Thursday, March 22, Session II

L7

A METHOD FOR ACCELERATED FREE ENERGY CALCULATIONS OF PROTEINS IN AN EXTENDED EXPERIMENTAL ENSEMBLE DERIVED FROM THE PROTEIN DATA BANK

Emanuel K. Peter, Jiří Černý

Institute of Biotechnology, CAS, BIOCEV, Průmyslová 595, 252 50 Vestec, Prague West, Czech Republic

In this talk, we present a method to simulate protein systems within an experimental ensemble derived from the Protein Data Bank (PDB). After the collection of data over a non-redundant set of over 24,000 high resolution protein X-ray structures and the analysis of the radial distribution functions (RDF) g(r), we used the corresponding potential of mean force (PMF) w(r) to accelerate MD simulations of proteins, while the underlying forcefield is corrected at the same time. After a structural analysis of the PMF-data, where we identified collective properties of groups of aliphatic, hydrophilic and aromatic aminoacids, we validated our method in simulations of di-peptides based on dialanine and compared the results with path-sampling simulations. We found a dependency from the position of the alternating residue and the chemical configuration of the sidechain of the aminoacid next to Ala. A comparison of PMF-based simulations using different AMBER-forcefields leads to approximately identical free energy partitions independent from the choice of the forcefield. We continued with simulations of mutated penta-alanines with 5 different point-mutations, where we observed that a N-terminal mutation had the largest effect on the free energy landscape of the peptide. In an application of our approach, we applied our methodology in folding-simulations of TrpCage and observed that the PMF-based sampling significantly improves the description of structural conformers along the folding pathway of this peptide. Finally, we give a perspective on the application of g(r)-data from the PDB for the accelerated simulation of DNA-conformers and protein-DNA complexes.

L8

3DPATCH: FAST SEQUENCE AND STRUCTURE RESIDUE-LEVEL INFORMATION CONTENT ANNOTATION IN A WEB BROWSER

David Jakubec[1,2,*], Jiří Vondrášek[1], Robert D. Finn[3]

[1]*Department of Bioinformatics, Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, 166 10 Prague 6, Czech Republic*
[2]*Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, 128 43 Prague 2, Czech Republic*
[3]*European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK*
**david.jakubec@uochb.cas.cz*

Amino acid residues manifesting high levels of conservation are often indicative of functionally significant regions of protein structures. Residues critical for protein folding, hydrophobic core stabilization, intermolecular recognition, or enzymatic activity often manifest lower mutation rates compared to the rest of the protein. Quantitative assessment of residue conservation typically involves querying a sequence against a database, finding similar sequences, aligning them to bring equivalent positions into register, and applying an information theory-based measure to individual columns in the multiple sequence alignment. Understanding how the sequence conservation profile relates in 3D requires its projection onto a protein structure, which can be a time-consuming process.

We developed 3DPatch, a client-side web application that simplifies the task of calculating protein sequence information content, 3D structure identification, and conser-

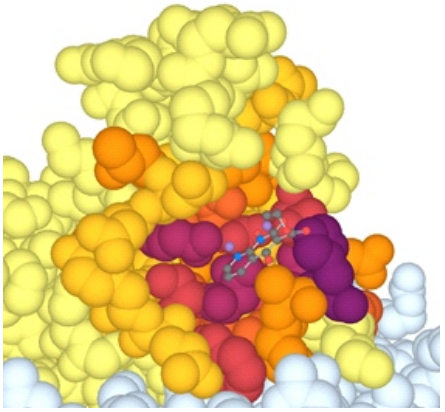

Figure 1. Detail of HIV-1 reverse transcriptase in complex with an inhibitor (PDB code 3lp0) marked-up with residue information content using 3DPatch. Higher conservation is indicated with darker colors. Catalytic/ligand binding pocket residues are clearly distinguished based on conservation level.

vation level-based mark-up (Figure 1). 3DPatch utilizes the power of profile hidden Markov models and speed of HMMER3.1 to provide accurate results in a matter of seconds. It was developed with easy integration into other peoples' websites in mind and supports most modern web browsers. 3DPatch is freely available at http://www.skylign.org/3DPatch/.

L9

SEARCH INTO UNEVOLVED PROTEIN SPACE

Tretyachenko V[1,2], Vymětal J[1,2], Bednárová L[2], Vondrášek J[2], Hlouchová K[1,2]

[1]Department of Biochemistry, Faculty of Science, Charles University, Hlavova 2030, 128 00 Prague 2, Czech Republic
[2]Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, Flemingovo náměstí 2, 166 10 Prague 6, Czech Republic

The protein sequences found in nature represent a tiny fraction of the potential sequences that could be constructed from the 20-amino-acid alphabet. To help define the properties that shaped proteins to stand out from the space of possible alternatives, we conducted a systematic computational and experimental exploration of random (unevolved) sequences in comparison with biological proteins. In our study, combinations of secondary structure, disorder, and aggregation predictions are accompanied by experimental characterization of selected proteins.

We found that the overall secondary structure and physicochemical properties of random and biological sequences are very similar. Moreover, random sequences can be well-tolerated by living cells. Contrary to early hypotheses about the toxicity of random and disordered proteins, we found that random sequences with high disorder have low aggregation propensity (unlike random sequences with high structural content) and were particularly well-tolerated. This direct structure content/aggregation propensity dependence differentiates random and biological proteins.

Our study indicates that while random sequences can be both structured and disordered, the properties of the latter make them better suited as progenitors (in both *in vivo* and *in vitro* settings) for further evolution of complex, soluble, three-dimensional scaffolds that can perform specific biochemical tasks.

