Extracellular proteoglycans

Shakibi, Saber; Onck, Patrick R.; Van, Erik; van der Giessen, Erik

Published in:
Biophysical Journal

DOI:
10.1016/j.bpj.2022.11.2279

Publication date:
2023

of complex systems; 2) extension and improvement of available molecule and force-field parameters; 3) management of feedback from the research community. To this end we created the vermouth python library as a unified framework for both program development and parameter management. Based on vermouth, tools such as martinize2, polyply, or fast-forward have been created, which facilitate setting up Martini simulations. In addition, improved parameters for small molecules and carbohydrates have been released with updates for other classes of biomolecules such as lipids underway. Lastly, by openly hosting all programs and parameters on GitHub, we give the research community an opportinuty to provide feedback and contribute to developments. Here we present how these individual developments make up the Martini 3 ecosystem for coarse-grained simulations and our current goals for future advances.

2039-Pos

An evaluation of force field accuracy for the mini-protein chignolin using Markov state models

Tim Marshall, Robert Raddi, Vincent Voelz.
Chemistry, Temple University, Philadelphia, PA, USA.

All-atom molecular dynamics (MD) simulations can provide detailed insight into a molecule's conformational ensemble in solution. While molecular force fields are parameterized to accurately model a protein's potential energy surface, it remains challenging in practice to evaluate how well force fields can capture ensemble-averaged experimental observables, since it requires simulation of the complete folding landscape. In this work, we employ massively parallel molecular simulations, performed using the Folding@home distributed computing platform, to investigate the ability of nine force fields (AMBER14SB, AMBER99, AMBER99SB, AMBER99SB-ildn, AMBER99SBnmr1-ildn, CHARMM22*, CHARMM27, CHARMM36 and OPLS-aa) with TIP3P explicit solvent to accurately reproduce experimental observables for chignolin, a beta-hairpin mini-protein with an experimental folding time of ∼600 ns. From over 200 microseconds of aggregate trajectory data, we constructed Markov state models (MSMs) to obtain estimates of thermodynamic and kinetic properties of chignolin in each force field. Quantitative assessment of the force fields was performed by comparing predicted and experimental folded populations, and the statistical agreement between predicted and experimental solution-state NMR observables. Our results rank force field accuracy for chignolin similarly compared to previous work using oligopeptides and ubiquitin (Beauchamp et al. 2012). This work highlights the utility of MSM approaches for force field evaluation, and provides a baseline for future studies using Bayesian inference methods to evaluate and parameterize force fields. It also provides an excellent dataset for testing new approaches to simulation-based sequence design of mini-proteins.

2040-Pos

Self-guided molecular dynamics simulation based on concerted movement

Xiongwu Wu, Bernard R. Brooks.
Laboratory of Cell Biology, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

Self-guided molecular/Langevin dynamics (SGMD/SGLD) simulation methods were developed to enhance conformational sampling through promoting low frequency motion of molecular systems and have been successfully applied in many simulation studies. The low frequency properties are derived effectively from local averages over the most recent trajectories. Using low frequency momentums and low frequency forces, SGMD/SGLD achieve accelerated diffusion and enhanced energy barrier crossing. In addition to the local frequency motion, we find concerted movement is important for rare event like protein folding and assembly. This work presents a new self-guided simulation method that accelerates conformational search through enhancing concerted movement. We abbreviate this method as SGMDc or SGLDc. Concerted motion is derived from spatial averages over chemical bonded atoms and/or nearby atoms. Combined with the local average scheme, one can derive concerted low frequency properties and use them as guiding forces to achieve accelerated conformational evolvement. Using peptide folding and pore formation in membrane, we demonstrate that SGMDc/SGLDc can accelerate rare events significantly.

