

The Molecular Biophysics of Evolutionary and Physiological Adaptation

Thesis by
Griffin Daniel Chure

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy in Biochemistry and Molecular Biophysics

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2020
Defended June 1, 2020

TABLE OF CONTENTS

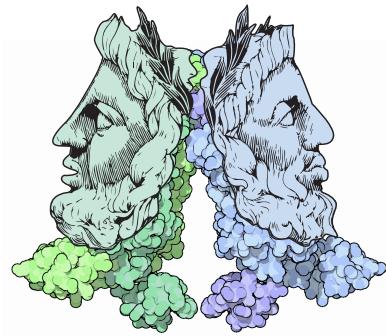
Acknowledgements	iii
Abstract	vii
Published Content and Contributions	ix
Table of Contents	xi
List of Illustrations	xiv
List of Tables	xxi
Chapter I: The Phenomenon of Adaptation Across Scales	1
1.1 Introduction	1
1.2 The Janus Face of Molecules	5
1.3 Using Free Energy to Examine Evolutionary Adaptation	10
1.4 Topic IV: The Physiological Adaptability of Transient Molecular Interactions	14
1.5 On Facing the Elements	21
1.6 On Molecular Biophysics and Evolutionary Dynamics	24
Chapter II: Through the Intramolecular Grape Vine: Signal Processing Via Allosteric Transcription Factors	28
Abstract	28
2.1 Introduction	29
2.2 Theoretical Model	31
2.3 Discussion	50
Chapter III: Unknown Knowns, Known Unknowns, and Unforeseen Consequences: Using Free Energy Shifts To Predict Mutant Phenotypes	61
3.1 Abstract	61
3.2 Introduction	62
3.3 Theoretical Model	64
3.4 Results	69
3.5 Discussion	82
3.6 Materials & Methods	85
Chapter IV: On The Physiological Adaptability of a Simple Genetic Circuit	91
4.1 Abstract	91
4.2 Introduction	92
4.3 Results	95
4.4 Fold-change dependence on carbon quality	101
4.5 Fold-change dependence on temperature variation	104
4.6 Discussion	106
4.7 Materials & Methods	110
Chapter V: 'Water, Water Everywhere, Nor Any Drop to Drink': How Bacteria Adapt To Changes in Osmolarity	118
5.1 Introduction	119

5.2 Results	121
5.3 Discussion	131
5.4 Materials & Methods	137
Chapter VI: Supplemental Information For Chapter 2: Signal Processing Via Allosteric Transcription Factors	146
6.1 Inferring Allosteric Parameters from Previous Data	146
6.2 Induction of Simple Repression with Multiple Promoters or Competitor Sites	151
6.3 Variable Repressor Copy Number (R) with Multiple Specific Binding Sites ($N_S > 1$)	154
6.4 Variable Number of Specific Binding Sites N_S with Fixed Repressor Copy Number (R)	155
6.5 Competitor Binding Sites	156
6.6 Properties of the Induction Response	157
6.7 Flow Cytometry	159
6.8 Single-Cell Microscopy	163
6.9 Fold-Change Sensitivity Analysis	168
6.10 Global Fit of All Parameters	170
6.11 Applicability of Theory to the Oid Operator Sequence	176
6.12 Comparison of Parameter Estimation and Fold-Change Predictions across Strains	177
6.13 Properties of Induction Titration Curves	179
6.14 Applications to Other Regulatory Architectures	183
6.15 <i>E. coli</i> Primer and Strain List	186
Chapter VII: Supplemental Information for Chapter III: Predictive Shifts in Free Energy Couple Mutations to Their Phenotypic Consequences	191
7.1 Non-Monotonic Behavior of ΔF Under Changing K_A and K_I	191
7.2 Bayesian Parameter Estimation For DNA Binding Mutants	195
7.3 Inferring the Free Energy From Fold-Change Measurements	209
7.4 Bayesian Parameter Estimation for Inducer Binding Domain Mutants	219
7.5 Additional Characterization of Inducer Binding Domain Mutants . .	226
7.6 Comparing Parameter Values To The Literature	229
7.7 Parameter Estimation Using All Induction Profiles	237
7.8 Generalizability of Data Collapse To Other Regulatory Architectures	240
7.9 Strain and Oligonucleotide Information	243
Chapter VIII: Supplemental Information for Chapter IV: The Physiological Adaptability of a Simple Genetic Circuit	254
8.1 Non-parametric Inference of Growth Rates	254
8.2 Approximating Cell Volume	256
8.3 Counting Repressors	258
8.4 Parameter Estimation of DNA Binding Energies and Comparison Across Carbon Source	274
8.5 Statistical Inference of Entropic Costs	278
8.6 Media Recipes and Bacterial Strains	284

