AmberTools

David A. Case,[a] Hasan Metin Aktulga,[b] Kellon Belfon,[c] David S. Cerutti,[d] Gerardo Andrés Cisneros,[e] Vinícius Wilian D. Cruzeiro,[f] Negin Forouzesh,[g] Timothy J. Giese,[h] Andreas W. Götz,[i] Holger Gohlke,[j,k] Saeed Izadi,[l] Koushik Kasavajhala,[m] Mehmet C. Kaymak,[b] Edward King,[n] Thomas Kurtzman,[o] Tai-Sung Lee,[h] Pengfei Li,[p] Jian Liu,[q] Tyler Luchko,[r] Ray Luo,[n] Madushanka Manathunga,[s] Matias R. Machado,[t] Hai Minh Nguyen,[a] Kurt A. O'Hearn,[b] Alexey Onufriev,[u] Feng Pan,[v] Sergio Pantano,[t] Ruxi Qi,[w] Ali Rahnamoun,[b] Ali Risheh,[g] Stephan Schott-Verdugo,[k] Akhil Shajan,[s] Jason Swails,[x] Junmei Wang,[y] Haixin Wei,[n] Xiongwu Wu,[z] Yongxian Wu,[n] Shi Zhang,[h] Shiji Zhao,[n,aa] Qiang Zhu,[n] Thomas E. Cheatham, III,[ab] Daniel R. Roe,[ac] Adrian Roitberg,[ad] Carlos Simmerling,[m] Darrin M. York,[h] Maria C. Nagan,*[,ae] and Kenneth M. Merz, Jr.*[,s]

a) Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854 USA; b) Department of Computer Science and Engineering, Michigan State University, East Lansing, Michigan 48824-1322, United States; c) FOG Pharmaceuticals Inc., Cambridge, MA 02140; d) Psivant, 451 D Street, Suite 205, Boston MA 02210; e) Department of Physics, Department of Chemistry and Biochemistry, University of Texas at Dallas, Richardson, TX 75801; f) Department of Chemistry and The PULSE Institute, Stanford University, Stanford, California 94305, USA; g) Department of Computer Science California State University, Los Angeles, CA 90032; h) Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA; i) San Diego Supercomputer Center, University of California San Diego, La Jolla, California 92093-0505, United States; j) Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; k) Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; l) Pharmaceutical Development, Genentech Inc, South San Francisco, United States; m) Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook NY 11794 and Department of Chemistry, Stony Brook University, Stony Brook NY 11794 n) Departments of Molecular Biology and Biochemistry, Chemical and Biomolecular Engineering, Materials Science and Engineering, and Biomedical Engineering, Graduate Program in Chemical and Materials Physics, University of California, Irvine, CA 92697; o) Ph.D. Programs in Chemistry, Biochemistry, and Biology–- The Graduate Center of the City University of New York, 365 5th Avenue, New York, NY, USA 10016; Lehman College, Chemistry, 250 Bedford Park Blvd West, Bronx, NY, USA 10468; p) Department of Chemistry and Biochemistry, Loyola University Chicago, Chicago, Illinois 60660, United States; q) Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China; r) Department of Physics and Astronomy, California State University, Northridge, Northridge, CA, 91330; s) Department of Chemistry and Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824-1322, United States; t) Protein Engineering Unit, Institut Pasteur de Montevideo, Montevideo, 11400, Uruguay; u) Departments of Computer Science and Physics, Virginia Tech, Blacksburg, VA, 24061 v) Department of Statistics, Florida State University, Tallahassee, FL, 32304; w) Cryo-EM Center, Southern University of Science and Technology, Shenzhen, 518055, China; x) Entos, 4470 W Sunset Blvd, Suite 107, Los Angeles, CA 90027 y) Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh,

PA, 15261; z) Laboratory of Computational Biology, NHLBI, NIH, Bethesda, MD 20892; aa) Nurix Therapeutics, Inc., San Francisco, California 94158, United States; ab) Department of Medicinal Chemistry, The University of Utah, 30 South 2000 East, Salt Lake City, Utah 84112; ac) Laboratory of Computational Biology, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA; ad) Department of Chemistry, The University of Florida, 440 Leigh Hall, Gainesville, FL 32611-7200; ae) Department of Chemistry, Stony Brook University, Stony Brook, NY 11794

Abstract

AmberTools is a free and open-source collection of programs used to set up, run and analyze molecular simulations. The newer features contained within AmberTools23 are briefly described in this Application note.

Introduction

The present status of the Amber (Assisted Model Building and Energy Refinement) suite of programs has been the product of decades of effort from a broad range of research groups starting with the group of the late Peter Kollman in the early 1980s.[1] Amber contains tools for energy minimization (EM), molecular dynamics (MD) simulations, free energy (FE) calculations, potential of mean force (PMF) capabilities and all the needed tools to set up the modeling effort. The software stack has been reviewed in the past[2-4] and the manual contains detailed descriptions of all the algorithms in Amber as well as a full list of contributors to Amber over the years (see https://ambermd.org). Besides the actual code, Amber is used to describe a series of highly regarded force fields[5] for proteins[6-12], carbohydrates[13, 14], nucleic acids[7, 8, 15] and lipids[16]. The present Application Note will only describe the latest additions to the open-source AmberTools23 and as such is not meant to give a through exposition of all the methods and capabilities of AmberTools and Amber.

