

Proteome allocation determines thermosensitivity of growth and structure of the evolutionary landscape

Ke Chen¹, and Bernhard O. Palsson²

Short Abstract — We reconstruct an integrated genome-scale protein-folding network for *Escherichia coli* termed FoldME. FoldME simulations delineate the multi-scale strategies cells use to resist unfolding stresses induced by high temperature or destabilizing mutations in a single gene. Furthermore, FoldME enables sampling of the solution space with diverse genotypes, thus reveals the intrinsic structure of the fitness landscape constrained by the cost of expression for the large energy production protein complexes. The results provide a system-level understanding of the regulatory relationship between global proteome allocation and bacterial stress response. The method is readily extended to study cell's complex responses to multiple stresses simultaneously.

Keywords — Proteome Allocation, Bacterial Growth Law, Evolutionary Landscape, Proteostasis, Thermoadaptation, Genome-scale Model.

Gene expression is intimately coupled to the growth physiology of the cell, and regulated by global allocation of the cellular resource and energy. Such relationship has been quantitatively captured by the empirical bacterial growth law [1], which successfully explain how bacterial growth is affected by a wide range of biological processes, including molecular crowding, protein overexpression, cAMP-signaling, overflow metabolism [2], and growth transition kinetics. However, these coarse-grained models lack the necessary details to address the underlying molecular mechanisms that drive regulation. It is not clear how to combine different regulatory mechanisms into one composite model, describe the behavior of more complex systems under diverse environmental perturbations, and provide insights into the structure of fitness landscape instead of simple linear physiological correlations. To address these challenges, an integrated genome-scale model is calling to connect our understanding of molecular stress response mechanisms with phenotypic adaptation.

We started by modeling bacterial thermoadaptation [3-5], because temperature is one of the most important environmental parameters that dictate the evolution. First, we calculated the temperature-dependent biophysical profile of the proteome using thermodynamic principles based on protein sequences and structures [6]. This profile serves as an internal “sensor” to reflect the environmental and genetic

perturbations. Second, we design a mathematical formulation to describe how the molecular chaperones respond to the folding request of each protein [7]. Third, we integrate the kinetic protein folding network into the genome-scale reconstruction of metabolism and protein expression for *Escherichia coli* [8], by enabling competition among the spontaneous and multiple chaperone-assisted folding pathways. As such, changes in the proteostatic state of the cell induced by environmental and genetic perturbations can be calculated based on first principles, evaluated by the protein quality-control machinery, and coupled to the whole cell's economics.

FoldME simulations reproduce the asymmetrical temperature response of *E. coli*, and the proteomic changes upon destabilizing mutation in a single gene [3]. The results highlight the system-level regulatory role of chaperones beyond efficient folding of any single protein. Rather, chaperones participate in the global proteome reallocation to balance between the need for folding and the complex machinery synthesizing the proteins in response to the perturbations. The ability of FoldME to delineate multi-level cellular responses to a variety of perturbations encouraged further sampling of the solution space using a large number of random genotypes. Preliminary results reveal a rugged fitness landscape defined by discreteness of ATP production strategies [9]. Overall, these results expand our view of cellular regulation, from targeted specific control mechanisms to global regulation through a web of nonspecific competing interactions that modulate the optimal reallocation of cellular resources. The methodology developed enables genome-scale integration of environment-dependent protein properties and a proteome-wide study of cellular stress responses.

REFERENCES

- [1] Scott M, et al. (2010) Interdependence of cell growth and gene expression: Origins and consequences. *Science* **330**, 1099-1102.
- [2] Basan M, et al. (2015) Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* **528**, 99-104.
- [3] Chen K, et al. (2017) Thermosensitivity of growth is determined by chaperone-mediated proteome reallocation. *PNAS* **114**, 11548-11553.
- [4] Chen P, Shakhnovich EI (2010) Thermal adaptation of viruses and bacteria. *Biophys J* **98**, 1109-1118.
- [5] Bennett AF, Lenski RE (2007) An experimental test of evolutionary trade-offs during temperature adaptation. *PNAS* **104**, 8649-8654.
- [6] Brunk E, et al. (2016) Systems biology of the structural proteome. *BMC Syst Biol* **10**, 26.
- [7] Powers ET, Powers DL, Gierasch LM (2012) FoldEco: A model for proteostasis in *E. coli*. *Cell Rep* **1**, 265-276.
- [8] O'Brien EJ, et al. (2013) Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol Syst Biol* **9**, 693.
- [9] Chen K, et al. “A rugged bacterial fitness landscape defined by the discreteness of ATP production strategy” in preparation.

Acknowledgements: This work was funded by NIH grants GM102098, and the Novo Nordisk Foundation award NNF10CC1016517. Computer simulations were performed using resources provided by NERSC under the DOE Contract DE-AC02-05CH11231.

¹Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093. E-mail: kec003@eng.ucsd.edu

²Department of Bioengineering, UCSD, La Jolla, CA 92093. Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kongens Lyngby, Denmark. E-mail: palsson@ucsd.edu