

Multiscale modeling of biomolecular systems: in serial and in parallel

Gary S Ayton, Will G Noid and Gregory A Voth

Considerable progress has been recently achieved in the multiscale modeling of complex biological processes. Multiscale models have now investigated the structure and dynamics of lipid membranes, proteins, peptides and DNA over length and time scales ranging from the atomic to the macroscopic. Serial multiscale methods that parameterize low-resolution coarse-grained models with data from high-resolution models have studied long time or length scale phenomena that cannot be investigated with atomically detailed models. Parallel multiscale methods that directly couple high- and low-resolution models have efficiently explored slow structural transitions and the importance of long-wavelength fluctuations for biological molecules. The success of such models relies upon new theories and methods for constructing accurate multiscale bridges that transfer information between models with different resolutions.

Addresses

Center for Biological Modeling and Simulation, University of Utah, 315 South 1400 East, Room 2020, Salt Lake City, UT 84112-0850, USA

Corresponding author: Voth, Gregory A
(voth@chemistry.chem.utah.edu)

Current Opinion in Structural Biology 2007, 17:192–198

This review comes from a themed issue on
Theory and simulation
Edited by Richard Lavery and Kim A Sharp

Available online 23rd March 2007

0959-440X/\$ – see front matter

© 2006 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.sbi.2007.03.004

Introduction

Many fundamentally important processes in biology are inherently multiscale. Biological processes (e.g. protein folding, nucleic acid packaging and membrane remodeling) that evolve on mesoscopic to nearly macroscopic length and time scales are intimately coupled to atomic and/or molecular level dynamics (e.g. fluctuations in side-chain conformation or lipid diffusion). Consequently, it is not surprising that many diverse computational methodologies have been developed for modeling biological processes with varying degrees of resolution. Atomically detailed modeling techniques (e.g. molecular dynamics [MD] [1,2]) remain a powerful tool for investigating biological structure and dynamics over nanosecond time and nanometer length scales, with femtosecond and Angstrom-level resolution. However, low-resolution

coarse-grained (CG) models provide the capability for investigating the longer time and length scale dynamics that are critical to many biological processes. CG models have now been developed for investigating lipid membranes [3[•],4–13,14[•],15[•],16,17[•],18[•]], proteins [19[•],20–35,36[•],27–39,40[•],41–43,44[•],45,46,47[•],48–50,51[•],52–56,57[•],58], peptides [59[•]–61[•],62–65], DNA [66[•],67–70,71[•],72[•],73]) and even the ribosome [74–76].

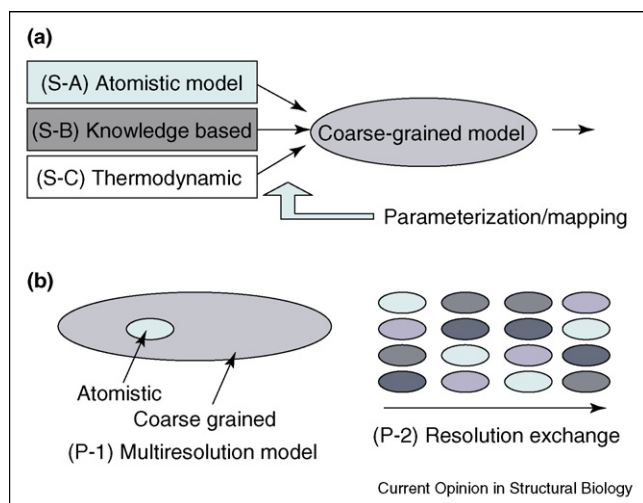
Coupling the CG and atomistic-level systems involves some degree of bridging of information across various length and time scales, the end goal ultimately being to integrate the different resolutions of the system into a single, unified, multiscale simulation methodology. The development of new theories and computational methodologies for connecting the disparate spatial and temporal scales relevant to cellular processes remains, arguably, one of the most significant challenges for the modeling of complex biological phenomena. As such, the aim of this review is to examine the multiscale methods currently employed to model various biological systems. For additional surveys of various CG models for biosystems, see [3[•]] for membranes and [19[•]] for proteins.

Serial and parallel multiscale simulation

The multiscale methods currently used to examine complex biomolecular systems can be roughly categorized according to the means by which information is transferred between different resolution models, ranging from atomistic, to CG and to even higher scales. This information transfer can proceed by either serial or parallel mechanisms, as illustrated in Figure 1.

In serial multiscale approaches, the different resolution models are employed in sequence. There is no direct interaction between atomistic-level molecules and CG particles. Serial multiscale approaches can be further classified according to the rigor characterizing this information transfer (Figure 1a). A ‘type S-A’ serial multiscale approach attempts to rigorously employ atomistic-level information to develop the reduced-resolution model. Snapshots of this type of approach are shown in Figure 2a, whereby a peptide is coarse grained at various levels. A ‘type S-B’ serial multiscale approach employs atomistic data obtained from various sources to assist directly in the parameterization. A ‘type S-C’ approach provides the least quantitative multiscale bridge and usually takes the form of a ‘top-down’ approach, in which, for example, desired thermodynamic data motivate the functional form and/or parameterization of the reduced-resolution model.

