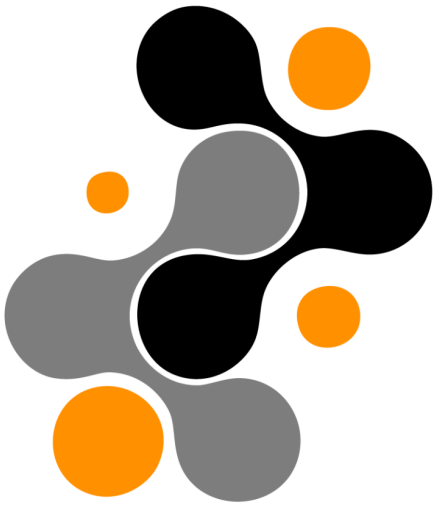


MD ANALYSIS UGM 2024

Abstracts Book

21-23 August, 2024

Lecture Theatre 1, Bush House,
Strand Campus, King's College London,
30 Aldwych, London, WC2B 4BG
United Kingdom

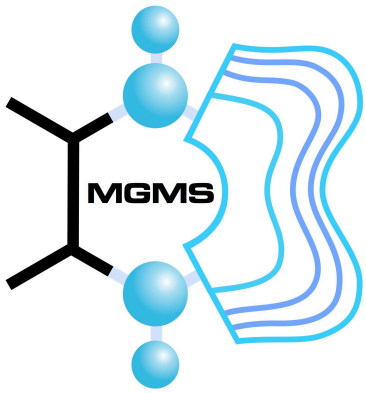
TABLE OF CONTENTS

Welcome to the 2024 MDAnalysis User Group Meeting. The following abstracts of presenting authors have been sorted into keynote, oral, and poster presentations. Poster presentation titles with three asterisks (***) indicate those presented as lightning talks as well. Abstracts are sorted alphabetically by the surname of the presenting author.

PARTNERS AND SPONSORS

We would like to give a special thanks to our partners and sponsors for this workshop, the Thomas Young Centre (TYC), Chan Zuckerberg Initiative (CZI), and Molecular Graphics and Modelling Society (MGMS). MDAnalysis also thanks NumFOCUS for its continued support as our fiscal sponsor.

Chan Zuckerberg Initiative



NUMFOCUS
[FISCALLY SPONSORED PROJECT]

KEYNOTE TALKS

From a Molecular Movie of a Protein to Quantitative Data

Antonia Mey, University of Edinburgh

KEYWORDS: Biomolecular Simulations, Drug Discovery

Biomolecular simulations are a great tool for understanding how proteins move and interact with small molecules at an atomistic level. But how do we get from a trajectory that provides a pretty movie of a wiggling protein to measurable observables such as a free energy of binding or a rate of interconversion between two metastable protein states? I will walk you through two applications of using molecular simulations and how to effectively analyse these to provide predictive tools for experiments. The first application is how we can use molecular simulations to compute free energies of binding for metalloproteins using a simulation technique called alchemical free energy methods. The second application will look at how we can effectively analyse ultra-large trajectory datasets such as 1 ms worth of molecular dynamics simulations of a kinase. The emphasis for both applications will be how well we can compare to experiments and how we can make use of tools such as MDAnalysis to check the integrity of our simulation data in an automated fashion.

Molecular Dynamics for Drug Discovery: Insights into Protein, Ligand, and Protein-Ligand Complexes

Francesca Stanzione (speaker, Senior Computational Chemist) and Ijen Chen

KEYWORDS: Biomolecular Simulations, Drug Discovery

Molecular Dynamics (MD) simulations are indispensable tools for elucidating the dynamic behaviour of biological molecules such as protein and their interaction with ligands. From drug discovery to structural biology, MD simulations offer a unique perspective, enabling researchers to explore the intricate interplay between proteins and ligands, revealing conformational changes, binding and activation mechanisms, and thermodynamics in exquisite detail. Despite significant progress, challenges remain in simulating large biomolecular systems over extended time scales with high accuracy.

Here, we explore the role of MD simulations in drug discovery and design, from activation mechanisms to rationalise the development of novel therapeutics, highlighting the importance of validation and benchmarking against experimental data to ensure the reliability and predictive power of MD simulations. Furthermore, we underscore the efforts to address some of the challenges, including the development of enhanced sampling techniques, force field refinement and the integration of multi-scale modelling approaches.

