

PERSPECTIVE

Advances and challenges in multiscale biomolecular simulations: artificial intelligence-driven paradigm shift

Wenfei Li | Wei Wang

Department of Physics, National Laboratory of Solid State Microstructure, Nanjing University, Nanjing, China

Correspondence

Wenfei Li and Wei Wang.
Email: wfli@nju.edu.cn and wangwei@nju.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 12574224, 12090052, 12347102

Abstract

Molecular simulation techniques have become an invaluable tool for elucidating the fundamental principles of life at the molecular level. After nearly five decades of development, biomolecular simulations have evolved to enable the quantitative characterization of complex biomolecular events, such as protein folding, conformational dynamics, and protein–protein interactions. These advancements have significantly influenced both fundamental and applied research. In recent years, the integration of machine learning, particularly deep learning algorithms, has further driven innovation in this field. This perspective aims to discuss the latest advancements in biomolecular simulation techniques and to explore emerging applications, development trends, and major challenges in biomolecular dynamics simulations.

KEYWORDS

artificial intelligence, biomolecules, molecular dynamics, protein dynamics

1 | FROM STATIC STRUCTURES TO FUNCTIONAL DYNAMICS

Life processes at the molecular level primarily manifest as the assembly, motion, and interactions of biomolecules. Consequently, deciphering the principles that underlie biological function demands not only high-resolution three-dimensional structures but also a mechanistic understanding of the dynamic processes that drive activity. Recent advances powered by an ever-growing repository of protein structures and breakthroughs in artificial intelligence (AI) have enabled platforms such as AlphaFold2 to achieve transformative progress in static protein structure prediction, effectively resolving the long-standing structure prediction challenge [1–6]. However, experimental characterization of biomolecular dynamics remains significantly underdeveloped compared to static structure analysis.

This deficiency has created a critical shortage of dynamics data, severely limiting AI applications in biomolecular dynamics prediction. Molecular simulation techniques, particularly molecular dynamics (MD), are garnering increased attention as powerful tools for characterizing biomolecular dynamics [7–11].

2 | ACCURACY-EFFICIENCY BOTTLENECK IN BIOMOLECULAR SIMULATIONS

Classical MD simulations generate atomistic trajectories of biomolecules, and thus the associated kinetic and thermodynamic properties, by integrating Newton's equations with empirical force fields [7, 12, 13]. However, the vast number of degrees of freedom and complex high-dimensional energy landscapes inherent

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Quantitative Biology* published by John Wiley & Sons Australia, Ltd on behalf of Higher Education Press.

