View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

methylammonium, acetate and water[4]. With the removal of the previous ambiguity brought by using water as both the hydrogen-bond donor and acceptor fragment[2], improved screening results are seen using targets in the DUD database in the extended SILCS-Pharm method. Validations of the protocol using eight different protein targets show equal or better screening results when compared with results using common docking methods such as DOCK or AutoDock, indicating the potential utility of the approach in rational drug design.

[1] Yu W, Guvench O, MacKerell AD, Jr. (2013) In:Understanding and exploiting protein-protein interactions as drug targets. Future Science Book Series. Future Science Ltd, London, UK, p90-102.
[2] Yu W, Lakkaraju SK, Raman EP, MacKerell AD, Jr. (2014) J Comput Aided Mol Des 28(5):491-507.
[3] Guvench O, MacKerell AD, Jr. (2009) PLoS Comput Biol 5(7):e1000435.
[4] Raman EP, Yu W, Lakkaraju SK, MacKerell AD, Jr. (2013) J Chem Inf Model 53(12):3384-3398.

63-Plat
Quantifying Macromolecular Conformational Transition Pathways

Oliver Beckstein, Sean L. Seyler, Avishek Kumar, Michael F. Thorpe.
Physics, Arizona State University, Tempe, AZ, USA.

Diverse classes of proteins such as molecular motors, enzymes, and active transmembrane transporters function through large-scale conformational changes. Computer simulations of these conformational transitions are challenging. A range of coarse-grained and biased simulation techniques have been used to generate individual transitions or ensembles of transition pathways but it has been difficult to compare pathways produced by different methods and so to assess their relative strengths. We introduce a comprehensive method (pathway similarity analysis, PSA) for quantitatively characterizing and comparing macromolecular pathways. The Hausdorff and Frechet metrics (known from computational geometry) are used to quantify the degree of similarity between piecewise-linear curves in configuration space. We tested PSA on a toy system to study the effect of temperature fluctuations (path roughness) and dimensionality. We compare a sample of publicly accessible transition pathway simulation servers and our own dynamic importance sampling (DIMS) MD method for the closed-to-open transitions of the apo enzyme adenylate kinase (AdK). PSA was applied to ensembles of hundreds of trajectories of the conformational transitions of the transporter Mhp1 and of AdK and diphtheria toxin, which were produced by DIMS MD and the Geometrical Pathways algorithm. Clustered PSA enabled the selection of a small subset of representative trajectories for further analysis. A strength of PSA is its use of the full information available from the 3N-dimensional configuration space trajectory, without requiring additional specific knowledge about the system. We show how trajectory analysis methods relying on pre-defined collective variables such as native contacts or geometric quantities can be used synergistically with PSA. We discuss the method's potential to enhance our understanding of transition path sampling methods, validate them, and ultimately help guide future research toward deeper physical insights into conformational transitions.

64-Plat
Computational Design of Repeat-Proteins with a Predefined Geometry

Sebastian Rämisch[1], Ulrich Weininger[2], Jonas Martnisson[1], Mikael Akke[2], Ingemar André[1].
[1]Department of Biochemistry and Structural Biology, Lund University, Lund, Sweden, [2]Department of Biophysical Chemistry, Lund University, Lund, Sweden.

New protein design methods are needed to further improve the development of protein-binding scaffolds. Repeat proteins are linear tandem arrays of structurally similar building blocks, and they are established platforms for engineering proteins inhibitors and biosensors. However, current sequence-based engineering approaches lack the possibility of customizing the overall shape of a binder to its target molecule. Structure-based protein design offers a possibility of optimizing the overall shape of engineered binding scaffolds to better match their targets. We developed a protocol for the computational design of shape-optimized binding scaffolds that can better match their targets. By combining sequence optimization of existing repeats and de novo design of capping structures, we designed leucine-rich repeat (LRR) proteins where the building blocks assemble with a novel geometry. We validated the geometric design approach by engineering an artificial donut-like ring structure constructed from ten self-compatible repeats. Characterization of the design constructs revealed that buried cysteines play a central role for stability and folding cooperativity in certain LRR proteins. This may be used to selectively stabilize or destabilize specific parts of a protein. The computational procedure may now be employed to develop repeat proteins with various geometrical shapes for applications where greater control of the interface geometry is desired.

