



Editorial: Advanced Sampling and Modeling in Molecular Simulations for Slow and Large-Scale Biomolecular Dynamics

Xiakun Chu^{1*}, *Yong Wang*^{2*}, *Pengfei Tian*^{3*}, *Wenfei Li*^{4,5*} and *Davide Mercadante*^{6*}

¹Department of Chemistry, State University of New York, Stony Brook, NY, United States, ²College of Life Sciences, Shanghai Institute for Advanced Study, Institute of Quantitative Biology, Zhejiang University, Hangzhou, China, ³Enzyme Research, Novozymes, Lyngby, Denmark, ⁴National Laboratory of Solid State Microstructure, Department of Physics, Nanjing University, Nanjing, China, ⁵Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou, China, ⁶School of Chemical Sciences, The University of Auckland, Auckland, New Zealand

Keywords: molecular dynamics, enhanced sampling, coarse-grained model, conformational dynamics, free energy landscape

OPEN ACCESS

Edited and reviewed by:

Massimiliano Bonomi,
Institut Pasteur, France

*Correspondence:

Xiakun Chu
xkchu2008@gmail.com
Yong Wang
isb@zju.edu.cn
Pengfei Tian
tianpengfei09@gmail.com
Wenfei Li
wfli@nju.edu.cn
Davide Mercadante
davide.mercadante@
auckland.ac.nz

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 15 October 2021

Accepted: 23 October 2021

Published: 12 November 2021

Citation:

Chu X, Wang Y, Tian P, Li W and
Mercadante D (2021) Editorial:
*Advanced Sampling and Modeling in
Molecular Simulations for Slow and
Large-Scale Biomolecular Dynamics*.
Front. Mol. Biosci. 8:795991.
doi: 10.3389/fmolb.2021.795991

Editorial on the Research Topic

Advanced Sampling and Modeling in Molecular Simulations for Slow and Large-Scale Biomolecular Dynamics

From conception, the impact of molecular dynamics (MD) simulations has grown dramatically (Karplus and McCammon, 2002), even though MD simulations are still affected by a critical timescale issue. Currently, timescales accessible to MD simulations are, on average, shorter than those of the investigated events, often resulting in insufficient sampling. Numerous efforts have been spent on accelerating MD simulations in order to ease the timescale problem. This Research Topic collects contributions focusing on developing and using advanced sampling techniques and modeling strategies to promote applications of MD simulations to study a diverse range of large-scale biomolecular systems.

Exhaustive sampling is especially important for intrinsically disordered proteins (IDPs), which show a more shallow and rugged energy landscape when compared to folded proteins (Papoian, 2008). Ding et al. presented case studies on two IDPs by an iterative screening-after-sampling strategy. In their study accelerated molecular dynamics was used to enhance the sampling of highly diverse conformational ensembles of IDPs, with Small-angle X-ray scattering (SAXS) to guide the sampling iteratively and obtain ensembles in good agreement with experimental data. Such integrative modeling might be more broadly useful for modeling IDPs ensembles. In general, the conformational dynamics of IDPs are strongly affected by the binding to other molecular partners or by aggregation. There are three excellent studies in this Research Topic, focusing on the conformational dynamics and aggregation properties of two Alzheimer's disease-related IDPs: the A β 42 peptide and the tau protein (Selkoe and Hardy, 2016). In the work by Xie and Guo, replica exchange with solute tempering (REST) (Liu et al., 2005) has been adopted to sample the binding of the intrinsically disordered A β 42 peptide to the Human serum albumin (HSA), elucidating the molecular mechanism of amyloid inhibition by HSA. Their results suggest that A β 42 binds to multiple sites on HSA, which shifts the conformational propensity of the peptide towards a more disordered state altering its aggregation propensity altogether. The reward behind the quest to mechanistically characterize fibrillar nucleation in proteopathies is enormous as it would

suggest effective strategies for therapeutic approaches to neurodegeneration. Ma et al. employed coarse-grained simulations to energetically define the fibril growth of the A β -peptide: discovering the binding site of new filaments to the protofibril and a downhill mechanism of filament addition. Together with the identification of an emerging mechanical property of the A β -peptide protofibril, their research adds valuable insights to our understanding of the nucleation of A β -peptide fibrils for the development of strategies to pharmaceutically tackle fibril growth. McCarty and co-workers performed non-equilibrium steered MD simulations to investigate the structural changes of the tau paired-helical filament (PHF) and straight filament (SF) under mechanical force. In particular, the authors identified weak spots of interchain interactions additionally providing the dissociation pathway of a single tau peptide from the protofibril, through metadynamics simulations (Laio and Parrinello, 2002). In addition, the free energy profile for tau dissociation was obtained by umbrella sampling simulations.

