Challenges and Perspectives in Biomolecular Simulations: From the Atomistic Picture to Multiscale Modeling

Michele Cascella*[a] and Matteo Dal Peraro[b]

Abstract: We review the state-of-the-art in computational molecular simulations for biological systems. We limit our discussion to three fields: all-atom simulations, coarse-grained models, and novel multiscale approaches. While molecular dynamics simulations are broadly used in the community, major efforts continue to be spent in pushing the boundaries in both size and time limits as well as in improvement of commonly-used force fields. Parallel to all-atom simulations, in recent times the development of coarse-grained methods has flourished. Such techniques are able to describe biophysical features of macromolecular complexes through the use of simplified model potentials. Finally, multiscale models are introduced, giving some perspective about possible future developments in this new field.

Keywords: Coarse-grained · Molecular dynamics · Molecular simulation · Multiscale modeling

Michele Cascella studied chemistry at the University of Rome 'La Sapienza' and then attended the International School for Advanced Studies in Trieste, where he gained a PhD in statistical and biological physics. After postdoctoral experience in the group of Prof. Ursula Röthlisberger at EPFL, in 2008 he was awarded a SNF Förderungprofessur grant. He is now Assistant Professor in computational chemistry at the Department of Chemistry and Biochemistry of the University of Bern. His research is focused on multiscale modeling of biopolymers, with particular interest in mechanisms of molecular motion and in the development of bioinspired molecular devices.

Matteo Dal Peraro graduated in physics at the University of Padua, obtained his PhD in Biophysics at the International School for Advanced Studies (SISSA, Trieste), and received postdoctoral training under the guidance of Prof. M. L. Klein (University of Pennsylvania). In 2007 he was nominated Assistant Professor at the EPFL-School of Life Sciences. His research at the Laboratory for Biomolecular Modeling focuses on the development of multiscale schemes able to extend the time- and size-scales of molecular simulations, in order to tackle problems such as the assembly of large macromolecular complexes and the design of improved remedies for pathogenic infections.

1. Introduction

Molecular simulations have proved to be a successful method to describe (bio)chemical events. Among different possible approaches to modeling, all-atom molecular dynamics (MD) has been shown to provide a reasonable compromise between accuracy and computational cost.[1] The increase in computational power allows routine application of such techniques to systems as large as 10^2 kDa in the multi-nanosecond timescale, using commodity or high-performance-computing cluster facilities. In present times, the scientific community has been very active in developing new force-fields and protocols, as well as new algorithms, to increase both the accuracy and the performance of current MD codes. Nowadays MD is broadly used to describe molecular interactions in biological systems. Nevertheless, it remains in general too expensive to thoroughly investigate dimensional and dynamical scales that are critical to most of the biological processes both *in vitro* and *in vivo*. In fact, all fundamental biological processes necessary to life (*e.g.* protein folding, signal transduction, DNA transcription, *etc.*), which are triggered by interactions at atomistic dimensionality, occur at very different timescales (from the femtosecond to the seconds, and more), and span over different sizes (from few tens to millions of atoms). Such scales imply a dimensionality of the corresponding phase-space so large that its complete sampling by MD remains unaffordable.[2]

To address such size/time scale issues various coarse-grained (CG) Hamiltonians have been recently developed for macromolecular systems.[3–11] In particular, the focus is on the development of force-fields-like Hamiltonians based on bead models, which are in principle of general applicability. Such models have been successfully applied to problems of great biological relevance, from membrane self-assembly and dynamics to protein folding.[12–17] Along these lines, the first CG models focused on simple hydrophobic-polar interactions, which led to a series of models for lipid–water mixtures.[8,9] Currently, this same concept is being extended to the field of protein–membrane interactions.[18,19] In recent years, a consistent number of publications both on method development and

Correspondence: Prof. Dr. M. Cascella[a]
Tel.: +41 31 631 4256
Fax: +41 31 631 3994
E-mail: michele.cascella@iac.unibe.ch
[a]Departement für Chemie und Biochemie
Universität Bern
Freiestrasse 3
CH-3012 Bern
[b]Laboratory for Biomolecular Modeling
Institute of Bioengineering, School of Life Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne

applications have emerged as relevant contributions to the investigation of biological systems. This trend strongly suggests that such approaches to molecular simulations could lead to important breakthroughs in the near future.[2] Nevertheless, CG models for proteins still lack the required universality, especially towards potential-function transferability, which is crucial for fundamental properties, such as secondary structure modifications during conformational transitions.