Figure 1



Schematic of the serial and parallel multiscale simulation decomposition for biomolecular systems. **(a)** A serial multiscale methodology in which different types of initial parameterizations are used to develop a CG model. Three different types of initial information can be used. A type S-A serial multiscale scheme has a systematic multiscale coupling between atomistic and CG representations. A type S-B serial multiscale scheme employs more general atomistic structural information, whereas type S-C employs thermodynamic and/or other top-down approaches to bridge the different scales. The type S-A approach gives the 'strongest' serial multiscale bridge and type S-C gives the weakest. **(b)** A parallel multiscale simulation. A type P-1 approach mixes different resolutions in one model, whereas type P-2 employs resolution exchange between concurrently running simulations.

By contrast, in parallel multiscale methodologies, all the different representations of the system are modeled concurrently and a direct information transfer couples different resolution models. In some ways, such methods are considerably more difficult to implement with the same level of rigor that serial multiscale methods approach. Figure 1b illustrates this scenario. As in the serial case, parallel multiscale modeling approaches may also be classified according to the mechanism by which the different resolution components interact. In this case, the parallel scheme can be implemented using a 'type P-1' or 'type P-2' methodology. The designation type P-1 and P-2 is employed here as opposed to the letter designations used in the serial case to emphasize that the two schemes are on a par, in terms of their respective multiscale character. Type P-1 parallel methods (see e.g. [66^{••},67]) are the analog of quantum mechanics/molecular mechanics (QM/MM) methods and combine atomically detailed models of a given subsystem of interest with a CG representation of the relevant environment. This situation is depicted in Figure 2b, whereby a type P-1 multiscale simulation of a transmembrane protein is shown. In this case, the coupling is continuous in time, and the atomistic and CG components of the system directly interact. The type P-2 parallel multiscale approach employs a resolution exchange methodology in which different representations

of a system evolve concurrently; however, after discrete time intervals, an exchange is attempted in which the resolutions describing different representations of the given system are swapped (see [59^{••}–61^{••}] for examples of this scheme). Here, high- and low-resolution models do not directly interact, but rather configurations swap resolution in a process that is analogous to parallel replica exchange.

Multiscale simulation: in serial

Lipid membranes

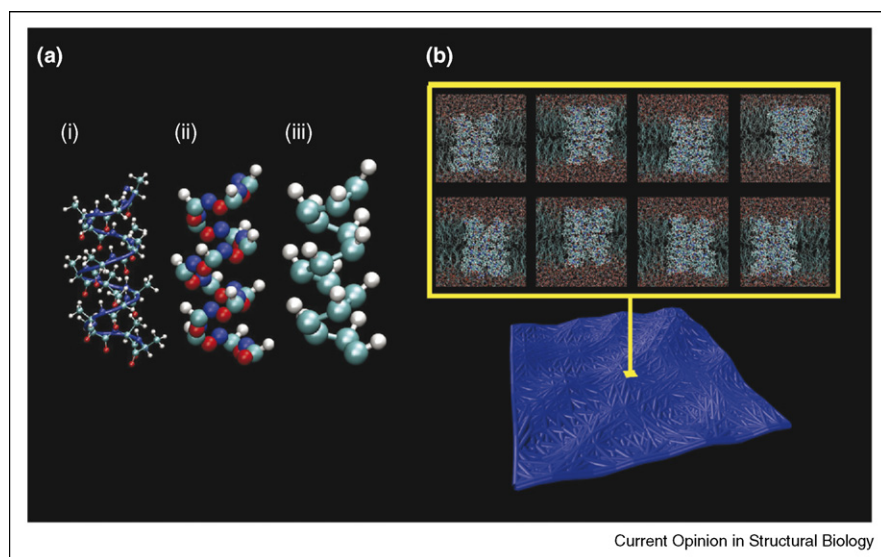
Lipid bilayers are critical to many biological phenomena and have been modeled using both atomistic-level MD simulations [2] and CG methods [3^{••},4–13,14[•],15[•],16]. In a serial multiscale sense, several recent CG lipid models deserve attention. The Marrink model [8], in particular, spontaneously formed stable bilayers for small systems and vesicles in larger systems [9], and can incorporate non-lipids (e.g. cholesterol) [10]. The multiscale connection between low- and high-resolution models using this type of approach relies on the success of the top-down 'building block' nature of the scheme. The reliance on thermodynamic information to parameterize the model makes it a type S-C serial multiscale approach (c.f. Figure 1). Similar type S-C implicit, or 'solvent-free', approaches have also been proposed (e.g. [11–13,14[•]]; [3^{••}] and additional references therein). The solvent-free approach is quite appealing because the computational cost of modeling a solvated bilayer is tremendously reduced if the effects of an aqueous solvent can be incorporated into the CG lipid interaction.

