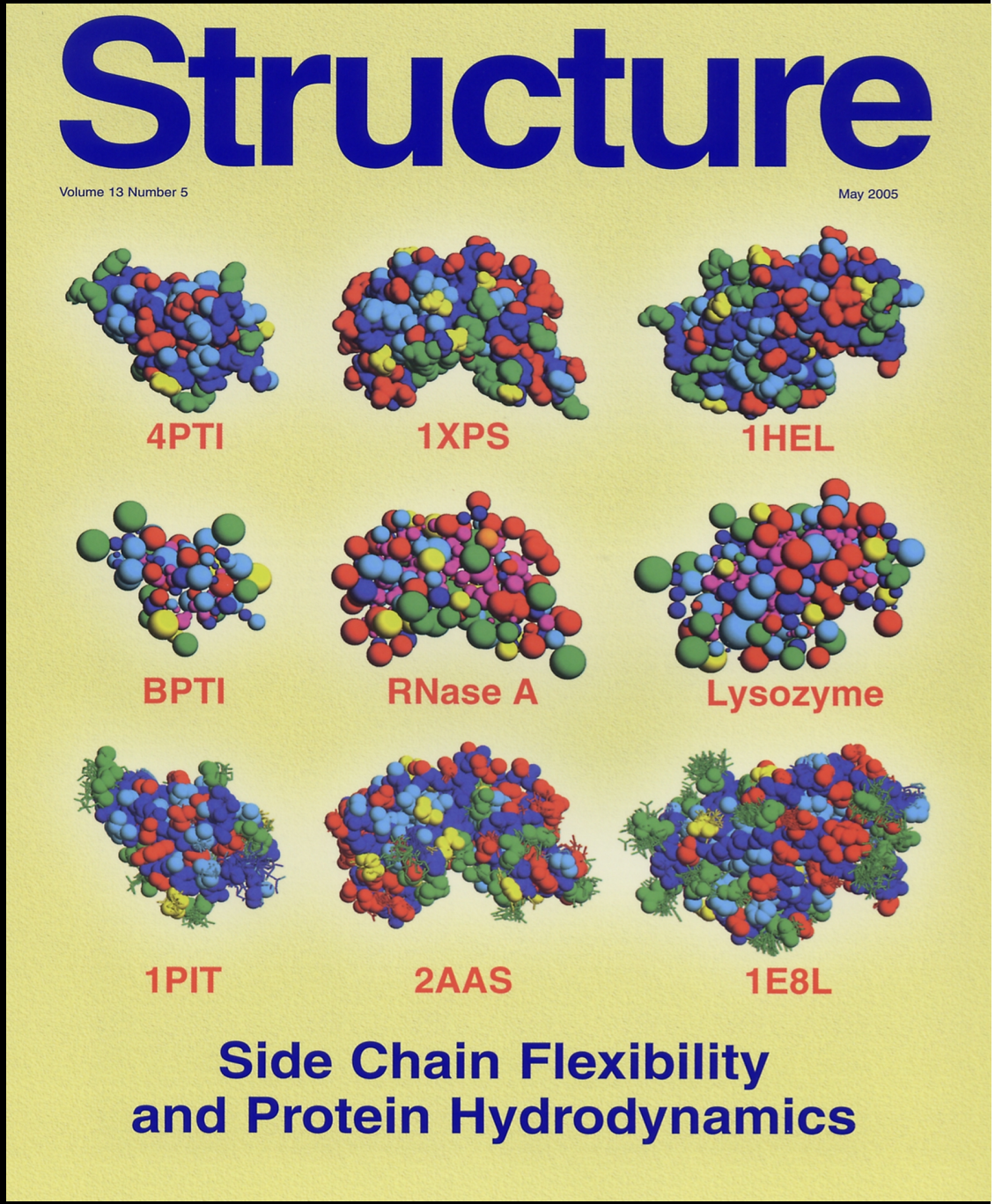
THE LAYER OF "TIGHTLY BOUND" WATER OF HYDRATION MUST ALSO BE TAKEN INTO ACCOUNT

Early programs developed by the Byron/Rocco groups:

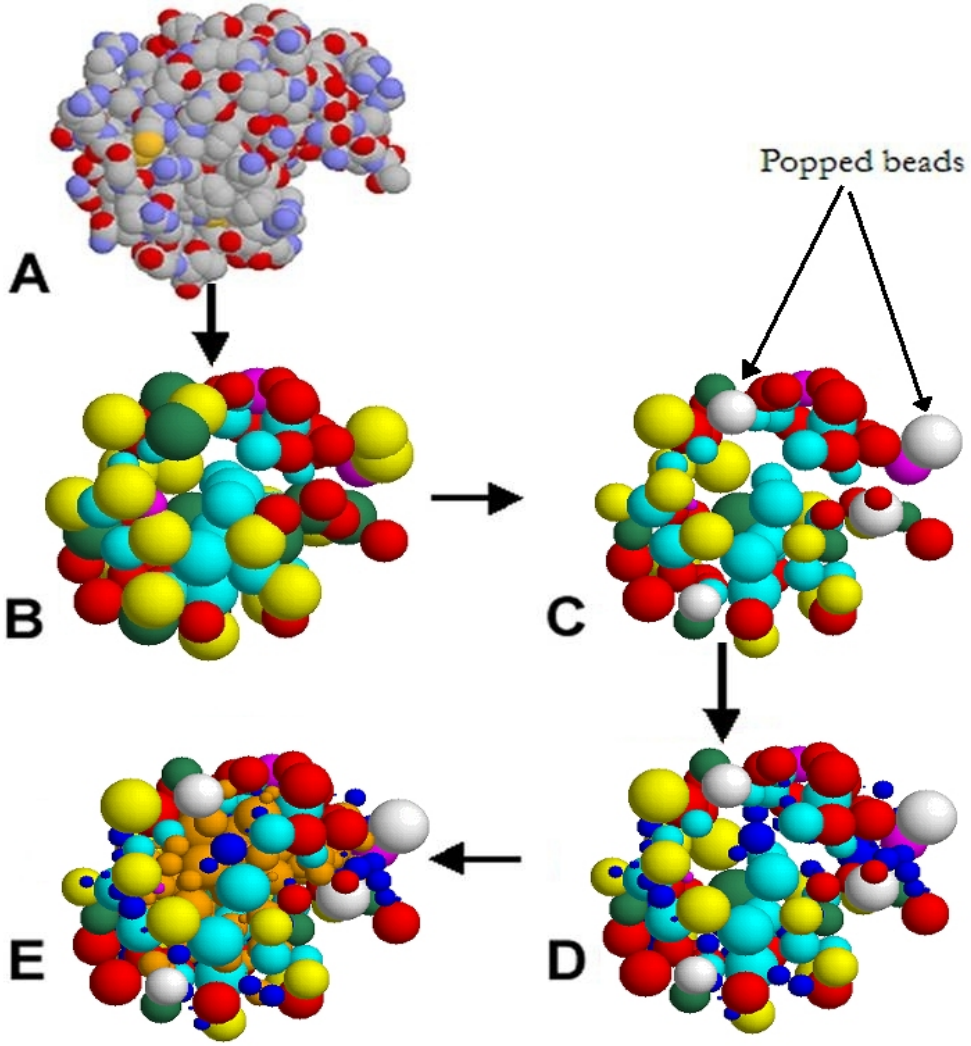
BEAMS (BEAds Modeling System) Spotorno et al., Eur. Biophys. J. 25, 373-384, 1997.

AtoB
Byron, Biophys J. 72, 408-415, 1997.

SOMO (SOlution MOdeler)
Rai et al., Structure 13, 723-744, 2005;



Method *SOMO* (SOlution MOdeller): generating medium-resolution bead models from atomic coordinates



Main features:

1 bead/side chain & 1 bead/main chain. Water of hydration, based on residues, is included in each bead.

A→B After ASA screening,exposed side -chains beads are placed.

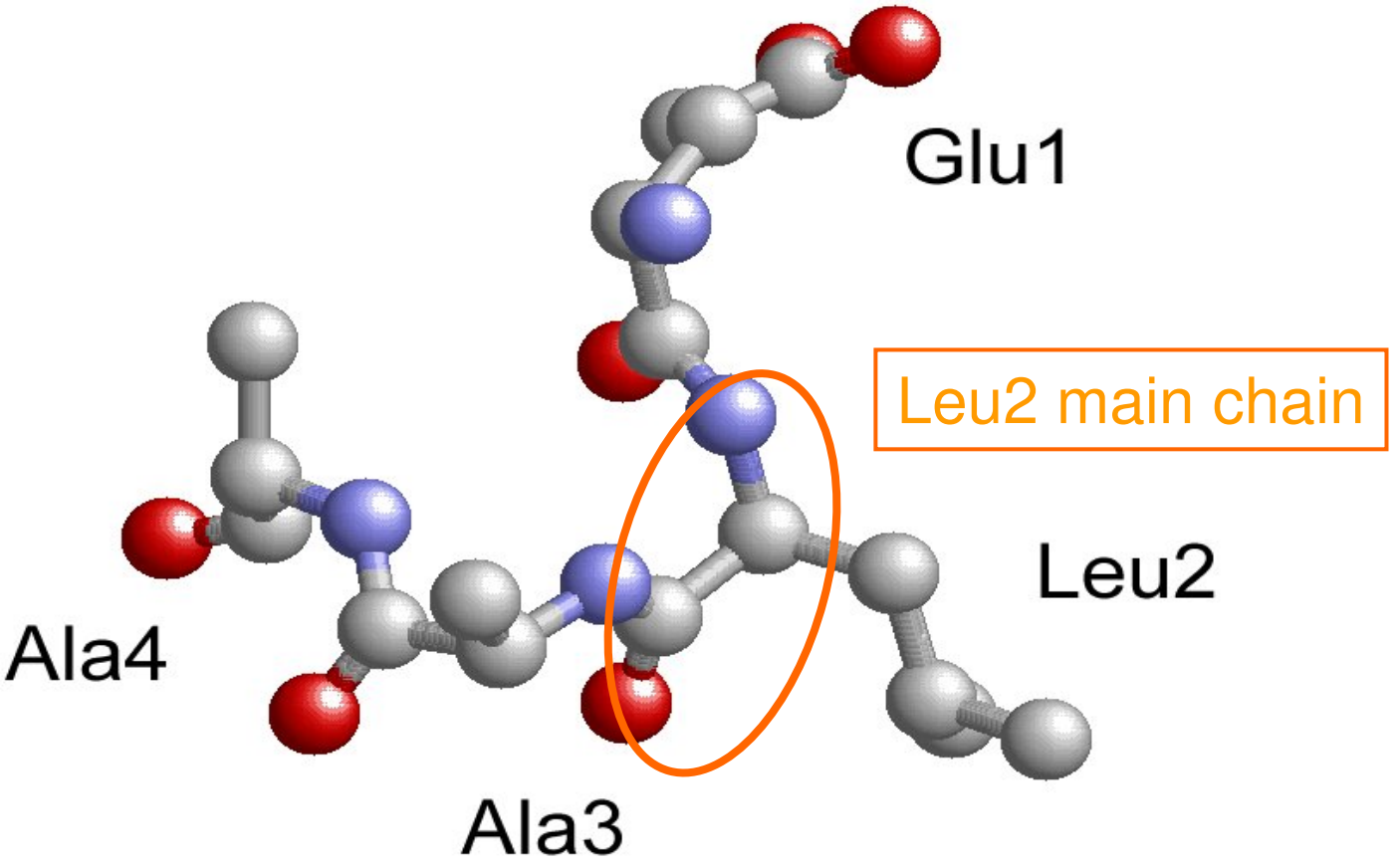
B→C Beads overlapping by more than a preset threshold can be fused together. Overlaps are then removed, reducing the radii and outward translating the centers of exposed beads.

C→D Exposed peptide bond beads are placed and overlaps removed.