Overview of Amber and AmberTools

Amber and AmberTools form a collection of programs that are designed to work together to facilitate system setup, MD simulations and trajectory analysis for biomolecules. It is useful to note that the Amber force fields mentioned above, can be used in a variety of molecular dynamics codes, outside of AmberTools and Amber. The Amber code releases, offset in years from the AmberTools releases, uniquely include the base MD code known as *pmemd*, which offers parallel and GPU-accelerated versions of the MD codes along with some free-energy based methods not implemented in AmberTools. Analogous MD function is available in *sander* in AmberTools. AmberTools is distributed under an open-source license, primarily the GNU General Public

License, with some portions covered by other, compatible, open-source licenses. The Amber force fields are in the public domain, and are distributed with AmberTools. The *pmemd* code is distributed as source code, but has a separate license that contains restrictions on use and redistribution; there is no license fee for non-commerical use of *pmemd*. Full details on licensing and distribution can be found at https://ambermd.org.

Typical workflow

The basic workflow for AmberTools is shown in the accompanying (see Figure 1), and describes preparation, simulation and analysis steps. **Preparation** starts at the top, since all MD simulations require some sort of starting three-dimensional structure, which for biomolecules is usually in the form of a PDB-format file; AmberTools has some model-building capabilities (e.g. *PACKMOL-Memgen*, see below), but other codes are generally used if experimental structures are not available. The *prepareforleap* step, which is recent and still under development, carries out tasks to map components in the input file to Amber nomenclature (especially useful for carbohydrates), add hydrogens, identify crosslinks, assign histidine protonation states, and similar tasks. Next is the *LEaP* program, which is a workhorse program that connects the nascent structure to Amber's built-in force fields for proteins, nucleic acids, carbohydrates, lipids and common solvents, and to bespoke force fields for other components like ligands and cofactors, that can be created by programs like *antechamber* and *mdgx* (for general organic molecules), and *pyMSMT* (for metal ions). The *LEaP* code creates two files: a "inpcrd" file that has complete three-dimensional coordinates, and a "prmtop" file that contains all other information needed for force field-based analyses of the system. The latter file can be examined and edited via *parmed*, which can also export similar files in GROMACS or CHARMM format.

The **simulation** phase is primarily the province of *sander* or *pmemd*. The "mdin" file contains a large number of parameters that control the type and length of the simulation to be carried out, the integration method, the use of a QM/MM (quantum mechanics/molecular mechanics) model, specification of enhanced-sampling and thermodynamic integration methods, and the like. Restraints on the system, often from NMR or X-ray data, but more recently from cryoEM and other sorts of integrative modeling can also be input at this point.

Snapshots of conformations are generally stored at regular intervals during a simulation, and then serve as input for an **analysis** phase. The *cpptraj* program is the workhorse code here, providing geometric and energetic analysis, clustering algorithms, and many other routines. Three other codes, *MMPBSA.py, FE_Toolkit* and *FEW* (Free Energy Workflow)[17] are devoted to estimating free energy changes. More complete descriptions of all of this, including a full list of programs, encompassing nearly 1,000 pages of text, is in the *Amber23 Reference Manual*.

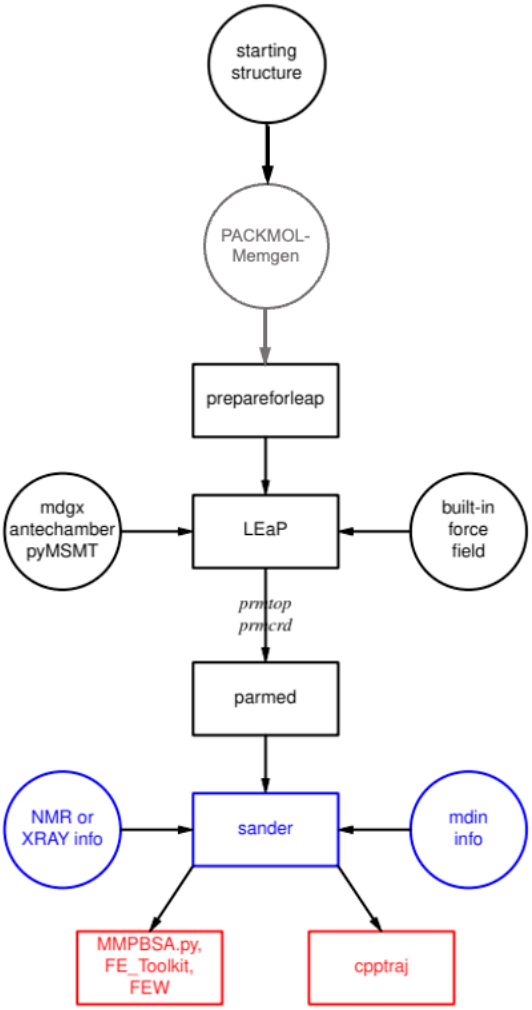

Figure 1. Common workflow in AmberTools. Flow goes from top to bottom. Black boxes are for preparation, gray indicates an optional preparation step specific for membrane systems, blue for simulation, and red for analysis.

AmberTools23 Updates

We have a number of significant new features for AmberTools23 which include automated building of membrane-protein−lipid-bilayer systems, enhancements to the polarizable Gaussian multipole method, extensions to the PBSA method, enhanced free energy capabilities, enhanced

QM and QM/MM capabilities, and a significant upgrade of Amber website and tutorials. Each of these additions are summarized below.