Despite the great potential of CG approaches, an atomistic-detailed description of molecular structures is still crucial when studying phenomena involving molecular recognition (*e.g.* receptor–ligand binding). These interactions are of fundamental importance to establish functional protein networks able to control complex metabolic pathways. Recently, considerable efforts have been put into the development of novel *multiscale* modeling (MSM) techniques that are able to couple atomistic Hamiltonians to CG models.[6,7,12,20–33] In particular, it has been shown that hybrid schemes (AA/CG), where only a portion of the system is treated at the atomistic-detailed level, can be implemented and successfully applied to proteins and protein–DNA complexes.[10,11,34,35] Different approaches present in the literature include parallel schemes, where information coming from models at different dimensionality are coupled together,[24–27] adaptive resolution schemes, where the dimensionality of the systems is dynamically tuned,[21,29,30,33] and multidimensional Hamiltonians, where different portions of the system are treated at different levels of resolution at the same time.[10,28,35]

In this review, we want to provide a short overview of the challenges and perspectives in atomistic, coarse-grained and multiscale simulations (Fig. 1), using re-cent highlights from the literature. For more extensive reviews on specific subjects, we address the reader to more extensive publications.[1,2,12,17,31,36]

2. Atomistic Simulations

The recent increase in computational power allows routine investigation, using atomistic resolution force fields, of biological systems of the order of hundred thousand atoms for times that typically reach the 10^2 nanosecond length. Therefore, current challenges in molecular simulations focus on the investigation of processes and timescales that were precluded just a few years ago. Such achievement has also become possible thanks to the improved implementation and scalability of molecular dynamics codes broadly used by the community.[37–41]

Actual frontiers in time and size of all-atom simulations in published works can be substantially stretched from those typically achievable by routine computation. Large all-atom simulations have been performed on lipids showing bilayer self-assembly and dynamics, being able to study not only equilibrium properties, but self-assembly mechanisms as well.[42] Grubmüller and coworkers have recently investigated the molecular mechanism of membrane electroporation.[43] The study comprised the investigation of two systems of 128 or 512 palmitoyl-oleoyl-phosphatidylcholine lipids in explicit water. The total size of the simulated systems reached about 20,000/90,000 atoms. The smaller system was used to study pore formation kinetics, and its reduced size allowed for repetition of 48 independent MD simulations for a total of about one millisecond. The bigger system, used to study pore stability, was simulated for a total of 50 ns. Repeated simulations at different external field intensities showed that pore formation is accelerated by asymmetrical changes in the dipole orientations between the monolayers both in the lipid and in the wetting water.

Schulten and coworkers have been pushing the study of very large biomolecular complexes within atomistic simulation up to the entire satellite tobacco virus.[44] Such simulation followed seeding studies, mainly at CG level, on viral capsid and other portions of viruses[45–47] and were able to report a first characterization of the physical properties of the whole virion. Such computations highlighted differences in capsid stability in the presence/absence of the RNA core, and proposed mechanisms of capsid assembly, stressing the importance of RNA.

Increase of size limits in atomistic simulations allows nowadays the study of membrane proteins by treating their native lipid bilayer environment at the same level of detail. The combined use of homology modeling and molecular dynamics enabled the study of cation transport through the channel of the muscle nicotinic acetilcholine receptor.[48] Repeated simulations at different membrane potentials showed that charge selectivity is achieved by electrostatic interactions, while translocation is governed by hydration and local structural fluctuations.

The usefulness of large-size simulations relies also on the fact that direct comparison with experimental measurements becomes more and more feasible. An outstanding example of direct experiment–simulation interplay is given by a recently published study on the mechanosensitive channel of small conductance.[49] MD simulations predict ion conductance similar to that experimentally measured, showing strong spatial correlation for phenotypes losing or gaining function. Gating mechanisms related to helix tilting are also accurately described.