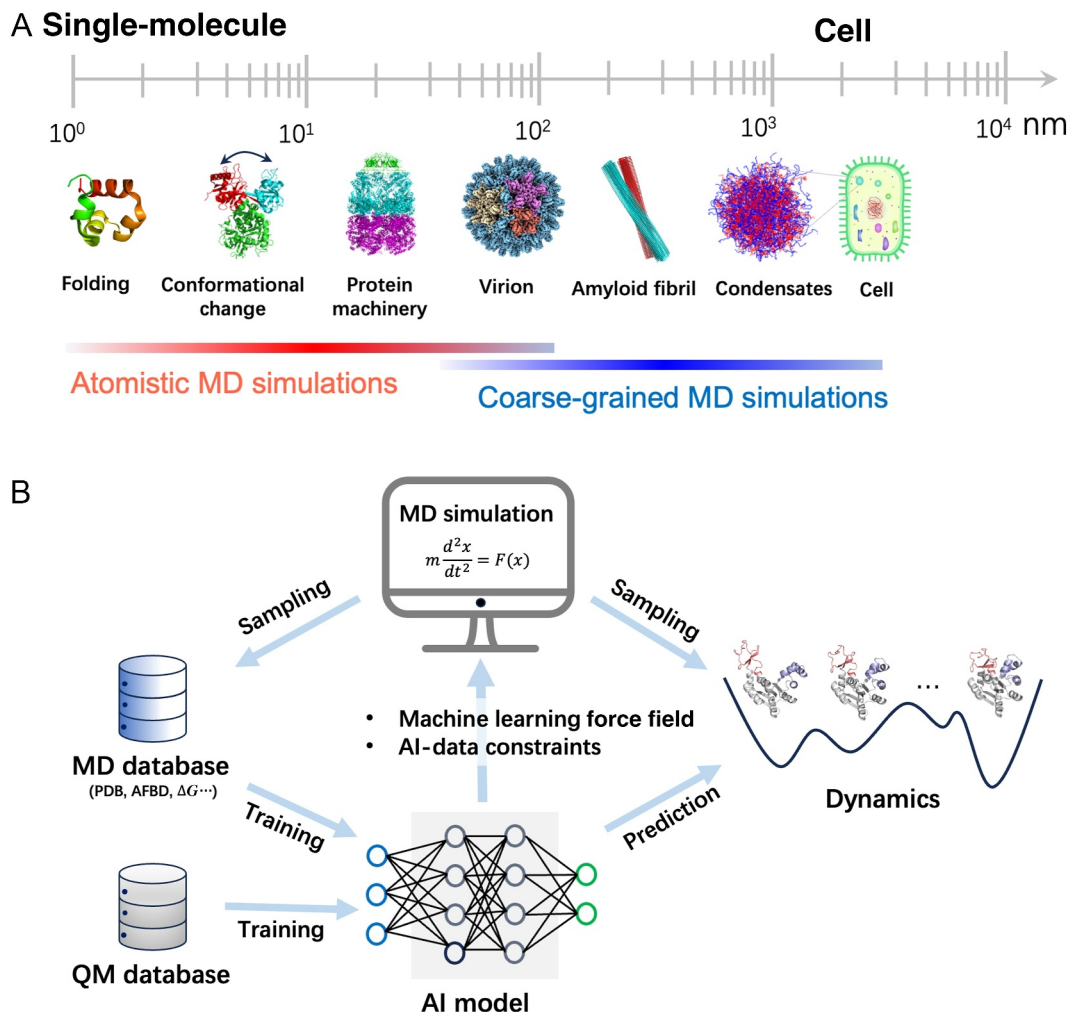


FIGURE 1 Schematic for the multiscale challenge of biomolecular simulations and its solution by AI-based methods. (A) Multiscale challenge of biomolecular simulations. (B) Modeling biomolecular dynamics by integrating molecular simulations and AI. Three strategies are illustrated: MLFF construction for high-accuracy MD simulations; AI-derived constraints guiding MD simulations; and AI models trained on MD data to predict protein dynamics. AI, artificial intelligence; MD, molecular dynamics; MLFF, machine learning force field.

to biomolecules render exhaustive sampling exceptionally challenging [14–16]. Functionally critical processes in proteins and other biomolecules span multiple spatiotemporal scales (Figure 1A) [17, 18]. Current atomistic MD simulations with enhanced sampling algorithms can access the timescales involved in the folding and conformational changes of proteins with moderate sizes. However, for protein dynamics involving collective behaviors of multiple biomolecules, such as the assembly and functional motions of large protein machineries, amyloid fibril aggregation, and biomolecular liquid–liquid phase separation (LLPS), which operate across much larger spatiotemporal scales, conventional atomistic simulations become computationally prohibitive. To overcome this computational barrier, coarse-grained (CG) models [15, 19–23], in which a group of atoms are simplified into single virtual particles, are introduced. Although the CG approximation provides a solution to overcome the

efficiency bottleneck of MD simulations, developing CG force fields that faithfully capture essential features of biomolecular interactions and dynamics remains profoundly difficult. Thus, enhancing the accuracy and efficiency constitutes a core mission and longstanding challenge in biomolecular simulations. Recently, AI-powered algorithms have emerged as novel pathways to address the accuracy–efficiency bottleneck (Figure 1B) [24–27].

3 | AI-DRIVEN MOLECULAR SIMULATIONS WITH QUANTUM MECHANICAL ACCURACY

The functional motions and interactions of biomolecules depend on their microscopic electronic structures, which in principle demand a quantum mechanical (QM) treatment. However, conventional QM methods remain

computationally prohibitive, limiting their application to small systems and short timescales. All-atom molecular force field model enhances computational efficiency by applying Born–Oppenheimer approximation, frozen-electron ground-state assumption, and empirical parametrization, but at the cost of reduced accuracy relative to *ab initio* approaches. Bridging this gap has long been a central goal in computational biophysics.

Deep neural networks provide powerful parameterization schemes without assuming predefined functional forms. By training machine learning force fields (MLFFs) on QM reference datasets, these models directly predict energies and forces from atomic coordinates without explicit electronic-structure calculations, achieving high-efficiency biomolecular computations at QM accuracy. In recent years, AI-driven quantum-accuracy simulation methods have made rapid progress [26, 28–31], enabling efficient and high-precision single-point energy evaluations, nuclear magnetic resonance (NMR) chemical shift predictions, protein folding free-energy estimation, and pKa calculations. Remarkably, the AI²BMD framework developed recently by Wang and colleagues has achieved QM-level simulations of small-protein folding and improved estimation of thermodynamic properties [26].

A key requirement for MLFFs is extensive QM training data, but high-throughput *ab initio* calculations on entire proteins remain intractable. This issue can be resolved by employing a divide-and-conquer strategy, that is, decomposing the large protein system into QM-treatable and overlapping fragments. These well-established fragmentation methodologies can be directly utilized to construct QM databases in training MLFFs [32–34]. It is worth noting that MLFFs leverage neural networks' exceptional fitting capacity but lack physically explicit energy functions, inherently limiting transferability. Expanding training datasets and establishing rigorous validation metrics are therefore essential tasks in future development. Additionally, reasonably modeling the key physical processes involving quantum effects, for example, catalytic reactions, charge transfer, and protonation dynamics of biomolecules, within the MLFF framework, are still challenging tasks.

