SOLVATION THERMODYNAMICS OF THE PROTEIN BACKBONE: IMPLICATIONS FOR COLLAPSE AND AGGREGATION

Justin Drake, University of Texas Medical Branch
2015-2016 Graduate Fellow

RESEARCH SUMMARY

Intrinsically disordered regions (IDRs) in proteins assume a diverse ensemble of highly dynamic, flexible structures. The development of drugs to treat IDR-mediated diseases requires extensive knowledge of the thermodynamic and structural properties of these disordered regions. Here, we wish to understand if the thermodynamic mechanisms that drive the proper folding of well-structured proteins similarly promote the collapse of IDRs to their native state as well as the aggregation of IDRs in a pathological state. Using molecular simulations and computational free energy methods, we calculated the solvation free energy, enthalpy, and entropy of oligoglycine, a protein backbone model, as a function of chain length. We find that the entropic penalty upon solvating successively longer oligoglycines does not explain the experimentally observed aggregation of short oligoglycines and collapse of longer ones. Rather, we propose that favorable peptide-peptide interactions outcompete favorable peptide-solvent interactions in a concentration and length-dependent manner.

For much of the 20th century, it was widely held that the well-defined three-dimensional structure of a protein dictates its cellular function. While true for a large group of proteins, at the turn of the century, seminal papers [1, 2] opened the door to an entirely new class of proteins that rely on highly dynamic, flexible (i.e. disordered) regions to carry out their function. These IDRs found within proteins play a major role in the protein signaling networks and regulation, and, unsurprisingly, have been implicated in many diseases [2]. If we are to be as successful in targeting drugs to IDRs as we are in targeting well-structured proteins, we need to have an extensive, complete understanding of the structural properties of IDRs and how they enable protein function. IDRs assume a diverse ensemble of structures that range from completely collapsed to highly extended depending, in part, on their sequence composition and preference to interact with the solvent as well as their length or number of amino acids [3]. However, the biophysical mechanisms that drive the collapse or extension of IDRs are not well understood.

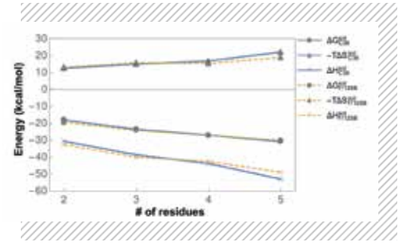

FIGURE 1: Solvation free energy (circle), enthalpy (x), and entropy (triangle) of successively longer oligoglycines calculated using free energy perturbation with the C36 (blue) and ff12SB (orange) force fields.

To uncover these mechanisms, we study the solvation thermodynamics of successively longer oligoglycine models ($Gly_{2\text{-}5}$), where the subscript denotes the number of glycine residues. Oligoglycine is an ideal protein backbone model [4] and is also found in many IDRs [5]. We use computational free energy methods to calculate the solvation free energy (ΔG^{sol}), entropy (ΔS^{sol}), and enthalpy (ΔH^{sol}) of $Gly_{2\text{-}5}$ with the commonly used CHARMM36 (C36) [3] and Amber ff12SB [7] force fields. We find that ΔG^{sol} is negative (i.e. favorable) and continues to decrease with the length of oligoglycine. The decrease in ΔG^{sol} with chain length is driven by a large, negative ΔHsol that is only moderately counterbalanced by an unfavorable entropic component, $-T\Delta S^{sol}$ (Fig. 1). Interestingly, experiments and simulations show that short oligoglycines (~Gly_5) become increasingly insoluble (aggregate) with chain length and long oligoglycines (>Gly_{10}) collapse in highly dilute solutions [4, 8, 9], but the mechanism that drives the aggregation of short oligoglycines and collapse is unclear. However, our results and those of others [3, 6, 9] suggest that favorable, non-specific peptide-peptide interactions outcompete the still favorable peptide-solvent interactions and that these weak peptide-peptide interactions are amplified in a length and concentration dependent manner. Whereas the initial collapse and folding of well-structured proteins are widely accepted as an entropically driven process, it may not play as dominant a role in the collapse of IDRs. Lastly, we find that despite C36 and Amber ff12SB generating different structural ensembles of various oligoglycine chains [10], both yield remarkably similar solvation thermodynamic profiles with respect to chain length (Fig. 1). Attempting to relate a thermodynamic mechanism to the structural properties of highly disordered polypeptides predicted by different force fields may be problematic.

WHY BLUE WATERS

To calculate solvation free energy, enthalpy, and entropy of our various oligoglycine models, we used a type of computational free energy method, termed "alchemical free energy perturbation (FEP)," that relies on molecular dynamics sampling in explicit solvent. While considered one of the most accurate methods to calculate free energy it is computationally demanding yet reasonably parallelizable. Our project required a total of 1,784 simulations for an aggregate simulated time of 92.3 microseconds. The high-throughput and ability to run a large number of simulations concurrently on Blue Waters enabled us to tackle such a large-scale project in a tractable amount of time.

PUBLICATIONS AND DATA SETS

Drake, J.A., and B.M. Pettitt, Solvation thermodynamics of oligoglycine with respect to chain length and flexibility: implication for aggregation and collapse. *Biophysical Journal*, (2016), dx.doi.org/10.1016/j.bpj.2016.07.13

Drake, J.A., and B.M., Pettitt, The protein backbone: how structural and thermodynamic properties scale with length. *251st American Chemical Society National Meeting and Exposition* (Philadelphia, PA, March 2016).

Drake, J.A., and B.M. Pettitt, The protein backbone: how structural and thermodynamic properties scale with length. *Sealy Center for Structural Biology and Molecular Biophysics 21st Annual Structural Biology Symposium* (Galveston, TX, April 2016).

Drake, J.A., and B.M. Pettitt, Protein backbone thermodynamics: Towards estimating conformational entropy. *Sealy Center for Structural Biology and Molecular Biophysics 21st Annual Structural Biology Symposium* (Galveston, TX, May 2015).

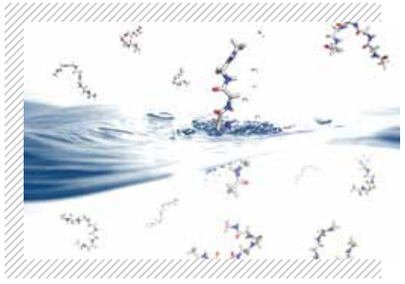

FIGURE 2: Artistic depiction of the physical process of the transfer of oligoglycine from the gas phase to aqueous solution (i.e. solvation thermodynamics).

Justin Drake is in his fifth year as a Ph.D. student in Biochemistry and Molecular Biology at University of Texas Medical Branch. After graduation, he would like to continue to use high-performance computing (HPC) to investigate complex biological systems either in academia or industry.

"Specifically, I am interested in continuing my research into the biophysical mechanisms that allow structural disorder to persist in proteins and how disease mutations alter the biophysical properties of disordered regions," says Drake. "The Blue Waters Graduate Fellowship has not only played a pivotal role in helping me achieve my doctorate but it has also exposed me to a variety of fields and topics related to HPC that I would have otherwise not experienced as part of my degree plan. I believe this breadth of knowledge will give me a competitive edge as I continue to the next stages of my career. In the future I also hope to further bridge the gap between biology and computational sciences and advocate how both fields can influence each other's further progression."