Coarse graining of biomolecules and beyond

Saturday, 7 October 2017

Book of Abstracts

Coarse graining of biomolecules and beyond
Molecular modeling workshop
Saturday, 7 October 2017

Hosted by:

CeNT CENTRE OF NEW TECHNOLOGIES

Organizing committee

Janusz Bujnicki, International Institute of Molecular and Cell Biology in Warsaw

Adam Liwo, University of Gdańsk

Piotr Setny, University of Warsaw

Joanna Sułkowska, University of Warsaw

Joanna Trylska, University of Warsaw

Local organizers

Vasilina Zayats, University of Warsaw

Anita Dudek, University of Warsaw

Contents

CGW2017
Program 3
Lectures 5
Posters 17

Name: cgw2017
Password: 2trivialknots4link

CGW 2017 Program

Saturday, Oct 7

08:00 – 08:30 Registration & morning coffee

Session I

08:30 - 09:00 Piotr Setny (T)

Discrete solvent based method for the prediction of protein hydration sties

09:00 - 09:30 Garegin Papoian (T)

AWSEM-MD: Overview of the Force Field and Selected Applications

09:30 - 10:30 Filip Leonarski (HS)

RedMDStream: Simulation Toolbox To Automatically Parameterize Coarse-grained Force Fields

10:30 - 11:00 cofee break

Session II

11:00 - 11:30 Szymon Niewieczerzał (T)

New implicit environment model for the study of membrane proteins manipulations by coarse-grained molecular dynamics

11:30 - 12:00 Adolfo Poma (T)

GoMartini: Study of Large Conformational Transition in Proteins with the Martini Force-Field

12:00 - 13:00 Mateusz Kurcinski, Maciej Ciemny, Tymoteusz Oleniecki, Karolina Dawid, Sebastian Kmiecik (HS)

Modeling of protein flexibility and protein-peptide interactions using CABS-flex and CABS-dock standalone applications

13:00 - 14:00 lunch

Session III

14:00 - 14:30 Mateusz Sikora (T)

Membrane Glycoproteins in MARTINI Forcefield – a Sticky Problem

14:30 - 15:00 Rafał Jakubowski (T)

Unfolding of membrane proteins with complex topology

15:00 - 16:00 Cezary Czaplewski, Paweł Krupa, Adam Sieradzan (HS)

UNRES, NARES, and SUGRES Coarse-Grained Models - An Introduction

16:00 - 17:00 poster session & coffee

Session IV

17:00 - 17:30 Michał Boniecki (T)

SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction

17:30 - 18:00 Filip Stefaniak (T)

Modeling of ribonucleic acid-ligand interactions

18:00 - 19:00 Mateusz Dobrychłop (HS)

PyRy3D: a software tool for modelling of large macromolecular complexe

19:00 - evening get-together

Legend:
T – talk
HS - hands on session

Lectures

Boniecki Michał J. 6
Czaplewski Cezary, Krupa Paweł, Sieradzan Adam 7
Dobrychłop Mateusz 8
Jakubowski Rafał 9
Kurcinski Mateusz, Ciemny Maciej, Oleniecki Tymoteusz, Dawid Karolina, Kmiecik Sebastian 10
Leonarski Filip 11
Niewieczerzał Szymon 12
Papoian Garegin A. 13
Poma Adolfo 14
Setny Piotr 15
SikoraMateusz 16
Stefaniak Filip 17

SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction

Michał J. Boniecki* [1], Grzegorz Łach [1], Wayne K. Dawson [1], Konrad Tomala [1], Paweł Łukasz [1], Tomasz Sołtysiński [1], Kristian M. Rother [1] and Janusz M. Bujnicki [1,2]
[1] International Institute of Molecular and Cell Biology, ul. Trojdena 4, 02-109 Warsaw, Poland
[2] Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, ul. Umultowska 89, 61-614 Poznań, Poland

*presenting author e-mail: mboni@genesilico.pl

The molecules of the ribonucleic acid (RNA) perform a variety of vital roles in all living cells. Their biological function depends on their structure and dynamics, both of which are difficult to experimentally determine, but can be theoretically inferred based on the RNA sequence. We have developed a computational method for molecular simulations of RNA, named SimRNA [1].
SimRNA is based on a coarse-grained representation of a nucleotide chain, a statistically derived energy function, and Monte Carlo methods for sampling of the conformational space. The backbone of RNA chain is represented by P and C4' atoms, whereas nucleotide bases are represented by three atoms: N1-C2-C4 for pyrimidines and N9-C2-C6 for purines. In fact, those three atoms are used to calculate local coordinate system that allows for positioning of 3D grid - actual representation of the base. 3D grid contains information about interaction of the entire base moiety (not only 3 atoms) including excluded volume.
All terms of the energy function were derived from a manually curated database of crystal RNA structures, as a statistical potential. Sampling of the conformational space was accomplished by the use of the asymmetric Metropolis algorithm coupled with a dedicated set of moves. The algorithm was embedded in either a simulated annealing or replica exchange Monte Carlo method. Recent tests demonstrated that SimRNA is able to predict basic topologies of RNA molecules with sizes up to about 50 nucleotides, based on their sequences only, and larger molecules if supplied with appropriate distance restraints. The user can specify various types of restraints, including restraints on secondary structure, distance and position.
SimRNA can be used for systems composed of several chains of RNA. It is also able to fold/refine structures with irregular (non-helical) geometry of the backbone (RNA pseudo knots, coaxial stacking, bulges, etc.). SimRNA is a folding simulations method, thus it allows for examining folding pathways, getting an approximate view of the energy landscapes, and investigating of the thermodynamics of RNA systems.
SimRNA is also available as a server: SimRNAweb [2].
Based on similar ideas and SimRNA framework, we have also developed coarse-grained method for modeling of RNA-protein complexes (not published)

[1] Boniecki MJ, Lach G, Dawson WK, Tomala K, Lukasz P, Soltysinski T, Rother KM, Bujnicki JM. Nucleic Acids Res. 2016 Apr 20;44(7)
[2] Magnus M, Boniecki MJ, Dawson W, Bujnicki JM. Nucleic Acids Res. 2016 Apr 19

Introduction to UNRES and NARES-2P packages for coarse-grained simulations of proteins and nucleic acids

Cezary Czaplewski[1]*, Paweł Krupa[1],[2]*, Adam K. Sieradzan[1]*, Ewa Gołaś1, Yi He[3], Dawid Jagieła[1], Agnieszka Karczyńska[1], Jooyoung Lee[4], Agnieszka Lipska[1], Adam Liwo[1], Mariusz Makowski[1], Magdalena A. Mozolewska[1],[5], Andrei Niadzvedtski[1], Stanisław Ołdziej[6], Harold A. Scheraga[3], Rafał Ślusarz[1], Tomasz Wirecki[1], Yanping Yin[3], Bartłomiej Zaborowski[1]

[1]Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland.
[2]Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, PL-02668 Warsaw, Poland.
[3]Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, N.Y., 14853-1301, U.S.A.
[4]School of Computational Sciences, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul 130-722, Republic of Korea.
[5]Institute of Computer Science, Polish Academy of Sciences, ul. Jana Kazimierza 5, Warsaw 01-248, Poland.
[6]Laboratory of Biopolymer Structure, Intercollegiate Faculty of Biotechnology, University of
Gdańsk and Medical University of Gdańsk, Abrahama 58, 80-307 Gdańsk, Poland
*presenting authors

This hands-on session will introduce the users to running calculations with the UNRES and NARES-2P packages[1] (www.unres.pl) for coarse-grained simulations. A short introduction to installation and running UNRES and NARES-2P in serial and parallel mode will be provided, followed by examples of simple calculations (energy evaluation and minimization, canonical molecular dynamics (MD)) and use of replica-exchange (REMD)[2] and multiplexed replica exchange molecular dynamics (MREMD)[3] and analyzing the results by weighted-histogram analysis method (WHAM)[4] and cluster analysis. Unrestrained simulations and simulations with restraints from experimental data (such as, e.g., NMR and SAXS) and those derived from templates[5] will be presented, as well as, simulations with dynamic disulfides[6], and steered MD[7].

References:

[1]A. Liwo et al., J. Mol. Model. 20 (2014) 2306.
[2]U.H.E. Hansmann, Y. Okamoto, J. Comput. Chem. 14 (1993) 1333–1338.
[3]Y.M. Rhee, V.S. Pande, Biophys. J. 84 (2003) 775–86.
[4]S. Kumar et al., J. Comput. Chem. 13 (1992) 1011–1021.
[5]P. Krupa et al. A. Liwo, J. Chem. Inf. Model. 55 (2015) 1271–1281.
[6]M. Chinchio et al., J. Chem. Theory Comput. 3 (2007) 1236–48.
[7]A.K. Sieradzan, R. Jakubowski, J. Comput. Chem. 38 (2017) 553–562.

Acknowledgements:
This work was supported by grants DEC-2013/10/M/ST4/00640, DEC-2015/17/N/ST4/03935, DEC-2015/17/N/ST4/03937, and DEC-2015/17/D/ST4/00509 from the National Science Center of Poland. Computational resources were provided by (a) the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) at the University of Warsaw under grant number GA65-20, (b) the Informatics Center of the Metropolitan Academic Network (CI TASK) in Gdańsk, (c) the Polish Grid Infrastructure (PL-GRID), and (d) our 796-processor Beowulf cluster at the Faculty of Chemistry, University of Gdańsk.

PyRy3D: a software tool for modelling of large macromolecular complexes

Mateusz Dobrychłop [1]*, Joanna M. Kasprzak [1,2], Mateusz Koryciński [1], Wojciech Potrzebowski [2], Mateusz Susik [3], Laura Pogorzelska [3], Rafał Niemiec [4], Witold Rudnicki [3], and Janusz M. Bujnicki [1,2]

[1] Laboratory of Structural Bioinformatics, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, ul. Umultowska 89, 61-614 Poznan, Poland
[2] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, ul. Ks. Trojdena 4, 02-109 Warsaw, Poland
[3] Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, Pawinskiego 5a, 02-106 Warsaw, Poland
[4] Faculty of Applied IT, University of Information Technology and Management, Sucharskiego 2, 35-225 Rzeszow, Poland

One of the major challenges in structural biology is to determine the structures of macromolecular complexes and to understand their function and mechanism of action. To maximize completeness, accuracy and efficiency of structure determination for large macromolecular complexes, a hybrid computational approach is required that will be able to incorporate spatial information from a variety of experimental methods into modeling procedure.
We developed PyRy3D, a method for building and visualizing coarse-grained models of large macromolecular complexes. The components can be represented as rigid bodies (e.g. macromolecular structures determined by X-ray crystallography or NMR, theoretical models, or abstract shapes) or as flexible shapes (e.g. disordered regions or parts of protein or nucleic acid sequence with unknown structure). Spatial restraints are used to identify components interacting with each other, and to pack them tightly into contours of the entire complex (e.g. cryo-EM density maps or ab initio reconstructions from SAXS or SANS methods). Such an approach enables creation of low-resolution models even for very large macromolecular complexes with components of unknown 3D structure. Our model building procedure applies Monte Carlo approach to sample the space of solutions fulfilling experimental restraints. We compared PyRy3D's performance with state-of-the-art methods for fitting atomic structure assemblies into EM maps (such as Situs, MultiFit, gamma-TEMPy, ADP-EM, Powerfit and gEMfitter). Our tool generated the most accurate models for most of the benchmark's 26 targets, which included complexes with mass ranging from 40 kDa to 840 kDa.
PyRy3D can be used as a command-line tool and via a UCSF Chimera plugin, which provides a graphical user interface for the program. It can be dowloaded from genesilico.pl/pyry3d/download. The tool is also available via a web server at pyry3d.icm.edu.pl. The workshop will demonstrate the capabilities of all three interfaces.

Unfolding of membrane proteins with complex topology

Rafał Jakubowski [1]*, Szymon Niewieczerzał [1] and Joanna I. Sułkowska [1,2]
[1] Center of New Technologies, University of Warsaw, ul. Banacha 2c, Warsaw
[2] Faculty of Chemistry, University of Warsaw, ul. Pasteura 1, Warsaw

Coarse grained (CG) models based on the native structure have undoubtedly increased possibilities of investigating biosystems, especially in the field of protein unfolding. While keeping an accurate event scenario, they allow to describe systems of complexity which are far beyond the accessibility of the all atom (AA) approaches. The usability of CG approaches grows even more for membrane proteins, where computational demands tend to be extremely high due to calculations necessary for a huge numbers of explicit membrane and water molecules atoms, which may dominate overall computational cost in AA representation. Therefore, using one of CG approaches is often the most reasonable and sometimes the only way to investigate stretching of usually large membrane proteins.
Here we report the unfolding pathways of membrane protein with non trivial slipknot topology - BetP, consisting of three internal knots, which we determined for the first time computationally. Investigated protein is a member of betain/choline/carnitine transporter family and performs three functions: betaine transport, osmosensing, and osmoregulation. We use the structure based model [1] and a newly developed energy function as an implicit membrane, which implemented with Go-like approach offers significant decrease of computational costs while keeping a good accuracy. Validation of our methods against experimental data gave very good result.
Beside finding specific, previously described metastable conformations [2], we also found a completely new slipknot untying events, which shed a new light on possible scenarios of untying process. Our results may help in understanding of nanomechanical properties of non trivial topology membrane proteins, while applied set of methods is in general suitable for investigations of a variety of membrane protein systems.
This research was funded by the Ministry of Science and Higher Educations of Poland by Ideas Plus II grant. Computational facilities of Interdisciplinary Centre for Modern Technologies, Nicolaus Copernicus University in Torun, Poland, are acknowledged.