65-Plat
Computational Prediction of G-Quadruplex Formation

Jacob S. Calvert[1], Alex Kreig[1], Saurabh Sinha[2], Sua Myong[1].
[1]Bioengineering, University of Illinois at Urbana-Champaign, Champaign, IL, USA, [2]Computer Science, University of Illinois at Urbana-Champaign, Champaign, IL, USA.

Guanine-rich regions of genomic DNA can spontaneously fold into secondary structures called G-quadruplexes (GQs). Akin to tiny switches, GQs regulate genetic processes through their folding and unfolding. Their interest to basic science, as well as their potential as therapeutic targets for human diseases, has motivated the creation of computational tools for their prediction. Currently, GQ folding predictors are based on results from structural and biochemical studies of GQ formed in single-stranded (ss) DNA. As a result, existing tools perform poorly when applied to the prediction of GQ formation in double-stranded (ds) DNA, the native context within which genomic GQs are found. Here, we present a probabilistic model of GQ formation, which is learned from large-scale human genomic pull-down experiments and applied to the analysis of gene ontological data. Advances in the characterization of GQs in dsDNA have enabled us to integrate results from small-molecule binding assays and single-molecule FRET microscopy into our model. In order to obtain training sets of sequences, we identified nearly 700,000 unique, potential GQs and categorized them according to pull-down experiment outcomes. Model parameters learned from these training sets agree with experimental evidence and, when asked to predict the folding of dsDNA GQ sequences, outperformed existing models of GQ folding. To further explore the model's utility, we screened potential GQ drug targets by selecting high-scoring sequences for gene ontological analysis. Our results indicate that highly scoring sequences are preferentially located near cancer-related genes. This tool can be applied to genomic sequences to locate the most strongly forming GQs, revealing valuable information for the design of GQ-targeting therapies, and represents the next step toward the practical, widespread use of GQs in medicine and technology.

66-Plat
Protein and RNA Structure Prediction by Integration of Co-Evolutionary Information into Molecular Simulation

Eleonora De Leonardis[1,2], Benjamin Lutz[3,4], Simona Cocco[5], Remi Monasson[6], Hendrik Szurmant[7], Martin Weigt[1,2], **Alexander Schug**[3].
[1]Computational and Quantitative Biology, Sorbonne Universités, Université Pierre et Marie Curie, Paris, France, [2]Computational and Quantitative Biology, CNRS, Paris, France, [3]Steinbuch Centre for Computing, Karlsruhe Institute of Technology, Karlsruhe, Germany, [4]Fakultät für Physik, Karlsruhe Institute of Technology, Karlsruhe, Germany, [5]Laboratoire de Physique Statistique, Ecole Normale Supérieure, associé au CNRS et à l'Université Pierre et Marie Curie, Paris, France, [6]Laboratoire de Physique Théorique, Ecole Normale Supérieure, associé au CNRS et à l'Université Pierre et Marie Curie, Paris, France, [7]Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA, USA.

Exploring the relationship of structure and function is crucial for the understanding of molecular life. Despite significant progress of experimental methods, the structural characterization of important structures for both proteins and RNA - typically preceding any detailed mechanistic exploration of their function - remains challenging. In recent years, increasingly ubiquitous availability of sequential information and novel statistical analysis has allowed to trace the co-evolution of residues and predict contact maps. These contact maps can be exploited in structure prediction tools. One maximum entropy based approach is called Direct Coupling Analysis (DCA)[1] and, e.g., was found to be sufficient for the blind prediction of a protein complex[2] and its active conformation[3] later confirmed by experiment. We have successfully adapted this method to the specifics of RNA (rnaDCA): While it is comparably simple to predict RNA secondary structure through analyzing possible Watson-Crick base pairings, predicting tertiary contacts has remained - despite significant effort - an elusive task met with limited success. In contrast, our novel rnaDCA is able to extract tertiary contacts from genomic data. We further demonstrate that these tertiary contacts are sufficient to systematically improve tertiary RNA prediction quality[4]. Considering the large gap of known ncRNA sequences to experimentally resolved tertiary structures, we are convinced that this will significantly impact all structural RNA related research.