One useful strategy to overcome sampling limitations is to introduce coarse-graining into biomolecular models. MARTINI is among the most widely used coarse-grained (CG) models for biomolecules (Monticelli et al., 2008). Mahmood et al. introduced a simple cutoff scheme to improve the definition of native contacts in the structure-based protein model of MARTINI (G \ddot{o} -MARTINI). By tuning the interaction strengths and cutoff distances, the MARTINI CG simulations can well reproduce the structural fluctuations from atomistic simulations for the membrane proteins investigated in the published study. The refined model has been successfully used to simulate the key steps leading to the assembly of the F-BAR protein involved in membrane remodeling. Using the CG UNited-RESidue (UNRES) model (Liwo et al., 1997), Stevens and He studied the large-scale conformational changes within the multidomain scaffolding protein PICK1. Considering the large size of PICK1, associated with extensive conformational flexibility, brute-force atomistic simulations on this system would inevitably lead to insufficient sampling. The physics-based CG model allowed the authors to reliably characterize the detailed interactions at the residual level and eventually uncover the forces driving the association of PICK1 subdomains. CG modeling has also been suggested to study and understand chromosomal organization and dynamics (Lin et al., 2021). However, determining the potential energy function suitably reproducing the behaviour of chromosomal and multi-chromosomal systems has been historically challenging. Zhang and Huang introduced a “bottom-up” approach, aiming to derive the CG potential from all-atom MD simulations. They studied dynamics and interactions in two-nucleosome systems, providing useful information to determine functional forms and parameters to sample the interaction between nucleosomes.

Coarse-graining the MD simulation trajectories is another useful strategy to reduce system complexity, while elucidating molecular dynamics. In addition to the sampling problem, extracting key structural and dynamic features from high-dimensional MD trajectories is a challenge in biomolecular modeling. Such a task relies on an appropriate definition of

feature space, within which the metastable states involved in biologically relevant conformational dynamics can be identified *via* clustering and dimensionality reduction analyses. The humanly understandable thermodynamic and kinetic information can then be reconstructed based on Markov State Models (MSM) (Chodera and Noé, 2014; Lane et al., 2011). In the work by Wang et al., an effective energy rescaling space trajectory mapping method has been developed to detect metastable states and construct kinetic transition networks. In their study the authors are able to successfully describe the major metastable states and the interstate transition kinetics involved in the folding of a dodecapeptide. MSM can be combined with enhanced sampling methods to further improve the performance of structural and dynamic characterizations as done by Fernández-Quintero et al. In this work, the authors demonstrated the correlation between the rigidification of the CDR-H3 loops of antibody fragments and the enhanced antigen specificity in different stages of affinity maturation by using metadynamics simulations in combination with MD simulations and MSM analyses.

This Research Topic also includes two binding case studies. Do et al. employed all-atom MD simulations with the aid of an enhanced sampling method called Gaussian-accelerated molecular dynamics (GaMD) (Wang et al., 2021) to determine the pathways and binding mechanism of caffeine to the human adenosine A2A receptor. By adding a harmonic boost potential, GaMD simulations allowed to capture the spontaneous ligand binding and release in the μ s time scale through smoothening the potential energy surface so as to reduce the energy barriers for slow conformational changes. This work provided a good example on how to implement enhanced sampling methods to study a protein-ligand binding mechanism. In investigating the binding between nanoparticles and the clathrin-associated protein adaptin-2 (AP2), Zhu et al. applied molecular modeling and simulations to understand the impact of nanoparticle morphology on binding specificity. They found that binding specificity is majorly dictated by electrostatic interactions as well as nanoparticle morphology. They also observed that nanoparticle binding significantly induces conformational changes in AP2. Overall, the authors provided a microscopic explanation for cargo recognition in clathrin-mediated endocytosis and possible mechanisms to design high-efficiency nano-biomaterials.

