Intrinsically Disordered Proteins I

1332-Pos

Effects of pH on Conformational Equilibria of Intrinsically Disordered Proteins

Albert H. Mao, Scott L. Crick, Caitlin L. Chicoine, Rohit V. Pappu.
Washington University in St. Louis, St. Louis, MO, USA.

Intrinsically disordered proteins (IDPs) adopt an ensemble of conformations under native physiological conditions. Despite their lack of folded structure, they perform important physiological functions and are predicted to constitute around 30% of the eukaryotic proteome. The success of disorder prediction using a protein's primary structure suggests that the propensity for disorder is encoded in the amino acid sequence. Recently, we found that net charge per residue segregates IDP sequences along a globule-to-coil transition, enabling construction of a sequence-space phase diagram that subdivides IDPs based on the polymeric character of their conformational ensembles. Here, we explore the effects of two perturbations that leave sequence composition unchanged: sequence permutation and pH titration. Using atomistic Monte Carlo simulations in ABSINTH implicit solvent, we find that permutants of an arginine-rich IDP sequence exhibit nearly identical polymeric attributes, yet differ in the details of their conformational ensembles. In contrast, fluorescence correlation spectroscopy shows a decrease in translational diffusion times with increasing pH, suggesting that electrostatic and conformational modulation of protonation equilibria is significant even for solvent-exposed titratable groups on flexible protein backbones. The effect is sufficient to induce a collapse transition in a polyarginine polypeptide at a pH well below the pKa of the arginine sidechain. We conclude that the polymeric character of an IDP is largely determined by its sequence composition, but emphasize the importance of accurate constant pH simulation technology in investigating details of its conformational equilibria.

Supported by grants 5T32GM008802 from the NIH National Institute of General Medical Sciences and MCB 0718924 from the National Science Foundation.

1333-Pos

Development of New Predictors of Intrinsically Disordered Proteins and Residues

Bin Xue[1,2], Christopher J. Oldfield[1], Weilun Hsu[1], Vladimir Uversky[1,2], A. Keith Dunker[1].
[1]Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA, [2]Institute for Intrinsically Disordered Protein Research, Indiana University School of Medicine, Indianapolis, IN, USA.

The concepts of intrinsic disorder (ID) and intrinsically disordered proteins (IDP) are being increasingly accepted by the scientific community. Although without unique 3D structures under physiological conditions, IDPs play important roles in many crucial biological processes, such as signaling, recognition, and regulation. However, identification of ID residues and IDPs is still challenging. Experimental methods are both time consuming and expensive. Computational methods are frustrated by the modest accuracy, especially on boundary regions and on short disordered segments. In attempts to improve the prediction accuracy, we developed four new predictors: CDF-all, PONDR-FIT, SPA, and Chopper. All these predictors showed some improvements over previous methods. CDF-all and PONDR-FIT employed artificial neural networks to refine the prediction results of six individual predictors (PONDR-VLXT, PONDR-VL3, PONDR-VSL2, IUpred, FoldIndex, and Top-IDP). CDF-all predicts the disordered status of the entire sequence, while PONDR-FIT gives the disorder tendency of each residue. SPA was specifically designed for short peptides by creating sequence ensembles and taking the average of PONDR-VLXT predictions over the ensemble. Chopper cuts the query sequence into short segments and applies SPA on each short segment. The final prediction of Chopper is a refined average of SPA. Since the development of the first disorder predictor in 1997, there are currently more than 50 predictors. Although new predictors are continuously coming, the prediction accuracy seems to be approaching a ceiling. Our approach as demonstrated by these four methods is to develop new prediction strategies. CDF-all and PONDR-FIT integrate several individual predictors. SPA and Chopper enable focus on regions giving lower accuracies. These new approaches each provide additional improvement over previous predictors, but none of these new approaches significantly surpasses the ceiling, which is estimated to be near 85% accuracy for two state, structure/disorder, predictions.

1334-Pos

Structural Disorder Within Henipavirus Nucleoprotein and Phosphoprotein

Sonia Longhi, Johnny Habchi, Laurent Mamelli, Hervé Darbon.
AFMB, UMR 6098, CNRS and Universities Aix-Marseille I and II, Marseille, France.