ORAL PRESENTATIONS

Mechanism of Ligand Binding to Target RNA Aptamer

Sana Akhter, University of Kansas

KEYWORDS: Biomolecular Simulations, Drug Discovery

With growing interest in targeting RNA for drug design, it has become crucial to understand the mechanisms of RNA-ligand interactions and characterize their binding thermodynamics and kinetics quantitatively. Experimental techniques such as X-ray crystallography are powerful for determining structures of RNA-ligand complexes, but it is challenging to probe their dynamic interactions and characterize the binding thermodynamics and kinetics for drug design. In this study, we have combined biochemical binding assays and accelerated molecular simulations to investigate ligand binding and dissociation in RNA, using the well-studied theophylline-binding RNA aptamer as a model system. All-atom simulations using a novel Ligand Gaussian accelerated Molecular Dynamics method (LiGaMD) have, for the first time to our knowledge, captured repetitive binding and dissociation of theophylline and caffeine to the RNA. This allowed us to characterize the free energy landscapes of the RNA-ligand interactions. The simulation predicted theophylline binding free energy and kinetic rate constants agreed with experimental data. LiGaMD simulations allowed us to identify distinct low-energy conformational states and multiple pathways of ligand binding to the RNA. LiGaMD simulations revealed a "conformational selection" mechanism for ligand binding to the flexible RNA aptamer. In summary, our complementary biochemical experiments and accelerated simulations have provided important mechanistic insights into ligand binding to the theophylline-binding model system. Our findings suggest that compound docking using structural ensemble of representative RNA conformations (“ensemble docking”) would be necessary for structure-based drug design of flexible RNA.

Using MDAnalysis for Machine Learning: Non-parametric Bayesian Kinetic Clustering

Oliver Beckstein, Arizona State University

KEYWORDS: Soft Matter, Biomolecular Simulations, Developing Toolkits using MDAnalysis

MDAnalysis is a widely-used Python library for processing data from particle-based molecular simulations by reading over 40 different file formats and making the data available through an object-oriented interface that is based on the ubiquitous NumPy array. Thanks to its flexibility it lends itself to the initial steps in machine learning workflows to extract features from simulation trajectories. We will introduce MDAnalysis with simple examples of applying classical machine learning approaches such as clustering and classifiers from scikit-learn to dihedral angles or RMSD matrices. We will then introduce a new method to perform non-parametric Bayesian analysis to quantify lipid binding to proteins (specifically, residency times of cholesterol with specific residues in GPCRs in coarse grained molecular dynamics simulations). Such an

analysis characterizes binding processes at different timescales (quantified by its kinetic off rate) and assigns to each trajectory frame a probability to belong to a specific process. In this way we can classify trajectory frames as belonging to a particular process in an unsupervised manner and obtain, for example, different binding poses or molecular densities based on the process. The non-parametric Bayesian analysis allows us to connect the coarse binding time series data to the underlying molecular picture and so not only infer accurate binding kinetics with error distributions from MD simulations but also describe the molecular events responsible for the varying kinetic rates.

Molecular Dynamics Analysis with N-ary Clustering Ensembles (MDANCE), A Novel Clustering Package Based on N-ary Similarity

Lexin Chen, University of Florida

KEYWORDS: Biomolecular Simulations, Developing Toolkits using MDAnalysis

Molecular Dynamics (MD) simulations offer a computational window into intricate biological processes, yet the challenge lies in extending their scope to encompass longer timescales and larger systems. Unfortunately, the post-processing analysis of MD trajectories has struggled to keep pace with this demand, particularly in the realm of clustering techniques. Clustering is essential for unraveling protein folding dynamics, constructing Markov State Models, enhancing Replica Exchange simulations, and discerning drug binding modes. The predicament arises from a stark trade-off: conventional algorithms such as k-means prove efficient but fall short in identifying subtle metastable states, while more robust methods like density-based clustering incur significant computational overhead. In this contribution, we unveil a transformative solution in the form of Molecular Dynamics Analysis with N-ary Clustering Ensembles (MDANCE) software. One of the key novelties in MDANCE is that it provides the tools to quickly implement and test novel clustering algorithms. MDANCE stands as an open-source powerhouse, serving both as a user-friendly tool for analyzing diverse MD simulations and as a versatile platform empowering developers with unparalleled flexibility in crafting and integrating clustering and post-processing tools. MDANCE transcends the limitations of traditional approaches by introducing novel, highly efficient clustering algorithms.

Molearn: Streamlining the Design of Generative Models of Biomolecular Dynamics

Matteo Degiacomi, Durham University

KEYWORDS: Biomolecular Simulations, Deep Learning

Characterising the structure and dynamics of proteins at the atomic level provides a fundamental understanding of the mechanisms underpinning life, and is the first step in numerous technological applications. Molecular dynamics (MD) simulations can help gaining such an insight, however exhaustive sampling of many key biological phenomena lay beyond

the reach of typical simulations. In this context, we have recently demonstrated that additional insight can be gained by combining MD simulations with Generative Neural Networks (GNNs). We have developed GNNs capable of producing protein conformations based on small pools of examples produced by MD. While this is a promising research direction, the issue is that developing a Machine Learning model to study biomolecular dynamics is a lengthy process. This requires setting up means of transforming conformational space data into tensor data submittable to a model, as well as assessing a model's quality (e.g., in terms of their energy or according to structural descriptors). In this presentation we introduce molearn, a Python framework streamlining the implementation of generative neural networks learning protein conformational spaces from examples obtained via experiments or MD simulations.