A solvent-free model using the reverse Monte Carlo method has also been developed recently [14[•]]; this particular model can, in fact, also be treated as a type S-A serial multiscale approach. Multiscale coarse-graining (MS-CG) [15[•],16] has also been employed to model pure bilayers [15[•]] and mixed lipid–cholesterol bilayers [16]. In the MS-CG method, CG force fields are systematically derived from atomistic MD simulations using a statistical implementation of the 'force matching' (FM) method [77]. The MS-CG link between the atomistic and CG models provides a robust type S-A serial multiscale bridge and accounts for three-body effects [78^{••}]. It also provides a possible route to making the dynamics of CG models more realistic [79].

Peptides and proteins

CG models of peptides and proteins have a long and distinguished history, since the seminal work of Levitt and Warshel [20], and Levitt [21]. The field of CG protein simulation as of 2005 has recently been surveyed in the review by Tozzini [19^{••}]. The original concept of using knowledge-based potentials [22], in combination with the quasi-chemical approximation [23], and employing potentials of mean force [24], has provided a framework

Figure 2



Two examples of a biomolecular multiscale simulation. **(a)** Snapshots of a type S-A serial multiscale simulation (c.f. Figure 1) employing MS-CG to examine a polyalanine pentadecamer [65] as the underlying atomistic-level template **(i)**. Two MS-CG schemes were employed: a three bead per backbone model **(ii)** and a coarser one bead per backbone model **(iii)**. For the one-bead model, the backbone (-NH-CH-CO-) is treated as a single CG site; the sidechain (-CH₃) is treated as another CG site. The higher resolution three bead per backbone peptide model treats the backbone groups of -NH-, -CH- and -CO- of each residue as CG sites. **(b)** Example of a type P-1 parallel multiscale simulation, in which an MSC simulation of a transmembrane protein is considered [17*]. An ensemble of eight different atomistic-level simulations of the same system (in this case, an influenza A M2 proton channel in a dimyristoylphosphatidylcholine lipid bilayer) is coupled to a corresponding mesoscopic membrane/solvent system [81]. The small light 'patch' region on the mesoscopic membrane at the bottom gives the relative length scales of the atomistic and mesoscopic systems.

for type S-B serial multiscale simulation methods for proteins.

Recent work has been aimed at developing more detailed orientation-dependent residue-residue interactions [25,62] and also including many-body interactions in CG force fields, such as with the united-residue (UNRES) force field [26,27]. It should be noted that the determination of pair interaction potentials from pair correlations within structural databases using approximations such as the quasi-chemical approximation [24] might be problematic, especially when many-body effects and multiparticle correlations are important [28,29]. Such approximations do not provide an exact multiscale link between atomistic and CG representations because the three-particle correlations are not directly considered. By contrast, the UNRES model developed by Scheraga and co-workers attempts to incorporate these effects by approximately evaluating the restricted free energy [30]. Additionally, some CG protein models incorporate direct experimental data using the reverse Monte Carlo method [31]; this type of methodology is thus a type S-A serial multiscale methodology. Likewise, MS-CG simulations of peptides have been able to systematically derive CG peptide interactions from underlying MD models at various CG resolutions [65]; these studies fall within the type S-A methodology. As seen from the

simulation snapshots in Figure 2a, the MS-CG approach can be applied at various resolutions. Within the serial multiscale approach, this particular aspect of the MS-CG approach is quite appealing.

Interaction potentials for CG proteins have also been determined using the 'consistency principle', also referred to as the 'principle of minimal frustration' [32,33], to optimize a funnel landscape for protein folding [34,35]. As an extreme example of such funnel-based potentials, Gō models consider only native contacts as favorable, providing a perfectly smooth landscape [33]. Protein folding simulations employing a Gō model in combination with discrete MD [64] have recently been used to investigate the transition state ensemble of the Src SH3 protein domain [36*]. Importance sampling techniques combined with MD identified an ensemble of atomically detailed structures near the folding transition state. CG representations of these structures were generated and their folding investigated using Gō-type interactions. This approach is quite promising and can be cast as a type S-B serial multiscale method.

Elastic network models (ENMs) also provide a CG protein model that has proved effective in structural biology [37–39]. This approach has been used in conjunction with a normal mode analysis (NMA)

[40[•],41–43,44[•],45]. (It should be noted that an NMA, when combined with, for example, cryo-electron microscopy (cryo-EM) low-resolution structures [44[•],45], is in itself a type S-B multiscale modeling approach.)

The fundamental assumption of these models is that the biologically significant fluctuations of folded proteins are low-frequency collective modes, which behave much like an elastic material. This behavior persists even for various levels of CG [46]. A variable-resolution ENM [47[•]] has been developed that models certain parts of the protein with atomistic-level detail (albeit with an ENM), whereas other parts are modeled with one site per residue. However, even though parts of the protein are concurrently described at different resolutions, because the entire system is modeled within an ENM, this approach is more of a type S-B serial approach, rather than a parallel multiscale methodology.

The MS-CG method has also been applied to systematically derive a CG model for actin [57[•]], employing large-scale atomistic-level MD simulation [58]. In this case, a fluctuation-matching MS-CG approach was employed and, as such, this method represents a type S-A multiscale approach. The pronounced collective modes in actin filaments enabled the fluctuation-matching MS-CG method to be implemented directly; other systems that exhibit similar robust collective motions might also be modeled using such an approach.