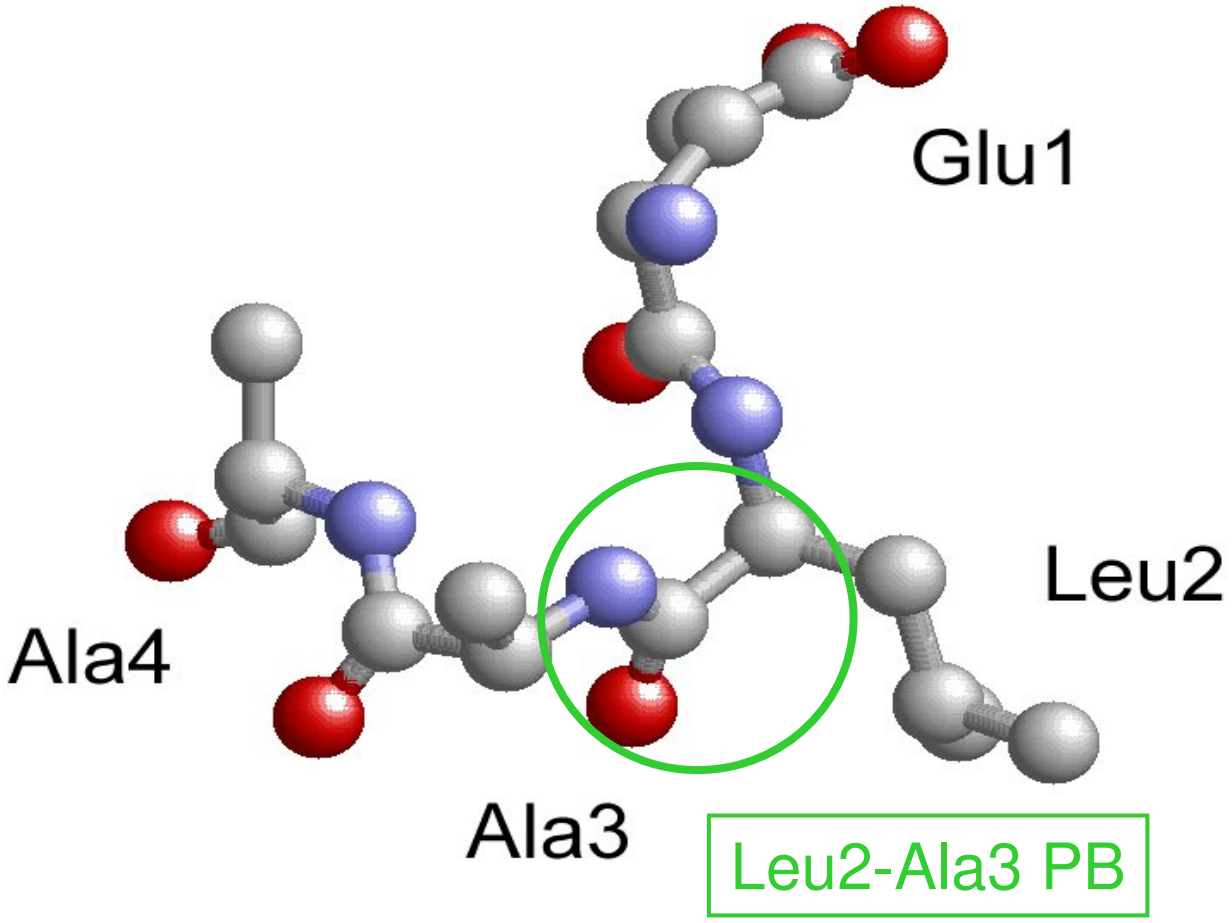
D→E Buried beads are placed and overlaps removed. They should be excluded from the computations of the hydrodynamic parameters.

Rai et al., Structure, May 2005

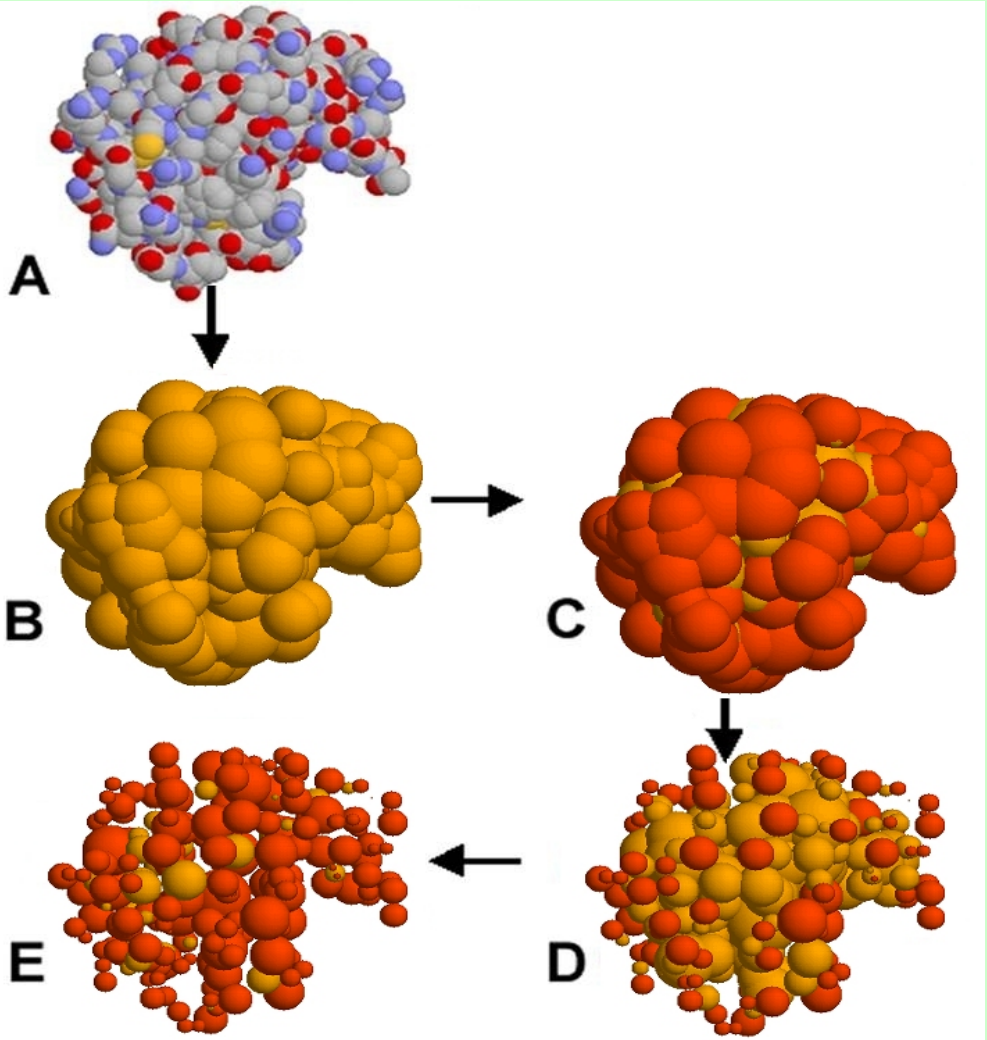
The “peptide bond” rule



The “peptide bond” rule



Improved AtoB (Grid) method: generating variable-resolution bead models from atomic coordinates



Main features:

1 bead/cube in a variable-size cubic grid. Water of hydration, based on atom values, included in each bead.

A→B All the beads are generated and placed (CM or CC).

B→C Beads are screened for surface accessibility (ASA; red, accessible; orange, buried).

C→D Overlaps between the exposed beads are then removed, reducing the radii and outward translating the beads' centers.

D→E Overlaps between the buried beads are then removed, and they are re-screened for accessibility. Buried beads are excluded from hydrodynamic computations.

Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac

Eur Biophys J
DOI 10.1007/s00249-009-0418-0

ORIGINAL PAPER

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

Emre Brookes · Borries Demeler · Camillo Rosano · Mattia Rocco

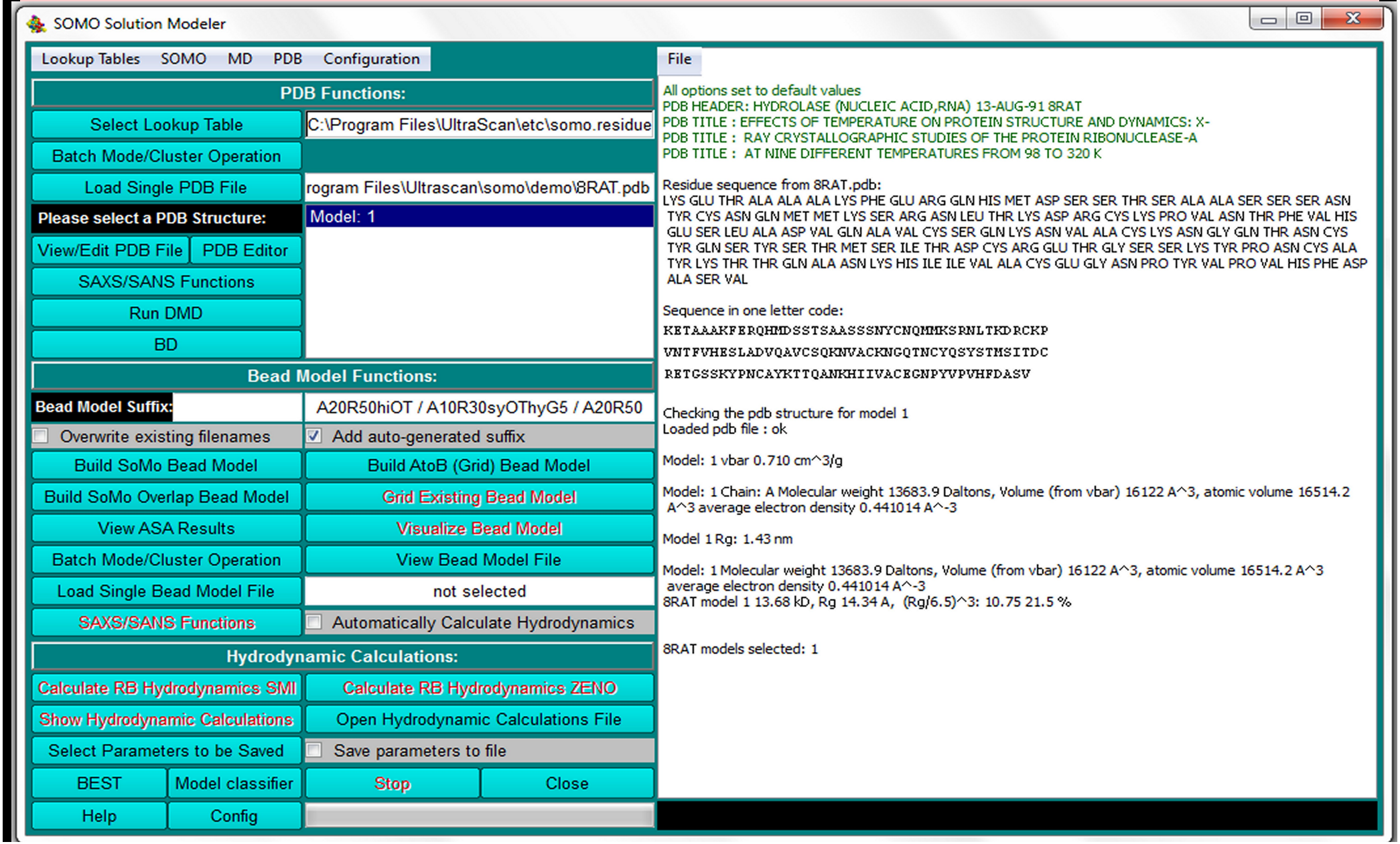
mabi.200900474

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[a]