[1] Selection of optimal variants of Go-like models of proteins through studies of stretching
JI Sulkowska, M Cieplak Biophys. J. 2008 95, 3174
[2] Jamming proteins with slipknots and their free energy landscape, JI Sulkowska, P Sułkowski, JN Onuchic, Phys. Rev. Lett. 2009 103 268103.

Modeling of protein flexibility and protein-peptide interactions using CABS-flex and CABS-dock standalone applications

Mateusz Kurcinski [1]*, Maciej Ciemny [1, 2]*, Tymoteusz Oleniecki [1, 3, 4]*, Karolina Dawid [1]*, Sebastian Kmiecik [1]*

[1] Biological and Chemical Research Centre, Faculty of Chemistry, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland
[2] Faculty of Physics, University of Warsaw, ul. Pasteura 5, 02-093 Warsaw, Poland
[3] College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Banacha 2C, 02-097 Warsaw, Poland
[4] Mossakowski Medical Research Centre Polish Academy of Sciences, Pawińskiego 5, 02-106 Warsaw, Poland

During session, we will present a suite of newly developed standalone applications for efficient modeling of protein flexibility and protein-peptide interactions. These standalone applications are extended and customizable versions of our widely used web servers: CABS-flex [1, 2] and CABS-dock [3, 4]. Both applications seamlessly merge CABS coarse-grained protein model [5] with reconstruction to all-atom resolution and fully automated plot analysis. The session will be focused on practical demonstrations on how to use: (i) CABS-flex standalone application for fast simulations of protein structure fluctuations; (ii) CABS-dock standalone application for flexible protein-peptide docking; (iii) our PyMOL plugin for convenient visualization and analysis of CABS-flex/CABS-dock modeling results. Numerous examples will be presented, including: modeling intrinsically disordered proteins, using flexibility simulations in predictions of protein aggregation properties [6], and docking with large-scale conformational changes of protein receptors [4].

[1] Jamroz, M.; Kolinski, A.; and Kmiecik, S. Nucleic Acids Res, 2013, 41:W427-31
[2] Jamroz, M.; Kolinski, A; and Kmiecik, S. Bioinformatics, 2014, 30(15):2150-4
[3] Kurcinski, M.; et al. Nucleic Acids Res, 2015, 43:W419-24
[4] Ciemny, M.P.; et al., Sci Rep, 2016, 6:37532
[5] Kmiecik, S.; et al., Chem Rev, 2016, 116:7898–7936
[6] Zambrano, R.; et al. Nucleic Acids Res, 2015, 43:W306-313

RedMDStream: Simulation Toolbox To Automatically Parameterize Coarse-grained Force Fields

Filip Leonarski [1,2]* and Joanna Trylska [1]

[1] Centre of New Technologies, University of Warsaw, Warsaw, Poland
[2] Current address: Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland

Coarse-grained models to simulate the dynamics of biological molecules have the advantage of reaching orders of magnitude longer simulation time scales than all-atom approaches. Therefore, global internal dynamics of biomolecules can be determined with lower computational cost. However, the coarse-grained models for a molecule of interest are typically not ready to use and require parameterization of the potential energy function for a specific application. The reason is that the coarse-grained models, potential energy terms, and parameters are typically not transferable between different molecules and problems. So parameterizing coarse-grained force fields, a task that is both tedious and time-consuming, is often necessary. Therefore, we have designed a systematic and objective method to help develop or adapt the coarse-grained models.
We will present RedMDStream, a method and software for developing, testing, and simulating biomolecules with coarse-grained molecular dynamics models [1]. The application implements an automatic procedure for the optimization of potential energy parameters based on metaheuristic methods such as evolutionary algorithm or particle swarm optimization [2]. The approach enables automatic testing of thousands of force field parameters and selecting the optimal one according to user criteria. In addition to an optimized force field, parameter correlations and significance of the potential energy terms can be determined.
As transferability of various coarse-grained force fields is even more problematic for highly charged nucleic acids, we will focus on RNA molecules – we will describe the parameterization of a one-bead coarse-grained model and force field for molecular dynamics simulations [3] of an RNA hairpin (repression of heat shock gene expression element called microROSE) and to RNA structure prediction problem.

[1] Leonarski F.; Trylska J., RedMDStream: Parameterization and simulation toolbox for coarse-grained molecular dynamics models, Biophys. J., 2015, 108, 1843-1847
[2] Leonarski F.; Trovato F.; Tozzini V.; Leś A.; Trylska J., Evolutionary algorithm in the optimization of a coarse-grained force field, J. Chem. Theory Comput., 2013, 9, 4874-4889
[3] Leonarski F.; Trylska J., Modeling Nucleic Acids at the Residue-Level Resolution, In: Computational methods to study the structure and dynamics of biomolecules and biomolecular processes - from bioinformatics to molecular quantum mechanics, Springer, 2014, 109-149

New implicit environment model for the study of membrane proteins manipulations by coarse-grained molecular dynamics

Szymon Niewieczerzał [1,2]*, Rafał Jakubowski [1] and Joanna I Sułkowska [1,2]
[1] Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097 Warsaw, Poland
[2] Departament of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

Single-molecule manipulation techniques provide a key insight into the unfolding mechanisms of proteins, investigating the transition state, and mechanical properties of protein molecules. Experimental methods have been improved over the years, increasing the force sensitivity, but Molecular Dynamics forms an important part in gaining more insight into investigated processes. In our study, we present a new impicit environment model dedicated to study mechanical properties of membrane proteins. The model is based on IMM1 method [1]. Membrane environment and solvent are characterized with different solvation free energies for each amino acid. We simplified the original method, and in our approach we do not consider the influence of the neighboring amino acids. We employed structure based model of the protein, which was intesively used in the past, in studies regarding mechanical manipulations on biological molecules [2,3]. The solvation free energies assigned to amino acids were based on experimental measurements [4]. The results obtained with this simple model (F-d curves, the unfolding scenario) show very high consistency with experimental results and more detailed all-atom molecular dynamics simulations, for the model system of bacteriorhodopsin. The accuracy is obtained in spite of representing each amino acid with one effective bead.

[1] Lazaridis, T. (2003). Effective energy function for proteins in lipid membranes. Proteins:
Structure, Function, and Bioinformatics, 52(2):176–192.
[2] Sułkowska, J. I. and Cieplak, M. (2007). Mechanical stretching of proteins—a theoretical survey of the protein data bank. Journal of Physics: Condensed Matter, 19(28):283201.
[3] Sułkowska, J. I. and Cieplak, M. (2008). Selection of optimal variants of gō-like models
of proteins through studies of stretching. Biophysical journal, 95(7):3174–3191.
[4] White, S. H. and Wimley, W. C. (1998). Hydrophobic interactions of peptides with membrane interfaces. Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes, 1376(3):339–352.

AWSEM-MD: Overview of the Force Field and Selected Applications

Garegin A. Papoian [1]*

[1] University of Maryland, College Park, MD, USA

AWSEM, which stands for the Associative memory, Water mediated, Structure and Energy Model, is a coarse-grained protein simulation package that is being developed and maintained by the Papoian lab at the University of Maryland and the Wolynes lab at Rice University [1]. It arose from one of the earliest applications of neural network theories to structural modeling of proteins [2], however, evolving over time into a Hamiltonian containing both physical and knowledge-based potentials [1]. In this talk, I will first give a brief historical overview of the development of various potentials that enter the AWSEM Hamiltonian followed by recent applications. In particular, I will highlight AWSEM's applications in structure prediction and dynamical behaviors of various macromolecular complexes, such as histone dimers, tetramers, octamers, histone chaperones and whole nucleosomes. Also, I will discuss some recent extensions and variations of AWSEM and the associated biophysical problems that would benefit most from these approaches.

[1] A. Davtyan, N. P. Schafer, W. Zheng, C. Clementi, P. G. Wolynes, and G. A. Papoian, "AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing", J. Phys. Chem. B, 116: 8494-8503, 2012.
[2] M. S. Friedrichs and P. G. Wolynes. Toward Protein Tertiary Structure Recognition by Means of Associative Memory Hamiltonians. Science, 246: 371–373, 1989.

GoMartini: Study of Large Conformational Transition in Proteins with the Martini Force-Field

Adolfo Poma*, Marek Cieplak and Panagiotis Theodorakis [1]

[1] Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland

The application of Coarse-Grained (CG) models in biology is essential to access large length and time scales needed for the description of several biological processes (e.g. self-assembly of beta-peptides into disease/functional amyloid fibrils, dissociation of protein-protein complexes, large movement of proteins under high mechanical stress, etc). The ELNEDIN [1] protein model is based on the well-known MARTINI CG force-field and incorporates additionally "harmonic bonds" of a certain spring constant within a defined cutoff distance between pairs of amino acid residues, in order to retain the native structure of the protein. In this case, the use of unbreakable harmonic bonds hinders the study of unfolding and folding processes. To overcome this barrier we have replaced the harmonic bonds with Lennard–Jones interactions based on the contact map of the native protein structure as is done in Go-like models. Our model [2] exhibits very good agreement with all-atom and the ELNEDIN simulations. Furthermore, our model is based on the van der Waals radii, instead of a cutoff distance, which results in a smaller number of interactions compared to ELNEDIN model. In conclusion, we anticipate that our model will provide further possibilities for studying biological systems beyond the scope of the ELNEDIN protein model.

[1] Periole, X.; Cavalli, M.; Marrink, S. J.; Ceruso, M. A. J. Chem. Theory Comput. 2009,
5, 2531–2543.
[2] Poma, A. B.; Cieplak, M.; Theodorakis,P. E. J. Chem. Theory Comput. 2017, 13, 1366.

Discrete solvent based method for the prediction of protein hydration sties.

Piotr Setny[1]*

[1] Centre of New Technologies, University of Warsaw, Poland

Proteins typically function in aqueous environment. Water not only embeds protein structures, but also penetrates them, filling small internal cavities or pockets formed during protein association with other binding partners. Isolated water molecules inhabiting such buried hydration sites typically maintain long-lived hydrogen bonds with their (macro)molecular environment and thus significantly contribute to protein stability, as well as their binding specificity and energetics. Accordingly, the knowledge of their placement is critical for any quantitative reasoning based on molecular structure of protein system under study.
Unfortunately, X-ray crystallography - the main source of experimental knowledge regarding water placement in proteins - provides often ambiguous data in this respect. Moreover, typical theoretical approaches are computationally expensive (molecular dynamics simulations), or too simplistic to capture complex effects such as the presence of several hydrogen-bonded water molecules in a single cavity, the occurrence of confined, yet disordered water, or water interaction with chemical groups that they were not explicitly parametrized for.
We will present a novel method for modeling of water in protein environment [1]. The method is based on discrete solvent representation and mean field description of its interactions. The method is applicable to macromolecules described by standard all-atom force fields. It indicates positions of buried hydration sites (including those filled by more than one water molecules) and differentiates them from sterically accessible to water but void regions with accuracy of 85%. Furthermore, it provides estimates of binding free energy for buried water molecules with good agreement with computationally more expensive double decoupling method (DDM).

[1] Piotr Setny JCTC 2015, 35, 5961

Membrane Glycoproteins in MARTINI Forcefield – a Sticky Problem.

Philipp S. Schmalhorst[1], Felix Deluweit[2], Roger Scherrers[2], Carl-Philipp Heisenberg[1], Mateusz Sikora[1]*

[1] Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria
[2] Wyatt Technology Europe, Hochstrasse 18, 56307 Dernbach, Germany

A majority of eukaryotic plasma membrane proteins is post-traslationally modified by comlplex polysaccharide chains in the process known as glycosylation, yet the precise function of these sugar moieties remains unknown. Loss of function studies in cell culture and model organisms suggest that glycans are not required for function or survival of individual cells, but essential for cell-cell interactions during tissue morphogenesis in early development. This is consistent with structure models of glycoproteins which often show that a large part of the “molecular surface” of glycoproteins is determined by their glycan moieties, suggesting a potential influence on protein-protein interactions. Indeed. glycosylation was found to modulate clustering of adhesion molecules – cadherins[1], yet molecular mechanism explaining this behaviour remains to be elucidated.
Here we explore the capabilities of MARTINI molecular dynamics coarse-grained force-field to study aggregation of glycosylated proteins. We demonstrate that the force-field predicts unphysical aggregation of saccharides, rendering simulations of glycoproteins impossible[2]. To understand the problem, we use the second virial coefficient of the osmotic pressure (B22) as an experimentally and computationally accessible measure of solute aggregation propensity. We perform light scattering experiments and provide MARTINI parameters for a large, biologically relevant A2 glycan. We further use Hamiltonian Exchange Umbrella Sampling technique to obtain potentials of mean force and finally B22 values for the glycan and four other small-to-medium sized saccharides and survey the agreement with experimental values, finding simulated values to be orders of magnitude too low. We address this issue by introducing a scaling parameter for inter-saccharide Lennard-Jones interactions, which makes saccharides' aggregative behavior more realistic and allows for simulations of glycoproteins.