1. Polarizable Gaussian Multipole model in the SANDER program

Polarizable Gaussian Multipole (pGM) model is a next generation induced dipole polarizable model aiming to balance accuracy and efficiency for molecular simulations of biomolecular systems.[18-22] We recently developed a new framework for efficient computation of analytical atomic gradient for the pGM model.[18] The pGM virial for constant pressure molecular dynamics simulations was also implemented in previous releases of Amber.[19] The accuracy and robustness of the pGM model have also been validated on various molecular properties.[20-22] In the AmberTools23 release, we further optimized the induced dipole iteration algorithm. Specifically, we introduced maximum relative error as the convergence criteria to ensure the energy conservation in molecular dynamics simulations. We also designed and implemented multi-order extrapolation (MOE) and local preconditioning conjugate gradient (LPCG) schemes to accelerate the induced dipole iteration.[23] Given the new developments, MD simulations with the pGM model is able to achieve similar level of energy conservation as those with the point charge additive models, within 2~3 induction iterations.

2. New features in the PBSA program

MM/PB(GB)SA[24] is an end-point method for calculating the free energies of molecules in implicit solvent, *i.e.*, Poisson Boltzmann (PB) and generalized Born (GB). Solvation interactions, especially solvent-mediated dielectric screening and Debye-Hückel screening, are essential determinants of the structure and function of biomolecules. Several efficient finite-difference

numerical solvers, both linear[25-27] and nonlinear,[28] are implemented in *pbsa* for various applications of the Poisson-Boltzmann method. The GPU support of those solvers is also implemented in *pbsa.cuda*.[29-31] In the 2023 release, improvements to the *pbsa* program include the integration of the Machine-Learned Solvent Excluded Surface (MLSES) model,[32] which provides a highly efficient and differentiable molecular surface for continuum solvation modeling of biomolecules. Various options for the MLSES model have been implemented, allowing users to optimize performance on both CPU and GPU platforms using Fortran, CUDA kernel, and LibTorch. This flexibility enables users to choose the best-suited hardware and software environments for their needs. Additionally, an MBAR/PBSA strategy has been developed, combining the PBSA continuum solvent model with the Multistate Bennet Acceptance Ratio (MBAR) approach. This coupling allows for more accurate modeling of electronic polarization, leading to improved accuracy in absolute binding free energy simulations of highly charged ligands.[33]

To date, the GB model in AmberTools could be specified with the following "igb" values: 1[34], 2[35], 5[36], 7, and 8[37]. In 2017, an accurate yet efficient grid-based surface GB model was introduced[38] which is currently available in AmberTools as a stand-alone application named GBNSR6 ($AMBERHOME/bin/gbnsr6).[39] GBNSR6 calculates the solvation free energy of an input structure on a single snapshot. In AmberTools23, GBNSR6 has been integrated into *MMPBSA.py*[40] such that it runs over multiple snapshots extracted from the trajectories of protein, ligand, and complex structures. To run this model "igb=66" is now available in *MMPBSA.py* which activates the GBNSR6 model. All input parameters of the stand-alone GBNSR6 program can be modified through the *MMPBSA.py* input file.

3. PyRESP and PyRESP_GEN

Accurate modeling of electrostatic and polarization effects is crucial in molecular simulations. Many polarizable force fields have been developed to account for these important effects. Among these models, polarizable Gaussian Multipole (pGM) model has emerged as a self-consistent approach in handling both short-range and long-range interactions.[18-22, 41] We have recently developed the *PyRESP* program[42] for electrostatic parameterizations for point charge additive models and induced dipole models, including the pGM model. By performing least-squares fittings to electrostatic potentials surrounding molecules, the *PyRESP* program extends functionalities of the ancestor *RESP* program that only perform parameterizations for point charge additive models.[43] However, the process of generating input files for *PyRESP* is tedious and error-prone. In the AmberTools23 release, we implemented a flexible and user-friendly program, *PyRESP_GEN*, to minimize the user efforts to set up a PyRESP run. In addition, we also optimized the restraint weights for the pGM models with and without permanent dipoles. For the pGM-perm model, the optimal strategy for electrostatic potential fitting is also proposed.

4. 3D-RISM

The 3D reference interaction site model (3D-RISM) of molecular solvation[44] is an implicit solvent model that calculates equilibrium density distributions and thermodynamics of explicit solvent models. The implementation in AmberTools permits MD, energy minimization and trajectory analysis through *sander*, while *rism3d.snglpnt* provides standalone trajectory analysis.[45] Recently, periodic boundaries were introduced, allowing application to crystal structure refinement and other periodic systems.[46] In addition, computational scaling for open boundaries was improved via

treecode summation for electrostatic interactions, providing a 2-4 times speedup for typical proteins and enabling application to large biomolecular complexes with more > 1 million atoms.[47]

5. LibTorch interface to Amber

We introduced a LibTorch interface to the 2023 release of AmberTools, which is a cutting-edge C++ runtime library developed by the PyTorch team.[48] This library enables flexible tensor computations and dynamic deep neural network modeling. Amber now provides two options for enabling the LibTorch library: built-in mode and user-installed mode. With the LibTorch integration, the *pbsa* program supports both CPU and GPU environments, making it highly versatile. Additionally, user instructions and tutorials have been provided for configuring the LibTorch library, making it more accessible to researchers and developers working in Amber and AmberTools.

6. Free Energy

Free energy methods have been a mainstay of Amber for decades.[49, 50] Besides our existing free energy technology base this latest release of AmberTools includes a collection of new software tools for the robust analysis of free energy simulations (*FE-ToolKit*)[51, 52], as well as workflow tools for production free energy simulation setup and analysis (*ProFESSA*)[53] using the GPU-accelerated Amber free energy engine with enhanced sampling features. This software is part of the Amber Drug Discovery Boost package.[54]

6.1. FE-ToolKit: FE-ToolKit contains two main utilities: *edgembar*[52] for analysis of alchemical free energy simulations (e.g., such as those used for prediction of ligand-protein absolute and

relative binding free energies in drug discovery[55]), and *ndfes*[51] for analysis of multidimensional free energy profiles (e.g., such as those used for prediction of minimum free energy pathways in studies of enzyme mechanisms[56, 57]).