Kinisi: Bayesian Analysis of Mass Transport from Molecular Dynamics Simulations

Josh Dunn, University of Bristol

KEYWORDS: Materials Science, Soft Matter, Biomolecular Simulations

Self-diffusion coefficients, *D*, are routinely estimated from molecular dynamics simulations by fitting a linear model to the observed mean-squared displacements (MSDs) of mobile species. MSDs derived from simulation suffer from statistical noise, which introduces uncertainty in the resulting estimate of *D*. An optimal scheme for estimating *D* will minimise this uncertainty, i.e., will have high statistical efficiency, and will give an accurate estimate of the uncertainty itself. We present a scheme for estimating *D* from a single simulation trajectory with high statistical efficiency and accurately estimating the uncertainty in the predicted value. The statistical distribution of MSDs observable from a given simulation is modelled as a multivariate normal distribution using an estimated covariance matrix, which we parameterise from the available simulation data. We then perform Bayesian regression to sample the distribution of linear models that are compatible with this model multivariate normal distribution, to obtain a statistically efficient estimate of D^* and an accurate estimate of the associated statistical uncertainty. This methodology is implemented in an open-source Python package called kinisi.

CodeEntropy Software Development

Sarah Fegan, UKRI-STFC Scientific Computing Department

KEYWORDS: Soft Matter, Biomolecular Simulations, Multiscale Modeling, Developing Toolkits using MDAnalysis

Entropy is an interesting thermodynamic quantity which contributes to the free energy of any system. The CodeEntropy software is an open source software program which implements the multiscale cell correlation method [1] for determining the entropy of a system from molecular dynamics trajectories. This method breaks the entropy into independent terms and calculates each at different length scales. The first release of CodeEntropy used MDAnalysis for reading in simulation data. The newest version is also using the MDAnalysis universe data structure for

accessing position and force data, and using MDAnalysis for extracting connectivity data and analysing dihedral angles.

References:
[1] Chakravorty, Arghya; Higham, Jonathan and Henchman, Richard H. (2020). J. Chem. Inf. Model., 60, 5540−5551.

Choline Based Plastic Crystals as Barocaloric Materials: Insights from Ab Initio Molecular Dynamics

Shivani Grover, University of Edinburgh

KEYWORDS: Materials Science, Ab Initio Molecular Dynamics

Solid-state refrigeration technologies based on barocaloric effects offer a cleaner alternative solution to conventional refrigeration technologies based on compression cycles of greenhouse gases. Thus, materials with large solid-state caloric effects induced by external field (mechanical, electric or magnetic field) need further investigation and development to realise their full potential.[1,2] In the present work, we explore the potential of choline-based plastic crystals, $(C_5H_{14}NO)_2MX_4$ where M=Co, Zn, Cu, and X=Cl, Br, I, as promising barocaloric materials due to a large entropy change (~100 $JK^{-1}kg^{-1}$) associated with the symmetry-breaking disorder-order phase transition. Using ab initio molecular dynamics as implemented in CP2K, we provide insights into the structural dynamics and hydrogen bonding between the choline and $CoCl_4$ that characterise the two phases. From our calculation of vibrational density of states, we estimate the entropy change across the phase transition. The choline and $CoCl_4$ disorder are quantified by computing the diffusion coefficients using MDAnalysis code.[3] Our model is validated with our experimental study, thus providing a complete understanding of the interplay between dynamic and disordering effects in choline based plastic crystals.

References:
[1] Moya, X., Kar-Narayan, S. & Mathur, N. D. Caloric materials near ferroic phase transitions. Nature Materials 13, 439-450 (2014).
[2] Li, B. et al. Colossal barocaloric effects in plastic crystals. Nature 567, 506-510 (2019).
[3] Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. and Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. J. Comput. Chem. 32, 2319-2327 (2011).

Computer Simulations of the Ribosome

Michal H. Kolar, University of Chemistry and Technology, Prague

KEYWORDS: Biomolecular Simulations, Multiscale Modeling

Computer simulations have become a valuable tool for understanding the behavior of biological macromolecules. We focus on the ribosome, one of the most complex and essential cellular machines, which is responsible for protein synthesis in all known organisms. The

ribosome simulations are especially helpful in addressing questions about conformational heterogeneity and energetics of various ribosome parts. Despite the progress in hardware and software development, computersimulations of the entire ribosome remain challenging, especially at the all-atom level. The main reason is the ribosome size and complex chemical compositions. Still, using world-class supercomputers, one may gather valuable data in reasonable time.Over the past few years, we have been using all-atom and coarse-grained molecular dynamics simulations to investigate several ribosome'scritical sites. Namely, we have studied the exit tunnel through which nascent proteins leave the ribosome, the decoding center where correct tRNAs are recognized or a portion of ribosome surface where translation factors bind.In the talk, we will highlight the results of these projects and discuss how effective the computer simulations are in approaching scientific questions about the ribosome.