Another possibility to combine experimental data and atomistic simulations is the so-called *molecular dynamics-flexible fitting* protocol,[50–52] which allows atomic models to be obtained from low-resolution structures (*e.g.* cryo-electron microscopy maps). This technique allowed characterization of ribosome-induced conformational changes in elongation factor Tu, which are functional to its GTPase activity.[53]

Although the CPU power allows systems of very large size to be addressed, a complementary computational strategy is to use the available simulation time to study smaller systems for very long timescales. Within this approach, different enhanced sampling techniques can be efficiently used to study rare events.[54–61] With
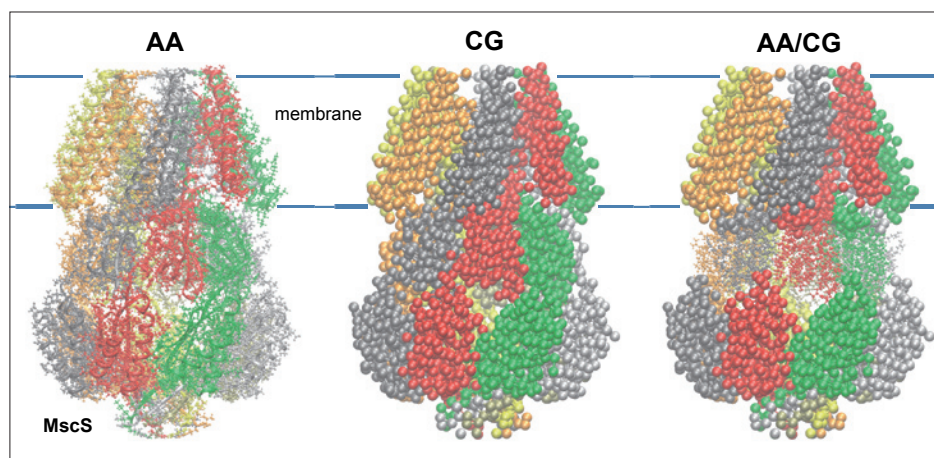

Fig. 1. Resolution models for molecular simulations. Pictorial views of the mechanosensitive channel (MscS, PDB code entry: 2OAU) at different resolution schemes: All-atom resolution (AA), two-bead coarse-grained (CG) and, hybrid all-atom/coarse-grained (AA/CG) models. Different colors correspond to the subunits of the channel. Atoms are drawn in ball-and-sticks, CG-beads in spheres. In AA, ribbon representation of the channel is superimposed to better clarify the secondary structure elements of the protein complex.

respect to folding from MD,[62] significant results have been reported for small protein domains like the villin headpiece,[63–65] WW domain[66] or the tryptophane cage.[67] Another outcome of this strategy is the revamped interest for the community for the improvement and development of current force fields. In fact, the quality of such potentials has been challenged by the stability of long-integrated equation of motion. Very recently, contributions have been made to improve the reliability of the AMBER force-field[68] and DNA models.[69]

3. Coarse-grained Simulations

In order to simultaneously extend the boundaries of time and size in molecular simulations, one can think to abandon the atomistic representation for a coarser description of molecular systems. From the original framework of elastic networks used by Go and Scheraga[3,70] to describe protein folding, more recent models make use of *beads* (representative of groups of atoms), which interact through effective force field-like potentials. Successful models using this idea were first introduced to describe polar/non-polar interactions, such as lipid self-assembly processes in a water environment.[8,9,42] Based on these and other works, Klein and coworkers have recently proposed a new generation of CG models able to study large aqueous surfactant solutions,[8,9,71,72] which can contain more than one million CG particles (*i.e.* ~10 millions atoms), and can be simulated for timescales approaching relevant experimental regimes. Marrink and coworkers have recently developed a second version of their CG model for lipids,[4] which has been used to investigate several relevant properties concerning biological membranes;[19,73–75] *e.g.,* they were able to closely reproduce experimental results concerning phase separation of mixtures of saturated/unsaturated lipids together with cholesterol, and show the molecular-detailed representation of raft-like lipid domains, which can be relevant to understand the self-assembly process of membrane–protein complexes.[76]