2041-Pos

Building a community-driven ecosystem for fast, reproducible, and reusable molecular simulation analysis using mdanalysis

Irfan Alibay[1], Jonathan Barnoud[2], Oliver Beckstein[3], Richard J. Gowers[4], Philip R. Loche[5], Hugo MacDermott-Opeskin[6], Micaela Matta[7], Fiona B. Naughton[8], Tyler Reddy[9], Lily Wang[10].
[1]Biochemistry, The University of Oxford, Oxford, United Kingdom, [2]Centro Singular de Investigaciòn en Tecnoloxìas Intelixentes, Universidade de Santiago de Compostela, Santiago de Compostela, Spain, [3]Department of Physics, Arizona State University, Tempe, AZ, USA, [4]Open Molecular Software Foundation, Irvine, CA, USA, [5]Laboratory of Computational Science and Modeling, Institute of Materials, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, [6]Research School of Chemistry, Australian National University, Canberra, Australia, [7]Department of Chemistry, King's College London, London, United Kingdom, [8]Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA, USA, [9]CCS-7 Applied Computer Science, Los Alamos National Laboratory, Los Alamos, NM, USA, [10]Open Force Field Consortium, Davis, CA, USA.

MDAnalysis (https://www.mdanalysis.org) is a widely used open-source Python library that enables scientists to handle and analyze simulation data in a transparent and simulation engine agnostic manner. It offers a variety of built-in analysis methods, efficient routines for coordinate manipulation, and supports over 40 file formats covering most popular simulation packages. With over 165 contributors and 17 years of development, the MDAnalysis library has established a mature, stable API and a wide user base. Combined with a focus on interoperability and extensibility, this has allowed users to rely on its components to create a wealth of new workflows and tools to explore the increasingly complex systems which they model.

Here we present the current state of the MDAnalysis library, including recent changes such as our interoperability-focused converters framework and new performance improvements to core routines. We also detail ongoing work to address modern challenges in the ever-evolving landscape of molecular simulations, notably meeting the tenets of FAIR (findability, accessibility, interoperability, and reusability). We outline a new framework for an ecosystem of MDAKits (MDAnalysis toolkits) which aims to enable scientists to create, and eventually publish, new software tools that adhere to FAIR principles. MDAKits are independent add-on packages that expand the core functionality of the MDAnalysis library, and meet a set of software standards to be reliably used by other scientists. We provide an overview of the various tools and processes that are being developed to support the MDAKit ecosystem, including cookiecutter templates, continuous testing workflows, a registry to publicize MDAKits, and incentives towards continuous improvement.

2042-Pos

Hypersound-perturbed molecular dynamics simulation for accelerating slow biomolecular interaction processes

Mitsugu Araki[1], Yasushi Okuno[2].
[1]Graduate School of Medicine, Kyoto University, Kyoto, Japan, [2]Kyoto University, Kyoto, Japan.

Capturing the dynamic processes of biomolecular systems in atomistic detail remains difficult despite recent experimental advances. Although molecular dynamics (MD) techniques enable atomic-level observations, simulations of ''slow'' biomolecular processes (with timescales longer than submilliseconds) are challenging because of current computer speed limitations. Therefore, we developed a method to accelerate MD simulations by high-frequency ultrasound (hypersound) perturbation. The binding events between the protein CDK2 and its small-molecule inhibitors were nearly undetectable in 100-ns standard MD simulations, but the method successfully accelerated their slow binding rates by up to 10-20 times. Hypersound-accelerated MD simulations revealed a variety of microscopic kinetic features of the inhibitors on the protein surface, such as the existence of different binding pathways to the active site. Moreover, the simulations allowed the estimation of the corresponding kinetic parameters and exploring other druggable pockets. This method can thus provide deeper insight into the microscopic interactions controlling biomolecular processes.

2043-Pos

Extracellular proteoglycans: A multiscale computational study

Saber Shakibi, Patrick R. Onck, Erik Van, der Giessen.
Micromechanics, Zernike Institute for Advanced Materials, University of Groningen, Groningen, Netherlands.

The extracellular matrix (ECM) plays an important role in cell migration and proliferation, and hence in the progression of malignant tumors. Healthy brain is one of the softest tissues, but in cancer, the stiffness of the brain ECM can increase by one or two orders of magnitude. Within the framework of studying mechanotransduction in brain tumors, we aim to construct a multiscale computational model of the brain ECM. In contrast to most tissues whose mechanical properties are governed primarily by collagen, the brain ECM is mainly composed of proteoglycans and hyaluronic acid. During cancer, expression of different proteoglycans is altered during remodeling of the brain ECM.