6.2. Edgembar: The *edgembar* program performs analysis of alchemical free energy simulations using the multistate Bennett acceptance ratio (MBAR) method,[58] the Bennett acceptance ratio (BAR) method,[59] exponential averaging,[60] thermodynamic integration,[61] or combinations of these approaches. Alchemical free energy simulations often calculate a network of relative free energy differences between two environments. For example, ligand binding applications in drug discovery use a network of alchemical transformations between ligands, termed a "thermodynamic graph", where each ligand represents a "node" in the graph and each "edge" represents an alchemical transformation between ligands bound to their target relative to in aqueous solution. Given the alchemical simulation outputs from the independent trials in both environments, edgembar will perform "network-wide" free energy analysis,[52] including the imposition of cycle closure and, optionally, experimental constraints. The analysis produces a comprehensive report of the results, including uncertainties and warnings. The report identifies potential problems with simulations that may require further attention. The issues include: a lack of convergence, the analysis of too few statistically independent samples, poor phase space overlap between adjacent alchemical states,[62] and poor reweighting entropy.[63]

6.3. Ndfes: The *ndfes* program evaluates multidimensional free energy surfaces from umbrella sampling.[51] The analysis can be performed with the variational free energy profile (vFEP) method, [64, 65] MBAR,[58] the weighted thermodynamic perturbation method (wTP),[66] and the generalized

weighted thermodynamic perturbation method (gwTP).[67] The wTP and gwTP methods estimate the free energy surface of an expensive target-level of theory from the sampling performed with inexpensive reference potentials.[67] The estimation *ab initio* QM/MM free energy surfaces in condensed phase environments has become more practical in the latest version of AmberTools with the combined introduction of the GPU-accelerated *QUICK* software[68]and *ndfes* analysis program.

6.4. ProFESSA: The *ProFESSA* workflow[53] uses the GPU-accelerated AMBER free energy engine. The workflow establishes a flexible, end-to-end pipeline for performing alchemical free energy simulations that brings to bear technologies including new smoothstep softcore potentials and optimized alchemical transformation pathways,[69] alchemical enhanced sampling (ACES) method,[70] and network-wide free energy analysis[52] with optional imposition of cycle closure and experimental constraints implemented in *FE-ToolKit*.

7. Quantum Mechanical/Molecular Mechanical methods

Amber has had a long tradition of QM/MM methods and implementations,[71] with the most recent additions being the *QUICK/sander* QM/MM implementation in AmberTools23.[68, 72-74] *QUICK/sander* has been extensively updated and its performance has been significantly improved.

7.1 Performance improvements/AMD Implementation: With the second generation electron repulsion integral code and other performance enhancements recently introduced into *QUICK*[68, 72-74], higher ps/day can be obtained in QM/MM simulations.[74] For instance, with respect to AmberTools21,[75] up to 2x speedups have been observed for benchmark simulations with different

QM regions of photoactive yellow protein on NVIDIA V100 GPU.[74] Furthermore, support for AMD GPUs has been enabled. Users can now make use of AMD data center cards such as MI50, MI100, MI200 and MI250 for simulations. According to benchmark studies, the performance on the MI100 is similar to that of NVIDIA V100.[74] The implementation runs properly on MI200 and MI250 cards, however, the performance is not yet optimized for these cards. The recommended AMD GPU for current version is MI100. An optimized version for MI2XX will be available to users in next AmberTools release.

7.2 Long-range Electrostatics: For the treatment of long range electrostatics in QM/MM, ambient composite Ewald method (CEw) developed by Giese and York[76] has been integrated. The performance penalty for turning on CEw in the GPU version is <25% for HF and <10% for DFT in comparison to standard QM/MM with 8 Å electrostatic cutoff. This allows users to carry out more accurate simulations at a slightly higher computational cost.

7.3 Dispersion: Among other minor features introduced into *QUICK*, dispersion corrections in DFT, data exporting capability into Molden format are notable. Grimme's dispersion corrections (D2, D3 with different damping)[77] can be used in QM/MM with appropriate functionals. Users can also export Cartesian coordinates, molecular orbitals, etc. of the QM region into Molden format for visualization purposes.

8. Automated building of membrane-protein−lipid-bilayer systems

PACKMOL-Memgen is a simple-to-use command line implementation of a generalized workflow for the automated building of membrane-protein−lipid-bilayer systems based on open-source tools

including *Packmol*, *memembed*, *pdbremix*, and AmberTools.[78] It allows setting up multiple configurations of a system in a user-friendly and efficient manner, which can serve as starting configurations in MD simulations under periodic boundary conditions. Since its introduction, support was added for additional lipid headgroups and to include solutes in the water or membrane phase and generate curved membrane surfaces or double bilayer systems. Additionally, SIRAH[79] coarse-graining routines can be used and membrane systems (water or mixed-solvent simulations) can now be set up[80] In the AmberTools23 release, *PACKMOL-Memgen* now handles all Amber-supported ions and the OPC3 water model as well as allows generating HMR systems, providing control for *pmemd.cuda*, and using *pdb2pqr* for protonating the protein.