Chapter IX: Supplemental Information For Chapter 5: How Bacteria Adapt To Changes in Osmolarity	288
9.1 Calibration of a Standard Candle	291
9.2 Classification of Cell Fates	301
9.3 Logistic Regression	305
9.4 Classification Of Shock Rate	312
9.5 Comparison of Survival Probability with van den Berg et al. 2016 . . .	316
9.6 <i>E. coli</i> Strains	318
References	321
Chapter X: Questionnaire	340
Chapter XI: Consent Form	341

Chapter 1

THE PHENOMENON OF ADAPTATION ACROSS SCALES



1.1 Introduction

From archaea thriving in hydrothermal vents on the ocean floor to aspen trees dominating a Coloradan mountainside, all forms of life are unified in their obedience to the whims of their environment. Over the past 3.5 billion years of evolution, life has evolved myriad clever ways to combat (and exploit) environmental fluctuations to amplify reproductive success. The mechanisms behind this adaptation are diverse and traverse the biological scales, ranging from nanosecond-scale conformational switching of proteins (Fig. 1.1(C)), to reconfiguration of metabolic networks to consume different sugars(Fig. 1.1 (B)), to evolutionary trajectories that only become visible over many generations (Fig. 1.1 (C)). While “adaptation” is typically only associated with organisms (at least colloquially), one can use the same language to describe the microscopic operations of molecules.

The idea of molecular adaption is not novel and demands a brief foray into the history of bacterial growth and the dawn of regulatory biology. In the late 1890’s, Emilé Duclaux and his graduate student Frédéric Diénert performed a series of experiments illustrating that the common yeast could only consume galactose after an incubation period with the sugar. This led to a general conclusion that “the production of diastases [enzymes] depends on the manner of nutrition” in which the cultures were grown (Loison, 2013), a phenomenon later coined *enzymatic adap-*

