

Is Side-Chain Entropy an Important Factor in Protein Folding?

Joshua Ballanco

November 10, 2008

Abstract

Protein folding is a process which involves thermodynamic contributions from forces which favor the folded state and other forces which favor the unfolded state. When the balance of these forces tips in favor of the folded state, a protein will fold spontaneously into a unique conformation representing a thermodynamic minimum. The difficulty in predicting just how, and when, a protein will fold stems from a fundamental lack of knowledge about the forces involved. This is especially true in the case of the conformational entropy of amino acid side-chains (side-chain entropy).

Because of the difficulty associated with either directly measuring side-chain entropy or calculating its value from molecular simulations, there is uncertainty as to the effect it has on protein folding. The side-chains of amino acids range in size and flexibility, and some of the longer side-chains contain polar or even charged groups. It's been proposed that side-chains might counteract a loss of entropy by a gaining interactions with other amino acids. Additionally, the hydrophobic residues commonly found in the core of a folded protein are not the most flexible, and so they will not be affected as significantly by the corresponding space constraints.

On the other hand, buried residues will loose some degree of side-chain entropy. Even though the cores of proteins rarely represent a closest-packed configuration, the side chains of long, charged residues found in internal salt bridges will undoubtedly be restricted in their motion. Furthermore, it is now understood that what is commonly referred to as the "native protein structure" is actually an ensemble of closely related structures. It is possible that side-chain entropy may not be important in determining the global fold of a protein, but may be vital in understanding the dynamics associated with function.

In this paper I will look at a variety of experiments which approach the question of side-chain entropy from differing perspectives. The first series of approaches involves investigating the structure and sequence of natural proteins, specifically focusing on side-chain entropy. The alternative approach is to perform *de novo* protein design either using or neglecting side-chain entropy in the calculations. Both of these approaches are, necessarily, indirect, and each leads to a different conclusion. I will look at the results in context, and discuss whether or not a real controversy exists. Finally, I will propose a course of investigation which might shed more light on the role of side-chain entropy in protein folding.

Topics

Thermodynamics

Protein Folding

Statistical Mechanics

Molecular Dynamics

X-ray Crystallography

de novo Protein Design

Introduction

Proteins are the work-horses of biology. They are responsible for the vast majority of the biochemical processes that take place in living organisms. It should come as no surprise, then, that there is keen interest in understanding how they function. Normally, this would be a question of brute-force experimentation, approaching and investigating proteins one at a time. Two aspects of proteins, however, make this a problem that is tantalizingly close, and yet still frustratingly far, from a complete solution.

The first useful aspect of proteins is the relationship between their structure and their function. Knowing a protein's structure goes a long way to understanding how it might function. This is particularly important when developing pharmaceutical agents to treat disease. Second, a protein's structure is completely determined by its amino acid sequence, which can in turn be determined from available genetic data. This was originally shown in 1961 by Anfinsen [3] in his work with ribonuclease. This notion, of 3-dimensional structure data encoded in a 1-dimensional array of elements, has been verified for proteins in all but a few rare cases.

Thus, since knowing a protein's structure was already more than half the battle, and because it should be possible to determine that structure with readily available information, work began almost immediately on elucidating the mechanism drives proteins to form their unique structures. This work has been ongoing for the 47-years since.

Shortly after Anfinsen's discovery, Brandts proposed a two-state model for protein folding where a protein exists in an equilibrium, $D \rightleftharpoons N$, between the denatured and native state [1]. This two-state model of protein folding is still the prevailing accepted model for folding, and it has some interesting implications. One of the first implications was recognized by Levinthal in 1968, and has since been known as Levinthal's paradox.

Essentially, what Levinthal realized was that Brandts' two-state model implied an absence of long-lived or marginally stable conformations. In other words, the denatured protein, which is essentially a random coil, must find the single lowest energy conformation from all of the possible conformations in one step. Levinthal's rough calculations implied that a robust search of all possible conformations would require more than the age of the universe to complete yet, paradoxically, proteins fold on a millisecond to second time scale. Obviously, proteins do not randomly sample conformation space. Instead, there must be a driving force that guides a denatured protein, no matter its conformation, toward the native state.

This understanding of Levinthal's Paradox lead Dill to propose the concept of a folding funnel [2], which has become an iconic symbol of the protein folding problem.

- Leventhal's paradox
- Dill's Funnel
- Difficulty in calculating Entropy from MD/X-ray structure

Side-Chain Entropy of X-ray Structures

One way of approaching the question of side-chain entropy is to look at a collection of existing protein structures.

- Review logic
- Outline methods
- Summarize conclusions

Side-Chain Entropy in Protein Design

Another approach to the question of side-chain entropy, and one that has only become feasible with recent advances, is to look at its role in the *de novo* design of proteins.

- Review logic
- Background on method
- Summarize conclusions
- Compare to new paper with contradictory results

Is Side-Chain Entropy Important?

What can we conclude about the importance of side-chain entropy in determining the folded structure of a protein?

- Design vs Structure
- Not yes or no, but how much?
- Both important and unimportant simultaneously

How Important Is Side-Chain Entropy?

With the possibility that side-chain entropy may be important for some aspects of protein folding and not for others, it is vital that we understand better the roles that side-chain entropy plays.

- Mechanical experiment for validation
- New techniques for simulating proteins

References

- [1] A Clay Clark. Protein folding: are we there yet? *Arch Biochem Biophys*, 469(1):1–3, Jan 2008.
- [2] Ken A Dill, S Banu Ozkan, M Scott Shell, and Thomas R Weikl. The protein folding problem. *Annual review of biophysics*, 37:289–316, Jan 2008.
- [3] Robert F Service. Problem solved* (*sort of). *Science*, 321(5890):784–6, Aug 2008.