DNA

Serial multiscale simulations have also been employed to investigate the packing of DNA into viral capsids, employing a low-resolution DNA model in which each bead corresponded to one turn of a double helix [68,69]. Because atomistic-level information was not employed in parameterizing the CG model, this method provides a type S-C serial multiscale approach; however, in principle, it should be possible to parameterize a CG DNA model using atomistic MD simulations and a fluctuation-matching approach [70]. This type of serial multiscale approach for DNA could potentially lead to type S-B serial multiscale simulations of nucleosomal array folding [71[•]]. Another example of a type S-B approach for modeling mesoscopic DNA fragments examined the collective low-frequency motions in a mesoscopic closed circular DNA molecule using an NMA that employed an initial energetic model at the base-pair level [72[•],73]. It might be possible to extend this approach by employing the reverse Monte Carlo method to model the protein–DNA interactions, as previously applied to model ion–DNA interactions [80].

Multiscale simulation: in parallel

A promising technique called ‘multiscale coupling’ (MSC) for type P-1 parallel simulations ‘embeds’ an atomistic MD simulation of a bilayer (with perhaps

non-lipid molecules, such as membrane-bound proteins) within a mesoscopic membrane/solvent model [17[•]]; simulation snapshots of this approach are shown in Figure 2b. The MS-CG methodology has also been applied to determine the effective interactions between atomistic and CG representations. Using this mixed all-atom (AA) MS-CG approach (denoted AA-CG) [18^{••}], simulations have been performed on an AA model of the gramicidin channel solvated within a MS-CG lipid bilayer [15[•]]. Moreover, because the interactions (both lipid–lipid and protein–lipid) were determined from the MS-CG methodology [77], the structures generated within the AA-CG simulation, in principle, occur according to the correct probability distribution.

Parallel multiscale simulations have also been applied to study protein function. The low-frequency motions determined by an ENM have been employed to ‘guide’ atomistic-level simulations [48]. An exciting type P-1 parallel multiscale simulation of a protein and DNA has recently been performed by Schulten and co-workers [66^{••},67]. In these multiscale investigations of a DNA–protein complex, the LacI complex has been modeled in atomic detail and coupled to a continuum model for a 75 bp DNA loop. The AA protein component (comprising over 250 000 atoms) was coupled to the continuum DNA model through elastic stresses and torques arising from the looping of the DNA model.

A ‘pseudo’ type P-1 parallel multiscale simulation has employed a hybrid fully atomistic/Gō model [49] to examine the folding of a 80-residue fragment of the λ -repressor [50]. This work draws from previous studies in which an AA Monte Carlo simulation used a Gō potential [50]. In these simulations, a Gō-type interaction is superimposed upon the atomistic-level force field, thus directing the system towards the native state. The multiscale bridge between CG and atomistic models occurs by incorporating a long wavelength ‘nudge’ towards the folded state resulting from the inclusion of the Gō interaction, which favors native contacts. Bridging mesoscopic and atomistic models of globular proteins has relied on specifying an interface region between the high-resolution (MD) and low-resolution (a Gō model employing only the C α carbons) representations [51[•]]. The overall success of this approach relies on systematically determining the interactions in the interface region; this scheme gives a ‘loose’ type P-1 methodology. Specifically, a more systematic means of incorporating the surrounding solvent is required.

Type P-2 parallel multiscale methodologies that exchange resolution between different representations of a given system [59^{••},60^{••}] have also been developed. As an example, the work of Liu and Voth [61^{••}] combines the MS-CG framework and resolution exchange. The MS-CG methodology interaction potential is

employed systematically to derive the low-resolution model from an AA model in a manner similar to that previously done for peptides [65]. In principle, this type of parallel multiscale scheme could also be applied to develop intermediate-resolution CG models using fluctuation matching [57], opening up the possibility of a fully parallel, multiscale simulation methodology capable of spanning quite large spatial and temporal scales. A serious issue with all parallel multiscale methodologies is the question of realistic dynamics, as well as the dynamical consistency of multiscale bridging across time scales.

Conclusions and outlook

This review classifies various approaches for modeling biomolecular processes within the context of an overall multiscale simulation perspective. Multiscale methodologies have been classified into two distinct categories: serial and parallel approaches. This classification of these methodologies into serial and parallel approaches facilitates an examination of these methods and the systems for which they have been employed. One observation arising from this review is fairly clear: both serial and parallel multiscale schemes provide increasingly valuable insight into the structure and dynamics of complex biomolecular processes.

A large number of serial multiscale approaches for lipids, proteins, peptides and DNA have already been developed. From this review, it might be observed that most modeling schemes fall into the type S-B approach. It seems clear that more systematic and direct links between the atomistic and reduced-resolution models must be defined in the future.

