

Molecular Dynamics Simulations

Subjects: Chemistry, Medicinal

Contributor: Outi Salo-Ahen

Molecular dynamics (MD) is a simulation technique that aims at deriving statements about the structural, dynamical, and thermodynamical properties of a molecular system. MD simulations have become increasingly useful in the modern drug development process. For example, in the lead discovery and lead optimization phases, MD facilitates the evaluation of the binding energetics and kinetics of the ligand-receptor interactions, therefore guiding the choice of the best candidate molecules for further development. In the future, the role of MD simulations in facilitating the drug development process is likely to grow substantially with the increasing computer power and advancements in the development of force fields and enhanced MD methodologies.

Keywords: Molecular Dynamics Simulations ; drug development

1. Introduction

The purpose of a computer simulation is to gain insight into the behavior of an actual physical system or process. To achieve that specific objective, a model system is developed that represents or emulates the given physical system. A suitable algorithm subsequently generates a time series or an ensemble of states ("observations") for the model system. Finally, an analysis is conducted by calculating various system properties from these states (certain properties may also be monitored during a simulation). In the present context, the word "simulation" usually refers to the process of generating states by numerically solving a set of differential equations for the selected degrees of freedom (state variables) of the given model system. Many computed system properties are measured by experiment as well so that an explanation of the observed experimental data is immediately available based on the model system and the simulation results. More importantly, one can also observe behavior that is inaccessible to experiment and test "what-if" scenarios (e.g., mutation studies). As a result, simulation techniques have become invaluable tools for modern research as they complement experimental approaches. With the continuing advance of computing power, such tools will only further increase in importance.

2. Classical Molecular Dynamics Simulations

Molecular dynamics (MD) is one such simulation technique ^{[1][2]}. It aims at deriving statements about the structural, dynamical, and thermodynamical properties of a molecular system. The latter is typically a biomolecule (solute) such as a protein, an enzyme, or a collection of lipids forming a membrane, immersed in an aqueous solvent (water or electrolyte). In the case of proteins and enzymes, the experimental protein structure as deposited in the Protein Data Bank (PDB) ^[3] serves as a starting point for MD simulations. If no structure is available, one must resort to modeling (predicting) the structure for which several techniques are available, such as homology or comparative modeling ^[4]. In atomistic "all-atom" MD (AAMD), the model system consists of a collection of interacting particles represented as atoms, describing both solute and solvent, placed inside a sufficiently large simulation box, where their movements are described by Newton's laws of motions. An algorithm such as velocity-Verlet or leap-frog ^[2] is employed to advance over the course of many time steps the state of the model system as a function of time (see Figure 1 for a schematic view of a basic MD algorithm). A single state consists of the combined values of the atoms' positions and velocities (or momenta). To advance the state, forces acting on particles are computed from a model or empirical potential energy ("force field"), a function of particle positions, which includes all types of "non-bonded" interactions, such as electrostatic and Lennard-Jones forces, but also various types of "bonded" potentials for preserving the structural integrity of the given biomolecular system. The latter include harmonic potentials for maintaining bonds, bond angles, and "improper" dihedrals as well as terms for dihedrals. Some force fields include the Morse potential for a more realistic representation of bonds, while others may account for explicit electronic polarization effects. An extensive discussion of force fields is given by Monticelli and Tieleman ^[5]. Of note, various force fields have been developed for different types of molecules. In the context of our review, force fields for proteins ^[5], biological lipids ^[6], and small molecules ^[7] are of particular relevance (see Table 1 for examples of current commonly used force fields). Force fields are also employed to compute energies in molecular mechanics (MM)

applications [8]. Such simulations are usually conducted under conditions of constant temperature and pressure to mimic laboratory conditions for which special algorithms are available. In addition, to emulate a very large molecular system, several techniques for artificially extending the size of the model system have been developed. The most common one is the periodic boundary condition (PBC). Here, an infinite number of replicas of the central simulation box surrounds the central box. Due to the long-range character of electrostatic interactions, special techniques such as Ewald-based methods are required to include the interactions between the particles in the central box and their replicas. In the course of a simulation, successive states at regular time intervals are stored in a trajectory for later analysis. Typically, in AAMD, the time step is 1–2 fs (1 fs = 10⁻¹⁵ s) and the system size is in the order of tens of thousands of atoms (including solvent). A larger number of software packages for MD of biomolecules are available, for example, GROMACS [9], AMBER [10], NAMD [11], and CHARMM [12].



Figure 1. Basic molecular dynamics simulation algorithm. Each particle moves according to Newton's second law or the equation of motion, $F = ma$ (where F is the force exerted on the particle, m is its mass, and a is its acceleration under a potential field), such that the particles in the system are captured in the trajectory [1][13]. r —position; v —velocity; t —time.