PySoftK 2.0: Tool for the Analysis of Interfaces, Interactions and Self-Assembly in Soft Matter Simulations

Raquel López-Ríos de Castro, Charité - Universitätsmedizin Berlin

KEYWORDS: Materials Science, Soft Matter, Biomolecular Simulations, Developing Toolkits using MDAnalysis

Molecular dynamics simulations have become an important tool in the study of soft matter and biological macromolecules. The large amount of high-dimensional data produced by such simulations does not immediately elucidate the atomistic mechanisms that underlie complex materials and molecular processes. Analysis of these simulations is complicated: the dynamics intrinsic to soft matter simulations necessitates careful application of specific (often complex) algorithms to extract meaningful molecular scale understanding. There is an ongoing need for high-quality computational workflows to facilitate this analysis in a reproducible manner with minimal user input. In this work, we introduce a series of new computational tools for analyzing soft matter interfaces, molecular interactions (including ring-ring stacking), and self-assembly. In addition, we include a number of auxiliary tools, including a useful function to unwrap molecular structures that are greater than half the length of their corresponding simulation box. These tools have now been added to the software package PySoftK [1] (https://alejandrosantanabonilla.github.io/pysoftk/), making application of these algorithms straightforward for the user. These new simulation analysis tools within PySoftK will not only allow all users to produce high-quality soft-matter simulations analysis, but will also contribute to the reproducibility of the analysis of soft matter and biomolecular simulations. This is key to bring about new predictive understanding in nano- and biotechnology.

References:
[1] Santana-Bonilla, A., López-Ríos de Castro, R., Sun, P., Ziolek, R. M., & Lorenz, C. D. (2023). Modular software for generating and modeling diverse polymer databases. Journal of Chemical Information and Modeling, 63(12), 3761-3771.

Building an Open Source Antiviral Drug Discovery Toolkit

Hugo MacDermott-Opeskin (speaker, ASAP Discovery), Jenke Scheen, Joshua Horton, Alexander M. Payne, Benjamin Kaminow, David Dotson, Michael M. Henry, Iván J. Pulido, Maria Castellanos, John D. Chodera, and ASAP Discovery Consortium

KEYWORDS: Drug Discovery

The history of drug discovery consists primarily of confidential planning, closed data, expensive or proprietary tooling and limited disclosure of endpoints and outcomes. The ASAP Discovery antiviral drug discovery consortium aims to change this with an open-science drug discovery campaign aimed at delivering globally accessible low-cost antivirals. As part of this effort, we have developed an open-source computational chemistry toolkit for antiviral drug discovery that covers many of the primary aspects of the design, make, test and analyze cycle. Built on an open source stack and integrating cutting edge components from Open Forcefield Open Free Energy Open Fold and MDAnalysis, the asapdiscovery toolkit enables rapid analysis of prospective compounds with machine learning, docking and free energy calculations as well as chemoinformatics and bioinformatics approaches. Additionally, the asapdiscovery toolkit integrates many common data services used by drug discovery teams across the industry. We detail the scientific and engineering challenges addressed by the asapdiscovery toolkit as well as examining future directions as we aim to generalize the toolkit and release subsets of functionality as standalone packages for use across the computational molecular sciences ecosystem.

The Effect of Missense Mutations on the Binding Pocket Dynamics of Skeletal Mysoin

Ivan Man, Queen Mary University of London

KEYWORDS: Biomolecular Simulations

Congenital myopathies are genetic disorders that lead to the impairment of skeletal muscle contractions. The overall prevalence of this disease is estimated to be 1:26000. The development of pharmaceutical treatments is challenging, and to date, no cure exists. Myosinopathies represent a class of congenital myopathies associated with mutations in different myosin isoforms. One of the most abundant myosin isoforms expressed in skeletal muscle is fast type IIa myosin (MYH2). A missense mutation in the MYH2 gene is often associated with muscle weakness, resulting in life-threatening conditions. Currently, there are 10 missense mutations identified for the MYH2 gene, which are found within the motor-neck domain. Following the discovery of Omecamtiv Mecarbil (OM), the first-in-class allosteric cardiac muscle activator, several myosin modulators have entered clinical trials, with Mavacamten being the first FDA-approved cardiac myosin inhibitor. Despite failing in a phase III clinical trial, OM has shown the possibility of selectively targeting and modulating cardiac myosin’s function. A recent computational study from this group has employed Molecular

Dynamics (MD) to rationalise the selectivity of OM in cardiac myosin over skeletal myosin. A hydrophobic subpocket adjacent to the OM binding pocket was identified specifically in the skeletal isoform, a finding that could be leveraged to design small molecules selective to skeletal myosin. One crucial aspect to consider when designing small molecules is the impact of mutations on protein dynamics, which could disrupt the target binding site. In this work, we present a computational investigation of the effects of four pathogenic MYH2 missense mutations using MD simulations. We found significant differences in the mutants compared to the wild-type form: i) the relay helix, a key element of myosin function and partially involved in OM binding, shows larger motions, ii) the dynamic correlations between the domains surrounding OM are weaker, and iii) the communication pathways between the mutation site and OM binding residues are less direct. Consistent with these findings, the OM binding pocket volume has larger fluctuations during the mutant simulations. These results can guide in silico drug design in skeletal myosin in the future.