9. mdgx

The *mdgx* program, which began as a *de-novo* re-implementation of the basic features needed for molecular dynamics and stayed in service for its uncommon capability of storing multiple topologies and coordinate sets in the space of a single runtime instance, has gained two noteworthy features. First, it can post-process Amber topology files to add *pmemd*-compatible representations of the GROMACS virtual sites.[81] While *mdgx* itself can perform limited MD simulations with such models, the performant *pmemd* GPU implementation can now incorporate the massless sites into its free energy methods. Virtual sites require parameters to be useful, but the *mdgx* program itself has tools for fitting their charges as well as bonded parameters in the context of these extra monopoles. Virtual site force fields are a logical extension of popular fixed-charge models, entailing incremental updates to the dynamics engine and incremental increases in the cost of the simulations. Second, through its ability to calculate multiple systems at once, *mdg* has an exploratory feature for running simple implicit solvent dynamics on many replicas of different

topologies on one GPU. By running independent trajectories on each GPU multiprocessor, *mdgx* scales simulations of small peptides and drug molecules to modern GPUs with tens of times the throughput of other GPU MD implementations when tasked with small systems. This capability has been applied to docked pose refinement.[82]

10. The Amber Website and Tutorials

The Amber website (https://ambermd.org) supports the user community with new release information, manuals, tutorials and information on force fields. Users are directed to the most recent manual version to learn about technical usage and appropriate literature references to communicate best practices in the field. The Amber tutorials have also been reorganized and span topics ranging from initial system setup to advanced methods (Figure 2). A tutorial overview page guides new users through the process of building, running and analyzing a system, and points them to key initial case studies. The recent tutorials overall are more modular and learning objectives are given. New tutorial development has focused on building different system types, tutorials for creating stable systems through relaxation of system positions for both explicit and implicit solvent, as well as a tutorial covering advanced thermodynamic integration methods such as using smoothstep softcore potentials,[69] enhanced sampling for softcore ligand energies and methods such as ACES.[70]

Modeling software is not useful without compatible force fields. Included in the release of AmberTools are the force fields developed by the Amber community. The main force fields page contains a list of recommended force fields and each type of molecule/ion has a separate page outlining nuances in choosing an appropriate force field.

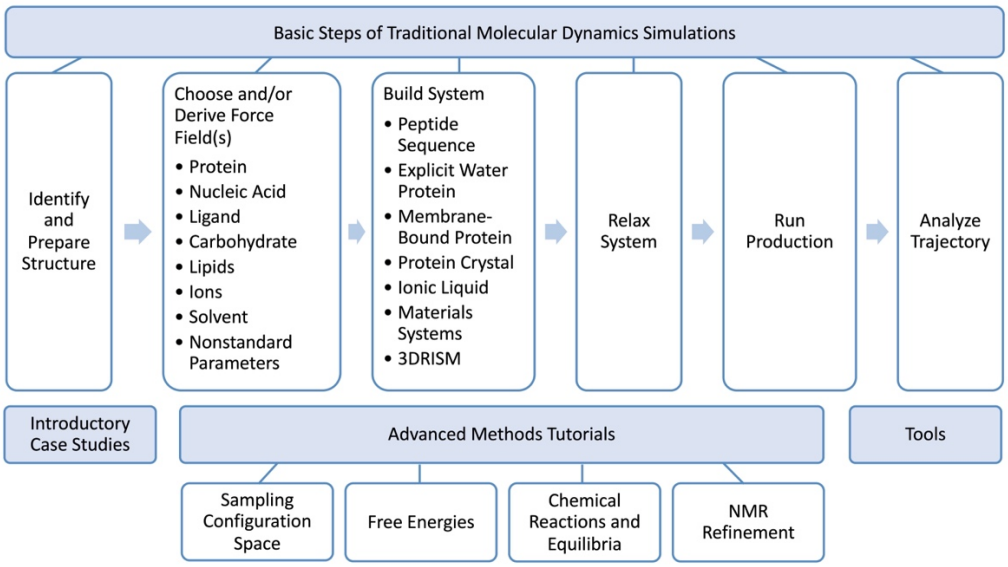

Figure 2. Overview of the Amber Tutorials. Tutorials are modular, cover the basic steps of a typical molecular dynamics simulation, introductory case studies, advanced methods and some tools that are commonly employed by Amber users.

Conclusions

The most significant additions to AmberTools23 are briefly summarized. AmberTools is freely available at https://ambermd.org. Full details on licensing, distribution and hardware supported can be found at https://ambermd.org.

ACKNOWLEDGMENTS

The authors are grateful for the financial support provided by the National Institutes of Health (No. GM62248 and GM107485 to DMY; GM130367 to RL; GM130641 to KMM and HMA; NIH U2COD026506 to KMM; GM108583 to GAC; GM149874 to TSL; GM107104 to CS; NHLBI Z01 HL001051-23 to DRR; GM147673 to JW; No.GM146633 to NF and R35GM144089 to TK), the National Science Foundation (2209718 to DMY; 2209717 to KMM, HMA and AWG; NSF OAC1835144 to KMM and AWG; OAC-2117247 to GAC; CHE-2050541 to MCN; CHE-2018427 to MCN and TAL; CHE-2102668 to TL; CTMC-1665159 to CS; CHE-1955260 to JW; 2216858 to OS, PSN. and NF), Loyola University Chicago (start-up

funds to PL), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (Project CRC 1208 (number 267205415) to HG (subproject A03)) and FOCEM (MERCOSUR Structural Convergence Fund) COF 03/11 to SP.

Data and Software Availability Statement

The AmberTools suite is free of charge, and its components are mostly released under the GNU General Public License (GPL). Please see https://ambermd.org for licensing and distribution.