Table 1. Examples of commonly used force fields in molecular dynamics simulations.

Type of Molecule	Force Field
Protein	AMBER [14], CHARMM [15], GROMOS [16], OPLS-AA [17]
Small organic molecule	General AMBER Force Field (GAFF) [18], CHARMM General Force Field (CGenFF) [19], Merck Molecular Force Field (MMFF) [20][21][22][23][24], OPLS3 [25], GROMOS96 [26][27][28][29]
Lipid	GROMOS (45A3, 53A6, 54A7/8) [30]; Berger lipid FF [31]; CHARMM (C36 lipid FF [32], C36-UA [33]); Slipids FF [34]; AMBER (LIPID14 FF) [35]

Early applications of AAMD were concerned with rather simple systems such as liquid argon consisting of just 864 atoms [36] and covered the ps (10⁻¹² s) time scale. The first MD simulation of a biomolecule was achieved by McCammon et al. in 1977; bovine pancreatic trypsin inhibitor (58 residues) was simulated for 9.2 ps [37]. Nowadays, μ s (10⁻⁶ s) timescale is easily achieved with relatively small proteins [38], wherewith larger systems, this is reachable if state-of-the-art computational power is available [39]. More typically, MD of larger biomolecular systems attain hundreds to thousands of ns (10⁻⁹ s). MD excels in focusing on the dynamical aspects of a given protein or enzyme in relation to its function. Applications include, for instance, estimation of affinities $\Delta_b G^\ominus$ (the standard Gibbs free energy of binding) of ligands for proteins using the so-called free energy perturbation (FEP) methods [40], the inclusion of charge fluctuations in constant pH AAMD [41], folding of small proteins [42], and the simulation of ion channels [43]. Additional examples are also listed in the recent reviews by Cavalli and collaborators [44] and Hollingsworth and Dror [45]. Recently, one of the largest atomistic simulations was reported by Rommie Amaro's group where they simulated an explicitly solvated influenza A viral envelope in a phospholipid bilayer [46], a system of ca. 160 million atoms, for approximately 121 ns. Another study by Jung et al. reported the first atom-scale simulation of an entire gene with a billion atoms [47].

It is worthwhile to stress that MD is meant to explore the configuration space (the set of all values of positions and momenta). In the widely used ligand-protein docking one also needs to sample the configuration space using a similar force field as in MD in order to optimize the location and binding mode (a "pose") of the ligand on the surface of a given protein (receptor) [48]. However, the sampled configuration space is frequently restricted to a specific region of the protein, while also only specific portions of the protein and/or ligand are allowed to fluctuate (e.g., only the protein's side chain are displaceable, while the main chain is kept rigid). The outcome of such docking studies is a small set of possible

complexes, typically ranked according to the force field employed. Dynamical information is not obtained from ligand-protein docking studies, but some measure for the affinity of the ligand for the protein may be obtained. In fact, the lack of a proper description of systems' true dynamics is one of the biggest caveats of docking [49]. Therefore, frequently after selecting one of the complexes (usually the highest-ranked complex), an MD simulation is conducted to explore the complex in much greater detail [50].

3. MD simulations in drug discovery and development

Over the past 50 years, MD simulation techniques have established their relevance in modern drug discovery and development processes. With the increasing computer power and development of new force fields and enhanced sampling methods, it is now possible to simulate even big and complex (bio)molecular systems in time scales that can give important insights into real-life molecular events and interactions. MD simulations can be successfully used to uncover the conformational dynamics of highly flexible target proteins, thus providing valuable insights for drug discovery and design. Moreover, MD-generated conformational ensembles can be helpful in ligand design against extremely flexible intrinsically disordered proteins. Enhanced MD approaches and force field development are key in tackling such challenging targets. MD simulations can also reveal important details of antibody-antigen interactions and help design novel antibody therapeutics with improved properties.

MD-based binding free energy calculations are widely used in the hit identification phase to improve the accuracy of ranking the putative hits. In addition, the design of drugs with improved binding kinetics benefits from various enhanced MD techniques that probe the drug residence times at the target site. Similarly, MD simulations can facilitate peptide docking by enhancing the conformational sampling and refinement of peptide-protein complexes in the development of peptide therapeutics.

MD simulations of membrane proteins embedded in a relevant cell membrane model have been challenging due to the big size of the simulation system, but especially coarse-graining has helped to reduce the computational cost. Inclusion of the membrane in MD simulations has revealed important insights into lipid-protein interactions, the function of membrane proteins, as well as the ligand entry and exit into the target binding site.

