Thermodynamics of gene regulation

Complexity mini-project proposal.
Supervisors: Mario Nicodemi (Physics and Complexity) & Sach Mukherjee (Statistics & Complexity)

Cis-regulatory modules (CRMs) are DNA sequences which act in the regulation of gene transcription. CRMs appear as a key evolutionary novelty of multi-cellular organisms and are remarkably conserved across hundreds of millions of years. An intriguing possibility is that interactions exist between CRMs, as well as with the promoter, which collectively permit complex transcriptional control. Recent experimental work, shows that such interactions do indeed take place, and play a key role in the regulation of a key mammalian gene called MyoD1. The central goal of this mini-project is to better understand the physical basis of such regulation, using MyoD1 as a model system.

Quantitative and, in particular, thermodynamic models for gene activity have usually focused on promoter sequence by relating expression to the equilibrium binding of transcription factors to DNA [1-5]. To address regulation due to interplay between entire CRMs requires going beyond such models. The main goals of the project are: (i) Develop a thermodynamic model of gene regulation at the level of entire CRMs. The aim will be to consider, at a coarse-grained scale, interactions of entire CRMs. Although information at the single base pair level would be inaccessible by such a description, it would provide the possibility to explore the collective properties of the system starting from basic physical principles. This would be an important step towards provide a physical interpretation of cis-regulatory interactions to elucidate the mechanisms whereby diverse expression levels are obtained from different CRM combinations and how they can be controlled by the cell. (ii) Analyze single-cell data pertaining to MyoD1 expression under a range of DNA constructs. Such data provide an opportunity to fit the model (i) and investigate the resulting implications.

References
[1] Shea, M. A. & Ackers, G. K. The OR control system of bacteriophage lambda. A physicalchemical model for gene regulation. J. Mol. Biol. 181, 211-230 (1985).
[2] Buchler, N. E., Gerland, U. & Hwa, T. On schemes of combinatorial transcription logic. Proc.
Natl Acad. Sci. USA 100, 5136-5141 (2003).
[3] Zinzen, R. P., Senger, K., Levine, M. & Papatsenko, D. Computational models for neurogenic

gene expression in the Drosophila embryo. Curr. Biol. 16, 1358-1365 (2006).
[4] Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. & Gaul, U. Predicting expression
patterns from regulatory sequence in Drosophila segmentation. Nature 451, 535-540 (2008).
[5] Gertz J., Siggia E.D. & Cohen B.A., Analysis of combinatorial cis-regulation in synthetic and
genomic promoters. Nature 457, 215-219 (2009).