JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim

Editorial

Applications of Free-Energy Calculations to Biomolecular Processes. A Collection

Cite This: *J. Chem. Inf. Model.* 2024, 64, 2129−2131

Read Online

ACCESS | Metrics & More | Article Recommendations

INTRODUCTION

The calculation of accurate free energies lies at the core of investigating biomolecular processes such as protein−ligand binding, protein−antibody interactions, enzymatic reactions, and solvation by means of computer simulations. Given their prevalence in cellular machinery, many of these processes have significant implications not only in various aspects of biophysics but also in the rational design of drugs. Consequently, it is highly desirable to predict their feasibility using theoretical tools. In recent decades, substantial strides have been made in determining free energies through the application of numerical simulations grounded in the fundamental principles of statistical mechanics. While the accessibility of computational resources, in particular the rapid development of fast and affordable graphics processing units, has contributed to these advancements, the primary catalyst for progress lies in the development of a diverse array of methods. These methods have effectively enhanced the efficiency and reliability of calculations related to free energy, opening new perspectives for the computational study of a variety of complex problems hitherto intractable. For this Collection, we have assembled 22 papers recently published in *Journal of Chemical Information and Modeling* and *The Journal of Physical Chemistry B* that serve as excellent examples of insight gained from free-energy calculations. Below, we briefly highlight the individual papers included in this Collection.

ALCHEMICAL FREE-ENERGY CALCULATIONS FOR DRUG DISCOVERY

Free-energy calculations are a valuable tool for drug design and in particular for lead optimization. Calculating the free energy of the ligand−protein binding process directly is an extremely daunting task involving diffusion, desolvation, and conformational changes of both ligand and protein. However, relative binding free energies (RBFEs) between congeneric ligands or absolute binding free energies (ABFEs) can be calculated relatively efficiently using a thermodynamic cycle that facilitates small perturbations of phase space. The evolution of alchemical free-energy calculations from their theoretical foundations and early applications to recent developments has recently been reviewed.[1] Nowadays, RBFE calculations are used both retrospectively for the validation of different methods and protocols and also prospectively for lead optimization and drug design.

Relative free-energy perturbation (FEP) and multisite lambda dynamics (MSλD) were applied to a series of inhibitors for the bromodomain-containing protein 4 (BRD4).[2] Guest et al. found that relative FEP and MSλD predicted binding free energies with comparable accuracy (∼0.6 kcal/mol for each method), but MSλD required 18-fold less simulation time for the entire molecule space.

Calculation of ABFE for a protein−ligand complex in solution is more challenging compared to RBFE calculations because of the increased needs in configurational sampling. The Jorgensen group applied both Monte Carlo and molecular dynamics (MD) simulations for sampling coupled with FEP calculations to estimate the ΔG_{bind} for the complex of a druglike inhibitor (MIF180) with the protein macrophage migration inhibitory factor (MIF). They found that calculation of accurate ΔG_{bind} for large ligands is both feasible and numerically equivalent using both sampling techniques.[3] In another study from the same group, a set of 16 inhibitors of the SARS-CoV-2 main protease (Mpro) with structural diversity, well-distributed affinities, and high-resolution crystal structures were subjected to MD/FEP.[4] The experimental and computed results showed encouraging correlation, indicating that FEP-based ABFE calculations can be employed for prospective hit identification. In another study from Chen et al., the ABFEP method was tested on eight congeneric compound series, including both neutral and charged ligands, binding to eight protein receptors, which demonstrated the validity of the method[5] and was further found to improve the hit rates as compared to docking scores or other methods like metadynamics. In a recent study from Ngo et al. calculating the binding affinity of known ligands to SARS-CoV-2 Mpro, it was demonstrated that, among free-energy methods, FEP using double annihilation is the most accurate method, followed by linear interaction energy (LIE), FPL, and MM-PBSA (FEP > LIE ≈ FPL > MM-PBSA).[6]

FEP calculations were also performed by Beuming et al. in the context of prospective AlphaFold2 protein structures, and the binding affinities arising from relative FEP calculations were comparable in accuracy to the corresponding calculations previously carried out using crystal structures.[7] Therefore, AlphaFold2-modeled structures could be accurate enough to

Published: April 8, 2024

https://doi.org/10.1021/acs.jcim.4c00349
J. Chem. Inf. Model. 2024, 64, 2129−2131

be used by physics-based methods such as FEP in typical lead optimization stages of a drug discovery program.