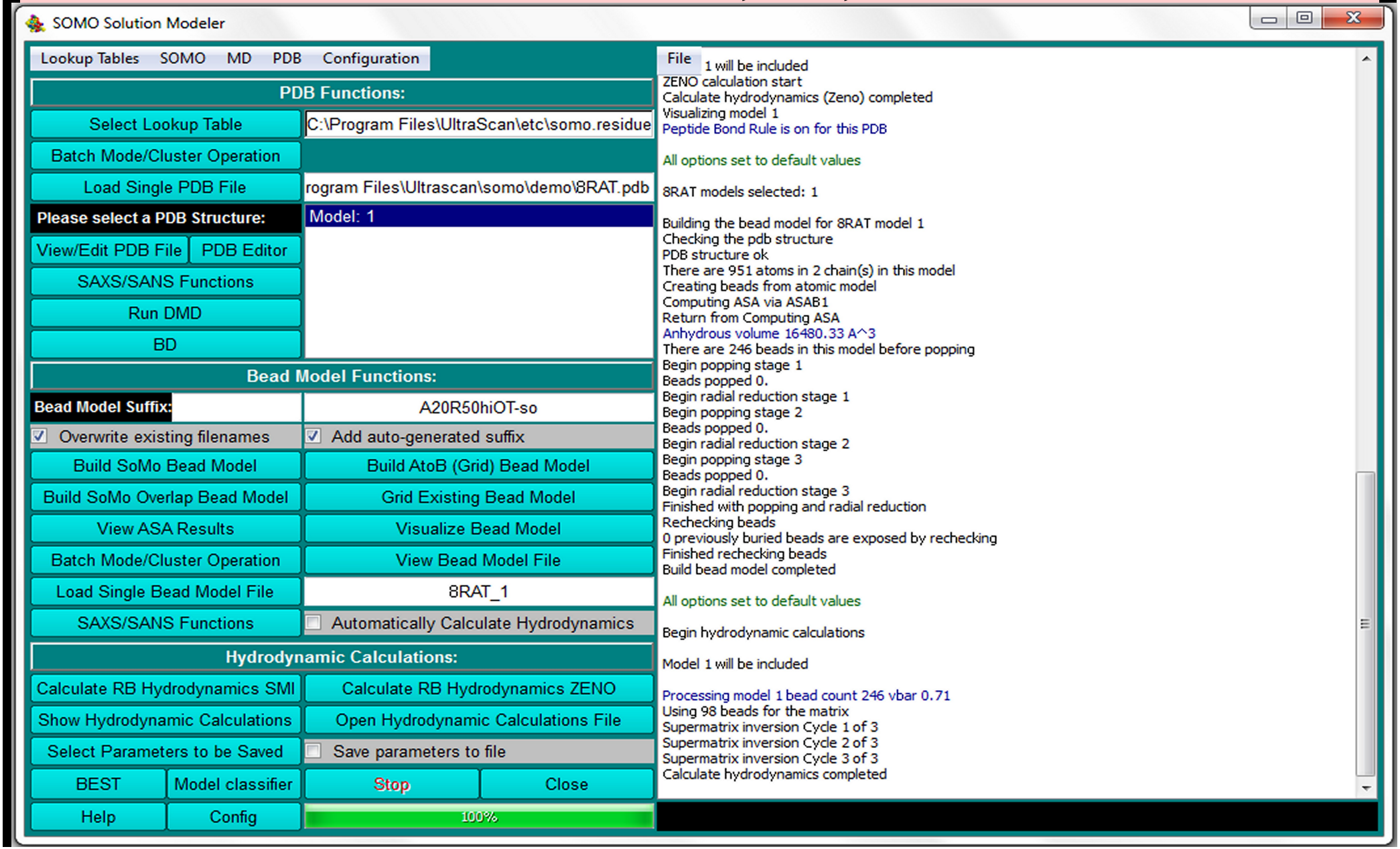
Author Proof

Emre Brookes, Borries Demeler, Mattia Rocco*

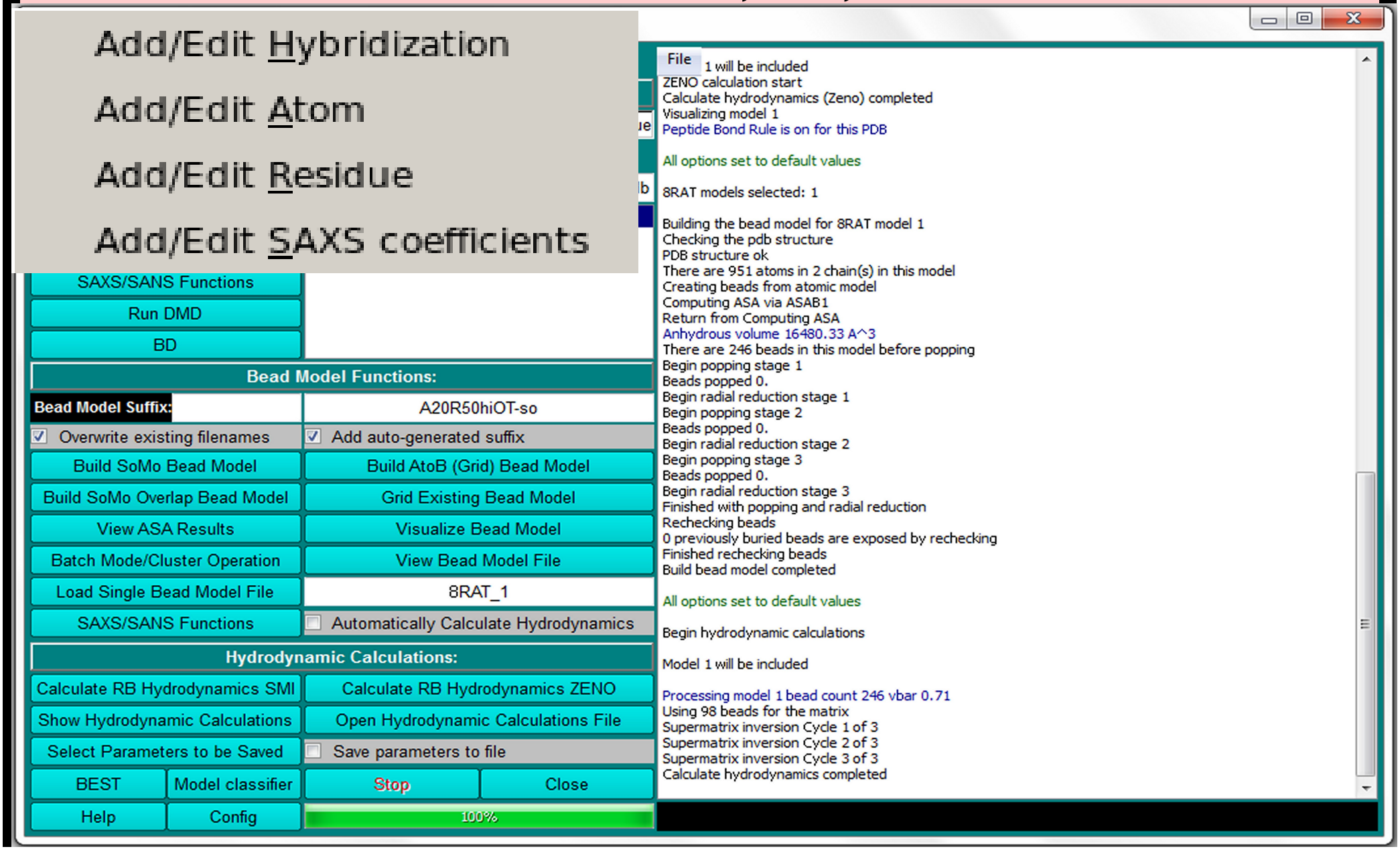
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



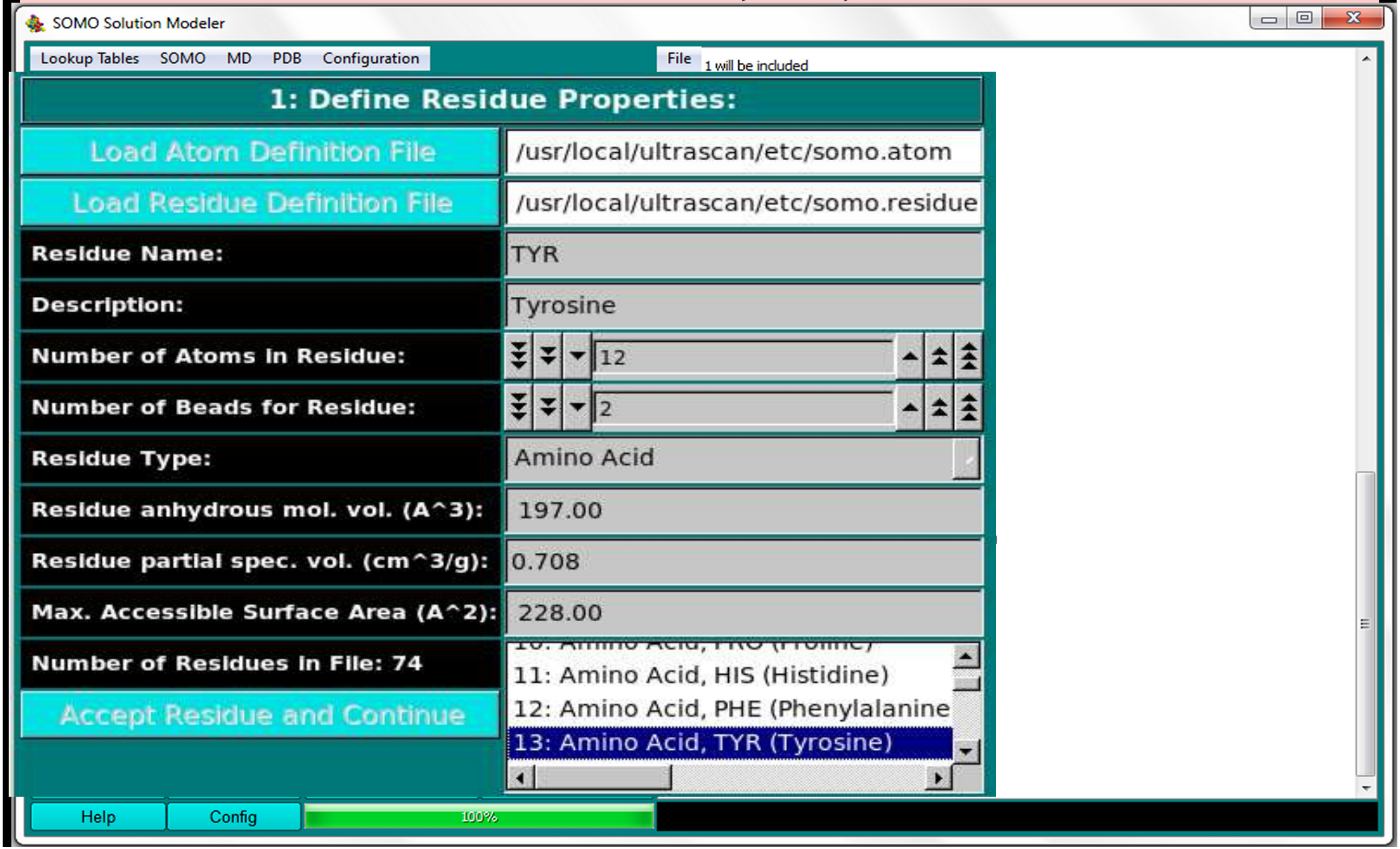
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