4 | AI-DRIVEN MULTISCALE BIOMOLECULAR SIMULATIONS

Due to the inherent multiscale nature of proteins and other biomolecules, even the all-atom MD simulations using empirical force fields are limited to small spatio-temporal scales (Figure 1A). For large-scale biomolecular assemblies and functional motions, CG approximations become essential to overcome computational complexity. In CG models, a group of atoms is mapped onto single virtual particles, drastically reducing system degrees of freedom and accelerating

simulations. Unlike all-atom force fields parameterized against QM data, CG force fields need to reproduce conformational ensemble properties of higher-dimensional molecular systems within reduced reaction coordinates. Consequently, CG force fields represent the potential of mean force projected onto CG reaction coordinates, including not only energetic terms but also contributions from conformational entropy of integrated fine-grained degrees of freedom. This makes the development of CG force fields extremely challenging and requires much more complex functional forms.

The success of MLFFs has provided novel solutions for constructing CG force fields attributing to the powerful function representation capability of deep neural networks [29, 35]. Currently, two primary strategies are employed to develop CG force fields using MLFFs, including the bottom-up and top-down strategies [35–37]. The bottom-up strategy directly trains the CG MLFFs using energy or force data from QM calculations or all-atom simulations. For example, CGNet and its more recent version CGSchNet utilize the force-matching technique to extract force field parameters, that is, by optimizing CG MLFFs by matching forces on CG particles computed from all-atom models [12, 35]. CG MLFFs derived through this bottom-up method have been successfully applied to the simulations of protein folding, which can reasonably reproduce the free energy surfaces and other statistical properties from all-atom simulations. In contrast, the top-down strategy optimizes deep neural network parameters to fit the global observables of biomolecules, for example, root mean square deviation from reference protein structures [36]. In addition to learning and representing the CG force field, generative AI models have also been used to reconstruct the all-atom structures from the CG structures [29] and to generate conformational ensembles for disordered proteins [38].

Because of the complex functional forms and limited transferability inherent to CG force-field models, the development of CG MLFFs requires more comprehensive conformational space coverage in training datasets. However, databases derived from atomistic MD simulations in the training of CG MLFFs inevitably encounter inadequate sampling. This challenge cannot be resolved simply by scaling up conventional atomistic MD simulations. A promising solution involves implementing adaptive iterative multiscale simulations [39, 40], in which the CG model and the atomistic model are utilized concurrently, and the CG MLFFs are improved with the iteration. The adaptive iteration is designed so that the conformational sampling is always performed by CG simulations, whereas energy evaluations are based on all-atom force fields. Early multiscale simulations have demonstrated the potential of this adaptive optimization methodology in solving the accuracy–efficiency bottleneck in biomolecular simulations [16, 18, 39].

5 | INTEGRATIVE MOLECULAR DYNAMICS SIMULATIONS WITH AI-DERIVED DATA

The transferability and accuracy of biomolecular force fields are not always consistent. To meet the transferability requirement, the results based on conventional biomolecular force fields often exhibit deviation from experimental data when applied to specific protein systems. On the other hand, substantial low-resolution experimental data have been accumulated, including single-molecule FRET, small-angle X-ray scattering, hydrogen-deuterium exchange, infrared spectroscopy, and NMR [41–44]. Integrating such multimodal experimental constraints into MD simulations through machine learning and other data-driven algorithms significantly enhances simulation accuracy for target proteins [45–49]. This kind of integrative MD simulations demonstrates particular strengths in characterizing dynamic structural information in biomolecular events and has gained widespread applications. Recent advances in AI have substantially expanded biomolecular data dimensions. AlphaFold2 exemplifies this progress by delivering not only static structure predictions but also auxiliary metrics including pLDDT, predicted aligned error (PAE), and residue contact maps. These AI-derived outputs reveal intrinsic structural and dynamic properties, such as structural order, local flexibility, and long-range interaction signatures. These high-quality AI-derived data possess comparable values to experimental data and can be used to effectively constrain molecular simulations. In recent studies, the PAE data have already been used to dictate the molecular simulations with the elastic network model [50], and the simulation results can provide improved modeling of the protein flexibility and low-frequency mode of functionally relevant collective motions. By integrating AlphaFold-predicted residue contact probability matrix into CG simulations, researchers from Cambridge university substantially improved the conformational ensemble modeling of intrinsically disordered proteins (IDPs) [51]. Such methodologies applying system-specific constraints to transferable force field models represent a balanced solutions that simultaneously address generality requirements and computational accuracy demands.

6 | MOLECULAR DYNAMICS POWERED AI-GENERATION OF PROTEIN CONFORMATIONAL ENSEMBLES

The above discussions have highlighted the critical role of AI in advancing molecular simulation techniques. It is noteworthy that conformational ensembles and other dynamic data generated through molecular simulations can conversely serve as valuable training resources for

AI models, particularly in protein dynamics predictions. With the success of static protein structure prediction, protein dynamics prediction becomes the new focus of research efforts [52–62]. However, training AI model for dynamics prediction is much more challenging due to the lack of experimental data. The prevailing strategy to address this gap involves leveraging MD simulation data as training sets. Recently, such simulation-derived data are increasingly becoming primary reference sources for training models that predict multi-conformational states and generate conformational ensembles, including diffusion-based and flow-based generative models [38, 63–66]. These models often need to be pretrained first on AlphaFold database and PDB structures, then fine-tuned based on MD simulations and even protein stability measurements. Incorporating energy information and other physical constraints in training the model can also improve the prediction [63, 65–67].