References

(1) Weiner, P. K.; Kollman, P. A. Amber: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. *J Comput Chem* **1981**, *2*, 287-303.
(2) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J Comput Chem* **2005**, *26*, 1668-1688.
(3) Pearlman, D.; Case, D. A.; Caldwell, J.; Ross, W. S.; Cheatham II, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. Amber, a Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput. Phys. Commun.* **1995**, *91*, 1-41.
(4) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *WIRESs Comput Mol Sci* **2013**, *3*, 198-210.
(5) Ponder, J. W.; Case, D. A. Force Fields for Protein Simulations. *Adv Protein Chem* **2003**, *66*, 27-85.
(6) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712-725.
(7) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A New Force Field Fo Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.
(8) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230-252.
(9) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
(10) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J Comput Chem* **2004**, *25*, 1157-1174.

(11) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14sb: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99sb. *J Chem Theory Comput* **2015**, *11*, 3696-3713.
(12) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. Ff19sb: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J Chem Theory Comput* **2020**, *16*, 528-552.
(13) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. Glycam06: A Generalizable Biomolecular Force Field. Carbohydrates. *J Comput Chem* **2008**, *29*, 622-655.
(14) Tessier, M. B.; Demarco, M. L.; Yongye, A. B.; Woods, R. J. Extension of the Glycam06 Biomolecular Force Field to Lipids, Lipid Bilayers and Glycolipids. *Mol Simul* **2008**, *34*, 349-363.
(15) Cheatham, T. E., 3rd; Case, D. A. Twenty-Five Years of Nucleic Acid Simulations. *Biopolymers* **2013**, *99*, 969-977.
(16) Dickson, C. J.; Walker, R. C.; Gould, I. R. Lipid21: Complex Lipid Membrane Simulations with Amber. *J Chem Theory Comput* **2022**, *18*, 1726-1736.
(17) Homeyer, N.; Gohlke, H. Few: A Workflow Tool for Free Energy Calculations of Ligand Binding. *J Comput Chem* **2013**, *34*, 965-973.
(18) Wei, H.; Qi, R.; Wang, J.; Cieplak, P.; Duan, Y.; Luo, R. Efficient Formulation of Polarizable Gaussian Multipole Electrostatics for Biomolecular Simulations. *The Journal of Chemical Physics* **2020**, *153*, 114116.
(19) Wei, H.; Cieplak, P.; Duan, Y.; Luo, R. Stress Tensor and Constant Pressure Simulation for Polarizable Gaussian Multipole Model. *The Journal of Chemical Physics* **2022**, *156*, 114114.
(20) Wang, J.; Cieplak, P.; Luo, R.; Duan, Y. Development of Polarizable Gaussian Model for Molecular Mechanical Calculations I: Atomic Polarizability Parameterization to Reproduce Ab Initio Anisotropy. *Journal of Chemical Theory and Computation* **2019**, *15*, 1146-1158.
(21) Zhao, S.; Wei, H.; Cieplak, P.; Duan, Y.; Luo, R. Accurate Reproduction of Quantum Mechanical Many-Body Interactions in Peptide Main-Chain Hydrogen-Bonding Oligomers by the Polarizable Gaussian Multipole Model. *Journal of Chemical Theory and Computation* **2022**, *18*, 6172-6188.
(22) Zhao, S.; Cieplak, P.; Duan, Y.; Luo, R. Transferability of the Electrostatic Parameters of the Polarizable Gaussian Multipole Model. *Journal of Chemical Theory and Computation* **2023**, *19*, 924-941.
(23) Huang, Z.; Zhao, S.; Cieplak, P.; Duan, Y.; Luo, R.; Wei, H. Optimal Scheme to Achieve Energy Conservation in Induced Dipole Models. *J Chem Theory Comput* **2023**, *19*, 5047-5057.
(24) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with Mm/Pbsa and Mm/Gbsa: Strategies and Applications in Drug Design. *Chem Rev* **2019**, *119*, 9478-9508.
(25) Luo, R.; David, L.; Gilson, M. K. Accelerated Poisson-Boltzmann Calculations for Static and Dynamic Systems. *Journal of Computational Chemistry* **2002**, *23*, 1244-1253.
(26) Wang, J.; Luo, R. Assessment of Linear Finite-Difference Poisson-Boltzmann Solvers. *Journal of Computational Chemistry* **2010**, *31*, 1689-1698.
(27) Wang, J.; Cai, Q.; Xiang, Y.; Luo, R. Reducing Grid Dependence in Finite-Difference Poisson–Boltzmann Calculations. *Journal of Chemical Theory and Computation* **2012**, *8*, 2741-2751.

