Proteins can fold spontaneously into well-defined three-dimensional structures and can carry out complex biochemical reactions such as molecular recognition, catalysis, and allosteric communication. The precision required for these properties is somehow achieved while also preserving evolvability – the capacity for adaptive variation in response to ever-changing selection pressures. How are proteins built in Nature to support all of these properties? To address this question, we have taken a statistical genomics approach – deducing the pattern of constraints on amino acid residues in proteins through analysis of amino acid coevolution a protein family$^{1,2,14}$. This approach reveals a novel decomposition of proteins into sparse groups of co-evolving amino acids termed “protein sectors”$^9$. The sectors comprise physically connected networks in the tertiary structure that often connect functional surfaces and can be modular – with different sectors in a single protein delivering different functional properties. Experiments demonstrate the connection of sectors to protein function$^{1,2,5-11}$ and importantly, this pattern was shown to be necessary and sufficient to design functional artificial proteins without the use of any direct structural or chemical information$^{3,4}$. These results suggest that sectors are the conserved units of folding, function, and adaptability in natural proteins. We are now working on two key problems: (1) understanding the physics of sectors$^{12}$, and (2) defining how the dynamics of the evolutionary process controls the emergence of this structural architecture in proteins$^{11}$.

This event is co-sponsored by the NSF-Simons Center for Quantitative Biology Scholars and Fellows.

*This is a required activity