As an essential element in MD simulations, the adopted force field is key to affect the precision of the simulation results. Wang and Li developed and tested force field parameters for some noncanonical amino acids (NAAs). NAAs have been widely applied in protein engineering, virus vaccine development, and medical therapeutics due to their strong site specificity, without the need to introduce significant perturbations to a protein structure. Based on quantum mechanics (QM) calculations and experimental data as a benchmark, the authors determined force field parameters for phenylalanine and tyrosine derivatives showing that the newly identified parameters well describe protein-ligand interactions with NAAs as substrates. Finally, to aid structural modeling useful for MD simulations, Xian et al. developed a structure manipulation (StructureMan) tool that proved to be

comprehensive and efficient when studying interactions in large biomolecular systems.

Overall, we believe that this Research Topic provides a well-rounded picture of the latest state-of-the-art developments useful to overcome historical limitations in modeling and sampling of large biomolecular systems and slow processes using classical molecular dynamics simulations.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Chodera, J. D., and Noé, F. (2014). Markov State Models of Biomolecular Conformational Dynamics. *Curr. Opin. Struct. Biol.* 25, 135–144. doi:10.1016/j.sbi.2014.04.002
- Karplus, M., and McCammon, J. A. (2002). Molecular Dynamics Simulations of Biomolecules. *Nat. Struct. Biol.* 9, 646–652. doi:10.1038/nsb0902-646
- Laio, A., and Parrinello, M. (2002). Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci.* 99, 12562–12566. doi:10.1073/pnas.202427399
- Lane, T. J., Bowman, G. R., Beauchamp, K., Voelz, V. A., and Pande, V. S. (2011). Markov State Model Reveals Folding and Functional Dynamics in Ultra-Long Md Trajectories. *J. Am. Chem. Soc.* 133, 18413–18419. doi:10.1021/ja207470h
- Lin, X., Qi, Y., Latham, A. P., and Zhang, B. (2021). Multiscale Modeling of Genome Organization With Maximum Entropy Optimization. *J. Chem. Phys.* 155, 010901. doi:10.1063/5.0044150
- Liu, P., Kim, B., Friesner, R. A., and Berne, B. J. (2005). Replica Exchange With Solute Tempering: A Method for Sampling Biological Systems in Explicit Water. *Proc. Natl. Acad. Sci.* 102, 13749–13754. doi:10.1073/pnas.0506346102
- Liwo, A., Oldziej, S., Pincus, M. R., Wawak, R. J., Rackovsky, S., and Scheraga, H. A. (1997). A United-Residue Force Field for Off-Lattice Protein-Structure Simulations. I. Functional Forms and Parameters of Long-Range Side-Chain Interaction Potentials from Protein crystal Data. *J. Comput. Chem.* 18, 849–873. doi:10.1002/(sici)1096-987x(199705)18:7<849::aid-jcc1>3.0.co;2-r
- Monticelli, L., Kandasamy, S. K., Periole, X., Larson, R. G., Tieleman, D. P., and Marrink, S.-J. (2008). The Martini Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theor. Comput.* 4, 819–834. doi:10.1021/ct700324x
- Papoian, G. A. (2008). Proteins With Weakly Funneled Energy Landscapes Challenge the Classical Structure-Function Paradigm. *Proc. Natl. Acad. Sci.* 105, 14237–14238. doi:10.1073/pnas.0807977105
- Selkoe, D. J., and Hardy, J. (2016). The Amyloid Hypothesis of Alzheimer's Disease at 25 Years. *EMBO Mol. Med.* 8, 595–608. doi:10.15252/emmm.201606210
- Wang, J., Arantes, P. R., Bhattarai, A., Hsu, R. V., Pawnikar, S., Huang, Y.-m. M., et al. (2021). Gaussian Accelerated Molecular Dynamics: Principles and Applications. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 11, e1521. doi:10.1002/wcms.1521

FUNDING

YW thanks the financial support from the start-up funds of Zhejiang University. WL thanks the support from the National Natural Science Foundation of China (11974173). DM thanks the financial support from the School of Chemical Sciences at The University of Auckland.

ACKNOWLEDGMENTS

We thank all the contributors and reviewers of this Research Topic.

Conflict of Interest: Author PT was employed by company Novozymes, Denmark.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Chu, Wang, Tian, Li and Mercadante. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.