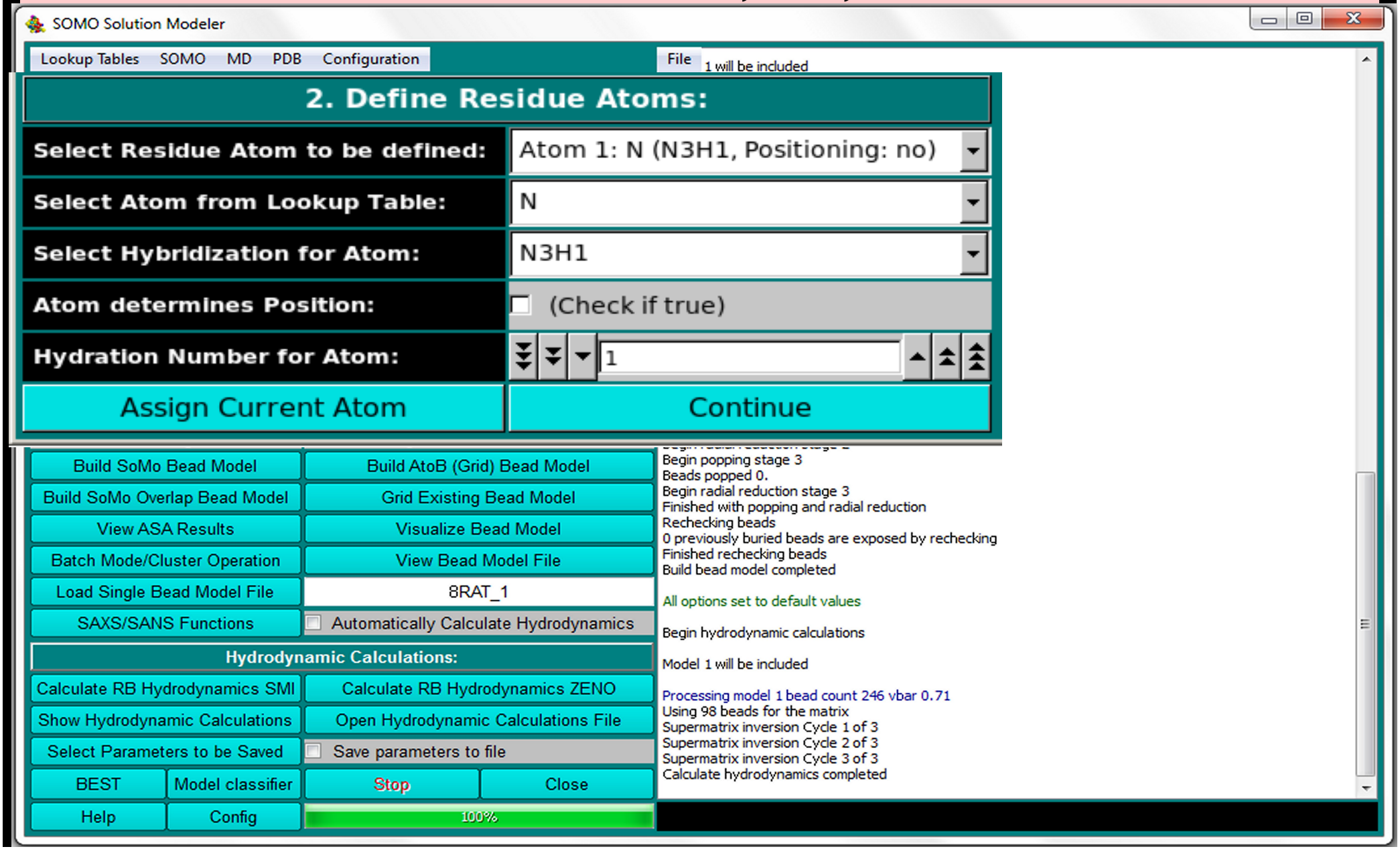
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



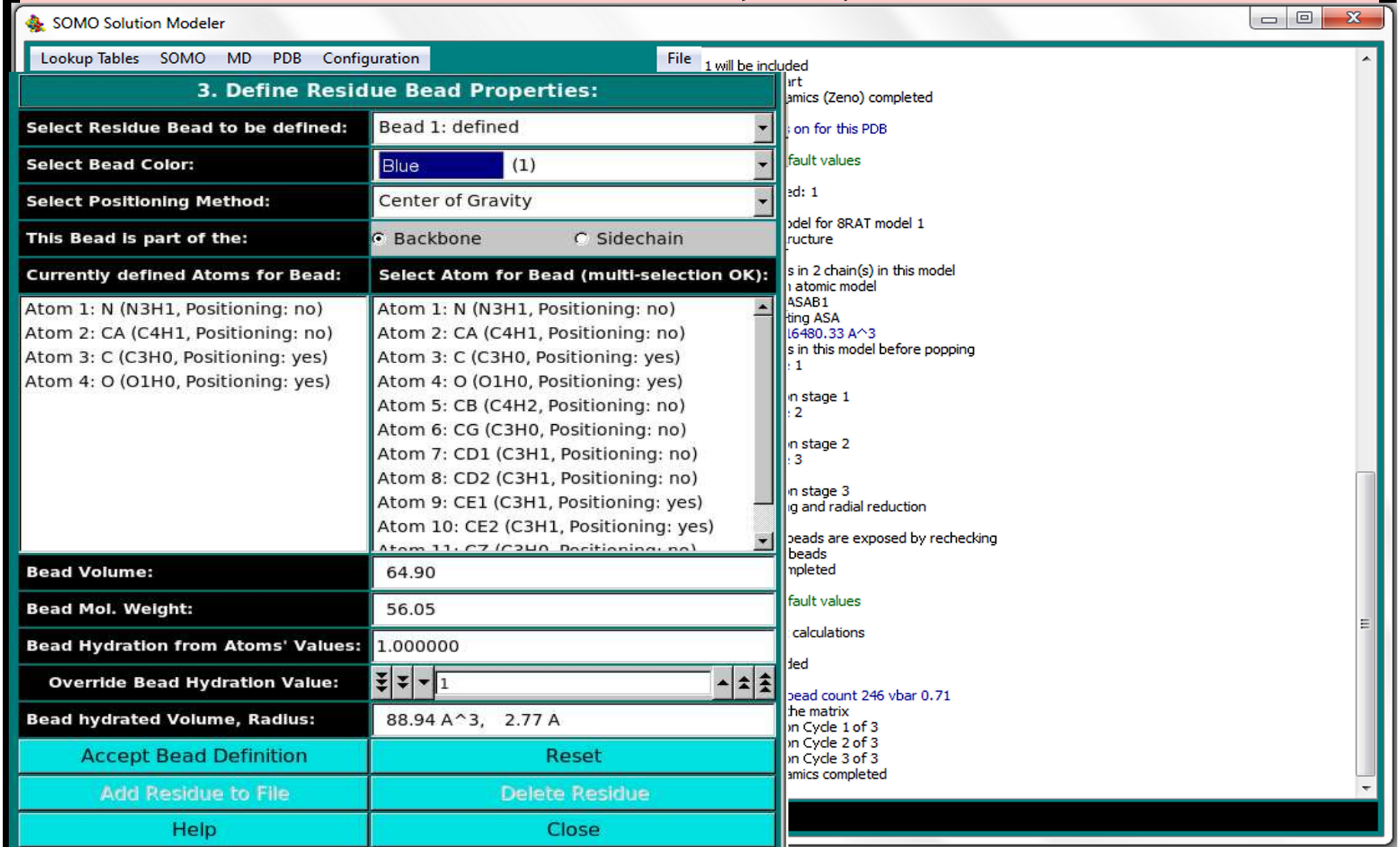
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



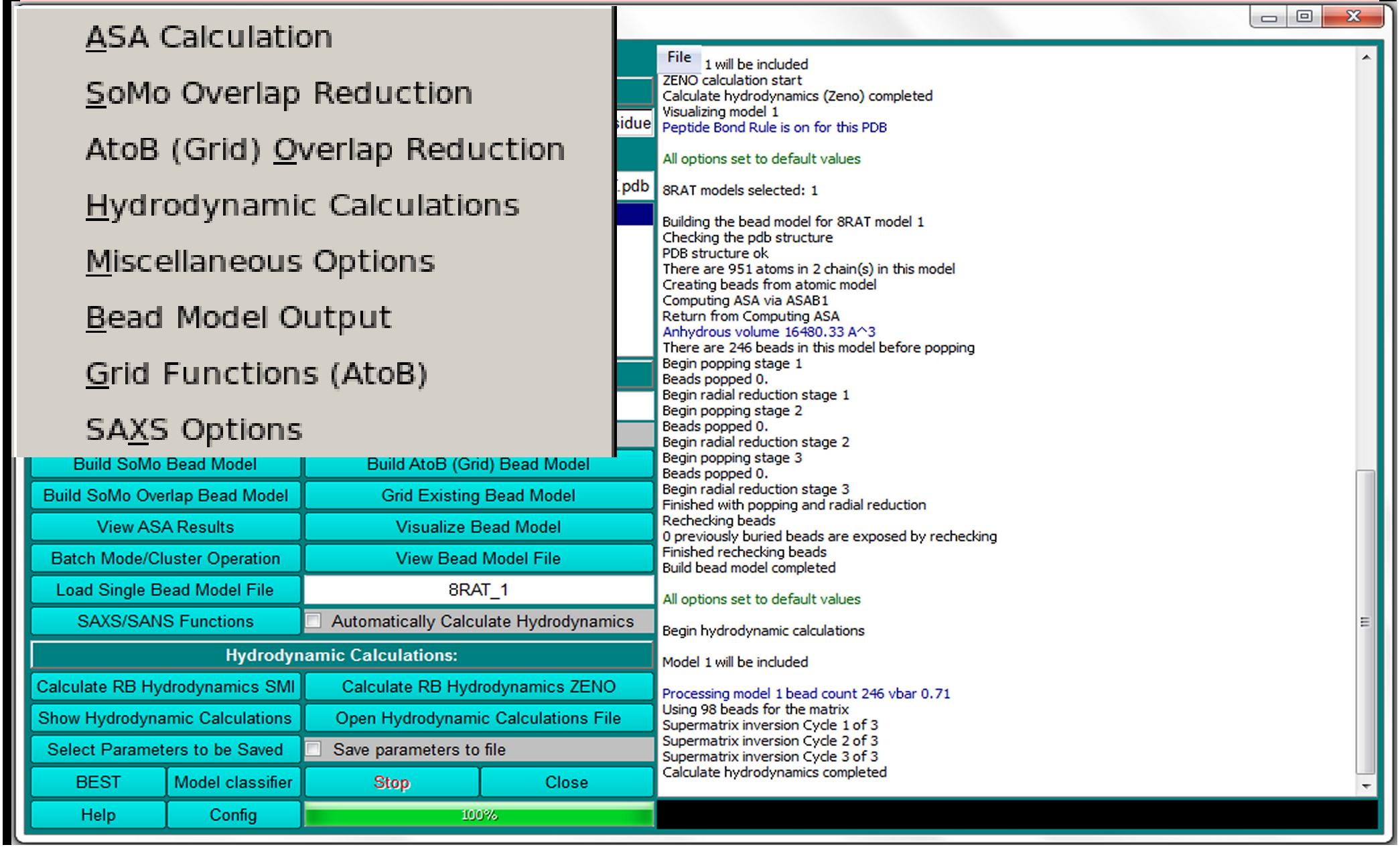
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