Parallel multiscale simulation methods can be divided into mixed resolution (type P-1) and resolution exchange (type P-2) methods. The combination of the type P-1 and P-2 parallel approaches can potentially become an increasingly robust multiscale simulation methodology for complex biosystems; current work in this area appears to be promising.

Acknowledgements

This research was supported by the National Institutes of Health (GM063796), the National Science Foundation Information Technology Research program (CHE-0218739) and Collaborative Research in Chemistry program (CHE-0628257). WGN acknowledges the support of the National Institutes of Health through a Ruth L. Kirschstein NRSA postdoctoral fellowship grant number 5F32GM076839-02. We thank Jihui Wei Chu for many invaluable discussions and assistance.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Karplus M, McCammon JA: **Molecular dynamics simulations of biomolecules.** *Nat Struct Biol* 2002, **9**:646-652.
 2. Feller SE: **Molecular dynamics simulations of lipid bilayers.** *Curr Opin Colloid Interface Sci* 2000, **5**:217-223.
 3. Brannigan G, Lin LCL, Brown FLH: **Implicit solvent simulation •• models for biomembranes.** *Eur Biophys J* 2006, **35**:104-124. In the 'solvent-free' approach, the hydrophobic/hydrophilic membrane-solvent interactions are effectively incorporated into a CG model that is stable in the absence of a surrounding solvent.
 4. Goetz R, Lipowsky R: **Computer simulations of bilayer membranes: self-assembly and interfacial tension.** *J Chem Phys* 1998, **108**:7397-7409.
 5. Stevens MJ: **Coarse-grained simulations of lipid bilayers.** *J Chem Phys* 2004, **121**:11942-11948.
 6. Shelley JC, Shelley MY, Reeder RC, Bandyopadhyay S, Klein ML: **A coarse grain model for phospholipid simulation.** *J Phys Chem B* 2001, **105**:4464-4470.
 7. Shelley JC, Shelley MY, Reeder RC, Bandyopadhyay S, Klein ML: **Simulation of phospholipids using a coarse grain model.** *J Phys Chem B* 2001, **105**:9792-9875.
 8. Marrink SJ, deVries AH, Mark AE: **Coarse grained model for semiquantitative lipid simulations.** *J Phys Chem B* 2004, **108**:750-760.
 9. Marrink SJ, Mark AE: **The mechanism of vesicle fusion as revealed by molecular dynamics simulations.** *J Am Chem Soc* 2003, **125**:11144-11145.
 10. Shih AY, Arkhipov A, Freddolino PL, Schulten K: **Coarse grained protein-lipid model with applications to lipoprotein particles.** *J Phys Chem B* 2006, **110**:3674-3684.
 11. Brannigan G, Brown FLH: **Solvent-free simulations of membrane bilayers.** *J Chem Phys* 2004, **120**:1059-1071.
 12. Cooke IR, Kremer K, Deserno M: **Tunable generic model of fluid bilayer membranes.** *Phys Rev E Stat Nonlin Soft Matter Phys* 2005, **72**:011506.
 13. Farago O: **'Water-free' computer model for fluid bilayer membranes.** *J Chem Phys* 2003, **119**:596-605.
 14. Lyubartsev AP: **Multiscale modeling of lipids and lipid bilayers.** • *Eur Biophys J* 2005, **35**:53-61. The reverse Monte Carlo algorithm is employed to determine an effective implicit solvent interaction potential for a CG lipid model. Simulations with this effective potential reproduced a stable bilayer, and also spontaneously formed bicelles and vesicles under appropriate conditions.
 15. Izvekov S, Voth GA: **A multiscale coarse-graining method •• for biomolecular systems.** *J Phys Chem B* 2005, **109**:2469-2473. It was shown how a biomolecular system (e.g. a lipid bilayer with surrounding solvent) can be systematically coarse grained using a statistical implementation of force matching.
 16. Izvekov S, Voth GA: **Multiscale coarse-graining of mixed phospholipid/cholesterol bilayers.** *J Chem Theor Comp* 2006, **2**:637-648.
 17. Ayton GS, Voth GA: **Multiscale simulation of transmembrane •• proteins.** *J Struct Biol* 2007, **157**:570-578. MSC occurs via modulated boundary conditions; as the mesoscopic membrane model 'pushes' on an ensemble of atomistic-level systems under periodic boundary conditions, the ensemble of atomistic-level systems can 'push back'. However, the atomistic-level system is never explicitly embedded within the mesoscopic model and in fact operates under the original periodic boundary conditions of the MD simulation. This method has revealed that membrane-bound proteins can couple to long-wavelength mesoscopic membrane undulation and/or plane strain modes.
 18. Shi Q, Izvekov S, Voth GA: **Mixed atomistic and coarse-grained •• molecular dynamics: simulation of a membrane bound ion channel.** *J Phys Chem B* 2006, **110**:15045-15048. The reduced resolution of the surrounding CG lipid environment accelerates the computation such that longer time scales could be reached.
 19. Tozzini V: **Coarse-grained models for proteins.** *Curr Opin Struct Biol* 2005, **15**:144-150. A review of approaches in the development of CG protein models, including elastic networks, Gō-like models and other approaches.