[1] Langer, M.D et al (2012) N-glycosylation alters cadherin-mediated intercellular binding kinetics Journal of Cell Science, 125(10):2478-2485.
[2] Stark, A.C et al (2013) Toward Optimized Potential Functions for Protein–Protein Interactions in Aqueous Solutions: Osmotic Second Virial Coefficient Calculations Using the MARTINI Coarse-Grained Force Field J. Chem. Theory Comput., 9 (9):4176–4185

Modeling of ribonucleic acid-ligand interactions

Filip Stefaniak [1]*, Janusz M. Bujnicki [1,2]

[1] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, ul. Ks. Trojdena 4, 02-109 Warsaw, Poland.
[2] Laboratory of Bioinformatics, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, ul. Umultowska 89, 61-614 Poznan, Poland.
e-mail: fstefaniak@genesilico.pl

Computational methods play a pivotal role in the early stages of small molecule drug discovery, and are widely applied in virtual screening, structure optimization, and compound activity profiling. Over the last decades in medicinal chemistry, almost all the attention has been directed to protein–ligand binding and computational tools were created with such targets in mind. However, with growing discoveries of functional RNAs and their possible applications, RNA macromolecules have gained considerable attention as possible drug targets. This flow of discovery was followed by adapting existing computational tools for RNA applications, as well as active development of new RNA-tailored methods. However, due to the different nature of RNA, especially its tendency to use morphological plasticity (conformational change in ligand binding), the modeling of RNA still remains a challenging task [1].

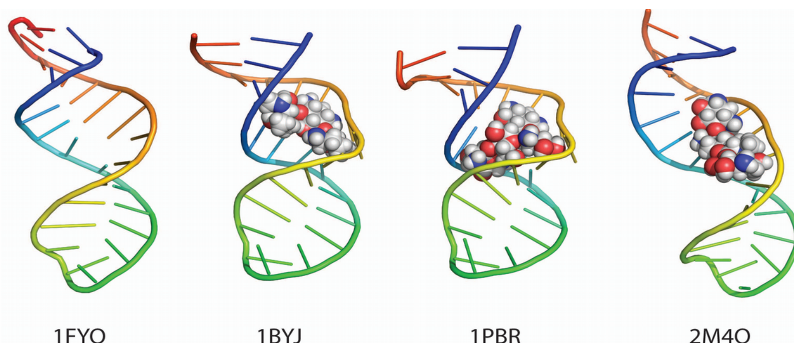


Figure . RNA flexibility in response to ligand-binding: NMR structures of decoding region A-Site.

The evolution of ‘protein-based’ drug discovery, and related computational methods, offers some clues on possible future directions and developments in modeling RNA interactions with small molecule ligands. I will present a new computational tool for predicting RNA-ligand interactions, which uses a coarse grained representation of both interacting partners. I will also present the plans for the future development of a predicting methods which takes into account the full flexibility of the RNA and ligand.

[1] Stefaniak F, Chudyk E, Bodkin M, Dawson WK, Bujnicki JM, Wiley Interdiscip Rev Comput Mol Sci 2015 Sep 14, doi: 10.1002/wcms.1226

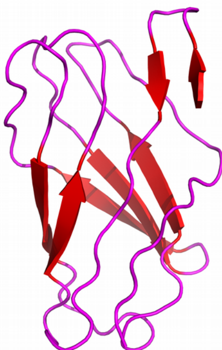
Posters

Bojarski Krzysztof K. 19
Ciemny Maciej P. 20
Cieplak Marek 21
Co Nguyen Truong 22
Dąbrowski-Tumański Paweł 23
Dawid Karolina 24
Dudek Anita 25
Faisca Patricia F.N 26
Ghosh Pritha 27
Gierut Aleksandra 28
Golon Łukasz 29
Grudinin Sergei 30
Jain Dham S. 31
Jarmolińska Aleksandra 32
Karczyńska Agnieszka 33
Kochańczyk Marek 34
Lipska Agnieszka 35
Liwo Józef, Czaplewski Cezary, Lubecka Emilia 36
Lubecka Emilia 37
Ludwiczak Jan, Dunin-Horkawicz Stanisław 38
Magnus Marcin 39
Marek Paulina 40
Mioduszewski Łukasz 41
Mozolewska Magdalena 42
Niemyska Wanda 43
Oleniecki Tymoteusz 44
Perlińska Agata 45
Plewczyński Dariusz 46
Ponczek Michał 47
Samsonov Sergey 48
Sikorska Celina 49
Suan Li Mai 50
Theodorakis Panagiotis 51
Wirecki Tomasz 52
Wiśniewska Marta D. 53
Zayats Vasilina 54

Effect of restraint type and strength on the quality of protein models obtained with homology-restrained UNRES simulations

Krzysztof K. Bojarski[1], Agnieszka Karczyńska[1],[2], Cezary Czaplewski[1], Adam Liwo[1],[2]
[1]Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 81-308 Gdańsk, Poland
[2]Center for In Silico Protein Structure and School of Computational Sciences, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul 130-722, Republic of Korea

Recently a new method was developed for protein-structure prediction, in which multiplexed replica-exchange simulations with the physics-based UNRES force field [1], with geometry restraints derived from homology models and target-specific knowledge-based Dynamic Fragment Assembly (DFA) pseudopotentials are carried out. The restraints are derived from common fragments extracted from server models of a given target protein. We applied this approach in the CASP12 experiment, using the models from BAKER-ROSETTASERVER, GOAL, QUARK and Zhang-Server.



Protein-peptide docking with large-scale conformational flexibility using the new CABS-dock standalone application

Maciej Paweł Ciemny [1,2], Tymoteusz Oleniecki [1,3,4], Mateusz Kurciński [1], Maciej Błaszczyk [1], Paulina H. Marek [1,5], Andrzej Koliński [1], Sebastian Kmiecik [1]

[1] Biological and Chemical Research Centre, Faculty of Chemistry, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland
[2] Faculty of Physics, University of Warsaw, ul. Pasteura 5, 02-093 Warszawa, Poland
[3] College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Banacha 2C, 02-097 Warszawa, Poland
[4] Mossakowski Medical Research Centre Polish Academy of Sciences, Pawińskiego 5, 02-106 Warszawa, Poland
[5] Faculty of Chemistry, Warsaw University of Technology, Noakowskiego 3, 00-664 Warsaw, Poland

Protein-peptide interactions may involve large-scale conformational changes of a target protein which are challenging to study both experimentally or computationally[1]. Here we present a new standalone application based on CABS coarse-grained protein model for flexible protein-peptide docking – the CABS-dock[2-4], so far available as a web server[4]. The method performs a blind global search for a binding site combined with an on-the-fly holding of a fully flexible peptide, while the target protein backbone fluctuates around its input conformation (in the default mode). Additionally, users can extend the degree of conformational flexibility of a protein receptor (for chosen regions) and enable large-scale conformational changes. That was the case of the modeling of the MDM2/p53 complex, modeled using CABS-dock with full flexibility of the intrinsically disordered regions of significant length[1]. The obtained CABS-dock results for MDM2/p53 system matched well the experimental data and provided new insights into the possible role of unstructured receptor regions. The standalone CABS-dock application allows for customization of the simulation parameters, providing constraints for user selected protein-peptide contacts, handing large-sized systems and provides a flexible framework for result analysis. CABS-dock is available as a standalone application at http://biocomp.chem.uw.edu.pl/CABSdockApp/ and as a web server at: http://biocomp.chem.uw.edu.pl/CABSdock.

References:
1.Ciemny, M. P.; Debinski, A.; Paczkowska, M.; Kolinski, A.; Kurcinski, M.; Kmiecik, S. Sci Rep 2016, 6, 37532.
2.Ciemny, M.; Kurcinski, M.; Kozak, K.; Koliński, A.; Kmiecik, S. In Methods in Molecular Biology; 2017; pp 69-94.
3.Blaszczyk, M.; Kurcinski, M.; Kouza, M.; Wieteska, L.; Debinski, A.; Kolinski, A.; Kmiecik, S. Methods 2016, 93, 72-83.
4.Kurcinski, M.; Jamroz, M.; Blaszczyk, M.; Kolinski, A.; Kmiecik, S. Nucleic Acids Res 2015, 43, W419-24.

Criteria for Folding in Structure-based Models of Proteins

Marek Cieplak

Institute of Physics, PAS, 02-668 Warsaw, Al. Lotników 32/46, Poland

In structure-based models of proteins, one often assumes that folding is accomplished when all contacts are established. This assumption may frequently lead to a conceptual problem that folding takes place in a temperature region of very low thermodynamic stability, especially when the contact map used is too sparse. We consider six different structure-based models [1] and show that allowing for a small, but model-dependent, percentage of the native contacts not being established boosts the folding temperature substantially while affecting the time scales of folding only in a minor way. We also compare other properties of the six models. We show that the choice of the description of the backbone stiffness has a substantial effect on the values of characteristic temperatures that relate both to equilibrium and kinetic properties. Models without any backbone stiffness (like the self-organized polymer) are found to perform similar to those with the stiffness, in! cluding in the studies of stretching. Finally, we present results of some recent applications of the structure-based models [2], including self-assembly of virus capsids [3].

[1] Wołek, K.; Cieplak, M. J. Chem. Phys. 2016, 144, 185102.
[2] Zhao, Y.; Cieplak, M. unpublished.
[3] Wołek, K; Cieplak, M. unpublished.

Dual effect of crowders on fibrillation kinetics of polypeptide chains revealed by lattice models

Nguyen Truong Co [1], Mai Xuan Ly [1]

[1] Institute of Physics Polish Academia of Science, Address al. Lotników 32/46 PL-02-668 Warsaw, POLAND

Neurodegenerative pathologies such as Huntington’s Alzheimer’s, Parkinson’s, and prion diseases are associated with formation of amyloid oligomers and fibrils that have crossβ-sheet structure. Understanding mechanisms governing fibrillation kinetics of peptides and proteins plays a key role in finding out the way for their effective treatment. So far most of investigation has been focused on exploring kinetics of oligomerization in ideal homogeneous milieu. However, all living processes take place in crowded environment which comprises DNA, protein, lipid, and sugar occupying 20%–30% volume of the typical cell cytoplasm. Therefore this factor should be taken into account.
We have developed the lattice model for describing polypeptide chains in the presence of crowders. The influence of crowding confinement on the fibrillation kinetics of polypeptide chains is studied using this model. We observed the non-trivial behavior of the fibril formation time Tfib that it decreases with the concentration of crowders if crowder sizes are large enough, but the growth is observed for crowders of small sizes. This allows us to explain the recent experimental observation on the dual effect of crowding particles on fibril growth of proteins that for a fixed crowder concentration the fibrillation kinetics is fastest at intermediate values of total surface of crowders. It becomes slow at either small or large coverages of cosolutes. It is shown that due to competition between the energetics and entropic effects, the dependence of Tfib on the size of confined space is described by a parabolic function

[1] . M. S. Li, D. K. Klimov, J. E. Straub, and D. Thirumalai, J. Chem. Phys.129, 175101 (2008)
[2] 7. M. S. Li, N. T. Co, C. K. Hu, J. E. Straub, and D. Thirumalai, Phys. Rev. Lett.105, 218101 (2010)
[3] Nguyen Truong Co, Chin-Kun Hu, and Mai Suan Li, Dual effect of crowders on fibrillation kinetics of polypeptide chains revealed by lattice models, J. Chem. Phys. 138, 185101 (2013)

The topology of your proteins

Pawel Dabrowski-Tumanski [1,2], Aleksandra Jarmolinska, [2,3], Wanda Niemyska [3,4] and Joanna Sulkowska [1,2]

[1] Faculty of chemistry, University of Warsaw, Pasteura 1, 02-093, Warsaw, Poland
[2] Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097, Warsaw, Poland
[3] Inter-Faculty Interdisciplinary Doctoral Studies in Natural Sciences and Mathematics, University of Warsaw, Banacha 2c, 02-097, Warsaw, Poland
[4] Faculty of Mathematics, University of Warsaw, Banacha 2, 02-097, Warsaw, Poland

Complex topology influences the various biophysical properties of polymers, such as persistence length, mean local fluctuations, response to stretching and others, which can be measured and interpreted in the coarse grained simulations. Moreover, the complex topology is decisive in correct description of e.g. protein-ligand interaction [1]. Therefore, awareness of the topology of the system is crucial for the understanding of its behaviour. In biology, even up to 6% of proteins contain some topological nontriviality, such as knot [1], link [2] or lasso [3]. Statistically, this means, that at least 2 participants of the Workshop deal with natively topologically complex structure. This number may be actually higher, as even the topologically trivial biopolymers may become entangled during the simulation. These (possibly unphysical) structures are however usually not easy to identify without meticulous analysis. Therefore, external tools analysing the topology of the structures are indispensable both for those who want to explain their results based on the topological analysis, as well as for those who want to avoid artificial entanglement in simulations.
In this talk, we discuss various entanglements found in proteins (knots, links and lassos) and present the servers designed to identify them: KnotProt, LinkProt and LassoProt [4-6]. We show how the servers may be utilized to analyse both single frames and whole trajectories, in the case of proteins and also any other (bio)polymer. We give examples of how topological analysis revealed the unphysical behaviour of the system, new reaction coordinates based on topology, and possibilities in design new materials by changing topology upon appropriate mutations. Finally, we present the databases conjugated with the servers collecting extensive information on the topology of each PDB entry and discuss their usability to distinguish incorrect PDB structures.