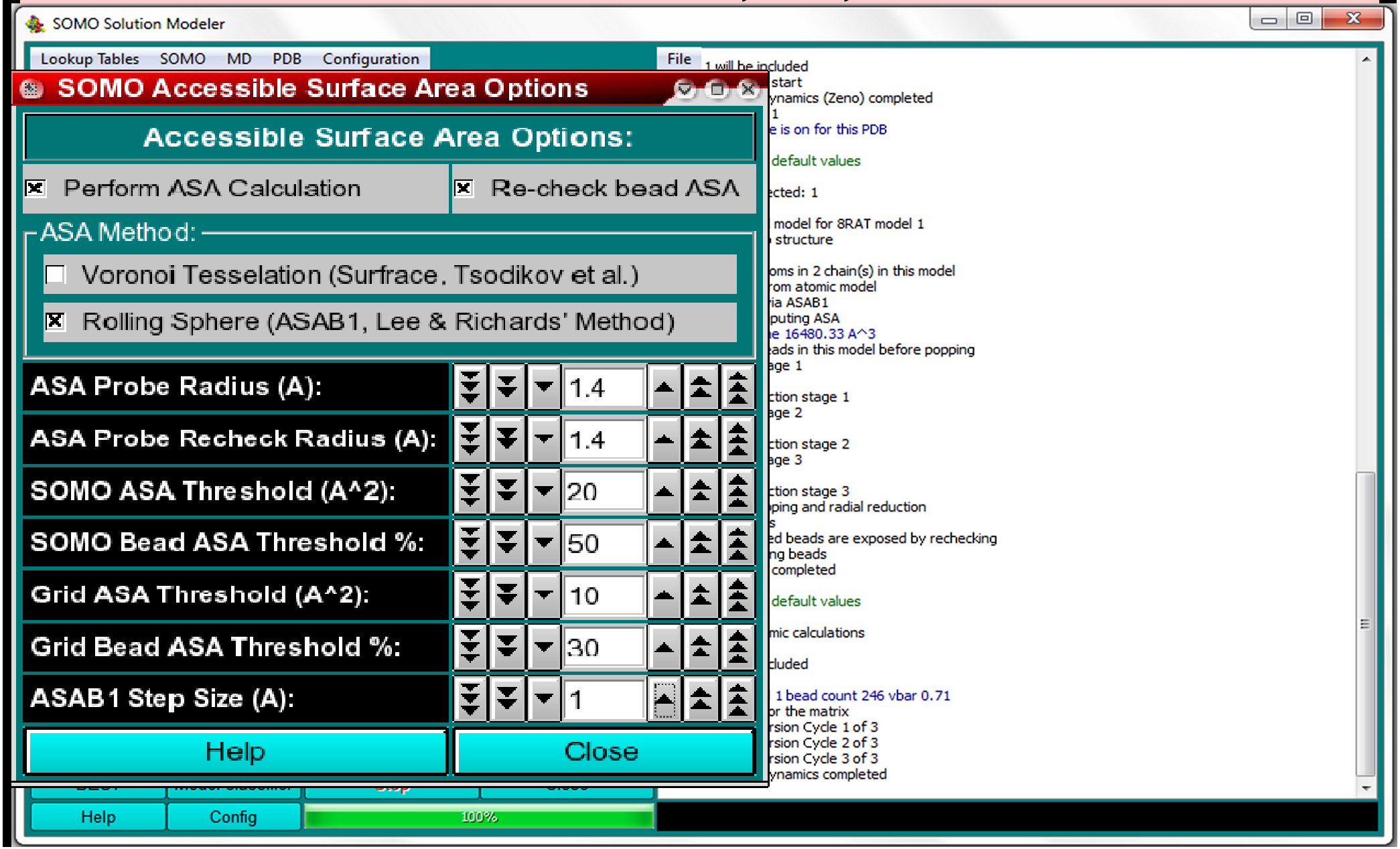
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



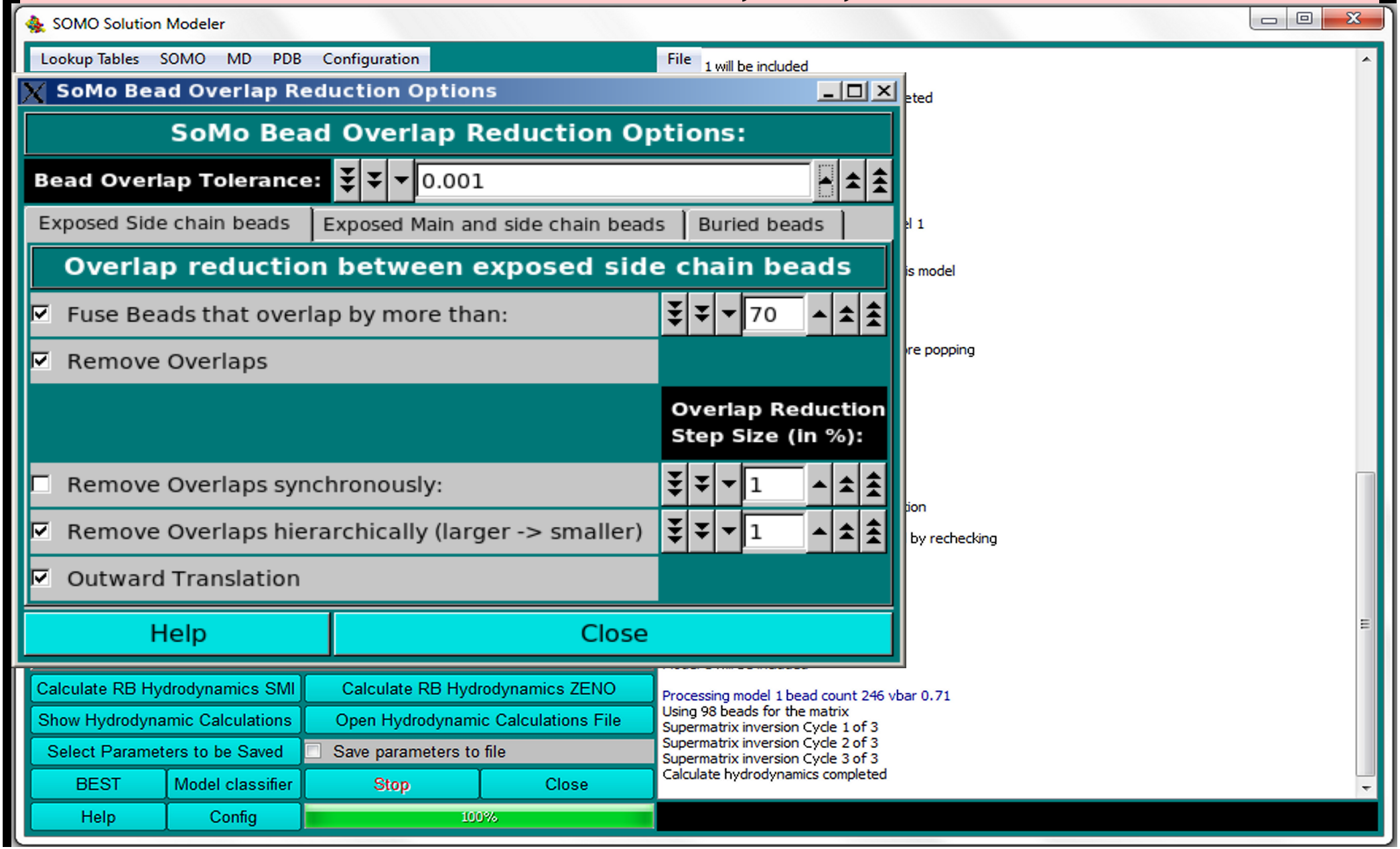
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



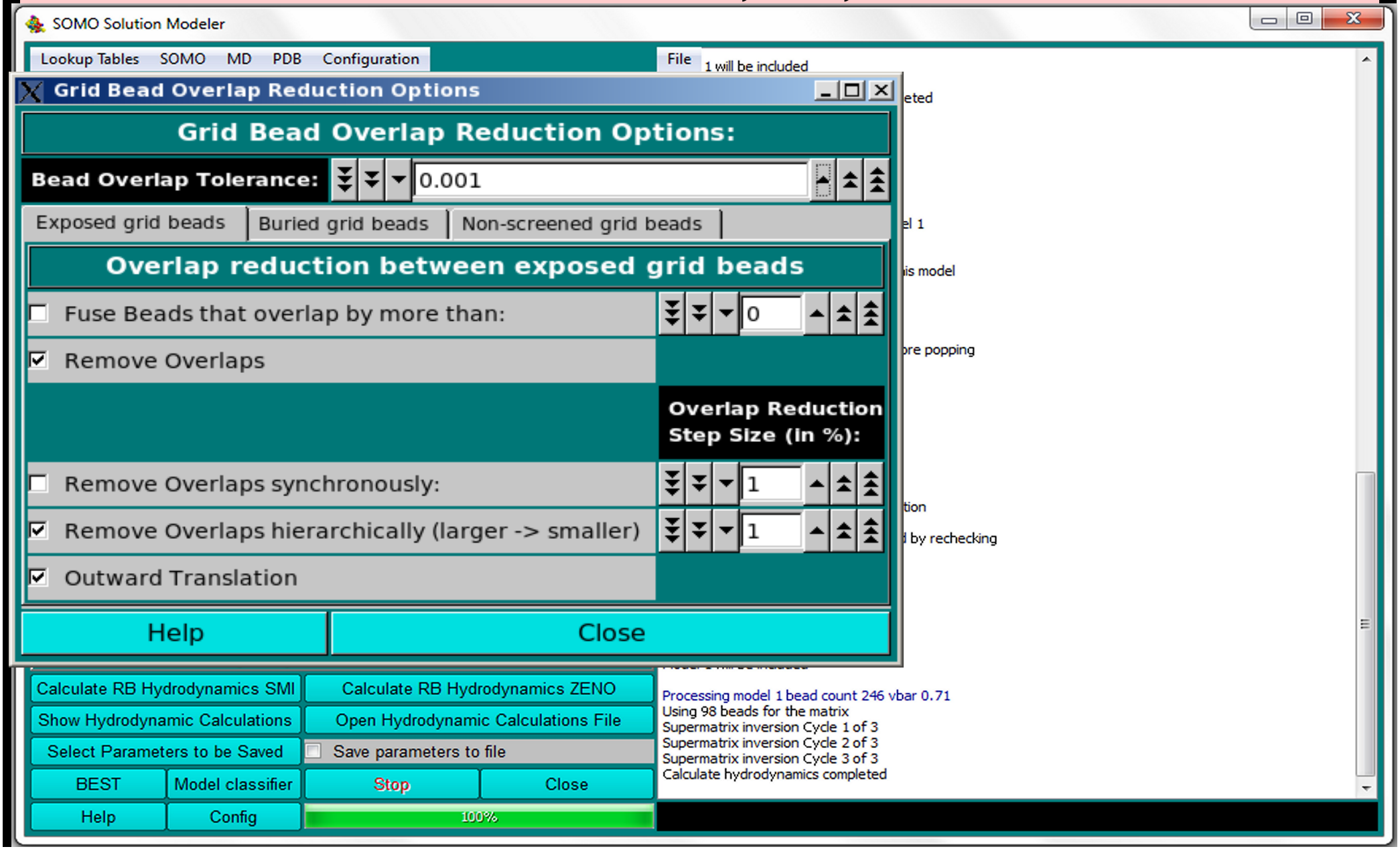
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