Based on these pioneering works, potentials based on similar philosophy have been recently extended to proteins[20] and nucleosomes.[77–79] The MARTINI force field has been indeed generalized for peptides,[5] using a 4-to-1 mapping (on average four heavy atoms are represented by a bead), and accounting for four types of interactions based on the electrostatic nature of residues. The model has been used in the past year to study very large membrane–protein complexes, and gave very promising results for the study of relevant structural rearrangements functional to many biological processes. For example,

the structural modifications upon transition to the conductive state have been unveiled within the microsecond timescale for the mechanosensitive channel.[19] For the voltage-gated Kv1.2 channel, Treptow *et al.*[18] proposed that the gating mechanism is principally promoted by transmembrane helical displacement of the channel. A slightly modified model based on the MARTINI force field is used by the group of Mark Sansom in the noteworthy effort of bringing CG simulations of membrane proteins to the high-throughput regime. A database collecting CG MD simulations (~150, each of the length of ~200 ns) of a large part of currently known structures of membrane proteins is available to the community and constantly updated at the following URL address, *http://sbcb.bioch. ox.ac.uk/cgdb*.

Even simpler approaches of coarse graining protein potentials have been proposed, *e.g.* those based on a single bead representation of amino acids[20] and nucleosomes.[77,78] One example is proposed by Tozzini *et al.*,[17] who mapped the dihedral degrees of freedom for the protein backbone in the all-atom description into a one-bead model. Despite its simplicity, this model is able to describe secondary structure transitions and large conformational rearrangements as in the case of substrate binding and product dissociation in HIV-1 protease.[79]

Bead-based potentials, which typically describe a portion or an entire amino acid, can be extended to describe structural domains or even single proteins in multi-molecular assemblies. Schulten and coworkers have used a sh-based coarse graining protocol for structural domains[20] to study the assembly of large systems, such as viral capsids,[45] bacterial flagellum,[80] BAR domains,[20] and large membrane leaflets on a multi-microsecond timescale.[81] Using neural networks to assign CG beads to structural domains it is possible to reproduce the shape of the protein with a minimal number of particles, and thus to study the dynamics of large networks of interacting assemblies. A similar philosophy, at a larger scale, has been recently used by Sali and coworkers to model the structural determinants of the nuclear pore complex,[82,83] integrating in the modeling procedure all the available experimental-based spatial restraints on the single subunits (*i.e.* stochiometry, X-ray and cryo-EM density maps, *etc.*) to reach a consistent model of the entire macromolecular complex.

4. Multiscale Simulations

Despite the possibility of pushing time- and size-scales in CG representations of biomolecules, it is not always

possible to abandon the description at the atomistic detail of molecular events. In fact, the interplay between subtle changes in the atomistic conformation (change of H-bond pattern, isomerization, *etc.*) and large changes in supramolecular structures are the basis of many biophysical signaling mechanisms. Multiscale modeling (MSM) schemes, where all-atom and CG levels are used in synergy, represent therefore a promising new path to provide tools for the study of such problems. Following the labeling introduced by Voth and coworkers,[12] they can be categorized in 'serial' schemes, where information at a given resolution is used to build up potential at another one, and 'parallel' schemes, where different resolutions coexist and influence each other. A key issue when building a 'parallel' multiscale approach lies in the consistency between different descriptions at different resolutions. Recently, a series of new parallel MSM schemes, which solve such problem in different ways, has been presented.[6,7,12,20–24,26–32] In hybrid schemes, where portions of the system are treated at different levels, it is crucial to retain an accurate description of long-range interactions, since they can significantly influence the polarization and the structural orientation at atomistic level, being functional to the biological activity.[84] Therefore, hybrid AA/CG schemes (Fig 1.) should somehow incorporate the long-range electrostatic potential between the CG part and the all-atom portion. In particular, long-range electrostatic interactions, crucial in many biological processes, are the most difficult to model by CG procedures, because they are influenced by the environment (pH, ionic strength, *etc.*) and because they have a strong dipolar nature, which severely limits description using simple radial potentials.

