Chapter 7

Discussion

The importance of the hydrophobic effect in protein folding and aggregation has been textbook knowledge for decades. For example, the undergraduate textbook "Introduction to protein structure" (Branden and Tooze, 1998), starts with the observation that the interior of proteins is hydrophobic. This observation is than followed by the statement: "The main driving force for folding water-soluble globular protein molecules is to pack hydrophobic side chains into the interior of the molecule, thus creating a **hydrophobic core** and a hydrophillic surface".

Moreover, Spolar et al. showed that important thermodynamic variables, the entropy and enthalpy of folding, and the related experimental observable, the heat capacity of the system, were determined by the interactions of the hydrophobic amino acids with the water (Spolar, Ha, and Record, 1989). Nevertheless, quantifying the role of the *temperature dependence* of the hydrophobic effect in protein folding, and in particular its contribution to cold denaturation and to the temperature dependence of the heat capacity, were studied in less detail.

In contrast to protein folding, little was known on the role played by hydrophobic contributions to aggregating proteins at the start of this project. Experimental results and theory showed that transforming a protein to a fibril occurred in three stages: oligomerization, nucleation, and elongation (Knowles et al., 2009). Knowles et al. also showed that different sequences behaved differently, and that the mechanism of aggregation could be determined by fitting a kinetic model to the experimentally observed rate of aggregation. This approach has gained wider adoption and has now been published as a method (Meisl et al., 2016). This approach does not provide insight on the molecular pathway from a folded protein to an aggregated fibril, nor on the final structure of the amyloid fibril. The leading hypothesis during this period was was that the structures of the final, fibrillar state consisted of twisted β-sheets. This hypothesis was strengthened by the fact that coarse grained computational models could reproduce the aggregated state, and the elongation process that allowed aggregates to grow. However, no models capture the trade-off between folding, oligomer formation, and elongation that exists in real proteins.

In this thesis, we have shown that cold denaturation, the heat capacity and the structure of the cold denatured state can be explained by incorporating the

temperature dependence of the hydrophobic effect. The coarse grained nature of our model allows us to find general principles and statistical correlations that can be applied to real proteins. We made a result from our model, prediction of the heat capacity of a given folded protein from the hydrophobic surface area, publicly accessible through a web server. Ultimately one would like to use simulations to make accurate, physics-based predictions. Two publications published concurrently with our work show that the general principles we found are also applicable to make quantitative predictions for the stability of real proteins (Sirovetz, Schafer, and Wolynes, 2015; Vajpai et al., 2013). I expect that in the near future these methods will be applied to more proteins to validate the results further. Additionally, these methods can likely be applied to predict protein stability as a function of temperature.

While much progress has been made during this project on the understanding of protein aggregation, there are still a lot of unknowns. In our computational model we used the commonly accepted hypothesis that amyloid fibrils consist of twisted β-sheets. This year a structure was found for α-synuclein that was consistent with this hypothesis. During my PhD, a paper describing a model that could model the trade off between folding, oligomer formation and aggregation was published by (Ni et al., 2015). We have used their model in combination with the temperature dependent hydrophobic effect described in chapter 3 to investigate if aggregation was enthalpically or entropically driven. We also investigated the role of the temperature dependent hydrophobic effect in the temperature dependence of the enthalpy and the entropy. Unlike in protein folding, our results agreed only partially with published experimental work. While our model did reproduce cold denaturation of aggregates, a phenomenon that was first observed during my PhD, we found that the slope of the enthalpy of aggregation with regard to temperature was always smaller than zero, irregardless of the parameters we choose. Unpublished data from an experimental collaborator showed that the experimental result we were comparing to could not be reproduced in his lab. This is an interesting result that will be investigated further in a follow-up project.

While quantitative predictions for the stability of single protein structures are constantly being evaluated and improved, for example through the CASP project, the predictions made so far for aggregating proteins are mostly qualitative. An important reason for this contrast is that the computational resources required to simulate such a big system in atomic detail are enormous. This problem previously existed for protein folding and was solved by the increase in computational power due to Moore's law, which states: the number of transistors in a dense integrated circuit doubles approximately every two years. While Moore's law still holds approximately, it seems to no longer translate directly to increased processing power. A breakthrough in algorithms, model abstraction, or in computer hardware (and possibly two or more of the above) is required to allow us to accurately predict the conditions under which a sequence will aggregate.

I hope that this thesis provides a starting point towards the goals outlined

above. To allow our work to be continued, we have attempted to make our research reproducible: We made the programs we developed open source and included the data we used to generate our results. Moreover, we also made the work that is directly applicable to experiments accessible through a web server.