MDGraphEmb: A Toolkit for the Encoding of Molecular Dynamics Data Using Graph Embedding

Ferdoos Hossein Nezhad, Brunel University London

KEYWORDS: Developing Toolkits using MDAnalysis

Molecular Dynamics (MD) simulations are routinely used to investigate protein dynamics and function. However efficient strategies to apply machine learning to simulation data are still missing. Graph embedding is a well-known powerful computing method that aims to automatically learn the low-dimensional node representation in a graph. The great computational advantage of using the embedding graph is that this representation can cast graph information into tabular data to build different classification tasks. We present MDGraphEmb, an object-oriented Python library that is built on top of MDAnalysis and MDSubSampler libraries. This toolkit facilitates the conversion of data from protein conformations into graphs and graph embedding. An application is presented for encoding and detection of important protein conformations from molecular dynamics simulations to classify functional states.

MDAutoMut: A Toolkit for the Automated Evaluation of the Impact of Mutations on Protein Dynamics

Namir Oues, Brunel University London

KEYWORDS: Biomolecular Simulations, Developing Toolkits using MDAnalysis, Evaluation of Mutations on Protein Dynamics

Motivation: Molecular dynamics (MD) simulations have become routine tools for the study of protein dynamics and function. Advances in methods for determining how mutations affect protein structure have made possible gaining insights on protein function and drug discovery.

However, there is still a critical gap in the lack of automated workflows for large scale evaluation of the impact of mutations on conformational changes in proteins.

Methods: We present MDAutoMut, an object-oriented Python library designed to automatically generate MD simulations, engineer mutations, and assess their impact on protein dynamics. The toolkit integrates with MDSubSampler[1] and MDAnalysis[2] and it extends their usability by providing a protocol to predict, via machine learning techniques, mutations that lead to desired changes in protein dynamics. MDAutoMut includes two core classes: the Protein class representing all information of the protein system and the Mutation class representing the mutation engine that links directly to PyRosetta[3] software. System preparation and MD simulations are done in GROMACS through gmxapi.

Results: MDAutoMut can run automated workflows to explore a list of single or double-site mutations and generate simulations of the mutated structures. Through integration with MDSubSampler, the workflow estimates changes in geometrical properties between the wild-type and the mutated structures. Mutations generating a desired change in protein dynamics can be automatically detected. The software is designed and implemented for three different user demographics. Novice users with limited experience in software development can execute preprepared scenario scripts available both as Python scripts and Jupyter notebooks. Advanced users can interact with a Unix-like command line interface offering flexibility and a wider choice of options. Scientific software developers can benefit from a set of reusable Python classes for ad-hoc customization and integration into existing workflows. In summary, the MDAutoMut library provides an automated approach for investigating mutations and their effects on conformational changes in proteins.

References:
[1] N. Oues, S.C. Dantu, R.J. Patel and A. Pandini., Bioinformatics (2023), 39, btad427
[2] N. Michaud-Agrawal, E.J. Denning, T.B. Woolf and O.Beckstein, J. Comput. Chem., (2011), 32, 2319-2327
[3] S. Chaudhury, S. Lyskov and J. J. Gray, Bioinformatics (2010), 26(5), 689-691

Availability: https://github.com/alepandini/MDAutoMut

Exploring Lipase Biocatalysis in Sugar-Based Natural Deep Eutectic Solvents for Production of Novel Polymeric Compounds

Özge Özkılınç, Università degli Studi di Udine

KEYWORDS: Biomolecular Simulations, Sugar-Based NADES, Lipase, Polyol Esters, Molecular Dynamics

Sugar-based natural deep eutectic solvents (NADESs) offer promising alternatives to conventional solvents due to their unique properties and tunability. Selecting the right solvent is

crucial for optimizing (bio)chemical processes to meet the needs of sustainable industry. However, how enzymes behave in these solvents is not fully understood. This study investigates the structural stability of lipase B from Candida antarctica (CalB) active sites in sugar-based NADES and examines the selectivity of catalysed esterification. Molecular dynamics (MD) simulations provide insights into microscopic interactions and structural organization within sugar-based NADESs. Our simulations focus on NADESs composed of choline chloride (ChCl) as a hydrogen bond acceptor (HBA) and various sugar bases (D-sorbitol, xylitol, D-arabitol, glucose, sucrose) as hydrogen bond donors (HBD). This comprehensive investigation enhances our understanding of how CalB catalyzes diesterification in sugar-based NADES, offering avenues for optimization and for diverse applications.

References:
[1] Buzatu, A.R.; Soler, M.A.; Fortuna, S.; Ozkilinc, O.; Dreavă, D.M.; Bîtcan, I.; Badea, V.; Giannozzi, P.; Fogolari, F.; Gardossi, L.; et al. Reactive Natural Deep Eutectic Solvents Increase Selectivity and Efficiency of Lipase Catalyzed Esterification of Carbohydrate Polyols. Catal. Today 2024, 426, 114373.