In addition to MD-powered AI models, other strategies have also been developed to predict protein dynamics based on AI framework. For instance, by integrating the biophysical principle of protein dynamics into the AlphaFold2 structure prediction pipeline [68–71], it is possible to generate not only multiple conformations of allosteric proteins but also the pathways of their conformational motions [72]. By training a deep-learning model with the transition state structures extracted from large scale CG simulations, one can also predict the pathways of protein conformational changes [73].

7 | EXPANDING THE APPLICATION CUTTING EDGE OF BIOMOLECULAR SIMULATIONS

The development in the field of biomolecular simulations is not only reflected in the technical aspect but also in the application aspect. In fact, the advancement of biomolecular simulation technology is constantly expanding its application boundaries. Previous biomolecular simulations mainly focused on proteins, nucleic acids, and phospholipid molecules. Recently, the key role of post-translational modifications of biological macromolecules, particularly glycosylation, in structural and functional dynamics has attracted attention. For instance, it has been discovered that the sugar chain of the spike protein of SARS-CoV-2 can assist in its immune escape [74]. In addition, glycosylation modification can regulate the LLPS and amyloid aggregation of proteins related to neurodegenerative diseases [75]. The dynamic nature of the sugar chain makes glycosylation difficult to characterize with traditional structural biology techniques, and it is thus called the “dark matter” in structural biology. Naturally, MD simulations are used to assist in the structural biology characterization of glycosylation. For instance, based on the sugar chain conformation database constructed

by massive MD simulation, significant progress has been made in the predictions of glycosylation sites, sugar abundance, and glycoprotein conformations. Accordingly, tools such as GlycoSHIELD and GlycoSHAPE have been developed [75, 76]. With the development of glycosylation force fields and molecular simulation techniques, molecular dynamics will play an increasingly important role in explaining the molecular mechanism related to the structure and function of glycosylated proteins, providing key tools for the study of biological dark matter.

In addition to the research objects, the spatial scale of the simulation system is also expanding. For instance, the recent development of CG biomolecular force fields has made it possible to conduct high-throughput simulations of IDPs. Consequently, the application of biomolecular simulations has been extended to the proteomic scales. Using the CALVADOS2 force field, researchers from the University of Copenhagen simulated tens of thousands of AlphaFold2-predicted IDP regions across the human genome [20], revealing location- and function-dependent patterns in IDP conformational compactness and linking cell-level functions to the conformational properties of protein molecules. In addition, the accessible sizes of biomolecular simulation systems are also constantly expanding. For instance, by using the CG model, molecular simulations of LLPS systems containing thousands of protein chains can be achieved. These large scale simulations enable quantitative characterization of two-phase coexistence curves, droplet fusion kinetics, and microphase separation [10, 21, 77–79], which is otherwise difficult to characterize in experiments. All-atom MD simulations have also been applied to study the detailed dynamics and interaction characteristics of biological condensates at short time scales [10, 78]. In addition, molecular simulation works at the cytoplasmic and full-virus scales are also constantly emerging [9, 80, 81].

8 | DATA SHARING AND SOFTWARE DEVELOPMENT IN BIOMOLECULAR SIMULATIONS

In the era of big data, the value of biomolecular simulation data is becoming increasingly prominent, and data sharing and open access have received significant attention. Recently, over 100 biomolecular simulation experts co-authored a paper in *Nature Methods* advocating FAIR principles (findable, accessible, interoperable, reusable) for biomolecular data sharing [82]. In line with the FAIR principle, depositing MD trajectory data alongside publications has become standard practice in recent years [83]. Platforms such as CERN's Zenodo now host a vast amount of MD trajectory files. The e-Science Center at Nanjing University also offers

data-sharing services for biomolecular simulations, along with other scientific computing supports. These biomolecular simulation data not only provide valuable insights into complex biomolecular dynamics but also become an important data source for various applications based on data-driven AI technology. How to share MD data more efficiently and smoothly is worth the efforts in future studies.

In tandem with methodological innovations, MD simulation software has also rapidly evolved [84]. In addition to those established packages (e.g., GROMACS, AMBER, CHARMM, NAMD, and GENESIS) [85–89], open-source frameworks, such as CafeMol, OpenMM, and its derivatives designed to facilitate GPU acceleration, CG simulations, and MLFFs are becoming popular [90–95]. Future progress requires deeper integration of simulation and structural prediction tools within unified open-source ecosystems, exemplified by platforms such as MindSpore SPONGE [96]. Meanwhile, advances in quantum computing and specialized hardware are expected to bring transformative breakthroughs for the biomolecular modeling field.