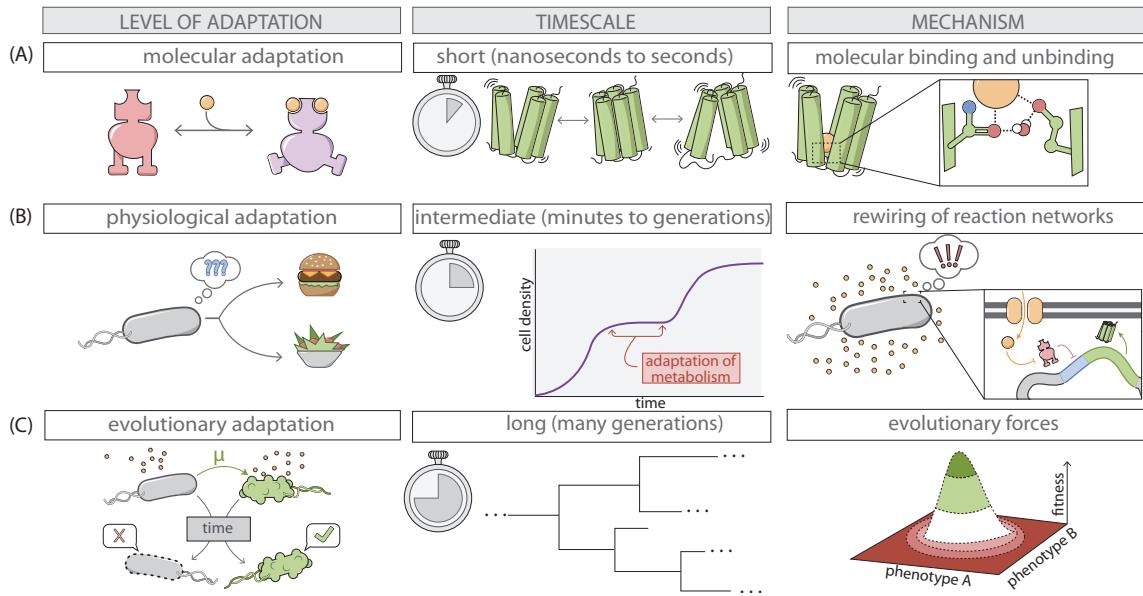


Figure 1.1: The spatial, temporal, and mechanistic scale of adaptation. (A) Molecular adaptation in this work is defined through the lens of allostery where the activity of a protein complex is modulated by the reversible binding of a small molecule. These binding and unbinding events lead to rapid changes in protein conformation whose behavior (both energetic and temporal) is comparable to that of thermal motion. (B) Physiological adaptation here is defined as the rewiring of biochemical reaction networks that lead to changes in cellular behavior (such as chemotaxis) or metabolic capacity (such as aerobic to fermentative metabolism). (C) Evolutionary adaptation is recorded in the variation in the genetic sequence of regulatory molecules. Variations in sequence influence the function of the proteins and RNAs they encode which ultimately define the cellular fitness.

tation. This is one of the first observations of the fact that, while an organism may be able to digest a certain sugar, it may not *always* be able to do so. Rather, there seemed to be certain conditions in which the production or formation of these enzymes could occur. In his doctoral thesis in 1900, Diénert proposed two mechanisms for the origin of enzymatic adaptation observed for galactozymase in *S. cerevisiae* (Loison, 2013). Either (a) the presence of galactose directly transformed enzymes already present in the cell into galactozymase or (b) that the galactose activated the production of galactozymase *de novo* (Diénert, 1900).

Nearly half a century later, Jacques Monod would rediscover the phenomenon of enzymatic adaptation, this time in the context of bacterial growth. In his 1941

work, *Sur un phénomène nouveau de croissance complexe dans les cultures bactériennes*, Monod for the first time reported on the phenomenon of diauxic growth, shown in Fig. 1.2 (A). He noted that for some mixtures of carbon sources, the culture grew “kinetically normal” meaning they grew exponentially to saturation (blue points, Fig. 1.2 (A)). However, some mixtures (such as sucrose and arabinose) led to biphasic growth where the culture would grow exponentially, undergo a period where growth had ceased, followed by again by another round of exponential growth (blue points, Fig. 1.2 (A)). Additionally, Monod showed that the onset of this diauxic shift could be tuned by varying the relative concentrations of the carbon sources, revealing a controllable chemical basis for the adaptation (Fig. 1.2 (B)).

Monod immediately made the connection between diauxic growth and enzymatic adaptation (Loison, 2013). Despite his work appearing 40 years after the pioneering work of Ducleaux and Diénert, there had been little progress towards a mechanistic, needless to say quantitative, explanation for the phenomenon. In fact, Monod was particularly disappointed by the teleological explanations where the cells simply changed their behavior to perform only the chemical reactions that were “needed” (Loison, 2013). The teleological approach to much of biology during this time period, especially in the French scientific community, severely bothered Monod. To him, this kind of approach belonged to a pre-scientific era and lacked the “postulate of objectivity” that other fields of science (particularly physics) had adopted (Loison, 2013). Near the middle of the 20th century, Monod published a 60-page treatise on the phenomena of enzymatic adaptation with the level of quantitative rigor he thought it deserved (Monod, 1947). In this work, he set out to progressively deconstruct and invalidate a series of hypotheses for the phenomenon of enzymatic adaptation. In doing so, he laid the groundwork for what would become (in his opinion) his greatest contribution to science, the nature of allosteric transitions (Loison, 2013), a topic that will feature prominently in the remainder of this thesis.

