

The Evolutionary “Design” of Proteins

Rama Ranganathan

The Green Center for Systems Biology, UT Southwestern Medical Center, Dallas TX, 75390.

Natural proteins can fold spontaneously into well-defined three-dimensional structures, and can display complex biochemical properties such as signal transmission, efficient catalysis of chemical reactions, specificity in molecular recognition, and allosteric conformational change. All of this is achieved while also preserving the capacity for rapid adaptive variation in response to fluctuating selection pressures, a central feature of evolving systems. What are the basic principles in the “design” of natural proteins that underlie all of these properties? To address this, we developed an approach (the statistical coupling analysis or SCA) for globally estimating the pattern of functional interactions between sites on proteins through statistical analysis of the evolutionary divergence of a protein family^{1,2}. This analysis indicates a novel decomposition of proteins into sparse groups of co-evolving amino acids that we term “protein sectors”⁹. The sectors comprise physically connected networks in the tertiary structure and can be modular – with different sectors representing different functional properties. Experiments in several protein systems demonstrate the functional and adaptive importance of the sectors^{1,3,4,7,8,10,11,12} and recently, the SCA information was shown to be necessary and sufficient to design functional artificial members of two protein families in the absence of any structural or chemical information^{5,6}. These results support the hypothesis that sectors represent the basic architecture underlying folding, function, and adaptive variation in proteins. We are now working on two key problems: (1) understanding the physical mechanisms underlying sectors, and (2) defining how the dynamics of the evolutionary process controls the emergence and structural architecture of sectors in proteins.

[1] Lockless, R. Ranganathan, *Science*, **286**, 295-9 (1999).

[2] Suel et al., *Nature Struct. Biol.*, **10**, 59-69 (2003).

[3] Hatley, et al., *PNAS*, **100**: 14445-14450 (2003).

[4] Shulman et al., *Cell*, **116**: 417-429 (2004).

[5] Socolich et al., *Nature*, 437: 512-518 (2005).

[6] Russ et al., *Nature*, 437: 579-583 (2005).

[7] Mishra et al., *Cell*, 131: 80-92 (2007).

[8] Lee et al., *Science* 322: 438-442 (2008).

[9] Halabi et al., *Cell* 138: 774-785 (2009).

[10] Smock et al., *Mol. Sys. Biol.*: 6: 414 (2010).

[11] Reynolds et al., *Cell*, 147: 1564-1575 (2011).

[12] McLaughlin et al., *Nature* 491: 138-142 (2012).