Proteoglycans are composed of a core protein to which a number of glycosaminoglycan chains are attached. One of the best characterized proteoglycans is

aggrecan, the key ingredient of cartilage. The biophysical properties of proteoglycans such as brevican, neurocan and versican, which are predominant in the brain ECM, however, are much less known. In this study, we use computational modeling to predict the biophysical properties of these proteoglycans. Due to the large scale of these macromolecules (with molecular weights of a few MDa) coarse-grained models are needed to model proteoglycans. In this work, we combine a one-bead-per-aminoacid (1BPA) model, developed by Onck and coworkers, to represent the core protein with our recently developed one-bead-per-saccharide (1BPS) model for modeling glycosaminoglycans. The combination of these two coarse-grained models allows us to predict the properties of proteoglycans as a function of the length of the side chains, grafting density, salt concentration and length of the backbone. Validation of the model is performed with respect to the biophysical properties of aggrecan.

2044-Pos

Development and benchmarking of an open, self-consistent force field for proteins and small molecules from the open force field initiative

Chapin E. Cavender[1], Pavan K. Behara[2], Simon Boothroyd[3,4], David L. Dotson[3,5], Joshua T. Horton[6], Joshua A. Mitchell[3], Ivan Pulido[7], Matthew W. Thompson[3], Jeffrey Wagner[3], Lily Wang[3], John D. Chodera[7], Daniel J. Cole[6], David L. Mobley[2], Michael R. Shirts[8], Michael K. Gilson[1].
[1]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA, [2]Department of Pharmaceutical Sciences, University of California Irvine, Irvine, CA, USA, [3]Open Force Field Consortium, Open Molecular Software Foundation, Davis, CA, USA, [4]Boothroyd Scientific Consulting Ltd, London, United Kingdom, [5]Datryllic LLC, Phoenix, AZ, USA, [6]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom, [7]Computational & Systems Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA, [8]Chemical & Biological Engineering, University of Colorado at Boulder, Boulder, CO, USA.

Contemporary protein force fields model canonical proteins well, but parameters for covalently modified residues are challenging to develop. Additionally, protein-specific parameters are derived separately from those for small molecules, limiting the transferability of small molecule parameters to protein contexts. A flexible force field that provides self-consistent parameters for proteins and small molecules would advance research in drug design and molecular mechanisms of protein structure and function. The Open Force Field (OpenFF) Initiative develops open, reproducible force fields for atomistic simulations, delivering automated infrastructure and systematic methodology for force field fitting and validation as well as version-controlled force field releases. OpenFF force fields assign parameters using direct chemical perception via the SMIRKS cheminformatics language, providing coverage of a large chemical space with fewer parameters than conventional atom-typed force fields. The most recent force field release, Sage (OpenFF 2.0.0), provides valence parameters trained against quantum chemical (QC) data and Lennard-Jones parameters trained against condensed phase properties of pure liquids and binary mixtures. Sage achieves competitive accuracy on benchmarks of the energies and geometries of QC minima, solvation free energies, and protein-ligand binding free energies. Here we describe the extension of Sage to proteins, resulting in a self-consistent force field that can simulate canonical and covalently modified proteins and small molecules. We used protein-specific SMIRKS to train proper torsions to two-dimensional QC scans of the backbone and sidechain dihedrals of capped peptides. The parameters were validated using benchmarks of published NMR observables for small peptides, folded proteins, and disordered proteins. Beyond force field parameters, the OpenFF software infrastructure now supports loading proteins from PDB files, iterating over hierarchies such as residues, and exporting parametrized systems to common molecular dynamics formats.

2045-Pos

The onset of whole-cell modeling using the martini force field

Jan A. Stevens[1], Fabian Grunewald[1], Marco van Tilburg[1], Zane R. Thornburg[2], Benjamin R. Gilbert[2], Troy A. Brier[2], Zaida Luthey-Schulten[2], Siewert J. Marrink[1].
[1]Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, Netherlands, [2]Chemistry, University of Illinois Urbana-Champaign, Urbana, IL, USA.