AUTHOR CONTRIBUTIONS

Wenfei Li: Writing—original draft; writing—review and editing; visualization; funding acquisition. **Wei Wang:** Project administration; writing—review and editing; funding acquisition.

ACKNOWLEDGMENTS

The authors acknowledge the valuable and inspiring discussions at the meeting of the Committee on Biomolecular Structure Prediction and Simulation, Bioinformatics Society of China. We also thank the support from the National Natural Science Foundation of China (Grant Nos. 12574224, 12090052 and 12347102).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

This work does not produce any new data.

ETHICS STATEMENT

This perspective article does not contain the studies related to human or animal subjects.

REFERENCES

- [1] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–9.
- [2] Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;630(8016):493–500.
- [3] Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021;373(6557):871–6.

- [4] Yang J, Anishchenko I, Park H, Peng Z, Ovchinnikov S, Baker D. Improved protein structure prediction using predicted inter-residue orientations. *Proc Natl Acad Sci*. 2020;117(3):1496–503.
- [5] Peng Z, Wang W, Han R, Zhang F, Yang J. Protein structure prediction in the deep learning era. *Curr Opin Struct Biol*. 2022;77:102495.
- [6] Martin J, Lequerica Mateos M, Onuchic JN, Coluzza I, Morcos F. Machine learning in biological physics: from biomolecular prediction to design. *Proc Natl Acad Sci*. 2024;121(27):e2311807121.
- [7] McCammon J, Gelin B, Karplus M. Dynamics of folded proteins. *Nature*. 1977;267(5612):585–90.
- [8] Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How fast-folding proteins fold. *Science*. 2011;334(6055):517–20.
- [9] Coshic K, Maffeo C, Winogradoff D, Aksimentiev A. The structure and physical properties of a packaged bacteriophage particle. *Nature*. 2024;627(8005):905–14.
- [10] Galvanetto N, Ivanović MT, Chowdhury A, Sottini A, Nüesch MF, Nettels D, et al. Extreme dynamics in a biomolecular condensate. *Nature*. 2023;619(7971):876–83.
- [11] Xue L, Yan N, Song C. Deciphering Ca^{2+} permeation and valence selectivity in CaV1 : molecular dynamics simulations reveal the three-ion knock-on mechanism. *Proc Natl Acad Sci*. 2025;122(22):e2424694122.
- [12] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem*. 2004;25(9):1157–74.
- [13] Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, De Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods*. 2017;14(1):71–3.
- [14] Grant BJ, Gorte AA, McCammon JA. Large conformational changes in proteins: signaling and other functions. *Curr Opin Struct Biol*. 2010;20(2):142–7.
- [15] Takada S, Kanada R, Tan C, Terakawa T, Li W, Kenzaki H. Modeling structural dynamics of biomolecular complexes by coarse-grained molecular simulations. *Acc Chem Res*. 2015;48(12):3026–35.
- [16] Li W, Zhang J, Wang J, Wang W. Multiscale theory and computational method for biomolecule simulations. *Acta Phys Sin*. 2015;64(9):098701.
- [17] Li W, Wang J, Zhang J, Wang W. Molecular simulations of metal-coupled protein folding. *Curr Opin Struct Biol*. 2015;30:25–31.
- [18] Li W, Yoshii H, Hori N, Kameda T, Takada S. Multiscale methods for protein folding simulations. *Methods*. 2010;52(1):106–14.
- [19] Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink SJ. The MARTINI coarse-grained force field: extension to proteins. *J Chem Theor Comput*. 2008;4(5):819–34.
- [20] Tesei G, Trolle AI, Jonsson N, Betz J, Knudsen FE, Pesce F, et al. Conformational ensembles of the human intrinsically disordered proteome. *Nature*. 2024;626(8000):897–904.
- [21] Dignon GL, Zheng W, Kim YC, Best RB, Mittal J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput Biol*. 2018;14(1):e1005941.
- [22] Tozzini V. Coarse-grained models for proteins. *Curr Opin Struct Biol*. 2005;15(2):144–50.
- [23] Darré L, Machado MR, Brandner AF, González HC, Ferreira S, Pantano S. SIRAH: a structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. *J Chem Theor Comput*. 2015;11(2):723–39.
