Abstract

To facilitate the design of protein sequences with desired properties, simulation techniques have been developed to allow large portions of amino acid sequence space to be evaluated by computer. These computational protein design methods apply optimization algorithms to sort through the enormity of sequence space and find desirable variants.

Simple modifications to the stochastic optimization algorithm FASTER enhanced its performance by two orders of magnitude without loss of accuracy, and rendered it more efficient than its major competitor by a factor of 10. These improvements allowed higher-quality amino acid solutions to be found more quickly, and accelerated the pace at which users could perform cycles of design and model adjustment.

This success prompted research into techniques for a protein design formulation that allows simulation in the context of multiple states simultaneously. This multi-state design can be used to wield explicit control over structural, binding, or catalytic specificity, and changes the scope of design goals that can be addressed by computation. Evaluation of multi-state FASTER indicated that it performed radically better than its major competitor in a variety of design contexts, and that in most cases it found solutions better than those that could ever be found using a lesser method.

Multi-state optimization using FASTER was applied to test the influence of various types of input structural data on the design of a small protein. To facilitate this evaluation, methods for the design and high-throughput stability screening of combinatorial libraries were developed. Screening of libraries based on single structures and structural ensembles indicated the success of multi-state modeling. Our results also

suggested that the exhaustive screening of designed libraries can help to elucidate the origins of design model failures. Finally, they showed that success of a design procedure does not hinge on its ability to correlate experimental and simulated measures of fitness, and prompted greater consideration of design methods that target explicitly conformational specificity.