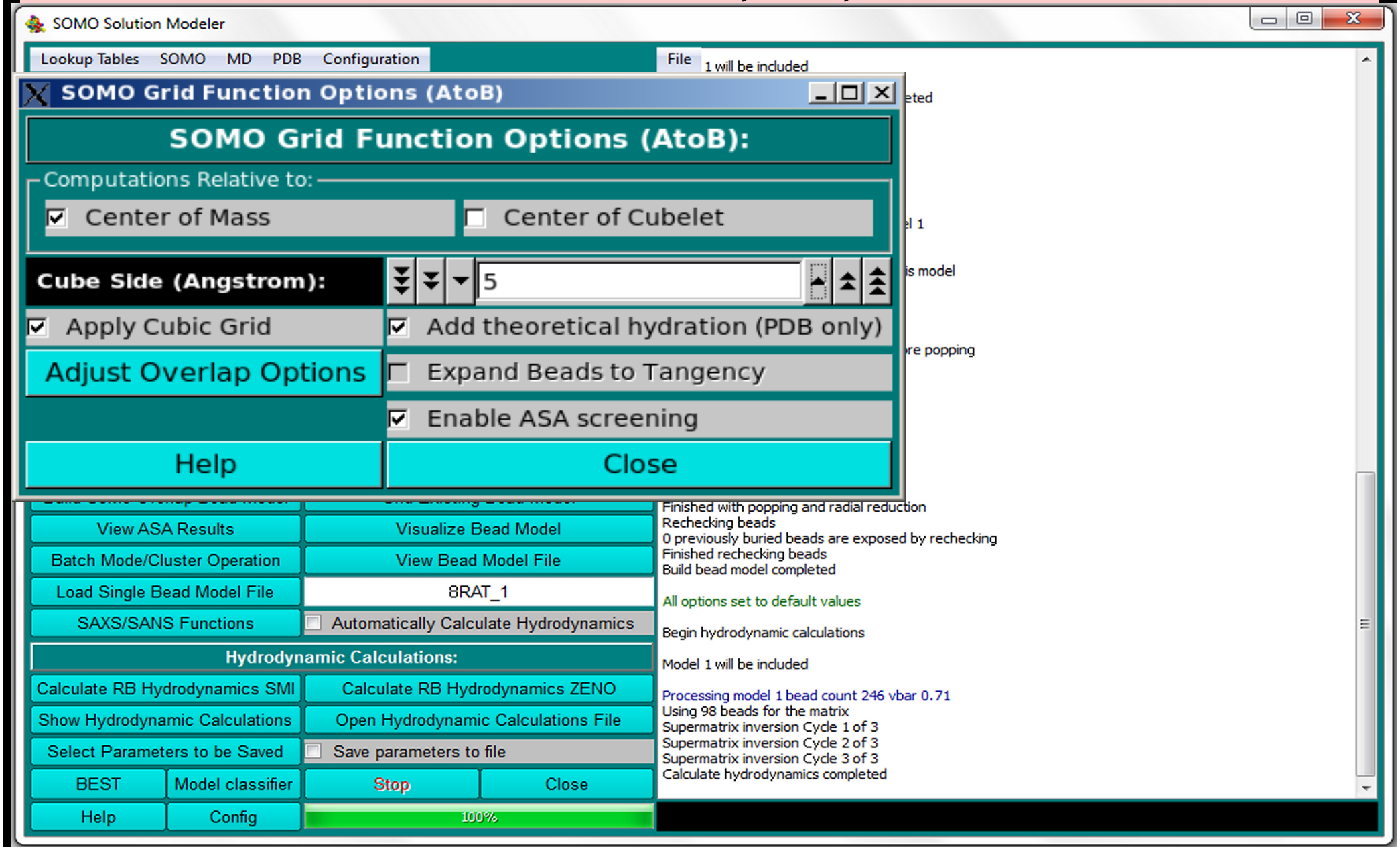
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac

SOMO Hydrodynamic Results

SOMO Hydrodynamic Results (Water at 20°C):	
Total Beads in Model:	246
Used Beads in Model:	98
Molecular Mass:	1.3681e+04 Da
Part. Specif. Volume:	0.710 cm^3/g
s(20,w):	1.93e+00 S
D(20,w), transl.:	1.18e-06 cm/sec^2
Stokes Radius:	1.81e+00 nm
Radius of Gyration:	1.48e+00 nm
Relaxation Time, tau(h):	7.81e+00 ns
Intrinsic Viscosity:	3.24e+00 cm^3/g

Load ASA Results File | View Bead Model File

Load Full Hydrodynamics Results File

Help | Close

Hydrodynamic Calculations:

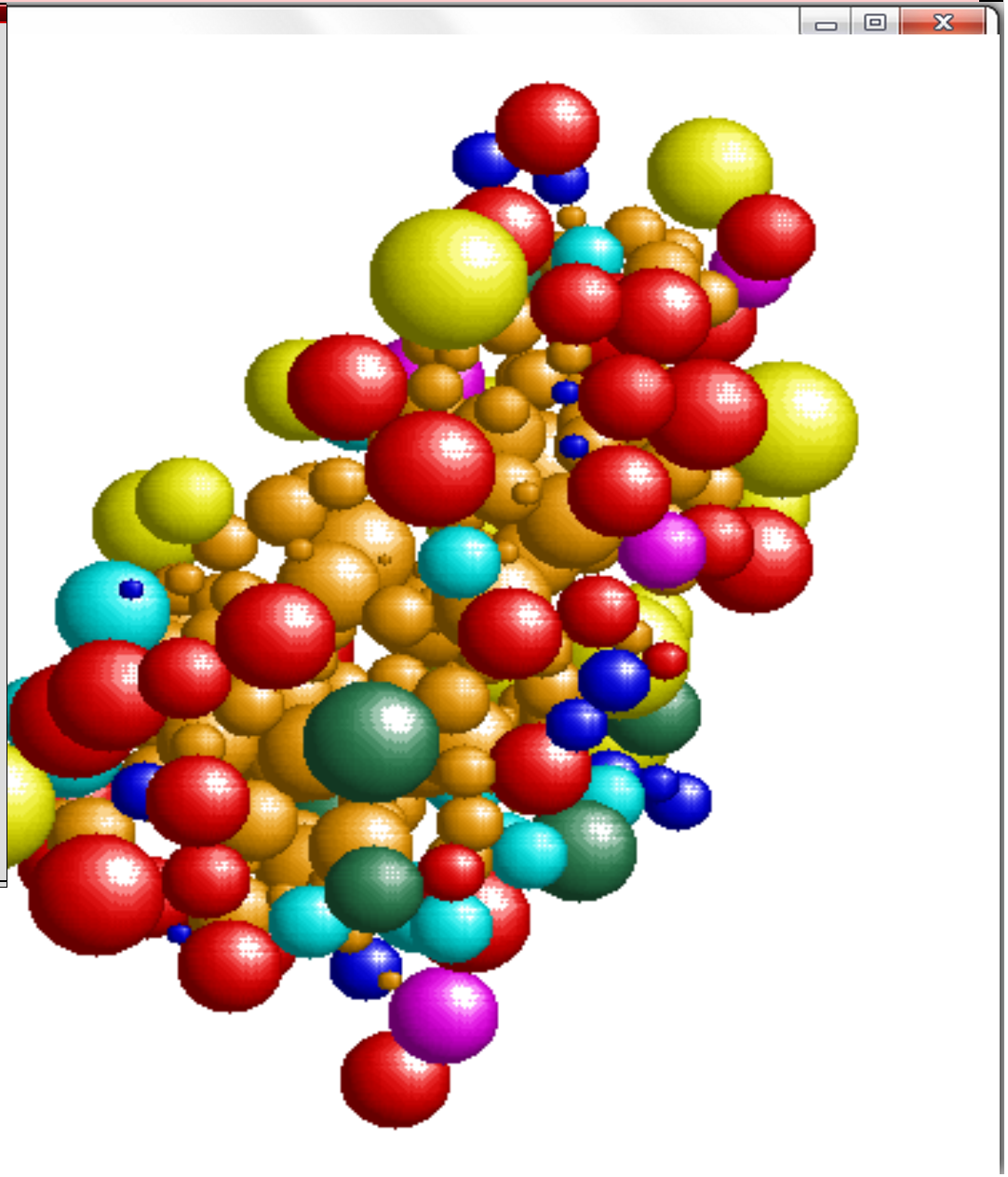
Calculate RB Hydrodynamics SMI | Calculate RB Hydrod

Show Hydrodynamic Calculations | Open Hydrodynamic C

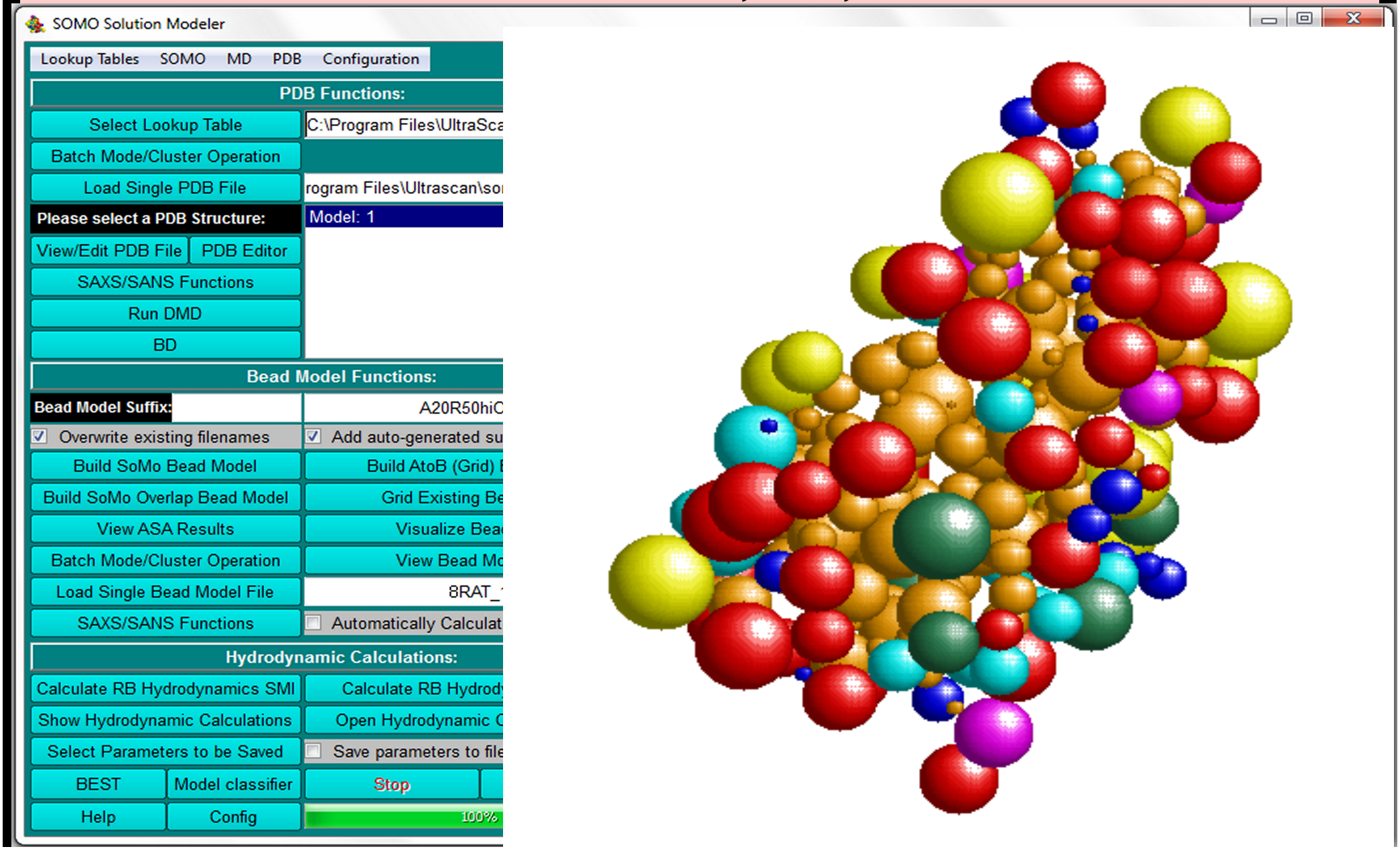
Select Parameters to be Saved | Save parameters to file

BEST | Model classifier | Stop

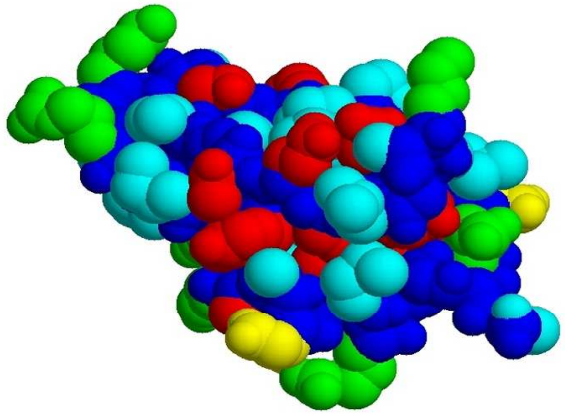
Help | Config | 100%



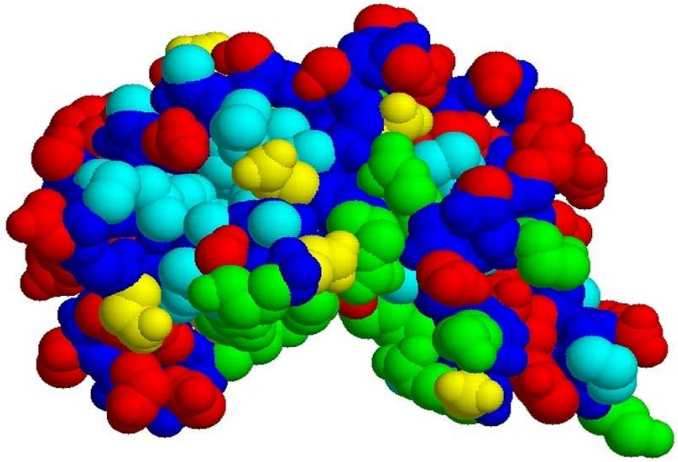
Layout and features of US-SOMO, integrated in UltraScan and available for Linux, Win, and Mac