METHODOLOGICAL DEVELOPMENTS FOR FREE-ENERGY CALCULATIONS

Alchemical free-energy calculations have been limited by technical challenges such as the manual creation of large numbers of input files to set up, run, and analyze. Zhang et al.[8] developed the CHARMM-GUI Free Energy Calculator, a user-friendly web-based tool to generate the AMBER-TI system and input files for high-throughput binding free-energy calculations. An automated workflow for setting up free-energy calculations and analysis (ProFESSA) using the GPU-accelerated AMBER free-energy engine with enhanced sampling features and analysis tools has also been developed by Ganguly et al. as part of the AMBER Drug Discovery Boost package (AMBER22 release).[9] To streamline the optimal perturbation path in alchemical free-energy calculations, Petrov developed an automatic perturbation topology builder based on a graph-matching algorithm, which can identify the maximum common substructure (MCS) of two or multiple molecules and provide the perturbation topologies suitable for free-energy calculations using the GROMOS and GROMACS simulation packages.[10]

Replica-exchange enveloping distribution sampling (RE-EDS) allows for the sampling of multiple end-states in a single simulation without the specification of any pathways. In a recent GROMOS/GAFF2 implementation of this scheme, Rieder et al. showed that the hydration free energies obtained using RE-EDS for multiple molecules are found to be in good agreement with both the experimental data and the results calculated using other free-energy methods.[11]

Another method from Azimi et al., the alchemical transfer method (ATM-RBFE) for the estimation of relative binding free energies of molecular complexes, was validated against absolute binding free energies of the SAMPL8 GDCC host–guest benchmark set and against protein–ligand benchmark sets, and it was found that the method yields self-consistent and converged relative binding free-energy estimates in agreement with absolute binding free energies, reference literature values, and experimental measurements.[12]

To address the sampling problem during an alchemical free-energy calculation, a nonequilibrium candidate Monte Carlo (NCMC) method was developed by Gill et al. and applied to improve sampling of ligand binding modes, demonstrating a 2 orders of magnitude improvement in binding mode sampling efficiency compared to a brute force MD simulation.[13]

MACHINE LEARNING AND FREE-ENERGY CALCULATIONS

On the machine learning front, a data-driven approach incorporating the hydration free energies obtained using 3D-RISM, termed 3D-RISM-AI, was developed by Osaki et al., leveraging several physicochemical descriptors.[14] In another study by Gusev et al., it was demonstrated that the combination of active learning with automated machine learning and free-energy calculations yields at least 20-fold speedup relative to naïve brute force approaches.[15] Finally, a new strategy from Akkus et al. to estimate binding free energies using end-state molecular dynamics simulation based on LIE and ANI-2x neural network potentials predicts the single-point interaction energies between ligand–protein and ligand–solvent pairs at the accuracy of the wb97x/6-31G* level.[16]

PROTEIN–PROTEIN INTERACTIONS

Many essential operations within the cell machinery hinge on the recognition and association of proteins. Disturbing the intricate web of interactions responsible for creating protein–protein complexes can result in a variety of pathologies. Binding affinities, which reflect the inherent tendency of molecular entities to unite, are crucial thermodynamic factors for unraveling the intricacies of recognition and association phenomena and uncovering potential malfunctions. This holds true for numerous proteins, where the process of oligomerization is a necessary step for them to carry out their biological functions.

Employing multi-μs biased and unbiased computer simulations, Dutta et al.[17] investigated how the spike protein of SARS-CoV-2 and of its successive variants interacts through its receptor binding domain with the human angiotensin conversion enzyme 2 protein.

In close relation with the latter work, Nguyen et al.[18] turned to steered molecular dynamics in conjunction with coarse-grained models to estimate the binding affinity of the monoclonal antibodies CR3022 and 4A8 to the SARS-CoV-2 receptor-binding domain (RBD) and SARS-CoV-2 N-terminal domain (NTD).

FREE ENERGIES DETERMINED FROM QM/MM AND REACTIVE MD SIMULATIONS

Quantum-mechanical/molecular-mechanical (QM/MM) simulations have evolved into an essential tool for the exploration of a gamut of chemical processes, chief among which are enzymatic reactions. In this particular research area, QM/MM simulations have not only allowed the underlying molecular mechanisms to be dissected, but they have also revealed the different factors influencing catalysis, thereby helping the design of enzyme inhibitors, or the analysis and enhancement of engineered enzymes. Still, despite their remarkable success, QM/MM methods, most notably those tailored for biomolecular systems, face significant challenges. In a comprehensive Perspective from Cui et al.,[19] these ongoing challenges are discussed exhaustively, alongside the current avenues to address them.