Voth and coworkers have shown that coarse-grained potentials can be rigorously determined from atomistic scale interactions through a force-matching approach. Such a procedure implies an initial all-atom classical MD run, over which the CG potential is fitted to reproduce all-atom forces.[7,23] The CG Hamiltonian obtained by such a scheme is not based on thermodynamic quantities, and is therefore general, unbiased towards the specific simulated thermodynamic conditions, and it is in principle able to correctly describe long-range interactions. The method has been successfully applied to construct models for different lipophilic/hydrophilic systems and distinct peptide structural moieties.[11,32] In practice, the quality of such a coarse-graining procedure is determined by the efficacy of the underlying all-atom MD to sample the relevant phase-space of the system. This may be more stringent when studying large macromolecular com-

plexes, which present long characteristic timescales.

We have recently proposed a topological scheme to reproduce the all-atom electrostatic field from structural CG information.[85] The proposed model reconstructs the orientation of the backbone dipoles using their statistical orientation in protein structures available in the Protein Data Bank, using the angle formed by three consecutive α-carbons as a simple order parameter (Fig. 2). This method is not computationally expensive, can be easily generalized to side-chains, and it is fully consistent with respect to protein motion, and therefore can be implemented in any MD algorithm. Moreover, long-range interactions produced within such a scheme are intrinsically anisotropic, suggesting its possible usefulness also in the development of more universal CG force-fields.

5. Concluding Remarks

Molecular simulations, pushed by both the constant increase in computational power and the birth of new methods and algorithms, nowadays present themselves as powerful tools to investigate living matter at subcellular dimensionality. While more expensive atomistic simulations push constantly the boundaries in terms of feasibility, multiscale methods, which couple coarse-grained models to atomistic resolution, are at present the most promising new strategy to allow a significant leap towards new scales and dimensionalities. On the experimental side, proteomics techniques are progressing very fast to enable the characterization of large systems, however what is still missing for a complete comprehension of the biological function is the atomistic dynamical picture of the macromolecular assemblies of which they are constituted. The only accessible way to access this level, having the opportunity to develop strategies to interfere with these networks, is through current structural biology techniques. For this reason, the possibility to computationally predict the network of molecular interactions is very desirable for the advance of life sciences.
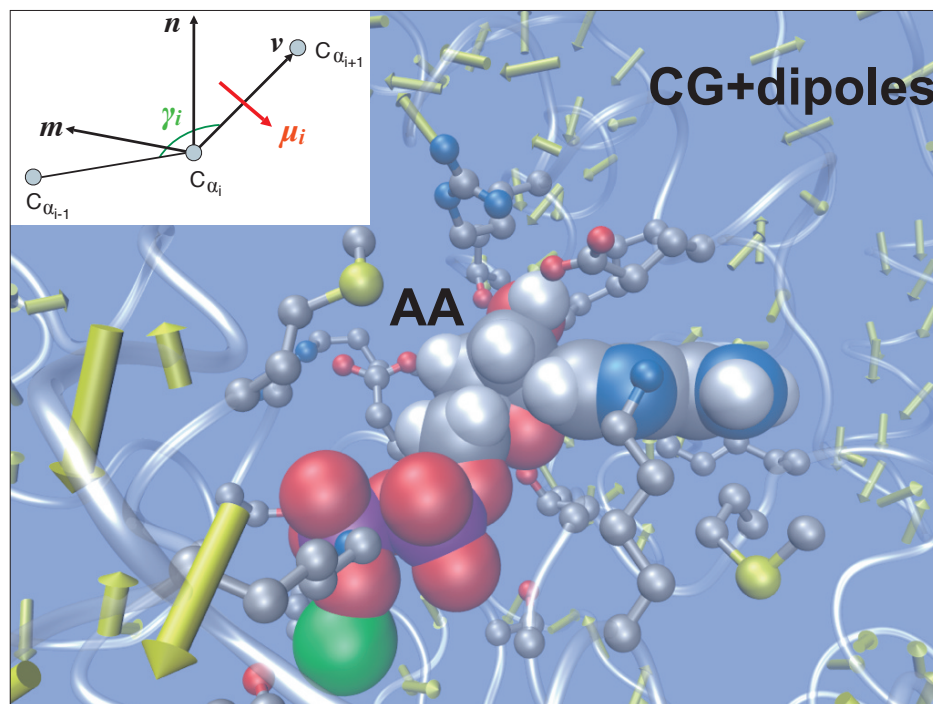
Received: January 12, 2009

