Theoretical and Physical Chemistry Institute

National Hellenic Research Foundation

Vass. Constantinou 48, Athens

LECTURE

"Efficient molecular simulation and AI approaches to determining structure-dynamics-function relationships of proteins as a step towards therapeutic discovery"

Dr. Z. F. Brotzakis

Department of Chemistry, University of Cambridge, Cambridge, United Kingdom

and

BSRC Al. Fleming, Athens, Greece

Thursday, September 7, 2023, 12:00

Seminar room, ground floor, NHRF

Efficient molecular simulation and AI approaches to determining structure-dynamics-function relationships of proteins as a step towards therapeutic discovery

In biology, the function or misfunction of the cell attributed to disease is greatly manifested as (de)regulation of molecular reactions occurring at the level of biomolecules such as DNA, RNA and proteins. To be able to understand and control such a molecular microverse against disease, I will showcase recently developed computational chemistry methods spanning enhanced sampling molecular dynamics and molecular forcefield optimization, to integrative structural biology and computer aided drug discovery.

The accuracy of molecular simulations (MD) in predicting microscopic ensembles where thermodynamic and kinetic macroscopic properties are derived, is delimited by the quality of the sampling and the underlying force-field. The first problem is addressed by showcasing Transition Path Sampling and optimal collective variable Metadynamics based simulations able to converge μs-ms timescale transitions of protein conformational transitions (L99A T4-Lysozyme) [1], binding (β-lactoglobulin) [2] and small-molecule protein (un)binding (Benzamidin-Trypsin) [3]. The corresponding microscopic mechanisms are resolved as well as thermodynamic/kinetic properties such as the conformational free energy landscape, (un)binding reaction coordinates, (un)binding transition rates and affinity, as well as transition state ensembles of these reactions.

However, such macroscopic properties depend on the forcefield accuracy. Two avenues for optimizing forcefields will be showcased. First, an approximate Bayesian computation framework able to calibrate TIP4P water force-field parameters in order to reproduce selected target structural experimental values such as the radial distribution function [4]. Followingly, I will present a Maximum Caliber and Path Reweighting based-approach to calibrate forcefield parameters in order to target experimental transition rate constants for various toy models [5]. A second way to address the forcefield problem will be presented. In particular, using hybrid methods introducing experimental data such as transition rate constants [6], Cryo-EM densities [7-8] and even Alpha-Fold predictions into molecular simulations on the fly in the form of constraints [9]. Such methods shift MD ensembles to agree with the experimental data according to the maximum entropy principle. In this manner, accurate microscopic ensembles are determined that span antibody-antigen complexes with SARS-CoV2 Spike, the Alzheimer's disease (AD) related tau-microtubule complex and misfolding disease related intrinsically disordered proteins. This knowledge gives unprecedented functional understanding of these proteins, such as that antibodies binding to Spike are destabilized by conformational entropy, identifying post-translational modifications of tau that destabilize the tau-microtubule complex which is related to the onset of AD, determining ensembles of multiple IDPs which is experimentally impossible to achieve due to their intrinsic dynamics.

Finally, I will address avenues to optimally navigate the chemical space of compounds with therapeutic function, for instance small molecules or proteins able to bind molecular targets by inhibiting aberrant protein reaction leading to disease such as Parkinsons disease (PD). In particular I will show that by combining molecular docking datasets with Gaussian Process regression it is possible to suggest inhibitors of α-synuclein aggregation, which is the hallmark pathology in PD [10].

Reference

1. Mendels D., Piccini G.M., Brotzakis Z.F., Yang Y. & Parrinello M. (2018) Folding a Small Protein Using Harmonic Linear Discriminant Analysis. *J. Chem. Phys*. 149(19), 194113
2. Brotzakis Z.F. & Bolhuis P.G. (2019) Unbiased Atomistic Insight in the Mechanisms and Solvent Role for Globular Protein Dimer Dissociation *J. Phys. Chem*. B 123 (9), 1883-1895
3. Brotzakis Z.F., Limongelli V. & Parrinello M. (2019) Accelerating the Calculation of Protein-Ligand Binding Free Energy and Residence Times using Dynamically Optimized Collective Variables. *J. Chem. Theory Comput.* 15(1) 743-750
4. Dutta R., Brotzakis Z.F. & Mira A. (2018) Bayesian Calibration of Force-fields from Experimental Data: TIP4P Water. J. Chem. Phys. 149: 154110
5. Bolhuis P.G., Brotzakis Z.F. & Keller B. (2022) Imposing kinetic constraints by force field optimization using Maximum Caliber for reweighted continuum path ensembles. *arxiv*, 2207.04558 (under review in J. Chem. Phys.)
6. Brotzakis Z.F., Vendruscolo M. & Bolhuis P.G. (2021) A method of incorporating rate constants as kinetic constraints in molecular dynamics simulations, *Proc. Nat. Acad. Sci*. USA 118 (2) e2012423118
7. Mikolajek+ H, Weckener+M, Brotzakis+ Z.F., Huo J et. al. (2022) Correlation between binding affinity and the conformational entropy of nanobodies targeting the SARS-CoV-2 spike protein. *Proc Nat. Acad. Sci. USA*, 119 (31):e2205412119
8. Brotzakis Z.F., Lisdstedt P., Taylor R., Rinauro D. J., Gallagher N., Bernardes G. & Vendruscolo M. (2021) A structural ensemble of a tau-microtubule complex reveals regulatory tau phosphorylation and acetylation mechanisms. *ACS Central Sci*. 7 (12) 1986-1995
9. Brotzakis Z. F., Zhang S & Vendruscolo M. (2023) AlphaFold Prediction of Structural Ensembles of Disordered Proteins doi**:** 10.1101/2023.01.19.524720
10. Horne, R., Possenti, A., Chia, S., Brotzakis Z.F., Staats, R.,Nowinska, M. & Vendruscolo, M (2021).A machine learning approach to identify small molecule inhibitors of senconary nucleation in α-synuclein Aggregation. *bioRxiv* (https://doi.org/10.1101/2021.11.10.468009) (in review stage in Nature Chemical Biology)