ClayCode: A Toolkit for Clay Simulation Setup and Analysis

Hannah Pollak, University of Edinburgh

KEYWORDS: Materials Science, Developing Toolkits using MDAnalysis

Clay minerals are a highly diverse group of materials with wide-ranging applications, from environmental sciences to materials science and drug delivery. Molecular dynamics simulations offer a powerful method for studying clay systems at the atomic level. However, the existing models often oversimplify clay structures, which leads to limitations in their accuracy and reliability. The primary challenge lies not in the computational resources or the methodology itself, but in the lack of user-friendly tools for preparing and analysing clay model systems.

To address this challenge, we introduce ClayCode, a user-friendly software equipped with tools for constructing realistic clay models and performing clay system-specific analyses. Written in Python, ClayCode utilises the MDAnalysis [1] library for several of its core functionalities. During the setup procedure, users can specify parameters such as stoichiometry, hydration, and ionic composition to generate representative clay systems. The models are assembled from a library of unit-cell building blocks and are assigned ClayFF [2] force field parameters. Upon completing the model construction, the software produces all the necessary input files for running simulations with the GROMACS [3] engine. Additionally, ClayCode includes analysis tools specifically designed to work seamlessly with models generated by ClayCode, enabling users to directly analyse trajectories within the ClayCode environment.

Through our work, we aim to streamline the workflow in molecular modelling of clay systems, from model construction to analysis, thereby contributing to the advancement of clay research.

References:
[1] Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., and Beckstein, O. (2011). MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. Journal of Computational Chemistry, 32(10), 2319-2327.
[2] Cygan, R. T., Liang, J. J., and Kalinichev, A. G. (2004). Molecular models of hydroxide, oxyhydroxide, and clay phases and the development of a general force field. Journal of Physical Chemistry B, 108(42), 1255-1266.
[3] Berendsen, H. J. C., van der Spoel, D., and van Drunen, R. (1995). GROMACS: a message-passing parallel molecular dynamics implementation. Computer Physics Communications, 91(1-3), 43-56.

Investigating Allosteric Inhibitory Mechanisms of the Soluble Epoxide Hydrolase

Evelyn Qiu, King's College London

KEYWORDS: Biomolecular Simulations

The soluble epoxide hydrolase (sEH) is a bifunctional enzyme with its C-terminal domain (CTD) responsible for hydrolysis of epoxy fatty acids (EpFAs). sEH is an attractive therapeutic target for many diseases such as cardiovascular diseases and pain as inhibiting sEH could restore higher levels of beneficial EpFAs. The majority of sEH inhibitors studied to date are orthosteric inhibitors that directly occupy the enzyme catalytic pocket. Although some of them have entered clinical trials, none has thus far been successful. Recently, several electrophilic fatty acids such as the endogenous 15-deoxy-Δ12,14-Prostaglandin J2 (15d-PGJ2) were found to allosterically inhibit sEH activity through covalently binding to C423 and C522 sites; however, the underlying mechanism of these allosteric inhibition remains unclear. This project aims to explore the allostery of sEH inhibition through a combination of experimental and computational approaches.

Replica exchange molecular dynamics (REMD) simulations were conducted and analysed with MDAnalysis. Results suggested that the binding of 15d-PGJ2 to C423 and C522 sites lead to changes in protein properties including protein conformation, residues flexibility and accessible surface. Further analysis performed with AlloHub highlighted several regions that may be potentially involved in the allosteric signalling pathway. In addition, experimental and computational work suggested sEH can be inhibited by metal ions such as Cu2+ in a similar allosteric manner.

Further simulations and experimental work have been planned to understand the complex allosteric communications of sEH inhibition. This would provide valuable information for design of new drugs that inhibit sEH using allosteric mechanisms.

Spatially Resolved Impedance Spectra from Molecular Dynamics Simulations: A Generalised Correlation Analysis Approach

Henrik Stooß, University of Stuttgart

KEYWORDS: Materials Science, Developing Toolkits using MDAnalysis, Physics

Equilibrium molecular dynamics simulations provide valuable insights into the fundamental properties of perturbed systems by utilising fluctuation-dissipation relationships. By quantifying the fluctuations of system variables, it becomes possible to anticipate the response of the system to external perturbations. Energy storage and conversion is a research field of considerable importance for the transition to green energies. Simulations on the atomic scale deliver an important perspective to advance materials research at the nano-scale. In this work, we present a generalised and memory-safe implementation for observable correlation analysis using the MDAnalysis library and demonstrate its applicability in extracting information from large (in space and time) scale systems. We apply this to a system bounded by electrodes at constant potential to measure the position dependent impedance spectrum.