- [24] Ingólfsson HI, Bhatia H, Aydin F, Oppelstrup T, López CA, Stanton LG, et al. Machine learning-driven multiscale modeling: bridging the scales with a next-generation simulation infrastructure. *J Chem Theor Comput*. 2023;19(9):2658–75.
- [25] Wang Y, Zhou R. A practical guide to recent advances in multiscale modeling and simulation of biomolecules. AIP Publishing LLC; 2023.
- [26] Wang T, He X, Li M, Li Y, Bi R, Wang Y, et al. Ab initio characterization of protein molecular dynamics with Al^2BMD . *Nature*. 2024;635(8040):1019–27.
- [27] Guan XY, Huang HY, Peng HQ, Liu YH, Li W, Wang W. Machine learning in molecular simulations of biomolecules. *Acta Phys Sin*. 2023;72(24):248708.
- [28] Unke OT, Stöhr M, Ganscha S, Unterthiner T, Maennel H, Kashubin S, et al. Biomolecular dynamics with machine-learned quantum-mechanical force fields trained on diverse chemical fragments. *Sci Adv*. 2024;10(14):eadn4397.
- [29] Poma AB, Caldas AH, Cofas-Vargas LF, Jones MS, Ferguson AL, Sandonas LM. Recent advances in machine learning and coarse-grained potentials for biomolecular simulations and their applications. *Biophys J*. 2025.
- [30] Lappala A. The next revolution in computational simulations: harnessing AI and quantum computing in molecular dynamics. *Curr Opin Struct Biol*. 2024;89:102919.
- [31] Liu J, Lan J, He X. Toward high-level machine learning potential for water based on quantum fragmentation and neural networks. *J Phys Chem A*. 2022;126(24):3926–36.
- [32] Li S, Li W, Ma J. Generalized energy-based fragmentation approach and its applications to macromolecules and molecular aggregates. *Acc Chem Res*. 2014;47(9):2712–20.
- [33] He X, Zhu T, Wang X, Liu J, Zhang JZ. Fragment quantum mechanical calculation of proteins and its applications. *Acc Chem Res*. 2014;47(9):2748–57.
- [34] Liu J, He X. Recent advances in quantum fragmentation approaches to complex molecular and condensed-phase systems. *Wiley Interdiscip Rev Comput Mol Sci*. 2023;13(3):e1650.
- [35] Charron NE, Bonneau K, Pasos-Trejo AS, Guljas A, Chen Y, Musil F, et al. Navigating protein landscapes with a machine learned transferable coarse-grained model. *Nat Chem*. 2025;89(8):1284–92.
- [36] Navarro C, Majewski M, De Fabritiis G. Top-down machine learning of coarse-grained protein force fields. *J Chem Theor Comput*. 2023;19(21):7518–26.
- [37] Kanada R, Tokuhisa A, Nagasaka Y, Okuno S, Amemiya K, Chiba S, et al. Enhanced coarse-grained molecular dynamics simulation with a smoothed hybrid potential using a neural network model. *J Chem Theor Comput*. 2023;20(1):7–17.
- [38] Zhu J, Li Z, Zheng Z, Zhang B, Zhong B, Bai J, et al. Precise generation of conformational ensembles for intrinsically disordered proteins via fine-tuned diffusion models. 2024. Preprint at bioRxiv: 2024.05.05.592611.
- [39] Li W, Takada S. Self-learning multiscale simulation for achieving high accuracy and high efficiency simultaneously. *J Chem Phys*. 2009;130(21):214108.
- [40] Zhong Q, Li G. Adaptively iterative multiscale switching simulation strategy and applications to protein folding and structure prediction. *J Phys Chem Lett*. 2021;12(12):3151–62.
- [41] Aznauryan M, Delgado L, Soranno A, Nettels D, Huang J, Labhardt AM, et al. Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS. *Proc Natl Acad Sci*. 2016;113(37):E5389–98.
- [42] James EI, Murphree TA, Vorauer C, Engen JR, Guttman M. Advances in hydrogen/deuterium exchange mass spectrometry and the pursuit of challenging biological systems. *Chem Rev*. 2021;122(8):7562–623.
- [43] Lerner E, Barth A, Hendrix J, Ambrose B, Birkedal V, Blanchard SC, et al. FRET-Based dynamic structural biology: challenges, perspectives and an appeal for open-science practices. *eLife*. 2021;10:e60416.
- [44] Tumbic GW, Hossan MY, Thielges MC. Protein dynamics by two-dimensional infrared spectroscopy. *Annu Rev Anal Chem*. 2021;14(1):299–321.