Fig. 2. Topologically-based multipolar reconstruction of the electrostatic potential in AA/CG schemes. The binding site of actin and a bound ADP molecules, treated at the all-atom dimensionality, are drawn in ball-and-sticks and spheres, respectively. The rest of the system, studied at CG level, is represented by a tube. Electrostatic dipoles, reconstructed following ref. [85] (see inset), are shown as yellow arrows.

[1] M. Karplus, J. A. McCammon, *Nat. Struct. Biol.* **2002**, *9*, 646.
[2] M. L. Klein, W. Shinoda, *Science* **2008**, *321*, 798.
[3] N. Go, H. Taketomi, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 559.
[4] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. de Vries, *J. Phys. Chem. B* **2007**, *111*, 7812.
[5] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, S. J. Marrink, *J. Chem. Theory Comput.* **2008**, *4*, 819.
[6] W. G. Noid, J. W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, H. C. Andersen, *J. Chem. Phys.* **2008**, *128* , 244114.
[7] W. G. Noid, P. Liu, Y. Wang, J. W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen, G. A. Voth, *J. Chem. Phys.* **2008**, *128*, 244115.
[8] J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, M. L. Klein, *J. Phys. Chem. B* **2001**, *105*, 4464.
[9] J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, P. B. Moore, M. L. Klein, *J. Phys. Chem. B* **2001**, *105*, 9785.
[10] E. Villa, A. Balaeff, K. Schulten, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6783.
[11] J. Zhou, I. F. Thorpe, S. Izvekov, G. A. Voth, *Biophys. J.* **2007**, *92*, 4289.
[12] G. S. Ayton, W. G. Noid, G. A. Voth, *Curr. Opin. Struc. Biol.* **2007**, *17*, 192.
[13] P. J. Bond, C. L. Wee, M. S. P. Sansom, *Biochemistry* **2008**, *47*, 11321.
[14] T. Carpenter, P. J. Bond, S. Khalid, M. S. P. Sansom, *Biophys. J.* **2008**, *95*, 3790.
[15] J. W. Chu, G. A. Voth, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13111.
[16] L. Thogersen, B. Schiott, T. Vosegaard, N. C. Nielsen, E. Tajkhorshid, *Biophys. J.* **2008**, *95*, 4337.
[17] V. Tozzini, *Curr. Opin. Struc. Biol.* **2005**, *15*, 144.
[18] W. Treptow, S. J. Marrink, M. Tarek, *J. Phys. Chem. B* **2008**, *112*, 3277.
[19] S. Yelimov, E. van der Giessen, P. R. Onck, S. J. Marrink, *Biophys. J.* **2008**, *94*, 2994.
[20] A. Arkhipov, Y. Yin, K. Schulten, *Biophys. J.* **2008**, *95*, 2806.
[21] L. Delle Site, C. F. Abrams, A. Alavi, K. Kremer, *Phys. Rev. Lett.* **2002**, *89*, 156103.
[22] A. Heyden, D. G. Truhlar, *J. Chem. Theory Comput.* **2008**, *4*, 217.
[23] S. Izvekov, M. Parrinello, C. J. Burnham, G. A. Voth, G. *J. Chem. Phys.* **2004**, *120*, 10896.
[24] P. Liu, Q. Shi, E. Lyman, G. A. Voth, *J. Chem. Phys.* **2008**, *129*, 214103.
[25] P. Liu, G. A. Voth, *J. Chem. Phys.* **2007**, *126*, 045106.
[26] E. Lyman, F. M. Ytreberg, D. M. Zuckerman, *Phys. Rev. Lett.* **2006**, *96*, 028105.
[27] E. Lyman, D. M. Zuckerman, *J. Chem. Theory Comput.* **2006**, *2*, 656.
[28] M. Neri, C. Anselmi, M. Cascella, A. Maritan, P. Carloni, *Phys. Rev. Lett.* **2005**, *95*, 218102.
[29] M. Praprotnik, L. Delle Site, K. Kremer, *Ann. Rev. Phys. Chem.* **2008**, *59*, 545.
[30] M. Praprotnik, L. Delle Site, K. Kremer, *J. Chem. Phys.* **2005**, *123*, 224106.
[31] P. Sherwood, B. R. Brooks, M. S. P. Sansom, *Curr. Opin. Struc. Biol.* **2008**, *18*, 630.
[32] I. F. Thorpe, J. Zhou, G. A. Voth, *J. Phys. Chem. B* **2008**, *112*, 13079.
[33] B. Ensing, S. O. Nielsen, P. B. Moore, M. L. Klein, M. Parrinello, *J. Chem. Theory Comput.* **2007**, *3*, 1100.
[34] Q. Shi, S. Izvekov, G. A. Voth, *J. Phys. Chem. B* **2006**, *110*, 15045.
[35] E. Villa, A. Balaeff, L. Mahadevan, K. Schulten, *Multiscale Model. Sim.* **2004**, *2*, 527.
[36] A. N. Lupas, *J. Struct. Biol.* **2008**, *163*, 254.
[37] NAMD Scalable Molecular Dynamics. *http://www.ks.uiuc.edu/Research/namd/*.
[38] Gromacs: fast, free and flexible MD. *http://www.gromacs.org*.
[39] LAMMPS Molecular Dynamics Simulator. *http://lammps.sandia.gov/*.
[40] The Amber molecular Dynamics Package. *http://ambermd.org/*.
[41] CHARMM Chemistry at HARvard Macro-molecular Mechanics. *http://www.charmm.orgl*.
[42] S. J. Marrink, A. E. Mark, *J. Am. Chem. Soc.* **2003**, *125*, 11144.
[43] R. A. Bockmann, B. L. de Groot, S. Kakorin, E. Neumann, H. Grubmuller, *Biophys. J.* **2008**, *95*, 1837.
[44] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson, K. Schulten, *Structure* **2006**, *14*, 437.
[45] A. Arkhipov, P. L. Freddolino, K. Schulten, *Structure* **2006**, *14*, 1767.
[46] V. S. Reddy, H. A. Giesing, R. T. Morton, A. Kumar, C. B. Post, C. L. Brooks, J. E.