Apart from the first steps of the drug discovery process, MD simulations have also shown to be useful in pharmaceutical development and formulations studies. For example, crystalline and amorphous drugs, drug-polymer formulations, or drug-loaded nanoparticles can be studied by MD simulations to complement experimental studies. Molecular-level insights into such systems can help in improving the solubility, stability, or other properties of drug formulations.

4. Conclusion

In conclusion, MD simulations can serve as a powerful tool in facilitating the early phases of the modern drug discovery and development process. In the future, MD simulations will likely gain even more importance due to the theoretical and technological advancements in the field.

References

1. Leach, A.R. *Molecular Modelling: Principles and Applications*, 2nd ed.; Pearson Education: Harlow, UK, 2001.
2. Van Gunsteren, W.F.; Berendsen, H.J.C. *Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry*. *Angew. Chem. Int. Ed. Engl.* 1990, 29, 992–1023.
3. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* 2000, 28, 235–242.
4. Kuhlman, B.; Bradley, P. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* 2019, 20, 681–697.
5. Monticelli, L.; Tieleman, D.P. Force fields for classical molecular dynamics. *Methods Mol. Biol.* 2013, 924, 197–213.
6. Lyubartsev, A.P.; Rabinovich, A.L. Force Field Development for Lipid Membrane Simulations. *Biochim. Biophys. Acta Biomembr.* 2016, 1858, 2483–2497.
7. Lin, F.Y.; MacKerell, A.D., Jr. Force Fields for Small Molecules. *Methods Mol. Biol.* 2019, 2022, 21–54.
8. Hubbard, R.E. *Molecular graphics and modeling: Tools of the trade*. In *Guidebook on Molecular Modeling in Drug Design*; Cohen, N.C., Ed.; Academic Press: New York, NY, USA, 1996; pp. 19–54.

9. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* 2008, 4, 435–447.
10. Salomon-Ferrer, R.; Case, D.A.; Walker, R.C. An overview of the Amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2013, 3, 198–210.
11. Phillips, J.C.; Hardy, D.J.; Maia, J.D.C.; Stone, J.E.; Ribeiro, J.V.; Bernardi, R.C.; Buch, R.; Fiorin, G.; Hémin, J.; Jiang, W.; et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* 2020, 153, 044130.
12. Brooks, B.R.; Brooks III, C.L.; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* 2009, 30, 1545–1614.
13. Alder, B.J.; Wainwright, T.E. Studies in Molecular Dynamics. I. General Method. *J. Chem. Phys.* 1959, 31, 459–466.
14. 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.
15. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.; Mittal, J.; Feig, M.; Mackerell, A.D., Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J. Chem. Theory Comput.* 2012, 8, 3257–3273.
16. Oostenbrink, C.; Villa, A.; Mark, A.E.; Van Gunsteren, W.F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* 2004, 25, 1656–1676.
17. Kaminski, G.A.; Friesner, R.A.; Tirado-Rives, J.; Jorgensen, W.L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* 2001, 105, 6474–6487.
18. Wang, J.; 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.
19. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* 2010, 31, 671–690.
20. Halgren, T.A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* 1996, 17, 490–519.
21. Halgren, T.A. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* 1996, 17, 520–552.
22. Halgren, T.A. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *J. Comput. Chem.* 1996, 17, 553–586.
23. Halgren, T.A.; Nachbar, R.B. Merck molecular force field. IV. Conformational energies and geometries for MMFF94. *J. Comput. Chem.* 1996, 17, 587–615.
24. Halgren, T.A. Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *J. Comput. Chem.* 1996, 17, 616–641.
25. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J.Y.; Wang, L.; Lupyan, D.; Dahlgren, M.K.; Knight, J.L.; et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* 2016, 12, 281–296.
26. Daura, X.; Mark, A.E.; Van Gunsteren, W.F. Parametrization of aliphatic CH_n united atoms of GROMOS96 force field. *J. Comput. Chem.* 1998, 19, 535–547.
27. Schuler, L.D.; Daura, X.; Van Gunsteren, W.F. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comput. Chem.* 2001, 22, 1205–1218.
28. Horta, B.A.C.; Fuchs, P.F.J.; van Gunsteren, W.F.; Hünenberger, P.H. New interaction parameters for oxygen compounds in the GROMOS force field: Improved pure-liquid and solvation properties for alcohols, ethers, aldehydes, ketones, carboxylic acids, and esters. *J. Chem. Theory Comput.* 2011, 7, 1016–1031.
29. Horta, B.A.C.; Merz, P.T.; Fuchs, P.F.J.; Dolenc, J.; Riniker, S.; Hünenberger, P.H. A GROMOS-compatible force field for small organic molecules in the condensed phase: The 2016H66 parameter set. *J. Chem. Theory Comput.* 2016, 12, 3825–3850.
30. Marzuoli, I.; Margreitter, C.; Fraternali, F. Lipid Head Group Parameterization for GROMOS 54A8: A Consistent Approach with Protein Force Field Description. *J. Chem. Theory Comput.* 2019, 15, 5175–5193.