References
[1] Weigt M et al., Proc Nat Acad Sci USA (2009) 106, 67-72; F. Morcos et al., Proc Nat Acad Sci (2011) 108, E1293-E1301
[2] Schug A et al., Proc Nat Acad Sci USA (2009) 106, 22124-22129
[3] Dago A et al., Proc Nat Acad Sci USA (2012), 109: E1733-42
[4] E. De Leonardis et al. Direct-Coupling Analysis of nucleotide coevolution facilitates RNA tertiary structure prediction, (submitted)

67-Plat

Structural Ensembles of Intrinsically Disordered Proteins using Molecular Dynamics Simulation

Sarah Rauscher, Vytautas Gapsys, Bert de Groot, Helmut Grubmüller.
Theoretical & Computational Biophysics, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany.

Intrinsically disordered proteins (IDPs) fulfill many biological roles and are important drug targets. However, they are poorly understood relative to the wealth of structural information available for globular proteins. Their structural characterization presents a formidable challenge to both theory and experiment: the structure of an IDP must be described as a structural ensemble of many interconverting conformations.

Here, we use molecular dynamics simulations to obtain structural ensembles of two IDPs: (1) a 79-residue FG-nucleoporin peptide, which is responsible for the selectivity of the nuclear pore complex and (2) a 24-residue RS-repeat peptide derived from serine/arginine-rich-splicing-factor-1, which is crucial in RNA splicing. Because force fields for polypeptides have been developed primarily to study folded proteins, it is not clear how accurately they can model disordered states. To address this issue, we performed simulations using four force fields: amber99sb*-ildn, amber ff03w, CHARMM22*, and CHARMM36 using replica exchange (RE) for a total of 150 microseconds per force field.

The structural ensembles we obtain for both the FG and RS peptides differ markedly between force fields with respect to hydrogen bonding, radius of gyration, and secondary structure, and are sufficiently converged to make such a comparison. Importantly, secondary structure content differs more on average between force fields than between the two peptide sequences. Thus, disordered peptides appear to be particularly sensitive to force field selection, much more so than globular proteins. We compared the structural ensembles obtained with each force field to both NMR and small angle x-ray scattering data for the RS peptide by computing ensemble averages of the experimental observables. The CHARMM22* force field provides the most accurate description of the force fields tested. In addition, we find that CHARMM22* provides the fastest sampling of configuration space.

Platform: Protein-Nucleic Acid Interactions I

68-Plat

Binding Competition Studied with Single Molecule FRET: The Integron Recombinase Outcompetes the Single-Stranded DNA Binding Protein for its Recombination Site

M. Svea Grieb, Varsha Natarajan, Marko Swoboda, Michael Schlierf.
B CUBE-Center for Molecular Bioengineering, TU Dresden, Dresden, Germany.

Bacteria can adapt impressively to environmental changes and acquire antibiotic multiresistance. Adaptation is mainly achieved by the integron, a genetic device that uses DNA site-specific recombination to collect and express gene cassettes. Integron recombination is unique as its recombinase, IntI, targets a single-stranded recombination site, called attC. attC sites fold themselves into an imperfect hairpin structure of 65 to 150 nucleotides, resembling a canonical recombination site. However, it is unknown how attC hairpins form in the bacterial cytoplasm. This is an interesting question, as high amounts of single-stranded DNA binding protein (SSB) are present in the cytoplasm to prevent DNA secondary structure formation, which interfere with cellular processes, like DNA replication. To investigate if IntI can counter the effect of SSB, we characterized the binding mechanism of IntI and SSB to an attC hairpin, both individually and simultaneously, using single-molecule Förster Resonance Energy Transfer (smFRET). We show that SSB is able to open a folded attC hairpin by binding to it in the non-cooperative SSB(65) mode. Contrary to binding unstructured ssDNA, SSB avoids the cooperative SSB(35) mode when binding hairpins, even in very high protein-to-DNA ratios. We performed a kinetic analysis to quantify the concentration and substrate dependent dynamics of SSB binding/unbinding. We further found that the recombinase IntI plays an antagonistic role, as it stabilizes the folded attC site. Experiments with target competition revealed that IntI is able to counterbalance SSB even at much lower concentration.

We conclude that SSB impedes attC hairpin formation at background IntI concentrations, and thus ensures the genetic stability of those palindromic sequences. However, after induction of IntI expression, the action of SSB is outcompeted, allowing IntI to bind to its target site.

