

Constraint network analysis: Exploiting the link between protein rigidity and thermostability

Sebastian Radestock, Holger Gohlke

Mathematisch-Naturwissenschaftliche Fakultät, Pharmazeutisches Institut, Christian-Albrechts-Universität, Kiel, Germany

The molecular basis of stability relates closely to contemporary issues in protein science such as the protein folding problem, protein-protein interaction and protein-ligand binding. In addition, protein stability has industrial importance. The identification or the development of enzymes with higher stability will increase the adoption of biocatalytic syntheses in industrial production.^[1] Understanding and exploiting the relationship between microscopic structure and macroscopic stability is important for developing strategies to improve protein stability in the reaction media used in industrial processes.

Thermostability of proteins has been repeatedly linked to an enhanced structural rigidity of the folded native state. Here, we directly probe the rigidity of protein structures from mesophilic and thermophilic organisms along a thermal unfolding trajectory. For this, protein structures were modeled as constraint networks, and the rigidity in these networks was quantified using the Floppy Inclusion and Rigid Substructure Topography (FIRST) method.^[2] By the dilution of non-covalent contacts in the network, FIRST has been employed to simulate thermal unfolding.^[3] In going from a rigid to a flexible network, a phase transition can be observed that defines the rigidity percolation threshold and corresponds to the folded-unfolded transition in protein unfolding. Here, thermal unfolding simulations were applied to a dataset of 20 homologous proteins from thermophilic and mesophilic organisms.

Using concepts from percolation theory and network science, a higher phase transition temperature was observed for approximately two-thirds of the proteins from thermophilic organisms compared to their mesophilic counterparts. Direct support was found for the “corresponding states” concept, which states that mesophilic and thermophilic enzymes are in corresponding states of similar flexibility at their respective optimal temperature. Our approach allowed for identifying structural features from which a destabilization of the structure originates upon thermal unfolding. These predictions show a good agreement with experiment. The information might thus be exploited in data-driven protein engineering by pointing to residues that should be varied to obtain a protein with higher stability.

[1] D. C. Demirjian, F. Moris-Varas, C. S. Cassidy, *Curr. Opin. Chem. Biol.* **2001**, *5*, 144-151.

[2] D. J. Jacobs, A. J. Rader, L. A. Kuhn, M. F. Thorpe, *Proteins* **2001**, *44*, 150-165.

[3] A. J. Rader, B. M. Hespeneide, L. A. Kuhn, M. F. Thorpe, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 3540-3545.