- [45] Boomsma W, Ferkinghoff-Borg J, Lindorff-Larsen K. Combining experiments and simulations using the maximum entropy principle. *PLoS Comput Biol*. 2014;10(2):e1003406.
- [46] Latham AP, Zhang B. Maximum entropy optimized force field for intrinsically disordered proteins. *J Chem Theor Comput*. 2019;16(1):773–81.
- [47] Fröhking T, Bernetti M, Bussi G. Simultaneous refinement of molecular dynamics ensembles and forward models using experimental data. *J Chem Phys*. 2023;158(21):214120.
- [48] Zhang B, Wolynes PG. Topology, structures, and energy landscapes of human chromosomes. *Proc Natl Acad Sci*. 2015;112(19):6062–7.
- [49] Zhu W, Li W, Wang W. Modeling hydrogen exchange of proteins by a multiscale method. *Chin Phys B*. 2021;30(7):078701.
- [50] Jussupow A, Kaila VR. Effective molecular dynamics from neural network-based structure prediction models. *J Chem Theor Comput*. 2023;19(7):1965–75.
- [51] Brotzakis ZF, Zhang S, Murtada MH, Vendruscolo M. AlphaFold prediction of structural ensembles of disordered proteins. *Nat Commun*. 2025;16(1):1632.
- [52] Wayment-Steele HK, Ojoawo A, Otten R, Apitz JM, Pitsawong W, Hömberger M, et al. Predicting multiple conformations via sequence clustering and AlphaFold2. *Nature*. 2024;625(7996):832–9.
- [53] Tang Y, Yu M, Bai G, Li X, Xu Y, Ma B. Deep learning of protein energy landscape and conformational dynamics from experimental structures in PDB. 2024. Preprint at bioRxiv: 2024.06.27.600251.
- [54] Xu J, Wang Y. Generating multistate conformations of P-type ATPases with a conditional diffusion model. *J Chem Inf Model*. 2024;64(24):9227–39.
- [55] Hou M, Jin S, Cui X, Peng C, Zhao K, Song L, et al. Protein multiple conformation prediction using multi-objective evolution algorithm. *Interdiscipl Sci Comput Life Sci*. 2024;16(3):519–31.
- [56] Li J, Wang L, Zhu Z, Song C. Exploring the alternative conformation of a known protein structure based on contact map prediction. *J Chem Inf Model*. 2023;64(1):301–15.
- [57] Xie T, Huang J. Can protein structure prediction methods capture alternative conformations of membrane transporters? *J Chem Inf Model*. 2024;64(8):3524–36.
- [58] Del Alamo D, Sala D, Mchaourab HS, Meiler J. Sampling alternative conformational states of transporters and receptors with AlphaFold2. *eLife*. 2022;11:e75751.
- [59] Monteiro daSG, Cui JY, Dalgarno DC, Lisi GP, Rubenstein BM. High-throughput prediction of protein conformational distributions with subsampled AlphaFold2. *Nat Commun*. 2024;15(1):2464.
- [60] Gu X, Aranganathan A, Tiwary P. Empowering AlphaFold2 for protein conformation selective drug discovery with AlphaFold2-RAVE. *eLife*. 2024;13:RP99702.
- [61] Stein RA, Mchaourab HS. SPEACH_AF: sampling protein ensembles and conformational heterogeneity with Alphafold2. *PLoS Comput Biol*. 2022;18(8):e1010483.
- [62] Lee M, Schafer JW, Prabakaran J, Chakravarty D, Clore MF, Porter LL. Large-scale predictions of alternative protein conformations by AlphaFold2-based sequence association. *Nat Commun*. 2025;16(1):5622.
- [63] Lewis S, Hempel T, Jiménez-Luna J, Gastegger M, Xie Y, Foong AY, et al. Scalable emulation of protein equilibrium ensembles with generative deep learning. *Science*. 2025;389(6761):eadv9817.
- [64] Jin Y, Huang Q, Song Z, Zheng M, Teng D, Shi Q. P2dflow: a protein ensemble generative model with se (3) flow matching. *J Chem Theor Comput*. 2025;21(6):3288–96.
- [65] Zheng S, He J, Liu C, Shi Y, Lu Z, Feng W, et al. Predicting equilibrium distributions for molecular systems with deep learning. *Nat Mach Intell*. 2024;6(5):558–67.
- [66] Passaro S, Corso G, Wohlwend J, Reveiz M, Thaler S, Ram Somnath V, et al. Boltz-2: towards accurate and efficient binding affinity prediction. 2025. Preprint at bioRxiv: 2025.06.14.659707.
- [67] Xiang C, Wang W, Peng Z, Yang J. Generating dynamic structures through physics-based sampling of trRosettaX2-predicted inter-residue geometries. 2025. Preprint at bioRxiv: 2025.05.28.656531.
- [68] Li W, Wolynes PG, Takada S. Frustration, specific sequence dependence, and nonlinearity in large-amplitude fluctuations of allosteric proteins. *Proc Natl Acad Sci*. 2011;108(9):3504–9.
- [69] Ferreira DU, Hegler JA, Komives EA, Wolynes PG. On the role of frustration in the energy landscapes of allosteric proteins. *Proc Natl Acad Sci*. 2011;108(9):3499–503.
- [70] Li W, Wang J, Zhang J, Takada S, Wang W. Overcoming the bottleneck of the enzymatic cycle by steric frustration. *Phys Rev Lett*. 2019;122(23):238102.
- [71] Onuchic JN, Luthey-Schulten Z, Wolynes PG. Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem*. 1997;48(1):545–600.
- [72] Guan X, Tang QY, Ren W, Chen M, Wang W, Wolynes PG, et al. Predicting protein conformational motions using energetic frustration analysis and AlphaFold2. *Proc Natl Acad Sci*. 2024;121(35):e2410662121.
- [73] Hu Y, Yang H, Li M, Zhong Z, Zhou Y, Bai F, et al. Exploring protein conformational changes using a large-scale biophysical sampling augmented deep learning strategy. *Adv Sci*. 2024;11(44):2400884.
- [74] Souza PF, Mesquita FP, Amaral JL, Landim PG, Lima KR, Costa MB, et al. The spike glycoprotein of SARS-CoV-2: a review of how mutations of spike glycoproteins have driven the emergence of variants with high transmissibility and immune escape. *Int J Biol Macromol*. 2022;208:105–25.
- [75] Ives CM, Singh O, D'Andrea S, Fogarty CA, Harbison AM, Satheesan A, et al. Restoring protein glycosylation with GlycoShape. *Nat Methods*. 2024;21(11):2117–27.
- [76] Tsai YX, Chang NE, Reuter K, Chang HT, Yang TJ, von Bülow S, et al. Rapid simulation of glycoprotein structures by grafting and steric exclusion of glycan conformer libraries. *Cell*. 2024;187(5):1296–311.
- [77] Bian Y, Lv F, Pan H, Ren W, Zhang W, Wang Y, et al. Fusion dynamics and size-dependence of droplet microstructure in ssDNA-mediated protein phase separation. *JACS Au*. 2024;4(9):3690–704.
- [78] Galvanetto N, Ivanović MT, Del Grosso SA, Chowdhury A, Sottini A, Nettels D, et al. Material properties of biomolecular condensates emerge from nanoscale dynamics. *Proc Natl Acad Sci*. 2025;122(23):e2424135122.
- [79] Yu Y, Liu Q, Zeng J, Tan Y, Tang Y, Wei G. Multiscale simulations reveal the driving forces of p53C phase separation accelerated by oncogenic mutations. *Chem Sci*. 2024;15(32):12806–18.
- [80] Lynch DL, Pavlova A, Fan Z, Gumbart JC. Understanding virus structure and dynamics through molecular simulations. *J Chem Theor Comput*. 2023;19(11):3025–36.
- [81] Yu I, Mori T, Ando T, Harada R, Jung J, Sugita Y, et al. Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *eLife*. 2016;5:e19274.
- [82] Amaro RE, Åqvist J, Bahar I, Battistini F, Bellaiche A, Beltrán D, et al. The need to implement FAIR principles in biomolecular simulations. *Nat Methods*. 2025;22(4):641–5.