In the context of alchemical transformations performed in the framework of QM/MM simulations, determination of the free-energy difference between the low level, MM, and the high level, QM, of theory constitutes a computational challenge. One possible route to meet this challenge consists of turning to nonequilibrium work switching. To optimize the computational cost associated with this undertaking, Scholler et al. explored several strategies of short versus long switches in conjunction with the Crooks fluctuation theorem and principal component analysis to identify convergence failures.[20]

An important aspect of free-energy calculations leaning on QM/MM methods is the search of transition pathways, most notably for enzymatic reactions. Combining QM/MM simulations with the string algorithm, Yagi et al. have determined the minimum-energy pathway for proton-transfer processes.[21] The original implementation of their methodology and the level of theory utilized made possible the mapping of the free-energy landscape underlying the reactive process of interest.

https://doi.org/10.1021/acs.jcim.4c00349
J. Chem. Inf. Model. 2024, 64, 2129−2131

At the other end of the spectrum of free-energy calculations involving a QM description of the reactive process at play, ab initio molecular dynamics is well-suited for simulating chemical reactions, because it dynamically solves the electronic Schrödinger equation, without depending on a predetermined, fixed bonding structure of the system. A significantly more computationally efficient approach is provided by multiscale reactive molecular dynamics, which, in combination with diabatic matching, has been employed by Li and Voth to investigate proton dissociation reactions of amino acids in both water and protein.[22] This methodology allows the response to the environment change to be characterized in terms of accurate pK_a shift determination.

CALL FOR PAPERS

Journal of Chemical Information and Modeling and *The Journal of Physical Chemistry B* will jointly publish a Virtual Special Issue on this topic in late 2024/early 2025. More information on submitting to this Virtual Special Issue can be found in the accompanying ACS Axial post.

Zoe Cournia orcid.org/0000-0001-9287-364X
Christophe Chipot orcid.org/0000-0002-9122-1698

AUTHOR INFORMATION

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.4c00349

Notes

Views expressed in this editorial are those of the authors and not necessarily the views of the ACS.

This Editorial is jointly published in *The Journal of Physical Chemistry B* and *Journal of Chemical Information and Modeling*.

