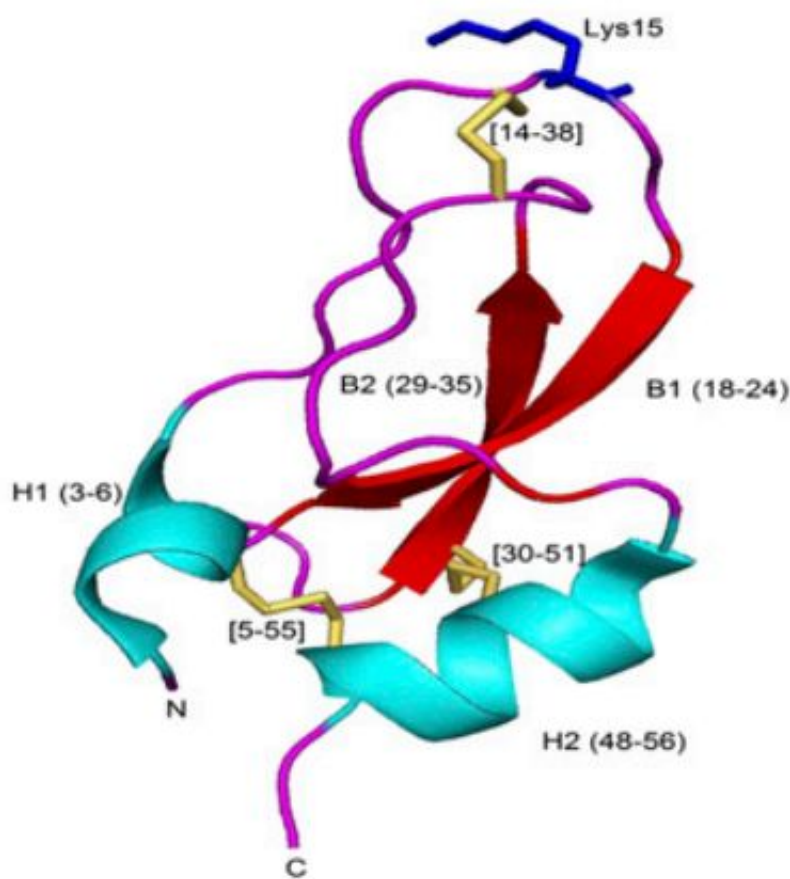# Simulation studies reveal the role disulfide bonds play in protein folding

August 17 2015, by Heather Zeiger

(A) Ribbon diagram of the native structure of the 58 residue BPTI containing three disulfide bonds (marked in yellow) between residues Cys5 and Cys55 [5–55], Cys14 and Cys38 [14–38], and Cys30 and Cys51 [30–51], respectively. The antiparallel β-sheet is in red. Credit: (c) 2015 *PNAS*, DOI: 10.1073/pnas.1503909112

(Phys.org)—Diseases like mad cow disease or Alzheimer's are a result of protein mis-folding. Proteins are comprised of an amino acid chain that folds into a three-dimensional structure. According to the Anfinsen Principle, the chemical character of the amino acid chain dictates how the protein folds into its three-dimensional structure. A key component to this folding is the formation of disulfide linkages.

Experimental interpretations have gone back-and-forth as to whether disulfide bond formation drives [protein folding](#) or if protein folding causes disulfide bond formation. Knowing the role disulfide bond formation plays in protein folding would be helpful in understanding when a protein mis-fold happens. Meng Qin, Wei Wang, and D. Thirumalai of the University of Nanjing and the University of Maryland have designed a novel computational method that mimics disulfide bond formation and subsequent protein folding in a known protein. They found that the protein conformation directs disulfide bond formation. Their work appears in the *Proceedings of the National Academy of Sciences*.

The protein's primary structure is the linear [amino acid chain](#). Among the possible amino acids are cysteine residues which contain a free thiol in its side chain. In the protein's final, or native, three-dimensional conformation, also known as its tertiary structure, some of these thiol groups are oxidized to form disulfide bonds. Reduction of these bonds back to thiols is one way that many researchers denature proteins. This means that the disulfide bonds are important to the protein tertiary structure; although exactly how they promote the formation of the tertiary structure is unknown.