Johnson, in 'Energetics of quasiequivalence: Computational analysis of protein-protein interactions in icosahedral viruses', Symposium on Quasi-Equivalence - Motion and Adaptability in Living Molecules in Honor of Donald L D Caspar, Tallahassee, Florida, Jan 10–11; Tallahassee, Florida, **1997**, pp 546–558.

[47] F. Tama, C. L. Brooks, *J. Mol. Biol.* **2005**, *345*, 299.

[48] H. L. Wang, X. L. Cheng, P. Taylor, J. A. McCammon, S. M. Sine, *PLoS Comput. Biol.* **2008**, *4*, e41.

[49] V. Vasquez, M. Sotomayor, J. Cordero-Morales, K. Schulten, E. Perozo, *Science* **2008**, *321*, 1210.

[50] F. Tama, O. Miyashita, C. L. Brooks, *J. Mol. Biol.* **2004**, *337*, 985.

[51] F. Tama, M. Valle, J. Frank, C. L. Brooks, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9319.

[52] L. G. Trabuco, E. Villa, K. Mitra, J. Frank, K. Schulten, *Structure* **2008**, *16*, 673.

[53] E. Villa, J. Sengupta, L. G. Trabuco, J. Lebarron, W. T. Baxter, T. R. Shaikh, R. A. Grassucci, P. Nissen, M. Ehrenberg, K. Schulten, J. Frank, *Proc. Natl. Acad. Sci. USA* **2009**, published online doi: 10.1073/pnes.0811370106.

[54] R. Affentranger, I. Tavernelli, E. E. Di Iorio, *J. Chem. Theory Comput.* **2006**, *2*, 217.

[55] A. Barducci, G. Bussi, M. Parrinello, *Phys. Rev. Lett.* **2008**, *100*, 020603.