(28) Cai, Q.; Hsieh, M.-J.; Wang, J.; Luo, R. Performance of Nonlinear Finite-Difference Poisson-Boltzmann Solvers. *Journal of Chemical Theory and Computation* **2010**, *6*, 203-211.
(29) Qi, R.; Botello-Smith, W. M.; Luo, R. Acceleration of Linear Finite-Difference Poisson–Boltzmann Methods on Graphics Processing Units. *Journal of Chemical Theory and Computation* **2017**, *13*, 3378-3387.
(30) Qi, R.; Luo, R. Robustness and Efficiency of Poisson–Boltzmann Modeling on Graphics Processing Units. *Journal of Chemical Information and Modeling* **2019**, *59*, 409-420.
(31) Wei, H.; Luo, R.; Qi, R. An Efficient Second-Order Poisson–Boltzmann Method. *Journal of Computational Chemistry* **2019**, *40*, 1257-1269.
(32) Wei, H.; Zhao, Z.; Luo, R. Machine-Learned Molecular Surface and Its Application to Implicit Solvent Simulations. *Journal of Chemical Theory and Computation* **2021**, *17*, 6214-6224.
(33) King, E.; Qi, R.; Li, H.; Luo, R.; Aitchison, E. Estimating the Roles of Protonation and Electronic Polarization in Absolute Binding Affinity Simulations. *Journal of chemical theory and computation* **2021**, *17*, 2541-2555.
(34) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* **1996**, *100*, 19824-19839.
(35) Onufriev, A.; Case, D. A.; Bashford, D. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712-3720.
(36) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins* **2004**, *55*, 383-394.
(37) Mongan, J.; Simmerling, C.; McCammon, J. A.; Case, D. A.; Onufriev, A. Generalized Born Model with a Simple, Robust Molecular Volume Correction. *J Chem Theory Comput* **2007**, *3*, 156-169.
(38) Forouzesh, N.; Izadi, S.; Onufriev, A. V. Grid-Based Surface Generalized Born Model for Calculation of Electrostatic Binding Free Energies. *J Chem Inf Model* **2017**, *57*, 2505-2513.
(39) Forouzesh, N.; Mukhopadhyay, A.; Watson, L. T.; Onufriev, A. V. Multidimensional Global Optimization and Robustness Analysis in the Context of Protein-Ligand Binding. *J Chem Theory Comput* **2020**, *16*, 4669-4684.
(40) Miller, B. R., 3rd; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. Mmpbsa.Py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* **2012**, *8*, 3314-3321.
(41) Huang, Z.; Zhao, S.; Cieplak, P.; Duan, Y.; Luo, R.; Wei, H. An Optimal Scheme to Achieve Energy Conservation in Induced Dipole Models. *Journal of Chemical Theory and Computation* **Submitted**.
(42) Zhao, S.; Wei, H.; Cieplak, P.; Duan, Y.; Luo, R. Pyresp: A Program for Electrostatic Parameterizations of Additive and Induced Dipole Polarizable Force Fields. *Journal of Chemical Theory and Computation* **2022**, *18*, 3654-3670.
(43) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The Resp Model. *The Journal of Physical Chemistry* **1993**, *97*, 10269-10280.
(44) Kovalenko, A.; Hirata, F. A Molecular Theory of Liquid Interfaces. *Phys Chem Chem Phys* **2005**, *7*, 1785-1793.

(45) Luchko, T.; Gusarov, S.; Roe, D. R.; Simmerling, C.; Case, D. A.; Tuszynski, J.; Kovalenko, A. Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber. *J Chem Theory Comput* **2010**, *6*, 607-624.
(46) Gray, J. G.; Giambasu, G. M.; Case, D. A.; Luchko, T. Integral Equation Models for Solvent in Macromolecular Crystals. *J Chem Phys* **2022**, *156*, 014801.
(47) Wilson, L.; Krasny, R.; Luchko, T. Accelerating the 3d Reference Interaction Site Model Theory of Molecular Solvation with Treecode Summation and Cut-Offs. *J Comput Chem* **2022**, *43*, 1251-1270.
(48) Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic Differentiation in Pytorch. In Conference on Neural Information Processing Systems, 2017.
(49) Kollman, P. Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395-2417.
(50) Song, L. F.; Merz, K. M., Jr. Evolution of Alchemical Free Energy Methods in Drug Discovery. *J Chem Inf Model* **2020**, *60*, 5308-5318.
(51) Giese, T. J.; Ekesan, S.; York, D. M. Extension of the Variational Free Energy Profile and Multistate Bennett Acceptance Ratio Methods for High-Dimensional Potential of Mean Force Profile Analysis. *J Phys Chem A* **2021**, *125*, 4216-4232.
(52) Giese, T. J.; York, D. M. Variational Method for Networkwide Analysis of Relative Ligand Binding Free Energies with Loop Closure and Experimental Constraints. *J Chem Theory Comput* **2021**, *17*, 1326-1336.
(53) 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.
(54) Lee, T.-S.; Tsai, H.-C.; Ganguly, A.; Giese, T. J.; York, D. M. Efficient and Automated Methods for Accurate Prediction of Protein-Ligand Binding Affinities in Amber Drug Discovery Boost. In *Acs Symposium Series*, Armacost, K. A., Thompson, D. C. Eds.; Vol. 1397; ACS, 2021; pp 161-204.
(55) Cournia, Z.; Chipot, C.; Roux, B.; York, D. M.; Sherman, W. Free Energy Methods in Drug Discovery—Introduction. In *Acs Symposium Series*, Armacost, K. A., Thompson, D. C. Eds.; Vol. 1397; ACS, 2021; pp 1-38.
(56) Gaines, C. S.; Giese, T. J.; York, D. M. Cleaning up Mechanistic Debris Generated by Twister Ribozymes Using Computational Rna Enzymology. *ACS Catal* **2019**, *9*, 5803-5815.
(57) Ganguly, A.; Weissman, B. P.; Giese, T. J.; Li, N. S.; Hoshika, S.; Rao, S.; Benner, S. A.; Piccirilli, J. A.; York, D. M. Confluence of Theory and Experiment Reveals the Catalytic Mechanism of the Varkud Satellite Ribozyme. *Nat Chem* **2020**, *12*, 193-201.
(58) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J Chem Phys* **2008**, *129*, 124105.
(59) Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Chem. Phys.* **1976**, *22*, 245-268.
(60) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420.
(61) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3*, 300-313.