Molecular dynamics (MD) is a well established simulation method, which has successfully been applied to study a wide range of biomolecular processes. Continuous improvements in both computational infrastructure and modelling methods have enabled scientists to study mesoscopic, multi-component systems using MD simulations. Understanding how biomolecular functions emerge from millions of interacting molecules is viewed as the next research objective. Since biomolecular processes function on a hierarchy of interconnected scales, decoding the complexity of cellular environments requires us to study them as a whole. We present our ongoing effort to construct a whole-cell model at a molecular resolution, using the Martini coarse-grained force field. We focus our efforts on modeling a genetically minimal cell: the JCVI-syn3A. This cell is engineered to have minimal complexity, making it an ideal starting point for this project. The cell's structural organization and composition is gathered in an integrative modeling approach, combining experimentally resolved data and stochastic lattice model simulations of the whole cell. Incorporating the necessary data into our MD model requires specialized tools, which will be part of the Martini3 ecosystem. The first version of our whole-cell model is presented, showing that performing MD simulations at this scale is feasible. Despite simplifying the cell composition in some key components, this model will provide a starting point for further improvements and pushing the boundaries of realistic whole-cell modeling. Computational microscopy of entire cells and cell organelles will provide valuable insights in a wide range of problems, ranging from drug design to understanding the internal organization of the cellular environments.

2046-Pos

Evaluation of the CHARMM36m force field combined with the OPC water model for protein simulations

Nicolai Kozlowski, Gabor Nagy, Helmut Grubmuller.
Theoretical and Computational Biophysics, Max Planck Institute for Multidisciplinary Sciences, Gottingen, Germany.

The accuracy of atomistic molecular dynamics (MD) simulations depends on the accuracy of the used force field. For proteins many force fields have been developed so far, each in combination with a particular water model. Typically, combining a protein force field with a water model for which is has not been developed is not expected to yield reliable results. However, recently the combination of the CHARMM36m force field with the OPC water model was shown to accurately estimate the compactness and secondary structure content of intrinsically disordered proteins. Whether CHARMM36m+OPC yields similar accuracy for globular proteins as well has, however, not yet been evaluated systematically. Here, we benchmark this combination on a set of six different globular proteins. To this end, we performed 50 x 1 μs MD simulations per protein using CHARMM36m+OPC and, for comparison, the same simulations using the well-established Amber99SB-ILDN+TIP4P force field. We compared the generated ensembles by means of RMSD, radius of gyration and secondary structure content and also compared the conformational dynamics. We found that CHARMM36m+OPC generates less compact ensembles and shows higher barriers for conformational transitions. Furthermore, we compare ensembles of both force fields to experimental crystal structures, B-factors, and NMR chemical shifts. Here, we found that both force fields compare equally well to experimental data, with CHARMM36m+OPC ensembles agreeing with chemical shifts slightly better, whereas Amber99SB-ILDN+TIP4P ensembles better agree with crystal structures. Our results show that CHARMM36m+OPC and Amber99SB-ILDN+TIP4P yield similar accuracy for globular proteins. Combined with earlier results for disordered proteins, these findings suggest that CHARMM36m+OPC should provide good accuracy for a broad range of disordered and folded proteins as well.

2047-Pos

Data-driven genome compartmentalization with nuclear landmarks

Zhuohan Lao.
Massachusetts Institute of Technology, Cambridge, MA, USA.

In eukaryotic cells, the three-dimensional (3D) organization of genomes plays a crucial role in genome function and has inspired the development of numerous innovative experimental techniques for its characterization. Although researchers have made a great deal of significant progress in deciphering the folding mechanisms of an individual chromosome, the principles of the dynamic large-scale spatial arrangement of the whole genome inside the nucleus are poorly understood. Following the maximum entropy principle pioneered by Zhang and Wolynes, we used polymer simulations to develop a predictive model for the whole genome at one hundred kilo-base resolution (100 KB) with nuclear bodies such as nuclear lamina, nucleoli, and speckles and parameterized a force field to study genome structure and dynamics using genome-wide chromosome conformation capture data. We discovered that self-organization based on a co-phased separation between chromosomes and nuclear bodies process can capture various features of genome organization, including the formation of chromosome territories, and phase separation of A/B compartments. Both sequencing-based genomic mapping and imaging assays that probe chromatin interaction with nuclear bodies are quantitatively reproduced with the simulated 3D structures. Essentially, we captured the heterogeneous distribution of chromosome positioning across cells, while