20. Levitt M, Warshel A: **Computer simulation of protein folding.** *Nature* 1975, **253**:694-698.
21. Levitt M: **A simplified representation of protein conformations for rapid simulation of protein folding.** *J Mol Biol* 1976, **104**:59-107.
22. Tanaka S, Scheraga HA: **Medium- and long-range interaction parameters between amino acids for predicting three-dimensional structures of proteins.** *Macromolecules* 1976, **9**:945-950.
23. Miyazawa S, Jernigan RL: **Estimation of effective interresidue contact energies from protein crystal structures: Quasi-chemical approximation.** *Macromolecules* 1985, **18**:534-552.
24. Sippl MJ: **Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins.** *J Mol Biol* 1990, **213**:859-883.
25. Buchete NV, Straub JE: **Anisotropic coarse-grained statistical potentials improve the ability to identify native-like protein structures.** *J Chem Phys* 2003, **118**:7658-7671.
26. Nianias M, Chinchio M, Oldziej S, Czaplowski C, Scheraga HA: **Protein structure prediction with the UNRES force-field using replica-exchange Monte Carlo-with-minimization; comparison with MCM, CSA, and CFMC.** *J Comput Chem* 2005, **26**:1472-1486.
27. Nianias M, Czaplowski C, Scheraga HA: **Replica exchange and multicanonical algorithms with the coarse-grained united-residue (UNRES) force field.** *J Chem Theory Comput* 2006, **2**:513-528.
28. Jernigan RL, Bahar I: **Structure-derived potential and protein simulations.** *Curr Opin Struct Biol* 1996, **6**:195-209.
29. Vendruscolo M, Domany E: **Pairwise contact potentials are unsuitable for protein folding.** *J Chem Phys* 1998, **109**:11101-11108.
30. Liwo A, Czaplowski C, Pillardy J, Scheraga HA: **Cumulant-based expressions for the multibody terms for the correlation between local and electrostatic interactions in the united-residue force field.** *J Chem Phys* 2001, **115**:2323-2347.
31. Matysiak S, Clementi C: **Optimal combination of theory and experiment for the characterization of the protein folding landscape of S6: how far can a minimalist model go?** *J Mol Biol* 2004, **343**:235-248.
32. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG: **Funnels, pathways, and the energy landscape of protein folding: a synthesis.** *Proteins* 1995, **21**:167-195.
33. Go N: **Theoretical studies of protein folding.** *Annu Rev Biophys Bioeng* 1983, **12**:183-210.
34. Kussell E, Shimada J, Shakhnovich EI: **A structure-based method for derivation of all-atom potentials for protein folding.** *Proc Natl Acad Sci USA* 2002, **99**:5343-5348.
35. Fujitsuka Y, Takada S, Luthey-Schulten ZA, Wolynes PG: **Optimizing physical energy functions for protein folding.** *Proteins* 2004, **54**:88-103.
36. Ding F, Guo W, Dokholyan NV, Shakhnovich EI, Shea JE: **Reconstruction of the Src-SH3 protein domain transition state ensemble using multiscale molecular dynamics simulations.** *J Mol Biol* 2005, **350**:1035-1050.
Structures belonging to the putative transition state ensemble of the Src SH3 protein domain, comprising a 56-residue β -barrel structure composed of two hydrophobic sheets, are examined with multiscale simulation. The folding free energy landscape of this domain was found from importance sampling MD. The protein was described in full atomistic detail using the CHARMM19 force field with TIP3P water. Unfolded structures were generated using high-temperature simulations. CG simulations of the Src SH3 domain employed a G ϕ potential to model interactions between different amino acids.
37. Tirion MM: **Large amplitude elastic motions in proteins from a single-parameter, atomic analysis.** *Phys Rev Lett* 1996, **77**:1905-1908.
38. Haliloglu T, Bahar I, Erman B: **Gaussian dynamics of folded proteins.** *Phys Rev Lett* 1997, **79**:3090-3093.
39. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I: **Anisotropy of fluctuation dynamics of proteins with an elastic network model.** *Biophys J* 2001, **80**:505-515.
40. Gohlke H, Thorpe MF: **A natural coarse graining for simulating large biomolecular motion.** *Biophys J* 2006, **91**:2115-2120.
This paper discusses a 'natural way' to select groups of atoms for coarse graining. The rigid regions of the biomolecule are employed using the 'floppy inclusions and rigid substructure topography' (FIRST) algorithm combined with the 'framework rigidity optimized dynamic algorithm' (FRODA). The 'rigid cluster normal mode analysis' (RCNMA) approach is also discussed, which combines rigidity with elastic network theory.
41. Wu Y, Ma J: **Refinement of F-actin model against fiber diffraction data by long-range normal modes.** *Biophys J* 2004, **86**:116-124.
42. Ming D, Kong Y, Wu Y, Ma J: **Substructure synthesis method for simulating large molecular complexes.** *Proc Natl Acad Sci USA* 2003, **100**:104-109.
43. Tirion M, ben-Avraham D, Lorenz M, Holmes KC: **Normal modes as refinement parameters for the F-actin model.** *Biophys J* 1995, **68**:5-12.
44. Falke S, Tama F, Brooks CL, Gogol EP, Fisher MT: **The 13 Å structure of a chaperonin GroEL-protein substrate complex by cryo-electron microscopy.** *J Mol Biol* 2005, **348**:219-230.
A C α carbon model was employed to give the normal modes for the GroEL chaperonin (an allosteric protein that binds to partially folded protein intermediates). All structures were found using cryo-EM at 13 Å.
45. Tama F, Miyashita O, Brooks CL: **Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis.** *J Mol Biol* 2004, **337**:985-999.
46. Sen TZ, Feng Y, Garcia JV, Kloczkowski A, Jernigan RL: **The extent of cooperativity of protein motions observed with elastic network models is similar for atomic and coarser-grained models.** *J Chem Theory Comput* 2006, **2**:696-704.
47. Kurkcuglu O, Jernigan RL, Doruker P: **Loop motions of triosephosphate isomerase observed with elastic networks.** *Biochemistry* 2006, **45**:1173-1182.
An atom-based ENM model is combined with a more coarse-grained model (i.e. one site per residue, with the C α atoms at the network nodes) to give a mixed CG elastic network model.
48. Zhang Z, Shi Y, Liu H: **Molecular dynamics simulations of peptides and proteins with amplified collective modes.** *Biophys J* 2003, **84**:3583-3593.
49. Go N, Taketomi H: **Respective roles of short- and long-range interactions in protein folding.** *Proc Natl Acad Sci USA* 1978, **75**:559-563.
50. Shimada J, Shakhnovich EI: **The ensemble folding kinetics of protein G from an all-atom Monte Carlo simulation.** *Proc Natl Acad Sci USA* 2002, **99**:11175-11180.
51. Neri M, Anselmi C, Cascella M, Maritan A, Carloni P: **Coarse-grained model of proteins incorporating atomistic detail of the active site.** *Phys Rev Lett* 2005, **95**:218102.
A methodology is presented whereby a CG model of a globular protein is coupled to an atomistically detailed MD biologically 'relevant' region of the protein. Only the CG centroid of the C α carbons is considered at the CG level. An interface region between the two domains bridges the two representations.
52. Brown S, Fawzi NJ, Head-Gordon T: **Coarse-grained sequences for protein folding and design.** *Proc Natl Acad Sci USA* 2003, **100**:10712-10717.
53. Zuckerman DM: **Simulation of an ensemble of conformational transitions in a united-residue model of calmodulin.** *J Phys Chem B* 2004, **108**:5127-5137.
54. Swigon D, Coleman BD, Olson WK: **Modeling the Lac repressor-operator assembly: the influence of DNA looping on Lac repressor conformation.** *Proc Natl Acad Sci USA* 2006, **103**:9879-9884.