POSTER PRESENTATIONS

*** Scrutinising the Conformational Ensemble of the Intrinsically Mixed-Folded Protein Galectin-3

Midhun Mohan Anila, Institute of Physics, Polish Academy of Sciences

KEYWORDS: Biomolecular Simulations, Galectin-3, Conformational Ensemble, Fuzzy Complexes, Molecular Dynamics Simulations, Martini 3

Galectin-3 is a protein involved in many intra- and extra-cellular processes. It has been identified as a diagnostic or prognostic biomarker for certain types of heart disease, kidney disease and cancer. Galectin-3 comprises a carbohydrate recognition domain (CRD) and an N-terminal domain (NTD), which is unstructured and contains eight collagen-like Pro-Gly-rich tandem repeats. While the structure of the CRD has been solved using protein crystallography, current knowledge about conformations of full-length galectin-3 is limited. To fill in this knowledge gap, we performed molecular dynamics (MD) simulations of full-length galectin-3. We systematically re-scaled the solute–solvent interactions in the Martini 3 force field to obtain the best possible agreement between available data from SAXS experiments and the ensemble of conformations generated in the MD simulations. The simulation conformations were found to be very diverse, as reflected, e.g., by (i) large fluctuations in the radius of gyration, ranging from about 2 to 5 nm, and (ii) multiple transient contacts made by amino acid residues in the NTD. Consistent with evidence from NMR experiments, contacts between the CRD and NTD were observed to not involve the carbohydrate-binding site on the CRD surface. Contacts within the NTD were found to be made most frequently by aromatic residues.

Formation of fuzzy complexes with unspecific stoichiometry was observed to be mediated mostly by the NTD. Taken together, we offer a detailed picture of the conformational ensemble of full-length galectin-3, which will be important for explaining the biological functions of this protein at the molecular level.

*** Structural Dynamics of a Metalloprotease Enzyme: Insights from Molecular Dynamics Simulations

Asal Azar, Durham University

KEYWORDS: Molecular Dynamics Simulations, MDAnalysis, Protein Dynamics, Structural Biology

Cu-dependent enzymes are crucial in numerous biological processes, yet the structural dynamics underpinning their function remain incompletely understood. Our research focuses on a Cu-dependent enzyme where the Cu-loaded form is a trimer, and the apo-form is a monomer. Trimer formation, dependent on Cu loading, is a slow process in vitro and appears to be irreversible. We hypothesize significant structural rearrangements between the apo- and holo-forms, particularly involving a C-terminal region of approximately 40 amino acids, which is suspected to be "loopy" in the apo-form and rigid in the holo-form. Utilizing molecular dynamics (MD) simulations and analyzing the results with the MDAnalysis package along with other Python tools, we have explored the structural transitions and stability of both forms. Preliminary results indicate notable conformational shifts that support our hypothesis of a dynamic C-terminal region. These insights enhance our understanding of the enzyme's activation mechanism and provide a framework for future structural studies.

*** Calculating Pair Distribution Functions in Anisotropic Geometries

Kira Fischer, University of Stuttgart

KEYWORDS: Materials Science, Developing Toolkits using MDAnalysis

The well known radial distribution function gives the spatial correlation between atoms, and provides insight into the microstructure and the formation of clusters. In anisotropic systems, such as interfaces and planar and cylindrical pores, the definition of the RDF breaks down. Therefore, we introduce the anisotropic pair distribution function for planar and cylindrical geometries and show how to calculate PDFs using the MDA Kit Maicos.

*** Molecular Dynamic Simulation of Methotrexate Drug with Silver Nanoparticle in Drug Delivery Across Cell Membrane

Kiran Gangarapu, Anurag University

KEYWORDS: Drug Discovery

This research employs molecular dynamics simulations to investigate the drug delivery potential of Methotrexate (MTX), conjugated with Silver nanoparticles (MTX-AgNPs) as drug carriers within a Lipopolysaccharide (LOP) membrane. The study focuses on understanding the dynamics of MTX drug, including their adhesion to and penetration into the LOP membrane. Simulation results reveal notable changes in the interaction between drug molecule and LOP upon conjugation with AgNPs. MTX, known for its versatility in anticancer therapy, exhibits altered diffusion within the lipid bilayer after conjugation with AgNP, leading to increased retention time at the LOP surface. The presence of MTX-AgNPs at the membrane surface enhances its impact on blocking estrogen binding at estrogen receptors, contributing to tumor cell growth arrest. Overall, these molecular dynamics simulation results offer valuable insights into the interplay of MTX-AgNPs with the LOP membrane and AgNPs, potentially addressing challenges in nanomedicine translation and optimizing the balance between targeting efficacy and minimizing potential nanotoxicity associated with metal nanoparticles.