[1] Christian, T.; Sakaguchi, R.; Perlinska, A. P.; Lahoud, G., Ito, T.; Taylor, E. A.; Yokoyama S.; Sulkowska, J. I.; Hou, Y. M. Nat. Struc. Biol. 2016, 23, 941-948
[2] Dabrowski-Tumanski, P.; Sulkowska, J. I.; PNAS 2017, 114, 3415-3420
[3] Niemyska, W.; Dabrowski-Tumanski, P.; Kadlof, M.; Haglund, E.; Sułkowski, P.; Sulkowska, J. I. Sci. Rep. 2016, 6, 36895
[4] Jamroz, M.; Niemyska, W.; Rawdon, E. J.; Stasiak, A.; Millett, K. C.; Sułkowski, P.; Sulkowska, J. I. NAR 2014, 43, D306-D314
[5] Dabrowski-Tumanski, P.; Niemyska, W.; Pasznik, P.; Sulkowska, J. I. NAR 2016, 44, W383-W389
[6] Dabrowski-Tumanski, P.; Jarmolinska, A. I.; Niemyska, W.; Rawdon, E. J.; Millett, K. C.; Sulkowska, J. I. NAR 2017, 45, D243-D249.

Structural Visualization And Analysis Of Protein-Peptide And Protein-Protein Complexes Using PyMOL Plugin

Karolina Dawid [2,1], Maciej Ciemny [3,1], Tymoteusz Oleniecki [4,1,5], Mateusz Kurciński [1], Maciej Błaszczyk [1], Andrzej Koliński [1] and Sebastian Kmiecik [1]

[1] Biological and Chemical Research Centre, Faculty of Chemistry, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland
[2] Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland
[3] Faculty of Physics, University of Warsaw, ul. Pasteura 5, 02-093 Warszawa, Poland
[4] College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Banacha 2C, 02-097 Warszawa, Poland
[5] Mossakowski Medical Research Centre Polish Academy of Sciences, Pawińskiego 5, 02-106 Warszawa, Poland

A clear molecular visualization of three-dimensional structures of protein-peptide and protein complexes is a necessity in their studies. Such visualizations can be performed using feature-rich PyMOL molecular graphics software; unfortunately, using PyMOL may be troublesome for many scientists inexperienced in programming. Our goal was to create a PyMOL plugin which provides users, especially beginners and those without any programming background, with simple analysis and visualization tools. Our plugin is adapted to work with PDB files and provides a dedicated support for results produced with our protein-peptide docking method, the CABS-dock [1, 2, 3]. It comes with a set of default visualization options to choose from, as well as easy-to-use analytical tools for different tasks, such as characterization of the binding interface or measurement of the similarity to the known reference structure. The plugin may be used in the graphical interface mode for regular use, or as a command-line tool.

[1] Kurcinski, M.; Jamroz, M.; Blaszczyk, M.; Kolinski, A.; Kmiecik, S. Nucleic Acids Res. 2015, 43, 419-424
[2] Blaszczyk, M. et al. Methods. 2016, 93, 72-83
[3] Ciemny, M. P.; Kurcinski, M.; Kozak, K. J.; Kolinski, A.; Kmiecik, S. Methods Mol Biol. 2017, 1561, 69-94

Investigating the structural aspects of influenza virus fusion peptide and its interactions with lipid membrane

Anita Dudek[1], Piotr Setny[1]

[1]Centre of New Technologies, University of Warsaw, Poland

Worldwide, influenza causes of about 500,000 human deaths per year. The viruses are still evolving by mutations that are contributory to the production of new antigens on their surfaces. Fusion of host cell membrane and viral envelope is essential for influenza virus infection and leads to release of viral RNA inside host cell. The knowledge of how the juxtaposition of membranes leads to fusion is rather sparse.

One of the proteins involved in recognition of target cells and fusion process is a glycoprotein called hemagglutinin (HA), which is found on the surface of influenza virus and enabling influenza virus to initiate membrane fusion. The particularly important N-terminal fragment is able to form tight helical hairpin, creating molecular “grappling hook” structure inside host cell membrane. Isolated N-terminal fragment - consisting of only 20 amino acids with relatively hydrophobic sequence - has itself fusiogenic properties and is called a fusion peptide (HAfp). Additional three C-terminal amino acids: W21-Y22-G23 in HAfp1-23 are highly conserved and increase fusion efficiency compared to HAfp1-20. HAfp harpin structure is stabilized also by positively charged N+ group on N-terminus. The actual atomistic mechanism of action of the three C-terminal amino acids and the N-terminal positive charge remain unknown.

In this work in collaboration with experimental group we investigated the properties of different HAfp structures: HAfp1-23 and HAfp1-20 both with N-terminal charge modifications introduced by us. Molecular dynamics (MD) simulations complemented with Replica Exchange of HAfp peptides were performed and compared with experimental data. We found that N-terminal positively charged group N+ together with three terminal amino acids promote deep, perpendicular configuration across membrane. Based on the analysis of lipid disorder we identified this configuration as most fusiogenic. In contrast HAfp1-20 peptide has a tendency to form a boomerang structure, which remains at the membrane surface and introduces smaller perturbation of the surrounding lipids.

Effects of steric confinement in the folding of knotted proteins

Patrícia F.N. Faísca [1], Miguel A. Soler [2] and Antonio Rey [3]

[1] Departamento de Física and BioISI – Biosystems and Integrative Sciences Institute, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, Ed. C8, Portugal
[2] Scuola Internazionale Superiore di Studi Avanzati (SISSA) , via Bonomea, 265 - 34136 Trieste ITALY
[3] Departamento de Química Física I, Facultad de Ciencias Químicas, Universidad Complutense, 28040 Madrid, Spain

The chaperonin complex GroEL–GroES is able to accelerate the folding process of knotted proteins considerably. However, the folding mechanism inside the chaperonin cage is elusive. In [1] we used a combination of lattice and off-lattice Monte Carlo simulations of simple Gō models to study the effect of physical confinement and local flexibility on the folding process of protein model systems embedding a trefoil knot in their native structure. This study predicts that steric confinement plays a specific role in the folding of knotted proteins by increasing the knotting probability for very high degrees of confinement. This effect is observed for protein MJ0366 even above the melting temperature for confinement sizes compatible with the size of the GroEL/GroES chaperonin cage. An enhanced local flexibility produces the same qualitative effects on the folding process. In particular, we observe that knotting probability increases up to 40% in the transition state of protein MJ0366 when flexibility is enhanced. This is underlined by a structural change in the transition state, which becomes devoid of helical content. No relation between the knotting mechanism and flexibility was found in the context of the off-lattice model adopted in this work.

[1] Soler, M.A, Rey, A, Faísca, P.F. N., Phys. Chem. Chem. Phys. 2016,18, 26391-26403

Modeling of dynamic interactions between RNA and small molecules and its practical applications

Pritha Ghosh [1], Chinju John [1], Ewa Skowronek [1], Yuliia Varenyk [1], Elżbieta Purta [1], Filip Stefaniak [1] and Janusz M. Bujnicki [1,2]

[1] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Warsaw, Poland

[2] Laboratory of Bioiformatics, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland

RNA molecules play pivotal roles in living organisms and are involved in a variety of biological processes. The experimental determination of high-resolution RNA structures, which is necessary for understanding their molecular function, is however difficult and expensive. Therefore, computational methods have been developed for predicting RNA 3D structures from the ribonucleotide sequence information. Structures and functions of RNAs are very often modulated by small molecules, like naturally occurring molecules, and small molecule drugs. Hence, the analysis of RNA-ligand interactions is important in biomedical research. Unfortunately, at present, it is almost impossible to computationally predict structures of RNA-ligand complexes that involve large conformational changes of the RNA upon ligand binding unless very similar structures are already known. This situation hampers equally basic studies of RNA sequence-structure-function relationships and applied research on the development of small molecule regulators of biomedically important RNAs. We aim to develop and experimentally validate a general-purpose computational method for modeling and simulations of conformational changes in RNA 3D structures, in response to ligand binding, and demonstrate its various biological applications.

The "Development of new computational methods for modeling RNA interactions with small molecule ligands, and its application to study and regulate the mechanism of action of viral and bacterial RNA molecules" project is carried out within the TEAM programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund.

PyLasso - a PyMOL plugin to identify lassos

Aleksandra Gierut [1], Wanda Niemyska [1], Pawel Dabrowski-Tumanski [1, 2] Piotr Sulkowski [3] and Joanna I. Sulkowska [1, 2]

[1] Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097, Warsaw, Poland
[2] Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland
[3] Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland

Lassos form a new class of entangled motifs recently found in proteins. These are configurations with one or two termini of a protein chain pierced through a covalent loop, which is closed e.g. by a disulfide bridge. Lassos have been identified in around 4% of known protein structures and are most common in viruses, plants and fungi [1]. However lassos are very difficult to identify via visual inspections, what makes difficult their further studies and applications, e.g. designing molecular machine such as rotaxanes.
The PyLasso is a multiplatform, user-friendly plugin, which detects all types of lassos [2]. The detection involves analysis of surfaces of minimal area spanned on closed loops, and the number and directions of segments piercing such surfaces. The plugin enables to detect different types of covalently closed loops. A user can also define news loops, either by selecting a part of the backbone forming a loop, or typing sequential numbers of two atoms that form a bridge. Detected lasso configurations, including the minimal surface and piercings, are represented graphically in the PyMOL environment.
The PyLasso contains more advanced options, e.g. allows manual parametrization of the algorithm to detect lassos, as well as an option to change a parameter responsible for surface smoothing. Moreover, it is equipped with a new method to detect entanglement, which provides additional information about the lasso's geometry. The analysis of detected loops involves both single frames and whole trajectories.

[1] Niemyska, W.; Dabrowski-Tumanski, P.; Kadlof, M.; Haglund, E.; Sułkowski, P.; Sulkowska, J.I; Scientific Reports 2016, 6, 36895
[2] Gierut, A.; Niemyska, W.; Dabrowski-Tumanski, P.; Sułkowski, P.; Sulkowska. J.I; Bioinformatics 2017, in press

Extending the UNRES Force Field to Treat Peptide Groups in Cis Conformation

Łukasz Golon [1], Agnieszka G. Lipska [1], Adam K. Sieradzan [1] and Adam Liwo [1,2]

[1] Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland
[2] School of Computational Sciences, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul 130-722, Republic of Korea

In proteins ca. 99.7% of peptide groups are found in the trans conformation, which means that the ω dihedral angle (Cα-C(O)-N(H)-Cα) is close to 180º because in cis conformation (i.e. when ω ≈ 0º) there is steric hindrance[1]. When a proline residue is at the the C-terminal side of a peptide group, it is much more likely that this peptide group is in the cis conformation as the steric hindrance is present in both conformations. Although in ribosomes all peptide bonds are formed in the trans conformation, there are special enzymes, called prolyl-cis/trans-isomerases, which catalyze the conversion of proline peptide bonds to cis conformation. There are no known enzymes that would catalyze this conversion for non-proline peptide bonds[1]. However, non-proline peptide bonds in the cis conformation often occur close to the active sites which has lead to a hypothesis that they may serve as a reservoir for energy which can be released to drive an accompanying chemical reaction or conformational change[2]. Cis-trans isomerisation is a very slow process, therefore to study it theoretically coarse-graining is required.
UNRES (UNited RESidue) is a coarse-grained force field for protein simulation which can be used to predict their structure or identify functionally important motions[3]. It gave rise to an entire family of force-fields based on multipole-multipole interactions known as the Unified Coarse-Grained Model (UCGM)[4]. In UNRES each amino acid residue is modeled using two interaction sites – one representing the side chain and another one representing the peptide group.
So far all peptide groups were modeled with UNRES only in the trans conformation. The aim of this work is to extend the UNRES model to peptide groups in cis conformation. To this end, we performed energy scans in several virtual-bond valence and dihedral angles for model molecules containing the amino acid side chains and peptide groups in cis conformation, using the PM7 semiempirical quantum chemistry method. We used COSMO [5] to estimate the effects of solvation. The potential energy maps resulting from these calculations will be integrated to derive potentials of mean force, to which model potentials for use in UNRES will be fitted.
Acknowledgements: This work was supported by University of Gdańsk Young Researcher Grant nr 538-8376-B374-16 and 538-8376-B374-17. Calculations were carried out at the Academic Computer Centre in Gdańsk (CI TASK).