55. Lu M, Poon B, Ma J: **A new method for coarse-grained elastic normal-mode analysis.** *J Chem Theory Comput* 2006, **2**:464-471.
56. Matysiak S, Clementi C: **Minimalist protein model as a diagnostic tool for misfolding and aggregation.** *J Mol Biol* 2006, **363**:297-308.
57. Chu JW, Voth GA: **Coarse-grained modeling of the actin filament derived from atomistic-scale simulations.** *Biophys J* 2006, **90**:1572-1582.
A multiscale CG procedure that employs fluctuation matching from detailed atomistic-level simulations is applied to actin filaments. The method is based on matching the fluctuations of the effective internal coordinates that describe the CG model with those extracted from AA-MD simulation. These fluctuations are calculated in a CG model using an NMA.
58. Chu JW, Voth GA: **Allostery of actin filaments: molecular dynamics simulations and coarse-grained analysis.** *Proc Natl Acad Sci USA* 2005, **102**:13111-13116.
59. Lyman E, Ytreberg FM, Zuckerman DM: **Resolution exchange simulation.** *Phys Rev Lett* 2006, **96**:028105.
This paper describes a resolution exchange simulation in which two independent simulations of a protein are run in parallel at different resolutions. The different resolutions can swap configurations via a chosen subset of coordinates. The CG model is employed to move over high free energy barriers that the atomistic-resolution model might never sample.
60. Lyman E, Zuckerman DM: **Resolution exchange simulation with incremental coarsening.** *J Chem Theory Comput* 2006, **2**:656-666.
Larger molecules are examined using resolution exchange. An incremental coarsening procedure is constructed by employing a 'ladder' of incrementally coarsened models. Dileucine and met-enkephalin are examined as test systems. Resolution exchange is also combined with parallel tempering by varying temperatures.
61. Liu P, Voth GA: **Smart resolution replica exchange: an efficient algorithm for exploring complex energy landscapes.** *J Chem Phys* 2007, **126**:045106.
A smart resolution replica exchange algorithm is developed. This method employs a low- and high-resolution component, whereby the low-resolution model is systematically constructed using multiscale CG.
62. Giessen AE, Straub JE: **Monte Carlo simulations of polyaniline using a reduced model and statistics-based interaction potentials.** *J Chem Phys* 2005, **122**:024904.
63. van Giessen AE, Straub JE: **Coarse-grained model of coil-to-helix kinetics demonstrates the importance of multiple nucleation sites in helix folding.** *J Chem Theory Comput* 2006, **2**:674-684.
64. Peng S, Ding F, Urbanc B, Buldyrev SV, Cruz L, Stanley HE, Dokholyan NV: **Discrete molecular dynamics simulations of peptide aggregation.** *Phys Rev E Stat Nonlin Soft Matter Phys* 2004, **69**:041908.
65. Zhou J, Thorpe IF, Izvekov S, Voth GA: **Coarse-grained peptide modeling using a systematic multiscale approach.** *Biophys J* 2007, in press.
66. Villa E, Balaeff A, Schulten K: **Structural dynamics of the lac repressor-DNA complex revealed by a multiscale simulation.** *Proc Natl Acad Sci USA* 2005, **102**:6783-6788.
A multiscale simulation is employed to examine a complex between the lac repressor protein (LacI) and a 107 bp DNA segment. The LacI complex with the operator DNA segments is modeled with atomistic MD; the 75 bp DNA loop employs a continuum model. The multiscale approach enables an examination of how the structural dynamics of the repressor change owing to the strained DNA loop. See [67] for further details on the method.
67. Villa E, Balaeff A, Mahadevan L, Schulten K: **Multiscale method for simulating protein-DNA complexes.** *Multiscale Model Simul* 2004, **2**:527-553.
68. LaMarque JC, Le TL, Harvey SC: **Packaging double-helical DNA into viral capsids.** *Biopolymers* 2004, **73**:348-355.
69. Petrov AS, Harvey SC: **Structural and thermodynamic principles of viral packaging.** *Structure* 2007, **15**:21-27.
70. Bruant N, Flatters D, Lavery R, Genest D: **From atomic to mesoscopic descriptions of the internal dynamics of DNA.** *Biophys J* 1999, **77**:2366-2376.
71. Sun J, Zhang Q, Schlick T: **Electrostatic mechanism of nucleosomal array folding revealed by computer simulation.** *Proc Natl Acad Sci USA* 2005, **102**:8180-8185.
The 'discrete surface charge optimization' (DiSCO) model combined with an elastic DNA approach is used to model nucleosome and linker DNA. The electrostatic potential was found from a non-linear Poisson-Boltzmann approach over the nucleosome surface.
72. Matsumoto A, Tobias I, Olson WK: **Normal-mode analysis of circular DNA at the base-pair level. 1. Comparison of computed motions with the predicted behavior of an ideal elastic rod.** *J Chem Theory Comput* 2005, **1**:117-129.
An NMA is applied to mesoscopic-length DNA circles and the local chemical environment is treated implicitly. The potential energy is minimized before the NMA and is composed of internal dimer step energies, which are expressed as a sum of elastic contributions over the 6 bp step parameters. The elastic constants are found from fluctuations. The backbone is incorporated by the superposition of a 5'-nucleotide fragment from the canonical B-DNA fiber diffraction model.
73. Matsumoto A, Tobias I, Olson WK: **Normal-mode analysis of circular DNA at the base-pair level. 2. Large-scale configurational transformation of a naturally curved molecule.** *J Chem Theory Comput* 2005, **1**:130-142.
74. Tama F, Valle M, Frank J, Brooks CL: **Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy.** *Proc Natl Acad Sci USA* 2003, **100**:9319-9323.
75. Wang Y, Rader AJ, Bahar I, Jernigan RL: **Global ribosome motions revealed with elastic network model.** *J Struct Biol* 2004, **147**:302-314.
76. Wang Y, Jernigan RL: **Comparison of tRNA motions in the free and ribosomal bound structures.** *Biophys J* 2005, **89**:3399-3409.
77. Izvekov S, Parrinello M, Burnham CJ, Voth GA: **Effective force fields for condensed phase systems from *ab initio* molecular dynamics simulation: a new method for force-matching.** *J Chem Phys* 2004, **120**:10896-10913.
78. Noid WG, Chu J-W, Ayton GS, Voth GA: **Multiscale coarse-graining and structural correlations: connections to liquid state theory.** *J Phys Chem B* 2007, in press.
A statistical mechanical framework for the MS-CG method is presented. The authors demonstrate that the FM method may be used to characterize a many-body multidimensional potential of mean force for CG sites. Moreover, they show that the MS-CG pair potential incorporates important three-body effects.
79. Izvekov S, Voth GA: **Modeling real dynamics in the coarse-grained representations of condensed phase systems.** *J Chem Phys* 2006, **125**:151101.
80. Lyubartsev AP, Laaksonen A: **Effective potentials for ion-DNA interactions.** *J Chem Phys* 1999, **111**:11207-11215.
81. Ayton GS, McWhirter JL, Voth GA: **A second generation mesoscopic lipid bilayer model: connections to field-theory descriptions of membranes and nonlocal hydrodynamics.** *J Chem Phys* 2005, **124**:064906.