MAR16-2016-020458

Abstract for an Invited Paper
for the MAR16 Meeting of
the American Physical Society

Molecular and cellular constraints on proteins

TANJA KORTEMME, UCSF

Engineering proteins with new sequences, structures and functions has many exciting practical applications, and provides new ways to dissect design principles for function. Recent successes in computational protein design provide a cause for optimism. Yet many functions are currently too complex to engineer predictively, and successful design of new biological activities also requires an understanding of the functional pressures acting on proteins in the context of cells and organisms. I will present two vignettes describing our progress with dissecting both molecular and cellular constraints on protein function. In the first, we characterized the cost and benefit of protein production upon sequence perturbations in a classic system for gene regulation, the *lac* operon. Our results were unexpected in light of the common assumption that the dominant fitness costs are due to protein *expression*. Instead, we discovered a direct linear relationship between cost and *lac* permease *activity*, not protein or mRNA production. The magnitude of the cost of permease activity, relative to protein production, has consequences for regulation. Our model predicts an advantage of direct regulation of protein *activity* (not just expression), providing a new explanation for the long-known mechanism of "inducer exclusion" that inhibits transport through the permease. Similar pressures and cost/benefit tradeoffs may be key to engineering synthetic systems with improved fitness. In the second vignette, I will describe our recent efforts to develop computational approaches that predict protein sequences consistent with multiple functional conformations. We expect such "multi-constraint" models to improve predictions of functional sequences determined by deep mutational scanning in bacteria, to provide insights into how the balance between functional conformations shapes sequence space, and to highlight molecular and cellular constraints that cannot be captured by the model.