[1] Jabs, A.; Weiss, M. S.; Hilgenfeld, R. J. Mol. Biol. 1999, 286, 291-304
[2] Touw, T. G., Joosten, R. P., Vriend, G. Acta Cryst., 2015, D71, 1604-16014
[3] Liwo, A., Kaźmierkiewicz, R., Czaplewski, C. et al. J. Comput. Chem., 1998, 19, 259-276
[4] Liwo, A., Baranowski, M., Czaplewski, C. et al. J. Mol. Model., 2014, 20, 2306
[5] Klamt, A., Schüürmann, G. J. Chem. Soc., Perkin Trans. 2, 1993, 0, 799-805

On the Normal Modes, Small-Angle Scattering and Convex Optimization for Protein Structure Prediction

Sergei Grudinin [1], Alexandre Hoffmann [1], Emilie Neveu [1], Petr Popov [1], Maria Garkavenko [2]

[1] Nano-D team, CNRS / Inria, Grenoble, France
[2] MIPT Moscow, Dolgoprudny, Russia

Although the fundamental forces between atoms and molecules are almost fully understood at a theoretical level, and computer simulations have become an integral part of research activities, the application of these methods to large biomolecules still faces important practical difficulties due to the combinatorial explosion of possible interactions involved. Developing efficient protein structure prediction algorithms thus remains a major scientific challenge in computational biology. I will give a brief overview of several computational methods for protein structure prediction developed in our group at Inria Grenoble.
I will first present a new conceptually simple and computationally efficient method for nonlinear normal mode analysis called NOLB [1,2]. It relies on the rotations-translations of blocks (RTB) theoretical basis developed by Y.-H. Sanejouand and colleagues [3]. I will demonstrate the motions produced with the NOLB method on several molecular systems and show that some of the lowest frequency normal modes correspond to the biologically relevant motions. For example, NOLB detects the spiral sliding motion of the TALE protein, which is capable of rapid diffusion along its target DNA. Overall, the NOLB method produces better structures compared to the standard approach, especially at large deformation amplitudes and is scalable, such that can be applied to very large molecular systems, such as ribosomes.
I will also present one particular applications of the NOLB NMA method. More precisely, I will show a computational scheme that uses the NOLB modes as a low-dimensional representation of the protein motion subspace and optimises protein structures guided by the small-angle X-ray scattering (SAXS) profiles [4,5]. For example, in the very recent CASP12 blind structure prediction exercise this scheme obtained best models for 3 out of 9 SAXS-assisted targets.
Finally, I will present a machine-learning approach to train free-shape distance-dependent potentials of several types (protein-protein, protein-ligand, and protein folding potential). Unlike knowledge-based methods based on Boltzmann statistics, we do not impose any functional form of the potential. Instead, we use an optimization approach, accepting that the target binding energy value is decomposed into a polynomial basis with unknown expansion coefficients. These are then deduced from the structural data (such as PDB structures) using a convex formulation of the optimization problem [6-9].

[1] Alexandre Hoffmann & Sergei Grudinin. J Chem Theory Comput., 2017, 13 (5), pp.2123-2134.
[2] https://team.inria.fr/nano-d/software/nolb-normal-modes/
[3] Durand, P., Trinquier, G., & Sanejouand, Y.H., A new approach for determining low-frequency normal modes in macromolecules. Biopolymers vol.34, p759, 1994.
[4] Sergei Grudinin, Maria Garkavenko, Andrei Kazennov. Acta Crystallographica D, 2017, D73, pp.449 – 464.
[5] https://team.inria.fr/nano-d/software/pepsi-saxs/
[6] Popov, P.; Grudinin, S. J. Chem. Inf. Model., 2015, 55 (10), pp.2242-2255.
[7] Grudinin, S.; Popov, P.; Neveu, E.; Cheremovskiy, G. J. Chem. Inf. Model., 2015, 56 (6), pp.1053-1062.
[8] Grudinin, S.; Kadukova, M.; Eisenbarth, A.; Marillet, S.; Cazals, F. J. Comput. Aided Mol. Des., 2016, 30 (9), pp.791-804.
[9] Neveu, E.; Ritchie, D.W.; Popov, P.; Grudinin, S. Bioinformatics, 2016, 32 (7), pp.i693-i701.

Development of a new alphabet for RNA sequence and secondary structure representation and database searches

Dharm S. Jain [1,3], Tomasz Wirecki [1], Marcin Magnus [1], and Janusz Bujnicki [1, 2]

[1] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Warsaw, Poland
[2] Laboratory of Bioinformatics, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland
[3] Warsaw University of Technology, Warsaw, Poland

In bioinformatics, sequence alignments have always been a subject of interest. For protein sequences, the conservation of position of the residues often suggests retention of structural and physical features. BLAST type tools have been developed to generate and score the local alignments. Towards the similar effort for studies on RNA sequences, there have been attempts at creating BLAST tools like BLASTR and BEAGLE. This project will be an effort towards developing a strategy by incorporating primary and secondary sequence information to create a protein like representation of RNA sequences that can be used in BLAST tools for proteins. Along with the encoding of RNA sequences, new BLOSUM like substitution matrices have to be calculated. In particular, this approach will allow one sequence to represent both primary and secondary data and at the same time also enable it to use the existing tools like PSI-BLAST built to work on protein sequences.

GapRepairer – Repair Protein Structures and Their Topology

Aleksandra I Jarmolinska [1,2], Michal Kadlof [1,3], Paweł Dąbrowski-Tumański[1,4] and Joanna I Sulkowska [1,4]

[1] Centre of New Technologies, University of Warsaw
[2] College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw
[3] Faculty of Physics, University of Warsaw
[4] Faculty of Chemistry, University of Warsaw

Protein structure is fundamental for its function. Topology is an important part of this structure. And topology can only be reliably studied for an unbroken backbone. As such, properly filling in unresolved regions in a protein's structure (appearing in more than 25% of PDB deposits) is instrumental for its in silico studies, such as molecular dynamics. Non-trivial topologies that have been found in proteins include knots[1], slipknots, lassos[2], and links[3] (in a common, not mathematical, sense). Such folds, that appear in around 6% of known structures, can most easily be broken by a careless, "straight-line" repair. None of the currently popular modeling tools take into account the topology of the protein - thus potentially introducing erroneous folds that can be hard to see even for researchers used to knotted proteins. GapRepairer is a server that fills this gap in the spectrum of structure modeling methods. It redefines homology (as in "homology modeling") to include topology, and presents in-depth topological analysis of both templates and final models. Additionally, it provides a friendly and easy, although not basic, interface to the Modeller engine. GapRepairer server along with tutorials, usage notes, movies and the database of already repaired structures is available at http://gaprepairer.cent.uw.edu.pl

Protein-protein docking with the UNRES force field.

Agnieszka Karczyńska[1,2], Paweł Krupa[1,3], Adam Liwo[1,2], Cezary Czaplewski[1,2]

[1] Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland
[2] Center for In Silico Protein Structure, Korea Institute for Advanced Study, 87 Hoegiro, Dongdaemun-gu, 130-722 Seoul, Republic of Korea
[3] Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, PL-02668 Warsaw, Poland

Simulations of the structure of protein-protein complexes are very important in molecular biology, biophysics, and molecular medicine, including drug design. The structure of protein oligomers can be predicted by molecular dynamics and its extensions such as, e.g., multiplexed replica exchange molecular dynamics (MREMD) or by using other methods of conformational search, carried out in unrestrained mode or assisted by knowledge-based information. Simulation time strongly depends on oligomer size and simulations with all-atom force fields are very time- and resource-consuming. Therefore, coarse-grained simulations, which enable us to shorten the simulation time, thanks to a number of simplifications, are a good option. In this study, the results of the simulations of the structure of protein-protein complexes with the use of the coarse-grained UNRES force field [1,2] will be presented.
The first stage of our protocol involves rigid-body docking using the ZDOCK program, and then extracting the models with various protein-protein docking modes, based on the ligand root mean square deviation (lrmsd) value. In the next step relatively long MREMD simulations, started from all plausible models found by ZDOCK are performed, in which the geometry of oligomers is restrained. In the last step, the results of the MREMD simulations are processed with the weighted-histogram analysis method (WHAM) and cluster analysis [3] to extract the most probable clusters of oligomer structures. The clusters are scored by their relative free energies determined by WHAM, whose component are the UNRES energies of the conformations constituting the respective cluster.

References
[1] Liwo, A. et al. J. Mol. Model. 2014, 20, 2306.
[2] Liwo A et al. ed. G. Voth, Taylor & Francis, Chapter 8, pp. 107-122.
[3] Liwo A et al. J. Phys. Chem. B, 2007, 111, 260-285.
Acknowledgments
Supported by grant UMO-2013/10/M/ST4/00640 from the National Science Center of Poland.

Spatkin: a simulator for rule-based modeling of biomolecular site dynamics on surfaces

Marek Kochańczyk [1], William S. Hlavacek [2] and Tomasz Lipniacki [1]

[1] Institute of Fundamental Technological Research PAS, Warsaw, Poland
[2] Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA

Rule-based modeling is a powerful approach for studying biomolecular site dynamics. Here, we present Spatkin, a general-purpose simulator for rule-based modeling in two spatial dimensions [1]. The simulation algorithm is a lattice-based method that tracks Brownian motion of individual molecules and the stochastic firing of rule-defined reaction events. Because rules are used as event generators, the algorithm is network-free, meaning that it does not require to generate the complete reaction network implied by rules prior to simulation. In a simulation, each molecule (or complex of molecules) is taken to occupy a single lattice site that cannot be shared with another molecule (or complex). Spatkin is capable of simulating a wide array of membrane-associated processes, including adsorption, desorption, and crowding. Models are specified using an extension of the BioNetGen language, which allows to account for spatial features of the simulated process.
The source code and precompiled binaries are available at the Spatkin Web site: http://pmbm.ippt.pan.pl/software/spatkin.

[1] Kochańczyk, M.; Hlavacek, W.S.; Lipniacki, T. Bioinformatics 2017 (accepted) DOI: 10.1093/bioinformatics/btx456

Calculation of potential energy maps for the new analytical potentials in physics-based coarse-grained force field

Agnieszka G. Lipska [1], Adam K. Sieradzan [1] and Adam Liwo [1]

[1] Faculty of Chemistry, University of Gdansk, Wita Stwosza 63, 80-308 Gdańsk, Poland

Coarse-grained (CG) force fields are widely used in biomolecular simulations because they offer a tremendous reduction of the computational cost compared with the all-atom models. However, designing the functional forms of the CG energy terms poses a serious problem and these terms are very often imported from all-atom force fields or assigned on a heuristic basis. The CG energy functions originate from the potentials of mean force (PMF) where the degrees of freedom that are not considered explicitly are integrated out [1]. Expansion of the PMF into Kubo cluster-cumulant functions [2], termed factor expansion1 enables us to split the PMF into specific terms.
In this work, we used the factor approach to revise the local potentials in the UNRES force field for proteins. Unlike the former approach, in which the units to derive the local potentials were composed of whole terminally-blocked residues, we used three units: (i) terminally-blocked glycine residue to represent protein backbone; (ii) H-Cα(R)-CONHCH3 and (iii) CH3CO-NH-Cα-H(R), where R denotes a side chain, to represent a side chain and the adjacent backbone fragment. As opposed to the previous formulation, this dissection uses exactly the UNRES sites and, consequently, provides a clear separation of the backbone-only local potentials from those involving side chains. Consequently, the number of different potential types is significantly reduced. The respective energy surfaces were calculated by using the semi-empirical PM7 quantum mechanics method. The calculations were performed on a grid in the θ (Cα…Cα…Cα), or (Cα…Cα…Cβ), χ(n) (rotational of side-chain dihedral angles) and either the λ(1) or λ(2) that describe the rotation of peptide group in the Cα…Cα…Cα frame. Based on the calculated potential-energy surfaces, the respective PMF will be calculated.
Supported by grants DEC-2016/21/N/ST4/03154 and DEC-2012/06/A/ST4/00376 from the National Science Center of Poland, Mistrz 7./2013 from the Foundation for Polish Science and BMN 538-8376-B391-16. Calculations were carried out at the Academic Computer Centre in Gdansk (CI TASK) and Interdisciplinary Center of Mathematical and Computer Modeling, University of Warsaw.