PDB

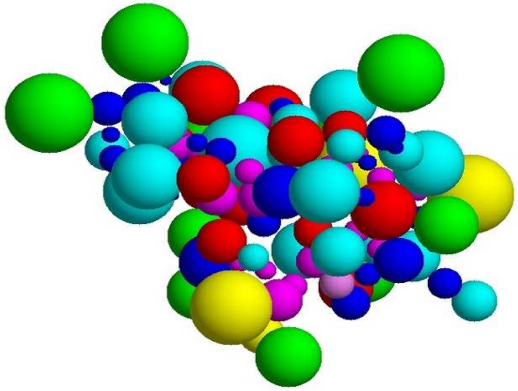


4PTI

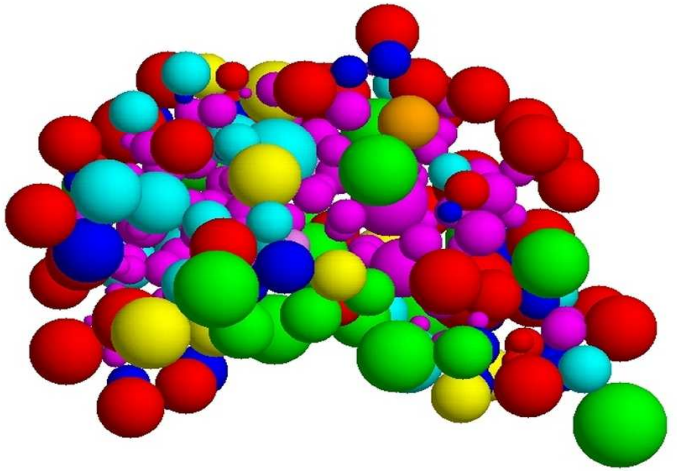


1XPS

SOMO



BPTI



RNase A

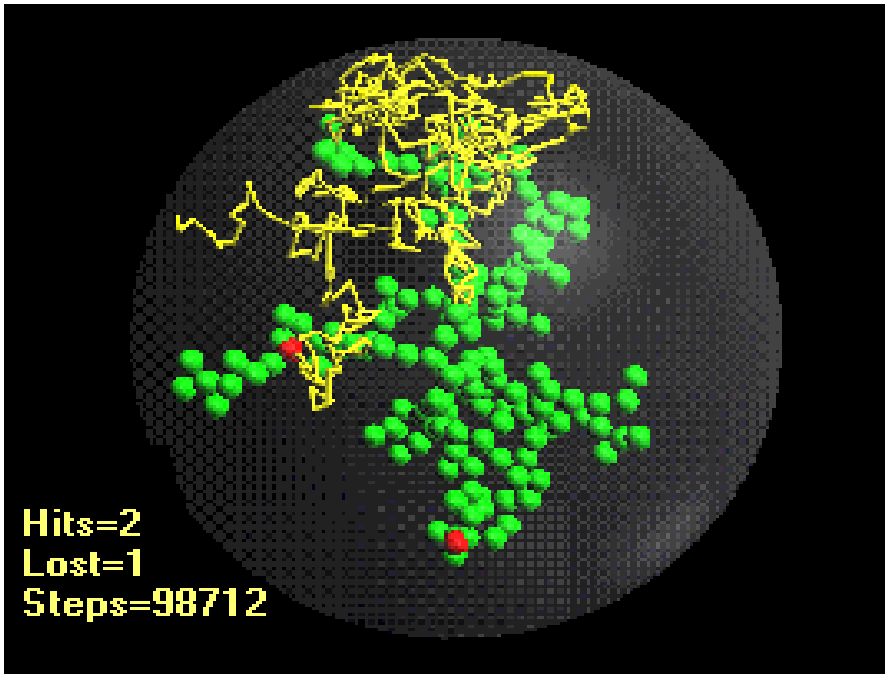
Other hydrodynamic computations methods

ZENO – Developed by Mansfield, Kang and Douglas (Stevens Institute and NIST)

BEST – Developed by S. Aragon, SFSU, CA

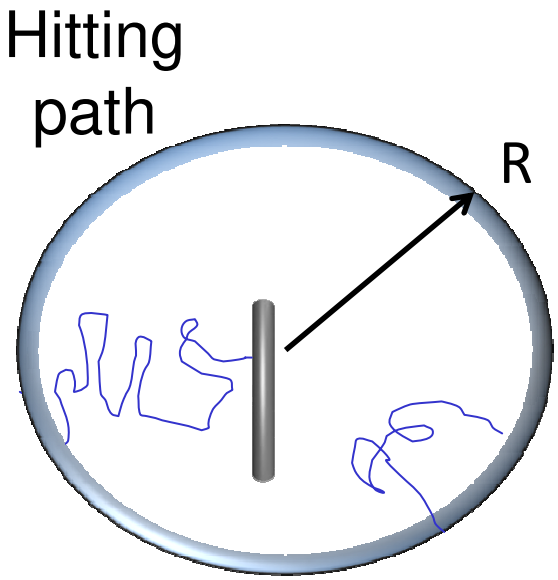
Scientific Principle of Program:

The Zeno computational method involves enclosing an arbitrary-shaped probed object within a sphere and launching random walks from this sphere. The probing trajectories either hit or return to the launch surface ('loss') as shown in the figure for a model soot particle aggregate, whereupon the trajectory is either terminated or reinitiated.



The fraction of random walk trajectories that hit the probed object determines its capacity C (hydrodynamic radius) and the electric polarizibility tensor α and [η] are estimated similarly.

ZENO: A Monte Carlo Numerical Path Integrator



$$\beta = \frac{\#\,hits}{\#\,attempts}$$

$$R_H = \beta R$$

$$fr = 6\pi\eta R_H$$

$$D = \frac{kT}{6\pi\eta R_H}$$

ZENO computes:

- electric Polarizability Tensor,
- self-capacity,
- intrinsic conductivity
- Intrinsic viscosity
- hydrodynamic radius
- translational diffusion coefficient
- translational friction coefficient
- radius of gyration
- structure factor...

of **arbitrarily shaped objects**

Douglas & Zhou & Hubbard, *PRE*, Vol 49, Page 5319, (1994).

Douglas & Garboczi, *Adv. Chem. Phys*, Vol 91, Pages 85-153,(1995).

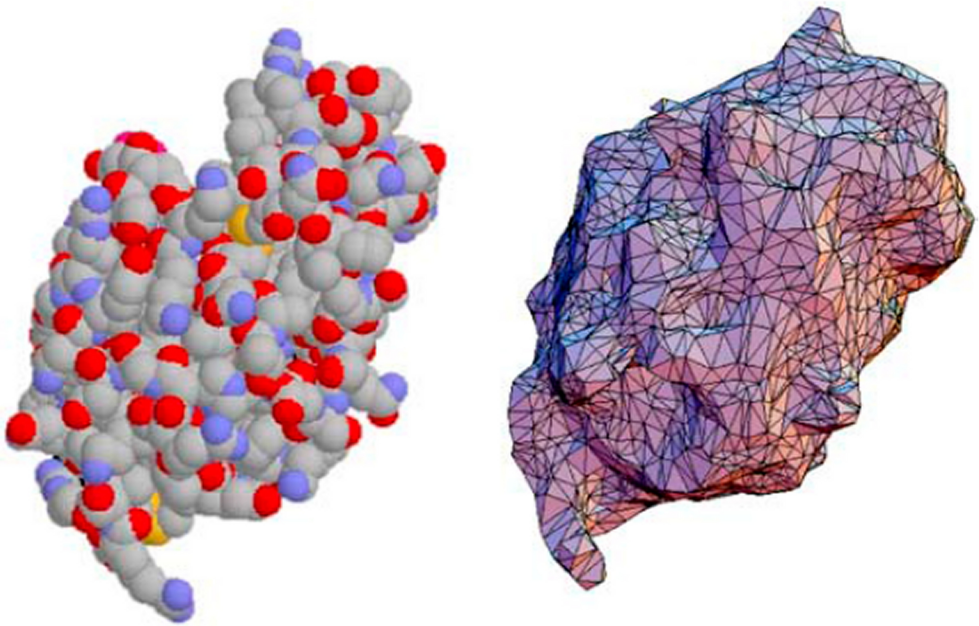
Mansfield et al., *PRE*, Vol 53, Vol 64, Pages 061401-16, 2001

Mansfield & Douglas, *PRE,* Vol 81, 021803 (2010)

http://web.stevens.edu/zeno/

NIST

BEST principles



BEST uses a very precise boundary element numerical solution of the exact formulation of the hydrodynamic resistance problem with stick boundary conditions to compute the full transport tensors in the center of resistance or the center of diffusion for an arbitrarily shaped rigid body, including rotation-translation coupling.

The input for BEST is a triangulation of the solvent-defined surface of the molecule of interest, given by Connolly's MSROLL. The triangulation is prepared for BEST by COALESCE, a program that allows user control over the quality and number of triangles to describe the surface. The computations are repeated for a series of triangulated structures with different number of plates, and extrapolated to zero plate size.

How to reliably test the methods?

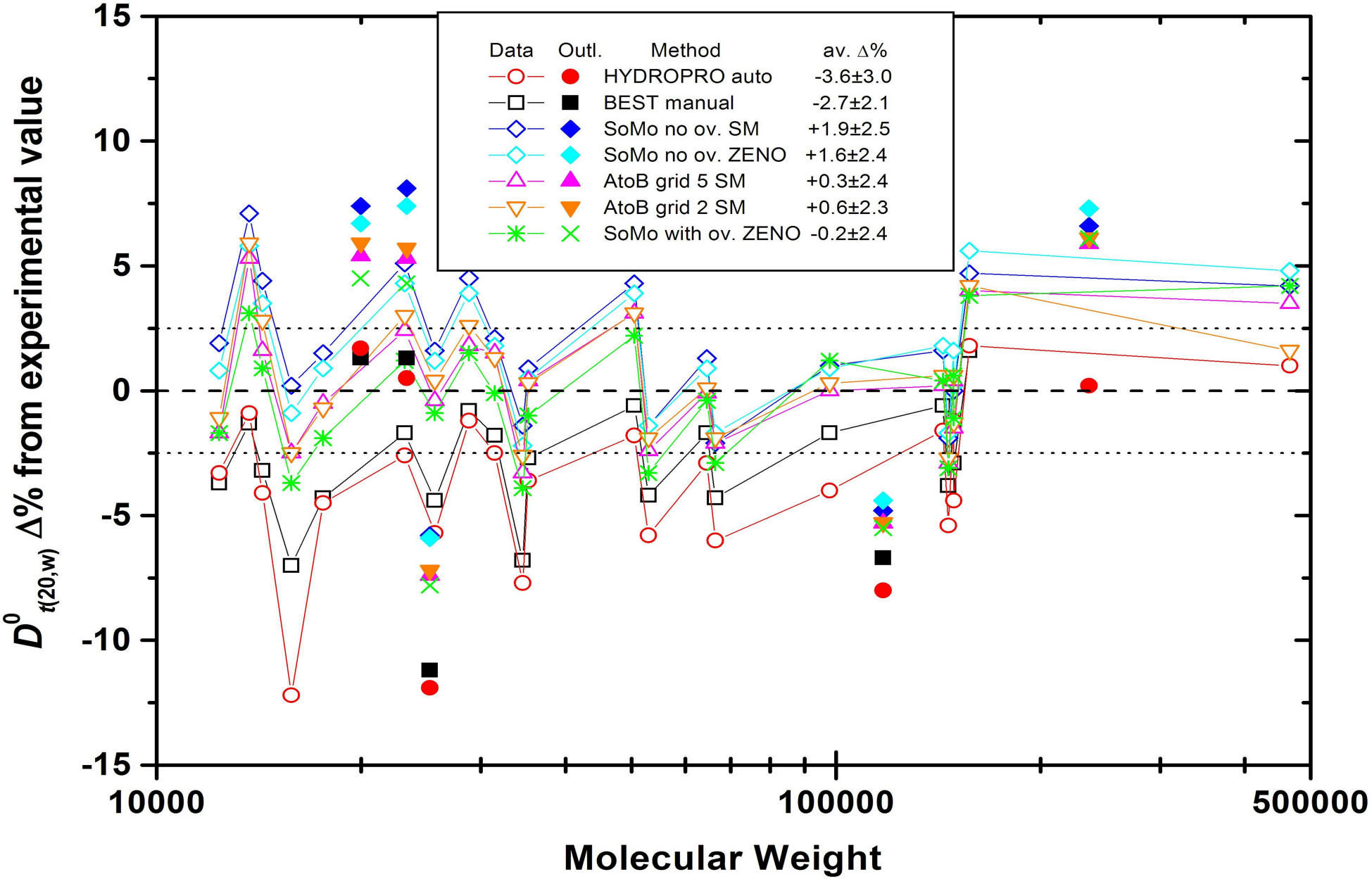
- We chose to investigate in depth only the translational friction data, namely $D^0_{t(20,w)}$ and $s^0_{(20,w)}$.