31. Berger, O.; Edholm, O.; Jähnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* 1997, 72, 2002–2013.
32. Venable, R.M.; Sodt, A.J.; Rogaski, B.; Rui, H.; Hatcher, E.; MacKerell, A.D., Jr.; Pastor, R.W.; Klauda, J.B. CHARMM a ll-atom additive force field for sphingomyelin: Elucidation of hydrogen bonding and of positive curvature. *Biophys. J.* 2014, 107, 134–145.
33. Lee, S.; Tran, A.; Allsopp, M.; Lim, J.B.; Héning, J.; Klauda, J.B. CHARMM36 United Atom Chain Model for Lipids and S urfactants. *J. Phys. Chem. B* 2014, 118, 547–556.
34. Jämbeck, J.P.M.; Lyubartsev, A.P. Another Piece of the Membrane Puzzle: Extending Slipids Further. *J. Chem. Theory Comput.* 2013, 9, 774–784.
35. Dickson, C.J.; Madej, B.D.; Skjevik, Å.A.; Betz, R.M.; Teigen, K.; Gould, I.R.; Walker, R.C. Lipid14: The Amber Lipid For ce Field. *J. Chem. Theory Comput.* 2014, 10, 865–879.
36. Rahman, A. Correlations in the Motion of Atoms in Liquid Argon. *Phys. Rev.* 1964, 136, A405–A411.
37. McCammon, J.G. Dynamics of folded proteins. *Nature* 1977, 267, 585–590.
38. Zou, Y.; Qian, Z.; Chen, Y.; Qian, H.; Wei, G.; Zhang, Q. Norepinephrine Inhibits Alzheimer's Amyloid- β Peptide Aggregation and Destabilizes Amyloid- β Protofibrils: A Molecular Dynamics Simulation Study. *ACS Chem. Neurosci.* 2019, 10, 1585–1594.
39. Neale, C.; García, A.E. The Plasma Membrane as a Competitive Inhibitor and Positive Allosteric Modulator of KRas4B Signaling. *Biophys. J.* 2020.
40. Abel, R.; Wang, L.; Harder, E.D.; Berne, B.J.; Friesner, R.A. Advancing Drug Discovery through Enhanced Free Energy Calculations. *Acc. Chem. Res.* 2017, 50, 1625–1632.
41. Donnini, S.; Ullmann, R.T.; Groenhof, G.; Grubmüller, H. Charge-Neutral Constant pH Molecular Dynamics Simulations Using a Parsimonious Proton Buffer. *J. Chem. Theory Comput.* 2016, 12, 1040–1051.
42. Duan, L.; Guo, X.; Cong, Y.; Feng, G.; Li, Y.; Zhang, J.Z.H. Accelerated Molecular Dynamics Simulation for Helical Proteins Folding in Explicit Water. *Front. Chem.* 2019, 7, 540.
43. Maffeo, C.; Bhattacharya, S.; Yoo, J.; Wells, D.; Aksimentiev, A. Modeling and Simulation of Ion Channels. *Chem. Rev.* 2012, 112, 6250–6284.
44. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 2016, 59, 4035–4061.
45. Hollingsworth, S.A.; Dror, R.O. Molecular Dynamics Simulation for All. *Neuron* 2018, 99, 1129–1143.
46. Durrant, J.D.; Kochanek, S.E.; Casalino, L.; Jeong, P.U.; Dommer, A.C.; Amaro, R.E. Mesoscale All-Atom Influenza Virus Simulations Suggest New Substrate Binding Mechanism. *ACS Cent. Sci.* 2020, 6, 189–196.
47. Jung, J.; Nishima, W.; Daniels, M.; Bascom, G.; Kobayashi, C.; Adedoyin, A.; Wall, M.; Lappala, A.; Phillips, D.; Fischer, W.; et al. Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations. *J. Comput. Chem.* 2019, 40, 1919–1930.
48. Brooijmans, N.; Kuntz, I.D. Molecular recognition and docking algorithms. *Annu Rev. Biophys. Biomol. Struct.* 2003, 32, 335–373.
49. Pantsar, T.; Poso, A. Binding Affinity via Docking: Fact and Fiction. *Molecules* 2018, 23, 1899.
50. Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* 2018, 9, 923.