[1] Liwo, A., Czaplewski, C., Pillardy, J.; Scheraga, H. A. J. Chem. Phys. 2001, 115, 2323–2347
[2] Kubo, R. J. Phys. Soc. Japan. 1962, 17, 1100–1120

UNRES, NARES, and SUGRES Coarse-Grained Models - An Introduction

Adam Liwo* [1,2], Cezary Czaplewski* [1], Adam K. Sieradzan [1], Paweł Krupa [3], Magdalena Mozolewska [4], Emilia Lubecka [5], Robert Ganzynkowicz [1], Agnieszka G. Lipska [1], Agnieszka Karczyńska [1,2], Łukasz Golon [1], Bartłomiej Zaborowski [1] Stanisław Ołdziej [6]

[1] Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland
[2] School of Computational Sciences, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul 130-722, Republic of Korea
[3] Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, PL-02668 Warsaw, Poland
[4] Institute of Computer Science, Polish Academy of Sciences, ul. Jana Kazimierza 5, Warsaw 01-248, Poland
[5] Institute of Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland
[6] Intercollegiate Faculty of Biotechnology, University of Gdańsk and Medical University of Gdańsk, Abrahama 58, 80-307 Gdańsk, Poland
*presenting authors

This talk will provide an overview of the theoretical foundations and implementation of the Unified Coarse-Grained Model of biological macromolecules [1] developed in our laboratory whose components are UNRES (proteins), NARES-2P (nucleic acids), and SUGRES-1P (polysaccharides). All three components share a common description, in which the virtual-bond chain is constructed by linking anchor points (C□ atoms in UNRES, sugar centers in NARES-2P, and the oxygen atoms of the glycosidic linkages for SUGRES-1P, respectively). The backbone interaction sites (united peptide groups in UNRES, united phosphate groups in NARES-2P, and united sugar rings in SUGRES-1P) are located halfway between two consecutive anchor points. In UNRES and NARES-2P, the other interaction sites are the united side chains or the united sugar-base groups, respectively, which attached to the anchor points. The respective energy functions originate from the potential of mean force of the system under study in water, which is expanded [2] into Kubo's cluster-cumulant functions [3], which correspond to effective energy terms. The analytical expressions for these energy terms, especially the multibody terms, are derived by expanding the cluster-cumulant functions into generalized cluster cumulants [2,3]. The conformational-search engine is based on Langevin dynamics [4] enhanced by the multiplexed replica exchange [5] to make the search more effective. Geometric restraints from experiments (such as NMR or SAXS) or from templates can also be used with the model.

[1] Liwo, A.; Baranowski, M.; Czaplewski, C.; Gołaś, E.; He, Y.; Jagieła, D.; Krupa, P.; Maciejczyk, M.; Makowski, M.; Mozolewska, M.A.; Niadzvedtski, A.; Ołdziej, S.; Scheraga, H.A.; Sieradzan, A.K.; Ślusarz, R.; Wirecki, T.; Yin, Y.; Zaborowski, B. J. Mol. Model. 2014, 20, 2306.
[2] Sieradzan, A.K.; Makowski, M.; Augustynowicz, A.; Liwo, A. J. Chem. Phys., 2017, 146, 124106.
[3] Kubo, R. J. Phys. Soc. Japan, 1962, 17, 1100-1120.
[4] Khalili, M.; Liwo, A.; Jagielska, A.; Scheraga, H.A. J. Phys. Chem. B, 2005, 109, 13798-13810.
[5] Czaplewski, C.; Kalinowski, S.; Liwo, A.; Scheraga, H.A. J. Chem. Theor. Comput., 2009, 5, 627-640.

Determination of the Potentials of Mean Force for rotation about the O · · · O virtual bonds in O(1->3)-bonded polyglucose chains

Emilia A. Lubecka [1] and Adam Liwo [2]

[1] Institute of Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland

[2] Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland

The new theory of the construction of physics-based Unified Coarse-Grained Model (UCGM) force field, with which to derive physics-based functional expressions in coarse-grained potentials for polymer chains, have been developing in our laboratory [1]. In particular, it has been found that the virtual-bond torsional potentials depend not only on the virtual-bond dihedral angles but also on the adjacent virtual-bond angles. Using this approach, we determined effective coarse-grained potentials, to be used in the single-center model of polysaccharides (SUGRES-1P), in which the anchor points are the glycosidic oxygen atoms, with the sugar interaction site positioned halfway between the two consecutive glycosidic oxygen atoms [2]. Recently, we have obtained the potentials for the O· · · O· · · O virtual-bond angles (θ) and for the dihedral angles for rotation about the O· · · O virtual bonds (γ) of 1 → 4-linked glucosyl polysaccharides, for all possible combinations of [α, β]-[D,L]-glucose [3]. We have used umbrella sampling molecular dynamics (MD) simulations with the all-atom AMBER16 [4] force field to calculate the free-energy surfaces of component glucosyl disaccharides. Now we are reporting the potentials of mean force corresponding to the virtual-bond angles and the virtual-bond dihedral angles of 1 → 3-linked [α, β]-D-glucose pairs. The 1,3-β-glycosidic bonds in linear, unbranched chains were found in callose, curdlan, paramylon, laminarin and pachyman [5]. The 1,3-glycosidic bonds between α-, and β- D-glucose and vice versa, are also known, and occur for example in salecan [6]. The determined potentials will be presented and compared with the available structural data of polysaccharides.

[1] Sieradzan, A.K.; Makowski, M.; Augustynowicz, A.; Liwo, A. *J. Chem. Phys.* **2017**, 146, 124106.
[2] Liwo, A.; Baranowski, M.; Czaplewski, C.; et al., *J. Mol. Model.* **2014**, 20, 2306-2320.
[3] Lubecka, E.A.; Liwo A.; *J. Chem. Phys.* [under review]
[4] Case, D.A; Betz, R.M.; Botello-Smith, W.; et al. **2016**, AMBER 2016, University of California, San Francisco.
[5] Bacic, A.; Fincher, G.B.; Stone, B.A. (Eds.). **2009**. *Chemistry, biochemistry, and biology of 1-3 beta glucans and related polysaccharides*. Academic Press.
[6] Xu, L.; Zhang, *J. Appl. Microbiol. Biotechnol.* 2016, 100, 9023-9036.

Increasing sequence diversity in protein design by combining Rosetta with molecular dynamics

Jan Ludwiczak [1, 2], Adam Jarmula [2] and Stanislaw Dunin-Horkawicz [1]

[1] Structural Bioinformatics Laboratory, Centre of New Technologies, University of Warsaw, Warsaw, Poland
[2] Laboratory of Bioinformatics, Nencki Institute of Experimental Biology, Warsaw, Poland

Protein design is a procedure for computing natural-like sequences that will fold into a specified structure. It has already been demonstrated that considering the backbone flexibility during the design process positively influences the diversity of the resulting sequences [1]. Rosetta Design, a commonly used software for protein design, allows for the effective exploration of the sequence space, while the molecular dynamics (MD) simulations can thoroughly sample the protein native state conformational space. By combining these two approaches, we developed an iterative design procedure, in which backbone conformational ensembles obtained by clustering of MD trajectories are used as templates for the design. We show that such a combined approach can generate significantly more diverse sequences than currently used procedures. The observed increase in the diversity is achieved without a loss in the quality of sequences, measured as overall resemblance of the designed sequences to natural sequences. In addition, we implemented a MD-based protocol [2] that can be used for assessing the stability of designed models and selecting the best candidates for experimental validation or generating the structural ensembles that can be used as an input for further design simulations. In sum, our results demonstrate that the MD ensemble-based flexible backbone design significantly outperforms the current state-of-the-art methods and thus should be a method of choice for the design of virtually all protein classes. Finally, to make the procedure accessible for the community we provide a set of easy-to-use scripts for performing the simulations and visualizing the results.

[1] Ollikainen, N.; Kortemme, T. PLoS Comput Biol. 2013, 9
[2] Bhardwaj, G.; Mulligan, VK.; Bahl, CD.; Gilmore JM.; et al. Nature. 2016, 538, 329–335

RNA 3D Structure Prediction Using Multiple Sequence Alignment Information

Marcin Magnus [1,2], Caleb Geniesse [2], Rhiju Das [2,3] and Janusz M. Bujnicki [1,4]

[1] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Warsaw, Poland
[2] Department of Biochemistry, Stanford University, Stanford, CA, USA
[3] Department of Physics, Stanford University, Stanford, CA, USA
[4] Laboratory of Bioinformatics, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland

The prediction of three-dimensional structures of complex RNAs is still a challenging task. Here, we present an approach called EvoClustRNA (https://github.com/mmagnus/EvoClustRNA) that takes advantage of evolutionary information in distant sequence homologs, based on a classic strategy in protein structure prediction. Using the empirical observation that RNA sequences from the same RNA family typically fold into 3D structures matching at near-atomic resolution, we test whether we might guide in silico modeling by seeking global helical arrangements for the target sequence that are shared across de novo models of numerous sequence homologs. EvoClustRNA performs a multi-step modeling process: First, for the target sequence, a subset of homologous sequences is selected using the RFAM database. Subsequently, independent folding simulations using ROSETTA/FARNA are carried out for each sequence. Structural fragments corresponding to the evolutionary conserved helical regions - determined from the alignment - are extracted from all obtained models and clustered. The model of the target sequence is selected based on the most common structural arrangement of helical. We tested our approach on a benchmark of RNAs of known structure and, most rigorously, on three blind RNA-Puzzles challenges. Our predictions ranked #1 (according to RMSD) of all submissions for the L-glutamine riboswitch (bound form) and #2 for the ZMP riboswitch and the Pistol ribozyme. Through this combination of parallel modeling of homologous sequences and selecting the final model based on the clustering of conserved fragments, we increase the performance of RNA structure prediction and the methodology provides useful structural information for biological problems.

High-resolution prediction of protein-peptide complexes using a combination of CABS-dock and Rosetta FlexPepDock tools

Paulina H. Marek [1,2], Maciej P. Ciemny [1,3], Mateusz Kurcinski [1], Maciej Blaszczyk [1], Andrzej Kolinski [1], Sebastian Kmiecik [1]

[1] Biological and Chemical Research Centre, Faculty of Chemistry, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland
[2] Faculty of Chemistry, Warsaw University of Technology, Noakowskiego 3, 00-664 Warsaw, Poland
[3] Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland

In recent years, peptides have gained much interest in pharmaceutical research and development. Rational design of peptide therapeutics usually starts with structure characterization of a protein-peptide complex.
In many cases it is difficult or impossible to use experimental approaches, thus reliable computational methods are needed. Practical applications require high-resolution predictions of sufficient accuracy for subsequent structure-activity relation analyses (i.e. studying the effect of in silico mutations). Here we present our results of using a combination of CABS-dock and Rosetta FlexPepDock tools for molecular docking of peptides to proteins. The presented protocol allows achieving high-resolution predictions (< 1.5 Å) [1-3].

[1] Kurcinski, M., et al. Nucleic Acids Res. 2015, 43, 419-424.
[2] Blaszczyk, M. et al. Methods 2016, 93, 72-83.
[3] Ciemny, M. P., et al. Methods Mol Biol 2017, 1561, 69-94.

Gluten: a fluid or a solid? Insights from coarse-grained molecular dynamics simulations

Łukasz Mioduszewski [1] and Marek Cieplak [1]

[1] Institute of Physics, PAS, Al. Lotników 32/46, 02-668 Warszawa

Gluten is obtained by washing away soluble components from wheat flour. The remaining mass consists mainly of storage proteins (over 75% mass [1]), which do not have one clearly defined tertiary structure [1] and can form megadalton-sized complexes [2]. This mechanochemical network of proteins is responsible for viscoelastic properties of wheat dough, despite being less than 20% of its mass [3]. It exhibits both solid-like behavior and liquid-like viscous drag from irreversible deformations. They can be characterized by dynamic Young modulus $G^* = G' + G''$, which describes response to small-amplitude oscillating deformation: G' for the in-phase (elastic) part and G'' for the out-of-phase (viscous) part. The poster presents a model that can recreate this elastic response of gluten. Existing theories of gluten elasticity point out the crucial role of interchain hydrogen and disulfide bonds [3]. When polymers are stretched, free space between them disappears and more hydrogen bonds are formed (causing resistance to further strain). The resistance is also provided by disulfide bonds that hold some gluten proteins (called glutenins) together, forming a sort of polymer gel [3]. These facts provide predictions that can be incorporated into a coarse-grained model of gluten. In that model amino acids are represented as pseudoatoms, connected harmonically to form protein chains. Additional interactions include Lenard-Jones potential that mimics hydrogen bonding, and a dynamic potential for disulfide bonds that enables their rupture and reforming. Computer simulations cannot start from the native structure, so chain conformations are generated randomly, and then evolve according to the simplified potential, forming large complexes. Fortunately, local details of structure are not thought to be very important in recreating rheological properties of many interacting polymers [2]. Preliminary results were obtained by periodically deforming the box containing gluten proteins and recording the response force. Amplitude of the force response seems to increase, indicating strain hardening.