Henipaviruses are newly emerged viruses within the Paramyxoviridae family. Their negative-strand RNA genome is packaged by the nucleoprotein (N) within a helical nucleocapsid that recruits the polymerase complex made of the L protein and the phosphoprotein (P). Using both computational and experimental approaches we herein show that Henipaviruses N and P proteins possess large intrinsically disordered regions. By combining several disorder prediction methods, we show that the N-terminal domain of P (PNT) and the C-terminal domain of N (NTAIL) are both mostly disordered, although they contain short order-prone segments. We then report the cloning, the bacterial expression, purification and characterization of Henipavirus PNT and NTAIL domains. By combining gel filtration, circular dichroism and nuclear magnetic resonance, we show that both NTAIL and PNT belong to the premolten globule sub-family within the class of intrinsically disordered proteins.

1335-Pos

Searching for the Native Molten Globules

Vladimir N. Uversky, A. Keith Dunker, Bin Xue, Christopher J. Oldfield.
Indiana University School of Medicine, Indianapolis, IN, USA.

Many biologically active proteins are either completely disordered or contain disordered regions of substantial size. These proteins are known as intrinsically disordered proteins, IDPs, among different names. The flexibility of these proteins and regions serves as the basis for their biological functions, where they are often involved in protein-protein interaction, regulation, recognition and signal transduction. These proteins are common in nature and are frequently associated with the pathogenesis of various human diseases Furthermore, intrinsic disorder-based protein-protein interactions, which are commonly accompanied by the disorder-to-order transitions, represent very attractive targets for novel drug development aiming at the specific inhibition of disease-associated protein-protein interactions. Therefore, it not surprising that IDPs have recently gained considerable attention. It is recognized now that ID comes in several flavors and that IDPs show an extremely wide diversity of their structural properties. The major pitfall of the current studies on IDPs in general and on the inhibition of the IDP-based protein-protein interactions in particular is that they are mostly focused on the extended IDPs almost completely ignoring a very substantial subset of IDPs, native molten globules. The goal now is to fill this gap and to identify and structurally and functionally characterize a set of native molten globules. The hypotheses that will be tested are the following: (i) native molten globules are very common in nature; (ii) native molten globules possess recognizable structural and functional properties and therefore can be found computationally and experimentally; (iii) native molten globules are frequently associated with human diseases.

1336-Pos

Exploring the Binding Diversity of Intrinsically Disordered Proteins

Wei-Lun Hsu.
Indiana University School of Medicine, Indianapolis, IN, USA.

Intrinsically disorder proteins often perform their functions through the binding of a short, loosely structured region from one protein onto the binding site of a protein partner. These short binding regions go by various names including eukaryotic linear motifs (ELMs), short linear motifs (SLiMs) and molecular recognition features (MoRFs). All of these represent a class of disordered protein that executes molecular recognition and binding functions typically via a disorder-to-order structural transition. Previous studies from our group showed 2 distinct examples of hub proteins, a disordered hub example (p53) and an ordered hub example (14-3-3), performing one-to-many signaling and many-to-one signaling, respectively. In the former, a single disordered region binds to multiple structured partners. In the latter, many distinct disordered regions bind to a single structured partner. Both alternatives use the MoRF mechanism described above. In this study, we tried to expand our previous work to find more examples and to determine whether the prevalence of the MoRF mechanism is common or rare through analyzing protein complexes deposited in PDB that consist of short nonglobular fragments bound to large globular partners. After examining and clustering the various MoRFs, we found 298 disordered hub examples and 246 ordered hub examples. Further experiments are providing detailed information about how intrinsic disorder facilitates binding to diverse partners. Exploring these examples is yielding a much clearer picture of the conformational changes that occur upon binding and showing that, in general, flexibility allows both subtle and complex structural variation thereby enabling different sequences to fit into the same binding site and the same sequence to morph into differently shaped binding sites. These results indicate that the observations made previously on the single examples of many-to-one and one-to-many signaling generalize to many hundreds of other well characterized protein-protein interactions.