*** Isotropic, Semi-isotropic, and Anisotropic Rotational Diffusion from Molecular Dynamics Trajectories

Simon Holtbruegge, Ruhr-University Bochum

KEYWORDS: Biomolecular Simulations

The random rotational dynamics of (bio-)molecules in solution play a crucial role in macromolecular assembly, protein-ligand binding, or nuclear magnetic resonance (NMR) relaxation experiments. The theoretical description of Brownian rigid-body rotation by Favro (1960) provides the foundation to extract the three rotation axes and the anisotropic diffusion tensor of a molecule from Molecular Dynamics trajectories using quaternion correlation functions. This method is extended to enable fitting diffusion tensors in semi-isotropic or isotropic approximation. Implementation details and best practices are discussed based on model simulations of ideal Brownian rotors and Molecular Dynamics simulations of Ubiquitin.

*** OpenMMDL: A Workflow for Molecular Dynamics Simulations of Protein-Ligand Complexes Setup, Simulation and Analysis

Valerij Talagayev, Free University of Berlin

KEYWORDS: Biomolecular Simulations, Drug Discovery, Developing Toolkits using MDAnalysis

Molecular dynamics (MD) simulations have become an indispensable tool in recent scientific research. Their ability to provide insight into the behaviour of protein-ligand complexes facilitates the rational design of novel drugs and therapeutics with increased precision and efficacy. OpenMM is a powerful MD simulation software package that allows the dynamics of molecular systems to be simulated with exceptional efficiency. The open source nature of OpenMM, together with its support for multiple platforms, has made it an important tool in the field of computational drug design. One of the main difficulties in using OpenMM for

protein-ligand complexes is the difficulty of setting up the simulation, which requires expertise in configuring force fields, selecting appropriate parameters and managing simulation protocols. In this work, I would like to present the workflow called OpenMMDL, an easy setup of OpenMM MD simulations of protein-ligand complexes and the analysis of interactions during the simulation.

The first part of the workflow consists of the OpenMMDL setup, which is a flask-based, easy-to-use interface that allows the user to use their files as input for the simulation. The interface allows the user to modify all the necessary parameters of the simulation, thus allowing the generation of a simulation script that is perfectly suited to the user. The second part of the workflow is called OpenMMDL Simulation and performs the MD simulation and post-processing of the simulation using MDAnalysis and MDTraj, providing the user with the output of the MD simulation of the complex, which can be used directly for further analysis. The final part of the workflow is called OpenMMDL Analysis, which uses PLIP, RDKit and MDAnalysis and allows the user to view the interactions of the protein and ligand during the MD trajectory, including the most common binding modes present in the MD simulation, as well as the transitions of the binding modes during the MD simulation.

*** Deep Learning for Binding Site Segmentation in Protein Ensembles

Yu-Yuan (Stuart) Yang, Queen Mary University of London

KEYWORDS: Biomolecular Simulations, Deep Learning, Drug Discovery, Developing Toolkits using MDAnalysis

Multiple protein conformations are used to accurately represent binding site structure and dynamics in ensemble docking or de novo drug design. Following conformational selection, these conformations are often generated through molecular dynamics (MD) simulations of the protein in the ligand- free (apo) state. However, it can be difficult to identify within these ensembles the conformations most likely to bind a ligand in the absence of prior information on the ligand-bound (holo) state. This research aims to develop a deep learning (DL) tool that can scan a putative binding site in a protein and recognise the shape and physicochemical features compatible with the binding of chemical fragments using semantic segmentation. The predictions from this tool can be utilised to identify holo-like conformations within an MD trajectory. Here, we present a prototype of the tool developed using a 3D U-Net architecture on MD simulations of the complex between cardiac myosin and the first-in-class cardiac myotrope omecamtiv mecarbil (OM). Preliminary results show that the model can segment the binding site with a good level of accuracy. Future work will aim to test the ability of the tool to differentiate holo-like and apo conformations of myosin.

*** Unraveling the Molecular Dance: Insights into TREM2/DAP12 Complex Formation in Alzheimer's Disease through Molecular Dynamics Simulations

Zhiwen Zhong, King's College London

KEYWORDS: Biomolecular Simulations, Deep Learning

Alzheimer’s disease (AD) is a widespread neurodegenerative condition affecting millions globally. Recent research has implicated variants of the triggering receptor expressed in myeloid cells 2 (TREM2) as risk factors for AD. TREM2, an immunomodulatory receptor on microglial surfaces, plays a pivotal role in regulating microglial activation by association with DNAX-activation protein 12 (DAP12). Despite its significance, the mechanism underlying the formation of the complex between the transmembrane domains (TMDs) of TREM2 and DAP12 remains unclear. This study employs multiscale molecular dynamics (MD) simulations to investigate three TMD complex models, including two derived from experiments and one generated by AlphaFold2. Conducted within a lipid membrane consisting of an 80:20 mixture of phosphatidylcholine (POPC) and cholesterol, our analysis reveals hydrogen-bonding interactions between K26 of TREM2 and D16 of DAP12 in all three models. We also simulated the four mutation in TREM2/DAP12 complex, namely K26A, K26X, W34A, W34X in TREM2, among them W34X is reported to related to AD. Our findings enhance our understanding of the molecular mechanism governing TREM2/DAP12 complex formation, providing a foundation for designing novel therapeutic strategies to address AD and other neurodegenerative diseases.