[1] P. Shewry, N. Halford, P. Belton and A. Tatham, "The structure and properties of gluten: an elastic protein from wheat grain." Phil. Trans. R. Soc. B: Biological Sciences, 357(1418): 133-142 (2002).
[2] H. Singh and F. MacRitchie, "Application of Polymer Science to Properties of Gluten." J. Cereal Sci. 33(3): 231-243 (2001).
[3] H. Wieser, "Chemistry of gluten proteins." Food Microbiol. 24(2): 115-119 (2007).

Applications of UNRES and NARES-2P Coarse-Grained Models in Structural Biology and Biophysics

Magdalena A. Mozolewska[1,2]*, Paweł Krupa[1,3]*, Adam K. Sieradzan[1]*

[1]Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland.
[2]Institute of Computer Science, Polish Academy of Sciences, ul. Jana Kazimierza 5, Warsaw 01-248, Poland.
[3]Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, PL-02668 Warsaw, Poland.
*presenting authors

This talk will cover recent applications of the coarse-grained UNRES[1] model of proteins to protein-structure prediction and the applications of the UNRES model and the NARES-2P[1] model of nucleic acids to study the dynamics, thermodynamics, kinetics, and free-energy landscapes of proteins and nucleic acids. The physics-based UNRES and NARES-2P coarse-grained force fields are an alternative to the bioinformatics tools, because they are independent of structural databases but stem from the physics of interactions. Owing to the simplified representation of the respective biopolymers, UNRES and NARES provide a 3−4 order of magnitude speed-up with respect to all-atom molecular dynamics (MD) simulations in water and hence enable us to simulate biomolecular systems extensively. The following applications of UNRES and NARES-2P will be discussed: (i) de novo[2] and (ii) assisted[3] prediction of structures of proteins and protein complexes; (iii) studies of the interactions of proteins with carbon nanotubes, (iv) studies of the structure and dynamics of the Isu1-Jac1[4] and Isu1-Jac1-Ssq1 chaperone systems in yeast, (v) thermal and force-induced unfolding of RNase A, and (vi) Steered MD studies of the telomeric DNA sequences[5].

References:
[1]A. Liwo et al., J. Mol. Model. 20 (2014) 2306.
[2]P. Krupa et al., Bioinformatics. 32 (2016) 3270–3278.
[3]M.A. Mozolewska, P. Krupa, B. Zaborowski, A. Liwo, J. Lee, K. Joo, C. Czaplewski, J. Chem. Inf. Model. 56 (2016) 2263–2279.
[4]M.A. Mozolewska, P. Krupa, H.A. Scheraga, A. Liwo, 83 (2015) 1414-1426.
[5]A.K. Sieradzan, P. Krupa, D.J. Wales, J. Phys. Chem. B. 121 (2017) 2207–2219.

Acknowledgements:
MAM, PK, and AKS were supported by grants UMO-2015/17/N/ST4/03935, 2015/17/N/ST4/03937, and UMO-2015/17/D/ST4/00509, respectively, from the National Science Center of Poland.

KnotProt: Knots and Slipknots exist in Proteins!

Wanda Niemyska [1,2], Ken Millett [3], Eric Rawdon [4], Piotr Sułkowski [5], Michał Jamróz [6], Pawel Dabrowski-Tumanski [1,6], Aleksandra Jarmolinska [1], and Joanna Sulkowska [1]

[1] Centre of New Technologies, University of Warsaw, Poland
[2] Faculty of Mathematics, University of Warsaw, Poland
[3] University of California, Santa Barbara, USA
[4] University of St. Thomas, Saint Paul, USA
[5] Faculty of Physics, University of Warsaw, Poland
[6] Faculty of Chemistry, University of Warsaw, Poland

First knots were found in proteins in 1994 [1], while first slipknots (knots formed by subchains even though their backbone chains as a whole are unknotted) in 2007 [2]. Till now we identified - using KnotProt - around 1500 entangled proteins. Recently considerable interest arose around this subject for a variety of reasons. First, it is believed that the presence of knots and slipknots in proteins is not accidental and therefore understanding their function is an important challenge. Second, recent work shows nearly perfect conservation of knotting fingerprints in some families whose members differ by hundreds of millions years of evolution (arising from distant organisms) and possess a low sequence identity [3]. Moreover, based on knotting fingerprints, it was shown that the locations of active sites in proteins are correlated with points characterizing their topology (e.g. positions of the knot core) [3].
This poster is devoted to present a KnotProt database which collects information about proteins with slipknots and knots [4]. This is the first database that classifies proteins with slipknots and knots, represents their entire complexity in the form of a "knotting fingerprint" [3], and presents many biological and geometrical statistics based on these results. The KnotProt database is based on protein chains deposited in Protein Data Bank (PDB) and contains around 1500 protein chains with knots or slipknots, among around 200000 protein chains in total. The KnotProt database will make knotting and slipknotting data easily available and should help researchers to understand biological reasons of protein knotting.

[1] Mansfield ML (1994) Are there knots in proteins? Nat Struct Biol 1:213–214.
[2] King NP, Yeates EO, Yeates TO (2007). Identification of rare slipknots in proteins and their implications for stability and folding. J. Mol. Biol.. Oct 2007. 373(1):153-66.
[3] Sulkowska JI, Rawdon EJ, Millett KC, Onuchic JN and Stasiak A (2012) Conservation of complex knotting and slipknotting patterns in proteins. Proc. Natl. Acad. Sci. U.S.A. 109, E1715–E1723.
[4] Jamroz M, Niemyska W, Rawdon, EJ, Stasiak A, Millett KC, Sułkowski P, Sulkowska JI (2014) KnotProt: a database of proteins with knots and slipknots. NAR 43, D306-D314.

Modeling protein flexibility using the new CABS-flex standalone application

Tymoteusz Oleniecki [1,3,4], Maciej Paweł Ciemny [1,2], Mateusz Kurciński [1], Maciej Błaszczyk [1], Andrzej Koliński [1], Sebastian Kmiecik [1]
[1] Biological and Chemical Research Centre, Faculty of Chemistry, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland
[2] Faculty of Physics, University of Warsaw, ul. Pasteura 5, 02-093 Warszawa, Poland
[3] College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Banacha 2C, 02-097 Warszawa, Poland
[4] Mossakowski Medical Research Centre Polish Academy of Sciences, Pawińskiego 5, 02-106 Warszawa, Poland

The conformational flexibility of protein structures is crucial for their functions. Simulations of protein flexibility remain computationally costly or intractable for most of protein systems using classical modeling tools. Here we present a new standalone version of our method – CABS-flex [1, 2, 3] – so far available as a web server [1]. The method combines a highly efficient, coarse-grained approach with all-atom modeling methods. The CABS-flex predictions reflect the flexibility of an investigated protein and provide a picture complementary to NMR conformational ensembles [2], as well as results obtained from molecular dynamics simulations [3]. The CABS-flex method was also successfully used for efficient simulations of protein flexibility in predictions of protein-peptide complexes [4, 5] and protein aggregation properties [6]. The standalone CABS-flex application allows for customization of the simulation parameters, handling large-sized systems and provides a flexible framework for result analysis. The standalone CABS-flex version is freely available at http://biocomp.chem.uw.edu.pl/CABSflexApp/ and server version at: http://biocomp.chem.uw.edu.pl/CABSflex/.

[1] Jamroz, M.; Kolinski, A.; Kmiecik, S. Nucleic Acids Res 2013, 41, W427-431.
[2] Jamroz, M.; Kolinski, A.; Kmiecik, S. Bioinformatics 2014, 30 , 2150-2154.
[3] Jamroz, M.; Orozco, M.; Kolinski, A.; Kmiecik, S. J Chem Theory Comput 2013, 9, 119-125.
[4] Ciemny, M. P. et al. Sci Rep 2016, 6, 37532.
[5] Ciemny, M. P.; Kurcinski, M.; Kozak, K.; Koliński, A.; Kmiecik, S. in Methods in Molecular Biology 2017, Vol. 1561 69-94.
[6] Zambrano, R. et al. Nucleic Acids Res 2015, 43, W306-313.

Novel folds amongst protein knots

Agata P. Perlinska [1,2], Aleksandra I. Jarmolinska [1,2] and Joanna I. Sulkowska [1,3]

[1] Centre of New Technologies, University of Warsaw, Poland
[2] Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Poland
[3] Faculty of Chemistry, University of Warsaw, Poland

Knots in proteins are increasingly being recognized as an important structural concept, and the folding of these peculiar structures still poses considerable challenges. From a functional point of view, most protein knots discovered so far are either enzymes or DNA-binding proteins. Our comprehensive topological analysis of the Protein Data Bank reveals several novel structures including knotted mitochondrial proteins and the most deeply embedded protein knot discovered so far. The latter appears in a three-domain protein, with knotted region located in the middle domain, which makes the knot tails more than 130 residues long. Using a C-alpha structure based model for folding simulations, we have found, that its folding pathway is unique among knotted proteins – a loose knot is formed first near the terminus, and then slides to its native position.
The mitochondrial knotted proteins include an apoptosis-inducing factor, and a ribosomal subunit, both of which are expressed in the cytoplasm. Since such proteins have to cross the membrane to get to their designated location, we analysed possible pathways of such transport. The literature and our simulation data show, that the translocation of these proteins occurs before the knotting. Moreover, the structural analysis shows that the knotted region in the apoptosis-inducing factor has an important structural role. It both stabilizes the flexible CTD domain and aligns the crucial DNA binding elements.

The higher-order organization of chromatin looping in Human genome

D. Plewczynski[]1], P. Szałaj[1], M. Sadowski[1], M. Kadlof[1], Z. Tang[2], Y. Ruan[2]

[1]Centre of New Technologies, University of Warsaw, Warsaw, Poland;
[2] Jackson Laboratory for Genomic Medicine, CT, USA

The recent genomic and bioimaging insights in the higher order chromatin organisation in human nucleus motivated us to propose novel simulation method aimed at the effective prediction of three-dimensional human genome structure. We applied an advanced long-read Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET) experimental strategy combined with computational modelling to comprehensively map higher-order chromosome folding and specific chromatin interactions mediated by CCCTC-binding factor (CTCF), cohesin and RNAPII with haplotype specificity and nucleotide resolution in different human cell lineages. We demonstrate the effectiveness of biophysical modeling in building 3D genome models at multiple levels, including the entire genome, individual chromosomes, and specific segments at megabase (Mb) and kilobase (kb) resolutions of single average and ensemble structures.

Perspectives of coarse graining to investigate the dynamics of blood coagulation proteins

Michał Błażej Ponczek [1]

[1] University of Lodz, Faculty of Biology and Environmental Protection, Department of General Biochemistry, ul. Pomorska 141/143, 90-236 Łódź

Blood clotting includes complex biological interactions with chemical and biochemical compounds, mostly proteins and non-protein chemical and biological elements. This complicated arrangement has been studied biochemically for several years. The protein complexes are the main functional part of the system. Coarse graining is the method which enables learning about the dynamics of that system.

AMBER-compatible coarse-grained model of glycosaminoglycans

Sergey A. Samsonov [1,2]

[1] Faculty of Chemistry, University of Gdańsk, ul. Wita Stwosza 63, 80-308 Gdansk, Poland [2] Biotechnology Center, Dresden University of Technology, Tatzberg 47/49, 01307 Dresden, Germany

Glycosaminoglycans (GAGs) is a class of long anionic periodic linear polysaccharides made up of repetitive disaccharide units containing glucuronic/iduronic acid and N-Acetylglucosamine/N-Acetylgalactosamine with different sulfation patterns. They play an important role in cell communication processes by interactions with their protein targets such as chemokines and growth factors in the extracellular matrix. Due to their high flexibility, charged and periodic nature, GAGs represent computational challenges, which could be addressed by using a coarse-grained (CG) modeling approach. In this work, we develop AMBER-compatible CG parameters for GAGs using all-atomic (AA) molecular dynamics (MD) simulations in explicit solvent by Bolztmann conversion approach. We compare GAGs global and local properties obtained with AA and CG approaches and conclude that our CG model is appropriate for MD applications for long GAG molecules at long time scales [1]. We characterize several protein/GAG complexes by applying MM-PBSA approach using AA and CG GAG models and show their good agreement in terms of binding and per residue decomposed energies. We use in-house developed docking tools and compare results obtained with AA and CG GAG models for several biologically relevant protein/GAG complexes.