69-Plat

Single-Molecule FRET for Dynamic Structural Biology: DNA Polymerase i Structure and Mechanism with Angstrom Precision

Marko Sustarsic[1], Timothy Craggs[1], Johannes Hohlbein[2], Andrew Cuthbert[1], Nicholas Taylor[1], Hendrik Kaju[3], Majid Mosayebi[3], Jonathan Doye[3], Achillefs N. Kapanidis[1].
[1]Department of Physics, University of Oxford, Oxford, United Kingdom, [2]Laboratory of Biophysics, Wageningen UR, Wageningen, United Kingdom, [3]Department of Chemistry, University of Oxford, Oxford, United Kingdom.

A key question in the function of DNA-processing enzymes is their mechanism for rapid location and binding of target sites in DNA. DNA polymerase I (Pol) is involved in high-fidelity DNA replication and repair, and must recognize gapped and flapped DNA substrates. Due to their transient and dynamic nature, the structures of these complexes and the recognition mechanism by which they are formed are largely unknown.

Combining single-molecule Förster Resonance Energy Transfer (smFRET) with dye-position modelling, we can obtain absolute distance measurements with angstrom precision. We have used 120 such distances in a rigid-body docking approach to determine the structure of the single-nucleotide-gapped DNA-Pol binary complex, and that of the gapped-DNA substrate alone.

In the complex structure, the DNA substrate exhibits a 120-degree bend. The upstream DNA is channeled into the Pol active site, and its position is in excellent agreement (RMSD < 3 Å) with the position of short DNA fragments seen in crystal structures of Pol. The downstream DNA is partially melted and held in close proximity to F771, a residue important for strand displacement.

The structure of the gapped-DNA substrate alone adopts a 20-degree bend and a twist. Our coarse-grained simulations show short excursions of the substrate to partially kinked conformations not accessible to duplex DNA. Together, these data are consistent with a two-step model whereby Pol recognizes and binds a partially kinked conformation of gapped DNA, which then further rearranges to the fully bent state seen in the complex structure.

Finally, using our method of internalization by electroporation, we have observed single-molecule FRET signals corresponding to unbound and Pol-bound gapped DNA in live bacteria, confirming the physiological relevance of the proposed substrate and complex structures and the recognition mechanism.

70-Plat

Heterochromatin Assembly and Dynamics on the Single Molecule Level

Sinan Kilic, Bachmann Andreas, Louise Bryan, **Beat Fierz**.
EPFL, Lausanne, Switzerland.

Multivalent interactions between effector proteins and histone post-translational modifications (PTMs) represent elementary processes in the chromatin signaling pathway. One such effector, heterochromatin protein 1 alpha (HP1a), is a major component of transcriptionally repressive heterochromatin. HP1a contains a chromodomain, a reader for H3 trimethylated at lysine 9 (H3K9me3), and forms dimers thereby allowing multivalent chromatin binding and playing a role in chromatin fiber compaction.

Macroscopic heterochromatin foci in the cell are stable structures, however, fluorescence photobleaching experiments demonstrate rapid dynamics of individual HP1a proteins. The mechanistic details between dynamic HP1 binding and its structural role in chromatin compaction is however not well understood.

We used single-molecule total internal reflection fluorescence microscopy (smTIRFM) to directly observe HP1a interaction dynamics with chemically defined chromatin fibers. In addition, we employed protein chemistry method to control the oligomeric state of individual HP1a proteins. Together, our measurements reveal that the HP1a residence time depends both on the local density of H3K9me3 marks, as well as on the oligomeric state of the protein. In addition, HP1a dimerization greatly accelerates its association kinetics with chromatin.

We thus propose that an increased chromatin binding rate is a key advantage of effector multivalency, allowing efficient competition between chromatin effectors and thus resulting in the maintenance of a stable chromatin state, which can quickly adapt to cellular signals.

71-Plat

RNA Unwinding by the Helicase Mtr4p and the TRAMP Complex Investigated via High-Resolution Optical Trapping

Eric M. Patrick[1], Sukanya Srinivasan[2], Eckhard Jankowsky[2], Matthew J. Comstock[1].
[1]Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA, [2]Department of Biochemistry, Center for RNA Molecular Biology, Case Western University, Cleveland, OH, USA.

We present single-molecule high-resolution optical trapping measurements of RNA unwinding by the Ski2-like nuclear helicase Mtr4p both alone and within the Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex. RNA helicases such