[56] M. Bonomi, M, D. Branduardi, F. L. Gervasio, M. Parrinello, *J. Am. Chem. Soc.* **2008**, *130*, 13938.

[57] G. Bussi, F. L. Gervasio, A. Laio, M. Parrinello, *J. Am. Chem. Soc.* **2006**, *128*, 13435.

[58] C. A. F. de Oliveira, D. Hamelberg, J. A. Mc-Cammon, *J. Chem. Theory Comput.* **2008**, *4*, 1516.

[59] M. Fajer, D. Hamelberg, J. A. McCammon, *J. Chem. Theory Comput.* **2008**, *4*, 1565.

[60] A. Laio, M. Parrinello, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562.

[61] A. Okur, D. R. Roe, G. L. Cui, V. Hornak, C. Simmerling, *J. Chem. Theory Comput.* **2007**, *3*, 557.

[62] M. Levitt, A. Warshel, *Nature* **1975**, *253*, 694.

[63] D. L. Ensign, P. M. Kasson, V. S. Pande, *J. Mol. Biol.* **2007**, *374*, 806.

[64] S. Piana, A. Laio, *Phys. Rev. Lett.* **2008**, *101*, 208101.

[65] S. Piana, A. Laio, F. Marinelli, M. Van Troys, D. Bourry, C. Ampe, J. C. Martins, *J. Mol. Biol.* **2008**, *375*, 460.

[66] P. L. Freddolino, F. Liu, M. Gruebele, K. Schulten, *Biophys. J.* **2008**, *94*, L75.

[67] S. Piana, A. Laio, *J. Phys. Chem. B* **2007**, *111*, 4553.

[68] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* **2006**, *65*, 712.

[69] A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham III, C. A. Laughton, M. Orozco, *Biophys. J.* **2007**, *92*, 3817.

[70] S. Tanaka, H. A. Scheraga, *Macromolecules* **1976**, *9*, 142.

[71] W. Shinoda, R. DeVane, M. L. Klein, *Soft Matter* **2008**, *4*, 2454.

[72] W. Shinoda, R. DeVane, M. L. Klein, *Mol. Simulat.* **2007**, *33*, 27.

[73] S. J. Marrink, J. Risselada, A. E. Mark, *Chem. Phys. Lipids* **2005**, *135*, 223.

[74] X. Periole, T. Huber, S. J. Marrink, T. P. Sakmar, *J. Am. Chem. Soc.* **2007**, *129*, 10126.

[75] H. J. Risselada, A. E. Mark, S. J. Marrink, *J. Phys. Chem. B* **2008**, *112*, 7438.

[76] H. J. Risselada, S. J. Marrink, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17367.

[77] F. Trovato, V. Tozzini, *J. Phys. Chem. B* **2008**, *112*, 13197.

[78] K. Voltz, J. Trylska, V. Tozzini, V. Kurkal-Siebert, J. Langowski, J. Smith, *J. Comput. Chem.* **2008**, *29*, 1429.

[79] J. Trylska, V. Tozzini, C. A. Chang, J. A. McCammon, *Biophys. J.* **2007**, *92*, 4179.

[80] A. Arkhipov, P. L. Freddolino, K. Imada, K. Namba, K. Schulten, *Biophys. J.* **2006**, *91*, 4589.

[81] A. Y. Shih, I. G. Denisov, J. C. Phillips, S. G. Sligar, K. Schulten, *Biophys. J.* **2005**, *88*, 548.

[82] F. Alber, S. Dokudovskaya, L. M. Veenhoff, W. H. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B. T. Chait, A. Sali, M. P. Rout, *Nature* **2007**, *450*, 695.

[83] F. Alber, S. Dokudovskaya, L. M. Veenhoff, W. Z. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B. T. Chait, M. P. Rout, A. Sali, *Nature* **2007**, *450*, 683.

[84] A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, M. H. M. Olsson., *Chem. Rev.* **2006**, *106*, 3210.

[85] M. Cascella, M. A. Neri, P. Carloni, M. Dal Peraro, *J. Chem. Theory Comput.* **2008**, *4*, 1378.