Qin, et al. used coarse-grained molecular simulations to determine the role of disulfide bonds in the folding of bovine pancreatic trypsin inhibitor (BPTI), a protein known to have three disulfide bonds in its native state. When these disulfide bonds are reduced, BPTI unfolds into

its primary structure. They used these simulations to quantitatively elucidate the structural basis of the folding pathway of BPTI.

The simulation takes into account key criteria that have been shown to be important for disulfide bond formation, namely proximity of sulfurs to each other and access to oxidizing agents. Qin, et al. then simulated the process of protein folding, investigating important intermediate steps in which one or two disulfide bonds are formed or broken during the process.

In its native state, BPTI's three disulfide bonds occur between Cys5 and Cys55, Cys14 and Cys38, and Cys30 and Cys51. The simulation studies, which corresponded to experimental results, showed that the Cys14-Cys38 bond forms quickly, forming a transient state in which the protein is compact but not yet in its native structure. The Cys14-Cys38 bond breaks when the structure rearranges to form the Cys5-Cys55 and Cys30-Cys51 bonds.

Qin, et al. then tested several factors including the effect of changing the redox conditions. Changing the redox conditions did not change the initial Cys14-Cys38 bond formation, which means that the driving force behind this bond formation is due to proximity of the cysteine residues, not environmental conditions. These residues are located on a β-hairpin loop, and the initial compaction of BPTI allows for the formation of the loop, putting the residues within bond distance. Importantly, additional studies confirmed that it is the formation of the loop that results in the disulfide bond, not the other way around.

Kinetic studies showed that the other disulfide bonds do not form until after the Cys14-Cys38 bond forms, then rearranges. The Cys14-Cys38 bond is a less stable bond because of its exposure to the oxidative solvent. The Cys5-Cys55 and Cys30-Cys51 bonds are more stable due to their location away from the solvent and once they form, they remain

intact.

The key finding in this study is the importance of the initial formation of the β-hairpin loop and terminal alpha helices that are in the final conformation, followed by the initial Cys14-Cys38 disulfide bond, which is then followed by the collapse of the polypeptide chain. Quantitative studies show that native state formation is reliant upon the initial β-hairpin formation.

This simulation also allowed for studies on how nonnative states affect protein folding. In BPTI, nonnative transition states do occur, particularly after the initial Cys14-Cys38 bond forms and the peptide chain collapses, but these states are randomly created and do not direct protein folding. Instead it is the formation of native interactions that direct protein folding.

This study provides a general framework for understanding a couple of aspects of protein folding. First, protein folding should be described by probable routes as opposed to *one* particular route. Secondly, the link between the initial protein collapse and folding is an important component to protein kinetics and yield. Finally, it is the formation of native conformations that drive protein formation and disulfide bond formation.

Abstract
The Anfinsen principle that the protein sequence uniquely determines its structure is based on experiments on oxidative refolding of a protein with disulfide bonds. Although such proteins are involved in diverse extracellular and intracellular functions, the problem of how protein

folding drives disulfide bond formation is poorly understood. Here, we have solved this long-standing problem by creating a general method for implementing the chemistry of disulfide bond formation and rupture in coarse-grained molecular simulations. As a case study, we investigate the oxidative folding of bovine pancreatic trypsin inhibitor (BPTI). After confirming the experimental findings that the multiple routes to the folded state contain a network of states dominated by native disulfides, we show that the entropically unfavorable native single disulfide [14–38] between Cys14 and Cys38 forms only after polypeptide chain collapse and complete structuring of the central core of the protein containing an antiparallel β-sheet. Subsequent assembly, resulting in native two-disulfide bonds and the folded state, involves substantial unfolding of the protein and transient population of nonnative structures. The rate of (14–38) formation increases as the β-sheet stability increases. The flux to the native state, through a network of kinetically connected native-like intermediates, changes dramatically by altering the redox conditions. Disulfide bond formation between Cys residues not present in the native state are relevant only on the time scale of collapse of BPTI. The finding that formation of specific collapsed native-like structures guide efficient folding is applicable to a broad class of single-domain proteins, including enzyme-catalyzed disulfide proteins.

© 2015 Phys.org