1337-Pos

Energy Landscape Analysis Reveals Residual Order in Histone Tail Dynamics

Davit Potoyan, Garegin Papoian.
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

Histone tails are highly flexible N terminal protrusions of histone proteins which help to fold DNA into dense superstructures known as chromatin. On a molecular scale histone tails are polyelectrolites with high degree of conformational disorder, allowing them to function as bio-molecular ''switches", regulating various genetic regulatory processes. Because of being intrinsically disordered, the structural and dynamical aspects of histone tails are still poorly understood. In this work we have carried out 3 microsecond all atom replica exchange molecular dynamics (REMD) simulations of four histone tails, H4, H3, H2B and H2A, to probe for their intrinsic conformational preferences. Our subsequent energy landscape analysis demonstrated that some tails are not fully disordered, but contain residual secondary structure elements. In particular, H4 formed beta hairpins, H3 and H2B adopted alpha helical elements while H2A was fully disordered. We also carried out polymer physics based analysis of the histone tails' conformational ensembles. We found an intriguing re-entrant contraction-expansion of the tails upon heating, which is caused by the way mobile counterions associate with the protein chains at various temperatures.

1338-Pos

Connecting Unfolded Protein Dynamics and Aggregation

Lisa J. Lapidus[1], Yujie Chen[1], Vijay R. Singh[1], Steven A. Waldauer[1], Eric Buscarino[1], Vince A. Voelz[2], Vijay S. Pande[2].
[1]Michigan State University, East Lansing, MI, USA, [2]Stanford University, Stanford, CA, USA.

Solving the ''protein folding problem'' will require a better understanding of the allowed conformations and dynamics of the unfolded state. Fundamentally, unfolded protein dynamics should be determined by the hydrophobic pattern of each sequence. As proof of this hypothesis, we have constructed a random wormlike chain model that is re-weighted to favor random conformations that have residues of similar hydrophobicity in close proximity. This model gives remarkable quantitative agreement with measurements of intramolecular contact formation for various sequences. Furthermore we have found that the intramolecular diffusion coefficient of various sequences in folding conditions vary by ~ 3 orders of magnitude. We find such dynamics qualitatively agree with those predicted by molecular dynamics in implicit solvent. The fastest sequences are intrinsically disordered proteins and peptides and the slowest sequences are well-behaved proteins with folding times of at least 1 ms. In between these two regimes are sequences prone to aggregation in which intramolecular diffusion is just fast enough to expose hydrophobes to solvent long enough to form bimolecular interactions.

1339-Pos

Investigation of the Intrinsically Disordered Protein IA3 by Multiple SDSL-EPR Techniques

Natasha L. Pirman, Gail E. Fanucci.
University of Florida, Gainesville, FL, USA.

Intrinsically disordered proteins (IDPs) are proteins that contain little to no secondary or tertiary structure and are often functional proteins that are essential in biological systems. Many IDPs undergo a conformational change where the lack of intrinsic structure is relieved upon binding to its target protein. Due to the very nature of unstructured proteins, characterization of the conformational propensities and function of these proteins present a major challenge. Site-directed spin labeling (SDSL) coupled with electron paramagnetic resonance (EPR) spectroscopy is a valuable tool in characterizing the mobility and conformational changes of proteins. This combined technique, however, is not often used to investigate intrinsically disordered proteins (IDPs). IA3 is a 68 residue IDP that has been extensively characterized by various biophysical techniques and was used in this study as a model system to show SDSL-EPR may be employed to characterize conformational changes in IDPs. The TFE-induced disordered-to-α-helical transition of IA3 was monitored by various SDSL-EPR techniques. CW-EPR experiments were performed at X-, Q-, and W-band resonant frequencies and reveal conformational changes can be observed at all three frequencies, with the W-band spectra revealing the most striking changes in dynamics upon inducing α -helical secondary structure by TFE. Low temperature CW-EPR and DEER distance measurements were also performed to evaluate the ability of these methods in characterizing the induced α-helical conformation.

1340-Pos

Coupled Folding and Binding of pKID with KIX Domain Investigated by Multicanonical Molecular Dynamics Simulation in Explicit Solvent

Koji Umezawa, Jinzen Ikebe, Haruki Nakamura, Junichi Higo.
Osaka University, Suita, Osaka, Japan.