[1] Samsonov, S.; Bichmann, L.; Pisabarro M.T. Coarse-grained model of glycosaminglycans. J. Chem. Inf. Mod. 2015, 55, 114-124

The performance of helical multi-walled carbon nanotubes for effective pollutant removal: combined experimental and theoretical study

Celina Sikorska[1], Monika Paszkiewicz[2,3], Danuta Leszczyńska[3], Piotr Stepnowski[2]

e-mail: celina.sikorska@ug.edu.pl
[1] Laboratory of Molecular Modeling, Department of Theoretical Chemistry, Faculty of Chemistry, University of Gdansk, Wita Stwosza 63, 80-308 Gdansk, Poland
[2] Department of Environmental Analytics, Institute for Environmental and Human Health Protection, Faculty of Chemistry, University of Gdansk, Wita Stwosza 63, 80-308 Gdansk, Poland
[3] Department of Civil and Environmental Engineering, Interdisciplinary Nanotoxicity Center, Jackson State University, 1400 John R. Lynch Street, Jackson, MS 39217, USA

Carbon-based nanoparticles (NPs) have been a major topic in chemical research over the past four decades [1-4]. Since carbon nanotubes (CNTs) combine three-dimensionality with unique physico-chemical properties, they are extremely promising NPs for the preparation of new advanced materials and biologically active molecules. The CNTs are used as sorbents in extraction techniques [5] and the popularity of these materials is associated with their tunable and unique properties. In particular, the high surface area of CNTs allows their use in a much smaller amount as compared to classic adsorbents, and the thermal stability of CNTs allows their easy regeneration and reuse. All potential useful properties of CNTs do stimulate us to pay more attention to those unique molecular systems.
In this contribution, a new type of multi-walled carbon nanotubes, referred to as helical, were used as the sorbent in dispersive solid-phase extraction to extract polycyclic aromatic hydrocarbons (PAHs) from water samples. Moreover, to understand the adsorbing characters of PAHs onto carbon nanotubes, we carried out a theoretical investigation on the interaction mechanisms between PAH molecules and MWCNTs using the PM6 method. The conclusion is supported by employing chemometric analysis as well as providing structural parameters and interaction energies for adsorption processes (PAH+CNTPAH-CNT).

[1]Sikorska C 2016 When a nanoparticle meets a superhalogen: a case study with C-60 fullerene Physical Chemistry Chemical Physics 18 18739-49
[2]Sikorska C, Gajewicz A, Urbaszek P, Lubinski L and Puzyn T 2016 Efficient way of designing fullerene derivatives based on simplified DFT calculations and QSPR modeling Chemometrics and Intelligent Laboratory Systems 152 125-33
[3]Sikorska C and Puzyn T 2015 The performance of selected semi-empirical and DFT methods in studying C60 fullerene derivatives Nanotechnology 26 455702
[4]Urbaszek P, Gajewicz A, Sikorska C, Haranczyk M and Puzyn T 2017 Modeling adsorption of brominated, chlorinated and mixed bromo/chloro-dibenzo-p-dioxins on C-60 fullerene using Nano-QSPR Beilstein Journal of Nanotechnology 8 752-61
[5]Paszkiewicz M, Tyma M, Jakubus A and Stepnowski P 2016 Recent Applications of Carbon Nanotubes as Sorbents for the Extraction of Pharmaceutical Residues Current Analytical Chemistry 12 268-79

This project was supported by the Polish National Science Centre (NCN) Grant No. NCN UMO-2012/07/D/NZ7/04342 . C.S. thanks the Foundation for Polish Science (START 2016 Program) for granting her with a fellowship.

Development of Lattice Models for Studying Protein Aggregation

Mai Suan Li [1] and Nguyen Truong Co [2,3]

[1] Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland
[2] Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland
[3] Institute for Computational Sciences and Technology, SBI building, Quang Trung Software City, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Vietnam

Understanding protein aggregation is important as it is associated with neurodegenerative diseases. Because the self-assembly of biomolecules is too slow the use of all-atom models becomes impractical in monitoring kinetics of this process. Motivated by this challenge our group has developed coarse-grained lattice models [1] which allow us to disclosed the key factors that govern fibril formation rate. The faster is the aggregation the higher hydrophobicity, the lower net charge and the higher population of the fibril-prone state of the polypeptide chain [2,3]. Lattice models allowed us to explain the non-trivial experimental observation on the dual effect of crowding particles on fibril growth of proteins that for a fixed crowder concentration the fibrillation kinetics is fastest at intermediate values of total surface of crowders [4]. A novel method for determining the size of critical nucleus of fibril formation of polypeptide chains was proposed. Based on the hypothesis that the fibril grows by addition of a nascent peptide to the preformed template, the nucleus size Nc is defined as the number of forming template peptides above which the time to add a new monomer becomes independent of the template size [5]. Lattice models are useful to capture the effect of confinement and surface roughness on self-assembly kinetics [6]. One of our current goals is to develop lattice models for lipid bilayer membrane and to study its impact on protein fibril formation.

[1] Li, M.S.; Klimov, K.D.; Thirumalai, D. J. Chem. Phys. 2008, 129, 175101
[2] Li, M.S.; Co, N.T.; Hu, C-K.; Straub, J.E.; Thirumalai, D. Phys. Rev. Lett. 2010, 105, 218101
[3] J. Nasica-Labouze, J. et al. Chem. Rev. 2015, 115, 3518-3563
[4] Co, N.T.; Hu, C-K.; Li, M.S. J. Chem. Phys. 2013, 138, 185101
[5] Co, N.T.; Li, M.S. J. Chem. Phys. 2012, 137, 095101
[5] Co, N.T.; Li, M.S. 2017 (In preparation)

GoMARTINI: Combining the MARTINI and structure-based coarse-graining approaches for the molecular dynamics studies of conformations transitions in proteins

A. Poma [1], M. Cieplak [1] and P. Theodorakis [1]

[1] Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland

The application of coarse-grained (CG) models in biology is essential to access large length and time scales required for the description of many biological processes. The ELNEDIN protein model is based on the well-known MARTINI CG force-field and incorporates additionally harmonic bonds of a certain spring constant within a defined cutoff distance between pairs of residues, in order to preserve the native structure of the protein. In this case, the use of unbreakable harmonic bonds hinders the study of unfolding and folding processes. To overcome this barrier we have replaced the harmonic bonds with Lennard–Jones interactions based on the contact map of the native protein structure as is done in Gō-like models [1]. This model exhibits very good agreement with all-atom simulations and the ELNEDIN. Moreover, it can capture the structural motion linked to particular catalytic activity in the Man5B protein, in agreement with all-atom simulations. In addition, our model is based on the van der Waals radii, instead of a cutoff distance, which results in a smaller contact map. In conclusion, we anticipate that our model will provide further possibilities for studying biological systems based on the MARTINI CG force-field by using advanced-sampling methods, such as parallel tempering and metadynamics.

[1] Poma, A. B.; Cieplak, M.; Theodorakis, P. E. J. Chem. Theory Simul. 2017, 13, 1366-1374

SimRNA Folding Simulations With Use Of Long-Range Restraints.

Tomasz Wirecki [1], Marcin Magnus [1], Michał Boniecki [1], Michael Solski [1] and Janusz Bujnicki [1, 2]

[1] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Warsaw, Poland
[2] Laboratory of Bioinformatics, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland

Ribonucleic acid (RNA) molecules play fundamental roles in a variety of cellular processes. RNA is not only a carrier of genetic information, but also plays a great role in gene expression regulation and catalysis of biochemical reactions. The ability of RNA to interact with other biomolecules depends on its ability to form complex, stable, three-dimensional structures. However, experimental determination of RNA 3D structures is laborious and challenging, therefore, it is important to develop computational tools for that purpose.
SimRNA is a computational method for de novo 3D RNA structure prediction developed in Janusz Bujnicki's laboratory [1]. SimRNA uses a coarse-grained representation of RNA molecules. The energy function is derived as a statistical potential and is used to guide a Monte Carlo procedure to investigate the conformational landscape of the simulated system. SimRNA is able to provide reliable predictions for sequence lengths up to 50 nucleotides only provided with the sequence. For longer RNA molecules additional restraints in the form of secondary structure (and other long-range restraints) may be needed. They can be obtained through dedicated computational methods or biological experiments (e.g., SHAPE). Obtaining secondary structure information is a standard procedure, whereas the methods for obtaining long-range 3D contacts information are scarce. There exist few experimental methods still in the developmental stage, and a few computational workflows e.g. evolutionary coupling analysis [2].
In this work we apply long-range restraints on SimRNA simulations. We are developing an efficient scoring method aiming to pick out decoys from simulations which fulfill the tertiary interactions in the most similar way to the experimental structure. We also work towards pointing out the most important long-range restraints, while trying to omit the ones that would hinder the simulation. Ultimately, we aim to derive a workflow for imposing long-range restraints and efficient scoring of the models.

[1] Boniecki, M. J.; Lach, G.; Dawson, W. K.; Tomala, K.; Lukasz, P.; Soltysinski, T.; Rother, K. M.; Bujnicki, J. M. Nucleic acids res. 2016, 44(7), e63-e63
[2] Weinreb, C.; Riesselman, A. J.; Ingraham, J. B.; Gross, T.; Sander, C.; Marks, D. S. Cell, 2016, 165(4), 963-975

Characterization of water patterns in catalytic subunit of protein kinases

Marta D. Wisniewska [1,2], Piot Setny [1]

[1] Centre of New Technologies, University of Warsaw, Poland
[2] Inter-Faculty Interdisciplinary Doctoral Studies in Natural Sciences and Mathematics, University of Warsaw, Poland

Water is present in every living cell. The water molecules play a dual role: firstly, they serve as an environment for biological processes or chemical reactions and secondly, they act as an internal water inside proteins. In the latter case they are called structural waters molecules. These structural water molecules occur in a number of protein families, one of them are protein kinases.
Protein kinases are unique: in spite of little sequence homology, their catalytic subunits retain conserved structure across the entire family and only few amino-acids or their motifs are well preserved. Protein kinases as phosphotransferases shift a phosphate group to the target protein. The catalysis performed by kinases is strictly regulated by their own activation. The activation process involves large conformational changes. Interestingly, some water molecules inside kinases are located within regions directly involved in the activation process. Moreover, that placement of structural water molecules is common among most members of kinase family, what suggests that, they may contribute to the activation process.
The exact role of those water molecules remains poorly investigated, in part due shortcomings in experimental techniques. For example, the data obtained through X-ray experiments give only the static snapshot, that lack information of dynamics. In turn, NMR measurements give insights into protein dynamics with low spatial resolution. Fortunately, computational approaches provide means to investigate both structural and dynamic properties of buried water.
Here, we present the results of computational analysis of buried water conservation among protein kinases family. We characterize the environment of preserved hydration sites, analyze water affinity to active/inactive enzyme, describe the potential role of bound water molecules, and their possible influence on kinase's activation.

This work is supported by the National Science Centre of Poland (NCN), grant UMO-2013/11/D/ST4/02821.

Role Of Slipknot Topology For Cation Binding

Vasilina Zayats[2], Joanna I. Sułkowska [1,2]

[1]Faculty of Chemistry, University of Warsaw, 02-093, Warsaw, Poland
[2]Centre of New Technologies, University of Warsaw, 02-097, Warsaw, Poland

Cation binding in proteins was reported to play important role for protein structure and function. In the superfamily of alkaline phosphatases a divalent cation is bound in the active site and suggested to play a functional role. All members of alkaline phosphatase superfamily possess +3.1 slipknot topology [1]. Slipknot formed when one loop threads across another loop. Crossings which form a slipknot located at the active site and conserved across whole superfamily [2]. The cation is located between residues forming a lasso around a slipknot loop. Such localization of the cation could be important for stabilizing a slipknot conformation. On the other hand, the slipknot topology could create a good environment and stable site for cation binding. Proteins of alkaline phosphatase superfamily represent three main groups of enzymes. Most of them are hydrolases, some belong to transferases and isomerases according to the CATH database. Several divalent cations are found in the active site of these proteins: zinc, calcium, magnesium and manganese. Interestingly, there is also protein which does not contain cation in its active site. Despite this protein as suggested does not have hydrolase activity, the fold type, slipknot topology and general architecture of the active site is preserved as in other proteins of alkaline phosphatase superfamily. In this work we would like to understand whether the cation binding is required to stabilize the conformation of the active site which is formed by the +3.1 slipknot or slipknot topology is able to provide a suitable conformation for binding of various divalent cations and at the same time stabilize overall protein structure.

Acknowledgements: This work was supported by European Molecular Biology Organization Installation Grant 2057 and Polish Ministry for Science and Higher Education Grant 0003/ID3/2016/64.

[1] K.C. Millett, E.J. Rawdon, A. Stasiak, J.I. Sulkowska, Biochem Soc Trans, 2013, 41, 533–537.
[2] J.I. Sulkowska, E.J. Rawdon, K.C. Millett, J.N. Onuchic, A. Stasiak, Proc Natl Acad Sci U S A. 2012, 109, E1715–E23.