L10

PROTEIN HOMODIMERIZATION FROM PERSPECTIVE OF STRUCTURAL BIOLOGY AND BIOPHYSICS

Hritz, J.

Central European Institute of Technology, Masaryk University, Brno, Czech Republic

Large number of proteins are forming homodimers. Here, we are presenting experimental and computational tools for the characterization of stable homodimers in terms of biophysical and structural properties of the dimer dissociation. First, fluorescence assays analyzed by ad-hoc mathematical models as presented here allow efficient and reliable determination of thermodynamic and kinetic parameters of dimer-monomer equilibria. To quantify microscopic dynamics between monomers and homodimers, we have designed sensitive fluorescent assays based on the Förster resonance energy transfer (FRET) and self-quenching (SQ) phenomena. The applicability of these approaches is shown here for the determination of dissociation constant (K_d) and dissociation and association rate constants (k_{off} and k_{on}, respectively) of 14-3-3 dimer-monomer equilibria. The most important biophysical factors altering this equilibrium are presented here.

Second, computational approach of Hamiltonian replica exchange molecular dynamics (H-REMD) combined with the distance restraints between the monomers is applied to reveal the structural details of the homodimer dissociation/association pathway. Its applicability is presented here for the case of the regulatory domain of human tyrosine hydroxylase 1 (RD-hTH1). Next, we analyzed the free energy profile and calculated the binding affinities, and compared the computational results with the experimental observations for three prepared human RD-hTH1 constructs.

This work was supported by the research grant from the Czech Science Foundation, grant no. GA. 15-34684L. The results of this research have been acquired within CEITEC 2020 (LQ1601) project.

1. Mei, G., Di Venere, A., Rosato, N., Finazzi-Agro, A., FEBS J., 2005, 272(1), 16-27.
2. Sluchanko N. N.; Gusev N. B., Journal of Alzheimer's Disease., 2011, 27(3), 467-476.
3. Liu D.; Bienkowska J.; Petosa C.; Collier R. J.; Fu H.; Liddington R., Nature, 1995, 376(6536), 191-194.
4. Yang X.; Lee WH.; Sobott F.; Papagrigoriou E.; Robinson C.V.; Grossmann J.G.; Sundström M.; Doyle D.A.; Elkins J.M., PNAS, 2006 103, 46, 17237-42.

5. Jandova, Z., Trosanova, Z., Weisova, V., Oostenbrink, C., Hritz*, J., BBA - Proteins and Proteomics, 2018, 1866, 442-450.

6. Nagy, G., Oostenbrink, C., Hritz*, J., PLoS ONE, 2017, 12(7), e0180633.

7. Hritz, J., Byeon, I-J., Krzysiak, T., Martinez, A., Sklenár, V., Gronenborn, A.M., Biophys. J. 2014, 107, 2185-2194

Friday, March 23, Session III

L10

TRANSFORMING BIOMOLECULAR NMR TO STAY AT THE FOREFRONT OF STRUCTURAL BIOLOGY

Konstantinos Tripsianes

CEITEC — Central European Institute of Technology, Masaryk University, Kamenice 5, Brno 62500, Czech Republic, kostas.tripsianes@ceitec.muni.cz

The automation of NMR structure determination remains a significant bottleneck towards increasing the throughput and accessibility of NMR as a structural biology tool to study proteins. The chief barrier currently is that obtaining NMR assignments at sufficient levels of completeness to accurately define the structures by conventional methods requires a significant amount of spectrometer time (several weeks), and effort by a trained expert (up to several months). We have recently addressed both bottlenecks by presenting a complete pipeline for NMR structure determination using a minimal set of NMR spectra. Key to our approach was the development of 4D-CHAINS algorithm that enables fully automated assignments of NMR chemical shifts, at high levels of completeness and with a minimum error rate, from only two complementary spectra. In combination with autoNOE-Rosetta, 4D-CHAINS provides a robust approach leveraging a highly automated process to obtain reliable structures in a matter of days. Besides illustrating the merits of our pipeline for timely NMR structural studies, novel concepts in automation will be discussed aiming to harness the powerful advantages of the next-generation NMR spectrometers with magnetic strengths of 1.2 GHz.

L11

CAPTURING DYNAMICALLY INTERACTING INHIBITOR BY PARAMAGNETIC NMR SPECTROSCOPY

Pavel Srb,[1] Michal Svoboda,[1] Ladislav Benda,[2] Martin Lepšík,[1] Ján Tarábek, [1] Václav Šícha,[3] Bohumír Grüner,[3] Klára Grantz-Šašková,[1] Jiří Brynda,[1,4] Pavlína Řezáčová,[1,4] Jan Konvalinka,[1] and Václav Veverka[1*]

[1]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czech Republic
[2]Institute des Sciences Analytiques, UMR 5280 CNRS / Université Claude Bernard Lyon 1 / ENS de Lyon, 5 rue de la Doua, 69100 Villeurbanne (Lyon), France
[3]Institute of Inorganic Chemistry of the Czech Academy of Sciences, Husinec-Řež u Prahy, Czech Republic
[4]Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic

Transient and fuzzy intermolecular interactions are fundamental to many biological processes. Despite their importance, they are notoriously uneasy to characterize. Paramagnetic NMR provides an opportunity to amplify rather small indices of intermolecular interactions often observed with diamagnetic ligands. Here, we present an intricate case of a partially flexible protein dynamically interacting with a ligand where data obtained by standard approaches fail to provide detailed structural interpretation. We demonstrate, that a combination of paramagnetic NMR experiments, advanced quantum chemical calculations and molecular dynamics simulations offer a route towards structural characterization of a class of inhibitors based on substituted metalacarboranes with HIV-1 protease.