- We have collected literature data, and carefully analyzed them for proper extrapolation and reduction to standard conditions.
- We selected PDB crystal structures from the same species as the experimental data.
- We performed the computations on those structures using SoMo, SoMo+Zeno, AtoB with two different grid sizes, BEST (all using the US-SOMO interface), and HYDROPRO (externally).

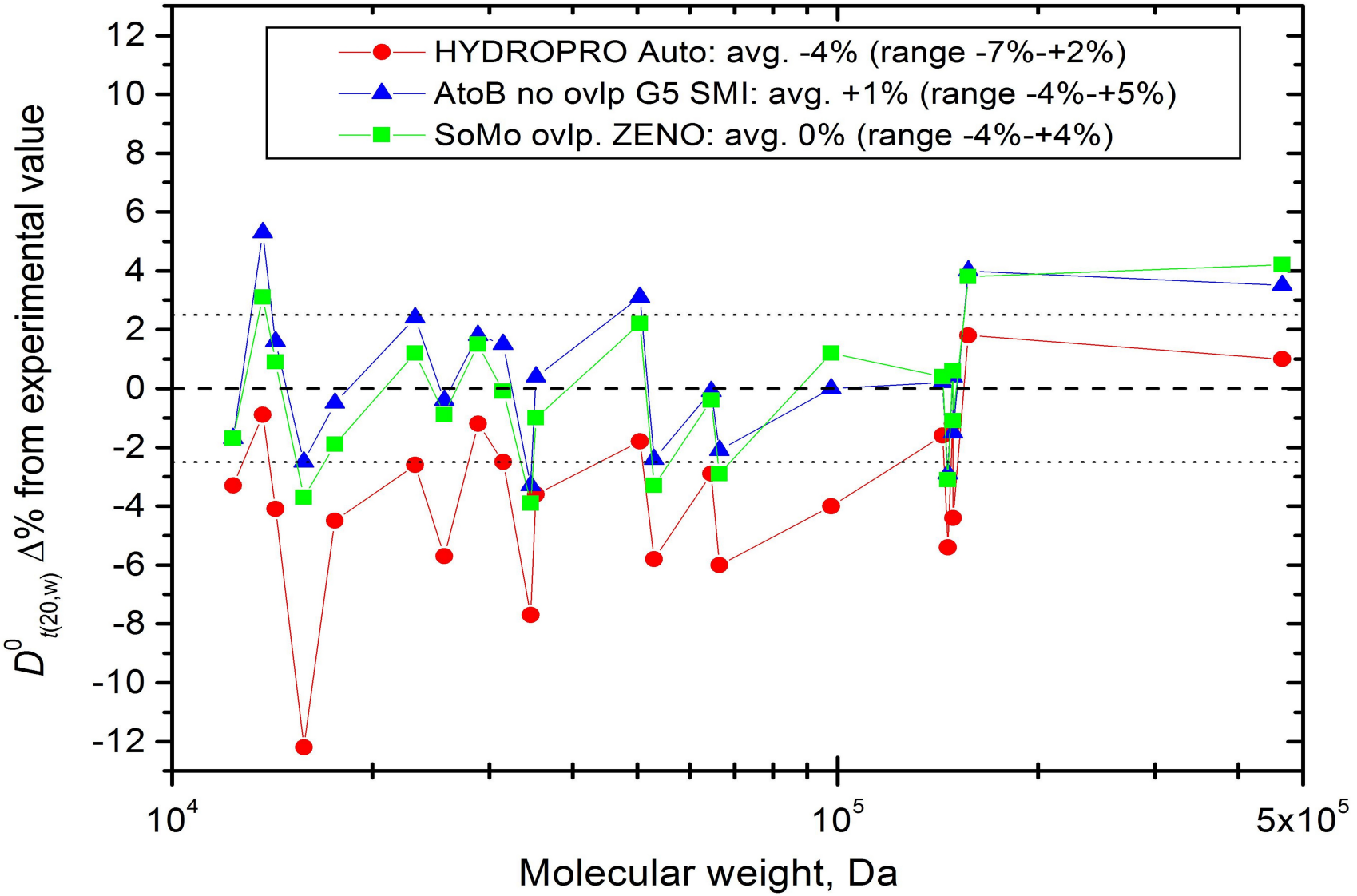
Test proteins used

#	Monomeric proteins	MW	#	Multimeric proteins	MW
1	Cytochrome c (1HRC)	12357.5	14	Superoxide dismutase (2SOD)	31442.2
2	Ribonuclease A (8RAT)	13683.8	15	β-Lactoglobulin (1BEB)	35224.7
3	α-Lactalbumin (1A4V+carb)	15784.7	16	α-Chymotrypsin (4CHA)	50473.5
4	Lysozyme (1AKI)	14306.7	17	Triosephosph. Isom. (1YPI)	52971.4
5	Myoglobin horse CO (1DWR)	17568.3	18	Hemoglobin CO (1HCO)	64559.7
6	Soybean Trypsin Inh. (1AVU)	19962.8	19	Citrate Synthase (1CTS)	97845.5
7	β-Trypsin (1TPO)	23335.9	20	Inorganic Pyrophosph. (1FAJ)	117339.0
8	Trypsinogen (1TGN)	23182.7	21	G3PD apo (2GD1)	143787.8
9	α-Chymotrypsin (4CHA)	25236.5	22	G3PD holo (1GD1)	146437.7
10	Chymotrypsinogen A (2CGA)	25659.0	23	LDH pig H + NAD (5LDH)	148942.6
11	Carbonic Anhydr. B (2CAB)	28820.5	24	LDH pig M + NAD (9LDH)	149063.5
12	Pepsin (4PEP)	34588.6	25	Aldolase (1ADO)	157136.0
13	H. Serum Albumin (1AO6)	66428.6	26	Catalase (4BLC)	235782.0
			27	β-Galactosidase (1BGL)	465557.0

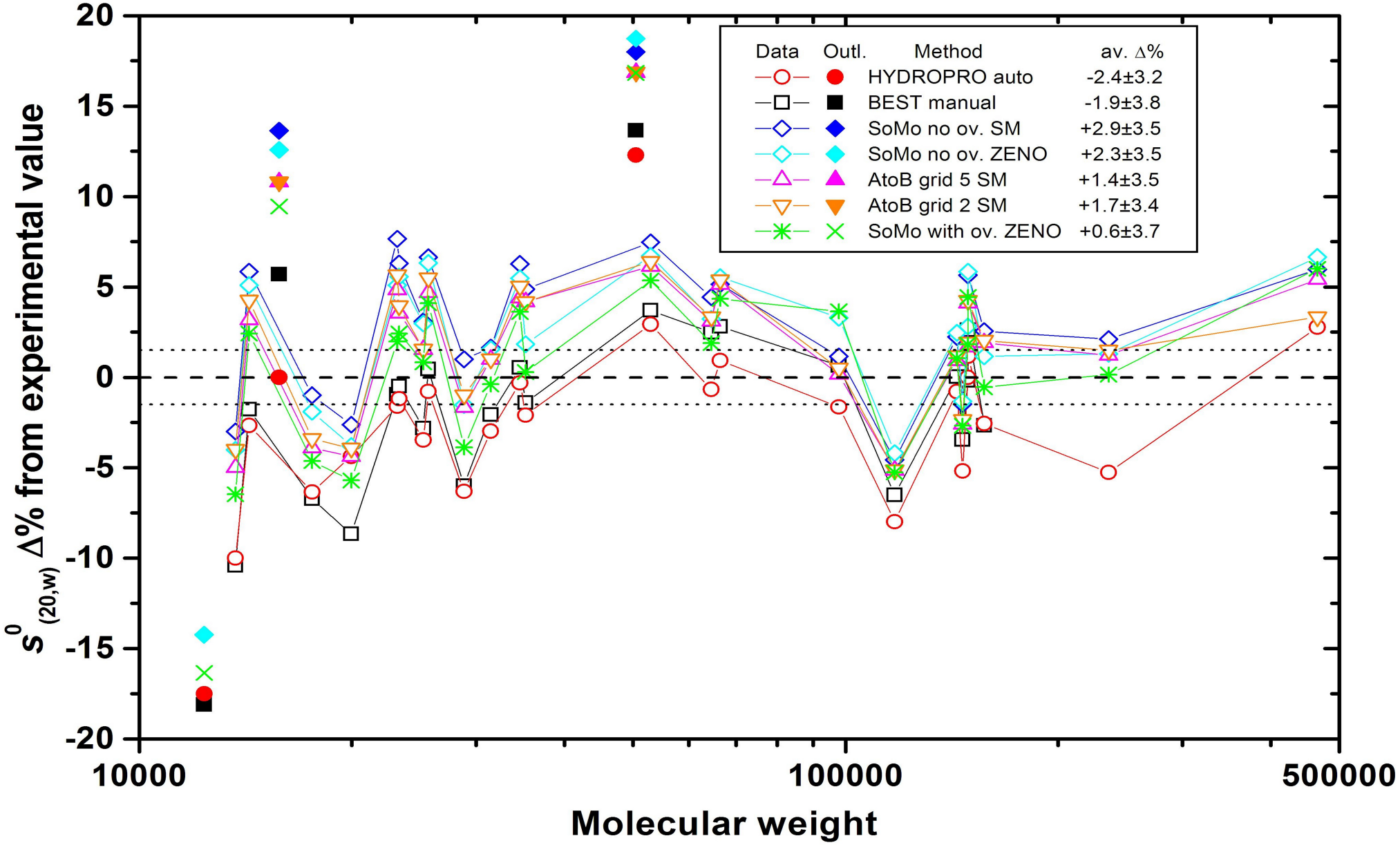
Overall performance of the different hydrodynamic modeling methods: D_t



Comparison between hydrodynamic calculations methods using 21 test proteins: translational diffusion coefficient



Overall performance of the different hydrodynamic modeling methods: *s*



Comparison conclusions:

- D_t is always better matched than s. This is likely due to poor psv knowledge/estimation.
- HYDROPRO and BEST both underestimate D_t and s. This is likely due to an excessive expansion of the surface in an attempt to account for hydration.
- SoMo with overlap removal overestimates D_t and s. This is likely due to an excessive shrinkage of the hydrated beads notwithstanding the outward translation.
- AtoB with a 5 Å grid appears to produce reasonable hydrated surfaces leading to very good D_t matching.
- The combination of SoMo models without overlap removal and Zeno computations produces the best D_t matching.