(62) Giese, T. J.; York, D. M. Development of a Robust Indirect Approach for Mm --> Qm Free Energy Calculations That Combines Force-Matched Reference Potential and Bennett's Acceptance Ratio Methods. *J Chem Theory Comput* **2019**, *15*, 5543-5562.
(63) Wang, M.; Li, P.; Jia, X.; Liu, W.; Shao, Y.; Hu, W.; Zheng, J.; Brooks, B. R.; Mei, Y. Efficient Strategy for the Calculation of Solvation Free Energies in Water and Chloroform at the Quantum Mechanical/Molecular Mechanical Level. *J Chem Inf Model* **2017**, *57*, 2476-2489.
(64) Lee, T. S.; Radak, B. K.; Pabis, A.; York, D. M. A New Maximum Likelihood Approach for Free Energy Profile Construction from Molecular Simulations. *J Chem Theory Comput* **2013**, *9*, 153-164.
(65) Lee, T. S.; Radak, B. K.; Huang, M.; Wong, K. Y.; York, D. M. Roadmaps through Free Energy Landscapes Calculated Using the Multi-Dimensional Vfep Approach. *J Chem Theory Comput* **2014**, *10*, 24-34.
(66) Li, P.; Jia, X.; Pan, X.; Shao, Y.; Mei, Y. Accelerated Computation of Free Energy Profile at Ab Initio Quantum Mechanical/Molecular Mechanics Accuracy Via a Semi-Empirical Reference Potential. I. Weighted Thermodynamics Perturbation. *J Chem Theory Comput* **2018**, *14*, 5583-5596.
(67) Giese, T. J.; Zeng, J.; York, D. M. Multireference Generalization of the Weighted Thermodynamic Perturbation Method. *J Phys Chem A* **2022**, *126*, 8519-8533.
(68) Cruzeiro, V. W. D.; Manathunga, M.; Merz, K. M., Jr.; Gotz, A. W. Open-Source Multi-Gpu-Accelerated Qm/Mm Simulations with Amber and Quick. *J Chem Inf Model* **2021**, *61*, 2109-2115.
(69) Tsai, H. C.; Lee, T. S.; Ganguly, A.; Giese, T. J.; Ebert, M. C.; Labute, P.; Merz, K. M., Jr.; York, D. M. Amber Free Energy Tools: A New Framework for the Design of Optimized Alchemical Transformation Pathways. *J Chem Theory Comput* **2023**.
(70) Lee, T. S.; Tsai, H. C.; Ganguly, A.; York, D. M. Aces: Optimized Alchemically Enhanced Sampling. *J Chem Theory Comput* **2023**, *19*, 472-487.
(71) Gotz, A. W.; Clark, M. A.; Walker, R. C. An Extensible Interface for Qm/Mm Molecular Dynamics Simulations with Amber. *J Comput Chem* **2014**, *35*, 95-108.
(72) Manathunga, M.; Jin, C.; Cruzeiro, V. W. D.; Miao, Y.; Mu, D.; Arumugam, K.; Keipert, K.; Aktulga, H. M.; Merz, K. M., Jr.; Gotz, A. W. Harnessing the Power of Multi-Gpu Acceleration into the Quantum Interaction Computational Kernel Program. *J Chem Theory Comput* **2021**, *17*, 3955-3966.
(73) Manathunga, M.; Miao, Y.; Mu, D.; Gotz, A. W.; Merz, K. M., Jr. Parallel Implementation of Density Functional Theory Methods in the Quantum Interaction Computational Kernel Program. *J Chem Theory Comput* **2020**, *16*, 4315-4326.
(74) Manathunga, M.; Aktulga, H. M.; Gotz, A. W.; Merz, K. M., Jr. Quantum Mechanics/Molecular Mechanics Simulations on Nvidia and Amd Graphics Processing Units. *J Chem Inf Model* **2023**, *63*, 711-717.
(75) Case, D. A.; Atulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham II, T. E.; Cisneros, G. A.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G. M.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izamailov, S. A.; Jin, C.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Machado, M.; Man, V.; Manathunga, M.; Merz, K. M.; Miao, Y.; Mikhailovskiii, O.; Monard, G.; Nguyen, H.; O'Hearn, K. A.; Onufriev, A.; Pan, F.;

Pantano, S.; Qi, R.; Ranhnamoun, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xue, Y.; York, D. M.; Zhao, S.; Kollman, P. A. *Amber 2021;* University of California: San Francisco, 2021.
(76) Giese, T. J.; York, D. M. Ambient-Potential Composite Ewald Method for Ab Initio Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulation. *J. Chem. Theory Comput.* **2016**, *12*, 2611-2632.
(77) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105-5154.
(78) Schott-Verdugo, S.; Gohlke, H. Packmol-Memgen: A Simple-to-Use, Generalized Workflow for Membrane-Protein-Lipid-Bilayer System Building. *J Chem Inf Model* **2019**, *59*, 2522-2528.
(79) Klein, F.; Soñora, M.; Santos, L. H.; Frigini, E. N.; Ballesteros-Casallas, A.; Machado, M. R.; Pantano, S. The Sirah Force Field: A Suite for Simulations of Complex Biological Systems at the Coarse-Grained and Multiscale Levels. *J Struct Biol* **2023**, *215*, 107985.
(80) Wang, Y.; Zhou, R. *A Practical Guide to Recent Advances in Multiscale Modeling and Simulation of Biomolecules*; AIP Publishing LLC, 2023. DOI: 10.1063/9780735425279.
(81) Larsson, P.; Kneiszl, R. C.; Marklund, E. G. Mkvsites: A Tool for Creating Gromacs Virtual Sites Parameters to Increase Performance in All-Atom Molecular Dynamics Simulations. *J Comput Chem* **2020**, *41*, 1564-1569.
(82) Hsu, D. J.; Davidson, R. B.; Sedova, A.; Glaser, J. Tinyifd: A High-Throughput Binding Pose Refinement Workflow through Induced-Fit Ligand Docking. *J Chem Inf Model* **2023**, *63*, 3438-3447.