REFERENCES

(1) Song, L. F.; Merz, K. M. Evolution of Alchemical Free Energy Methods in Drug Discovery. *J. Chem. Inf Model* **2020**, *60*, 5308−5318.
(2) Guest, E. E.; Cervantes, L. F.; Pickett, S. D.; Brooks, C. L.; Hirst, J. D. Alchemical Free Energy Methods Applied to Complexes of the First Bromodomain of BRD4. *J. Chem. Inf Model* **2022**, *62*, 1458−1470.
(3) Qian, Y.; Cabeza de Vaca, I.; Vilseck, J. Z.; Cole, D. J.; Tirado-Rives, J.; Jorgensen, W. L. Absolute Free Energy of Binding Calculations for Macrophage Migration Inhibitory Factor in Complex with a Druglike Inhibitor. *J. Phys. Chem. B* **2019**, *123*, 8675−8685.
(4) Ghahremanpour, M. M.; Saar, A.; Tirado-Rives, J.; Jorgensen, W. L. Computation of Absolute Binding Free Energies for Noncovalent Inhibitors with SARS-CoV-2 Main Protease. *J. Chem. Inf Model* **2023**, *63*, 5309−5318.
(5) Chen, W.; Cui, D.; Jerome, S. V.; Michino, M.; Lenselink, E. B.; Huggins, D. J.; Beautrait, A.; Vendome, J.; Abel, R.; Friesner, R. A.; Wang, L. Enhancing Hit Discovery in Virtual Screening through Absolute Protein-Ligand Binding Free-Energy Calculations. *J. Chem. Inf Model* **2023**, *63*, 3171−3185.
(6) Ngo, S. T.; Tam, N. M.; Pham, M. Q.; Nguyen, T. H. Benchmark of Popular Free Energy Approaches Revealing the Inhibitors Binding to SARS-CoV-2 Mpro. *J. Chem. Inf Model* **2021**, *61*, 2302−2312.
(7) Beuming, T.; Martin, H.; Diaz-Rovira, A. M.; Diaz, L.; Guallar, V.; Ray, S. S. Are Deep Learning Structural Models Sufficiently Accurate for Free-Energy Calculations? Application of FEP+ to AlphaFold2-Predicted Structures. *J. Chem. Inf Model* **2022**, *62*, 4351−4360.
(8) Zhang, H.; Kim, S.; Giese, T. J.; Lee, T. S.; Lee, J.; York, D. M.; Im, W. CHARMM-GUI Free Energy Calculator for Practical Ligand Binding Free Energy Simulations with AMBER. *J. Chem. Inf Model* **2021**, *61*, 4145−4151.
(9) Ganguly, A.; Tsai, H. C.; Fernandez-Pendas, M.; Lee, T. S.; Giese, T. J.; York, D. M. AMBER Drug Discovery Boost Tools: Automated Workflow for Production Free-Energy Simulation Setup and Analysis (ProFESSA). *J. Chem. Inf Model* **2022**, *62*, 6069−6083.
(10) Petrov, D. Perturbation Free-Energy Toolkit: An Automated Alchemical Topology Builder. *J. Chem. Inf Model* **2021**, *61*, 4382−4390.
(11) Rieder, S. R.; Ries, B.; Schaller, K.; Champion, C.; Barros, E. P.; Hunenberger, P. H.; Riniker, S. Replica-Exchange Enveloping Distribution Sampling Using Generalized AMBER Force-Field Topologies: Application to Relative Hydration Free-Energy Calculations for Large Sets of Molecules. *J. Chem. Inf Model* **2022**, *62*, 3043−3056.
(12) Azimi, S.; Khuttan, S.; Wu, J. Z.; Pal, R. K.; Gallicchio, E. Relative Binding Free Energy Calculations for Ligands with Diverse Scaffolds with the Alchemical Transfer Method. *J. Chem. Inf Model* **2022**, *62*, 309−323.
(13) Gill, S. C.; Lim, N. M.; Grinaway, P. B.; Rustenburg, A. S.; Fass, J.; Ross, G. A.; Chodera, J. D.; Mobley, D. L. Binding Modes of Ligands Using Enhanced Sampling (BLUES): Rapid Decorrelation of Ligand Binding Modes via Nonequilibrium Candidate Monte Carlo. *J. Phys. Chem. B* **2018**, *122*, 5579−5598.
(14) Osaki, K.; Ekimoto, T.; Yamane, T.; Ikeguchi, M. 3D-RISM-AI: A Machine Learning Approach to Predict Protein-Ligand Binding Affinity Using 3D-RISM. *J. Phys. Chem. B* **2022**, *126*, 6148−6158.
(15) Gusev, F.; Gutkin, E.; Kurnikova, M. G.; Isayev, O. Active Learning Guided Drug Design Lead Optimization Based on Relative Binding Free Energy Modeling. *J. Chem. Inf Model* **2023**, *63*, 583−594.
(16) Akkus, E.; Tayfuroglu, O.; Yildiz, M.; Kocak, A. Accurate Binding Free Energy Method from End-State MD Simulations. *J. Chem. Inf Model* **2022**, *62*, 4095−4106.
(17) Dutta, S.; Panthi, B.; Chandra, A. All-Atom Simulations of Human ACE2-Spike Protein RBD Complexes for SARS-CoV-2 and Some of its Variants: Nature of Interactions and Free Energy Diagrams for Dissociation of the Protein Complexes. *J. Phys. Chem. B* **2022**, *126*, 5375−5389.
(18) Nguyen, H.; Lan, P. D.; Nissley, D. A.; O'Brien, E. P.; Li, M. S. Electrostatic Interactions Explain the Higher Binding Affinity of the CR3022 Antibody for SARS-CoV-2 than the 4A8 Antibody. *J. Phys. Chem. B* **2021**, *125*, 7368−7379.
(19) Cui, Q.; Pal, T.; Xie, L. Biomolecular QM/MM Simulations: What Are Some of the "Burning Issues. *J. Phys. Chem. B* **2021**, *125*, 689−702.
(20) Scholler, A.; Kearns, F.; Woodcock, H. L.; Boresch, S. Optimizing the Calculation of Free Energy Differences in Non-equilibrium Work SQM/MM Switching Simulations. *J. Phys. Chem. B* **2022**, *126*, 2798−2811.
(21) Yagi, K.; Ito, S.; Sugita, Y. Exploring the Minimum-Energy Pathways and Free-Energy Profiles of Enzymatic Reactions with QM/MM Calculations. *J. Phys. Chem. B* **2021**, *125*, 4701−4713.
(22) Li, C.; Voth, G. A. Accurate and Transferable Reactive Molecular Dynamics Models from Constrained Density Functional Theory. *J. Phys. Chem. B* **2021**, *125*, 10471−10480.