Intrinsically disordered proteins (IDPs) do not form any rigid tertiary structures alone. Most of them bind corresponding proteins and fold into an ordered structure to play important roles in biological functions such as signal transduction cascades. The phosphorylated kinase induced domain (pKID) is one of IDPs. The pKID adapts an ordered helical structure while binds to the KIX domain. This structural property of pKID is called ''coupled folding and binding''. From nuclear magnetic resonance (NMR) studies a kinetics model of the coupled folding and binding has been proposed: there are four states of pKID (disordered, encounter complex, intermediate, and ordered helical structure). However, the details of these states at atomic level were still unclear.

In order to obtain the free energy landscapes and the stable complex structures at various temperatures for the system, where pKID, the KIX domain, water and ions are included and they can interact with each other, multicanonical molecular dynamics (McMD) simulation was performed with the all-atom model in explicit solvent. McMD simulation is one of generalized ensemble methods, which can search conformational space much wider than a conventional molecular dynamics simulation, as well as the replica-exchange molecular dynamics (REMD). In this study, starting from completely disordered and unbound states of pKID, the helices of pKID were reproduced as approaching to the KIX domain at adequate low temperature. Additionally, we found a different binding mode from the NMR model. Interestingly, this binding mode is similar to that of the activation domain of the mixed lineage leukemia (MLL) transcription factor upon binding to the KIX domain.

1341-Pos

On the Origins of Intrinsically Disordered Proteins

Jintao Liu, Faeder R. James, **Carlos J. Camacho**.
University of Pittsburgh, Pittsburgh, PA, USA.

A large number of proteins are sufficiently unstable that their full three dimensional structure cannot be resolved. The origins of this intrinsic disorder are not well understood, but its ubiquitous presence undercuts the principle that a protein's structure determines its function. Here, we present a quantitative theory that makes novel predictions regarding the role of intrinsic disorder in protein structure and function. In particular, we discuss the implications of analytical solutions of a series of fundamental thermodynamic models of protein interactions in which disordered proteins are characterized by positive folding free energies. We validate our predictions by assigning protein function using the Gene Ontology classification, in which Protein Binding, Catalytic Activity and Transcription Regulator Activity are the three largest functional categories, and performing genome-wide surveys of both the amount of disorder in these functional classes and binding affinities for both prokaryotic and eukaryotic genomes. Specifically, without assuming any a priori structure-function relationship, the theory predicts that both Catalytic and low-affinity Binding ($K_d > 10^{-7}$ M) proteins prefer ordered structures, while only high-affinity Binding proteins (found mostly in eukaryotes) can tolerate disorder. Relevant to both Transcription and signal transduction, the theory also explains how increasing disorder can tune the binding affinity to maximize the specificity of promiscuous interactions. Collectively, these studies provide insight into how natural selection acts on folding stability to optimize protein function.

1342-Pos

Concerted Involvement of Long-Range Electrostatic Interactions and Fly-Catsing in Recognition of Intrinsically Disordered Proteins

Debabani Ganguly, Jianhan Chen.
Kansas State University, Manhattan, KS, USA.

Intrinsically disordered proteins (IDPs) are the functional proteins that exist as dynamical ensemble under physiological conditions. IDPs play crucial roles in cellular signaling and regulation, and often fold upon binding to specific targets. In particular, the ability to interact with multiple targets appears to be a hallmark of IDPs. Two ideal mechanisms, namely fly-casting/induced folding and conformational selection, are possible for coupled folding and binding to the specific targets. Evidence recently accumulated to suggest that induced folding might be prevalent in IDP recognition. Our recent analysis reveals that both IDPs and the vicinity of their binding sites on the substrate surface are enriched with charges. Our hypothesis is the long-range electrostatic interactions between these charged residues triggers the promotion of the unfolding of residual structure in unbound IDPs and enhance the binding efficiency via the fly-casting effects. Go-like coarse-grained protein model has been used to investigate the interaction mechanisms of IDP complex pKID:KIX. Initial results appear to support the concerted involvement of long-range electrostatic interactions and fly-casting.