- [83] Tiemann J, Szczuka M, Bouarroudj L, Oussaren M, Garcia S, Howard R, et al. MDverse: shedding light on the dark matter of molecular dynamics simulations. *eLife*. 2023;12:RP90061.
- [84] Carrasco-Busturia D, Ippoliti E, Meloni S, Rothlisberger U, Olsen JMH. Multiscale biomolecular simulations in the exascale era. *Curr Opin Struct Biol*. 2024;86:102821.
- [85] Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *J Comput Chem*. 2005;26(16):1701–18.
- [86] Brooks BR, Brooks CL III, Mackerell JAD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. *J Comput Chem*. 2009;30(10):1545–614.
- [87] Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz JKM, et al. The Amber biomolecular simulation programs. *J Comput Chem*. 2005;26(16):1668–88.
- [88] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26(16):1781–802.
- [89] Jung J, Mori T, Kobayashi C, Matsunaga Y, Yoda T, Feig M, et al. GENESIS: a hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations. *Wiley Interdiscip Rev Comput Mol Sci*. 2015;5(4):310–23.
- [90] Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, et al. OpenMM 7: rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol*. 2017;13(7):e1005659.
- [91] Anderson JA, Glaser J, Glotzer SC. HOOMD-blue: a Python package for high-performance molecular dynamics and hard particle Monte Carlo simulations. *Comput Mater Sci*. 2020;173:109363.
- [92] Kenzaki H, Koga N, Hori N, Kanada R, Li W, Ki O, et al. CafeMol: a coarse-grained biomolecular simulator for simulating proteins at work. *J Chem Theor Comput*. 2011;7(6):1979–89.
- [93] Shu Z, Wu M, Liao J, Chen C. Fsatoool 2.0: an integrated molecular dynamics simulation and trajectory data analysis program. *J Comput Chem*. 2022;43(3):215–24.
- [94] Davtyan A, Schafer NP, Zheng W, Clementi C, Wolynes PG, Papoian GA. AWSEM-MD: protein structure prediction using coarse-grained physical potentials and bioinformatically based local structure biasing. *J Phys Chem B*. 2012;116(29):8494–503.
- [95] Liu S, Wang C, Latham AP, Ding X, Zhang B. OpenABC enables flexible, simplified, and efficient GPU accelerated simulations of biomolecular condensates. *PLoS Comput Biol*. 2023;19(9):e1011442.
- [96] Huang YP, Xia Y, Yang L, Wei J, Yang YI, Gao YQ. SPONGE: a GPU-accelerated molecular dynamics package with enhanced sampling and AI-driven algorithms. *Chin J Chem*. 2022;40(1):160–8.

How to cite this article: Li W, Wang W. Advances and challenges in multiscale biomolecular simulations: artificial intelligence-driven paradigm shift. *Quantitative Biology*. 2026; e70024. <https://doi.org/10.1002/qub2.70024>