The diauxic growth transitions shown in Fig. 1.2 illustrate adaptive processes

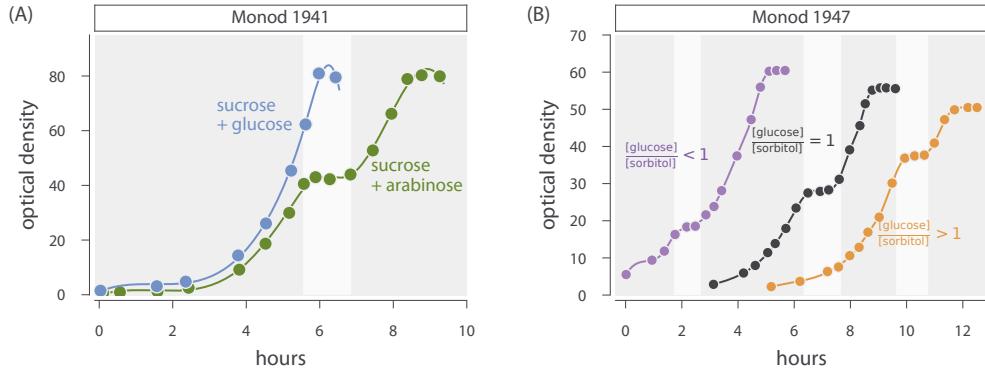


Figure 1.2: The phenomenon of enzymatic adaptation revealed in bacterial growth curves. (A) Optical density measurements of *Bacillus subtilis* cultures grown in a mixture of sucrose and either glucose (blue points) or arabinose (green points). Biphasic growth can be observed in the sucrose/arabinose mixture where the pause in growth (white vertical line) corresponds to enzymatic adaptation. Data digitized from Monod (1941). (B) Diauxic growth curves of *Escherichia coli* cells grown on a mixture of glucose and sorbitol in different proportions. Data digitized from Monod (1947). Periods of enzymatic adaptation are highlighted by white vertical lines.

across the biological scales, as were schematized in Fig. 1.1. While it was not known to Monod at the time, we now know that many cases of enzymatic adaptation are driven by the regulation of gene expression. As the bacterial culture approaches the auxic shift, the presence or absence of the substrate is sensed by regulatory molecules that control whether the genes encoding the enzymes for metabolism of the substrate are expressed. This represents the level of **molecular adaptation** where, given binding or unbinding of the substrate molecule, the activity of the regulatory protein is modulated. The amino acid sequence of these proteinaceous regulators are the product of billions of years of **evolutionary adaptation** and define how the regulatory senses and responds to these signals. Finally, the precision with which these genes are regulated are determined by their sensitivity to physiological states, capturing the level of **physiological adaptation**. j

The central aim of this dissertation is to explore the biophysical mechanisms by which these levels of adaptation – molecular, physiological, and evolutionary – are interconnected. Furthermore, in the spirit of Monod, we seek to make our explo-

ration quantitative and leverage the tools of statistical physics to provide precise predictions from pen-and-paper theory that can be rigorously tested through experiment. The remaining sections of this chapter will outline the major topics of this thesis and place them in a historical context alongside the work of Monod. Finally, we will close with a discussion of how these types of models can be used to explore the predictability of evolution.

1.2 The Janus Face of Molecules

Monod is perhaps most famous for his discovery of allostery, to which he famously referred to as “the second secret of life” (Monod et al., 1965; Ullmann, 2011). It is fair to say that this “secret” has been now been declassified. Allosteric regulation can be found in all domains of life across varied types of biological processes. Allostery can be found governing the behavior of ion channels (Einav and Phillips, 2017), enzymatic reactions (Einav et al., 2016), chemotaxis (Keymer et al., 2006), G-protein coupled receptors (Canals et al., 2012), quorum sensing (Swem et al., 2008), and transcriptional regulation (Huang et al., 2018), to name a few of many examples. Despite the objective complexity in the molecular structures of all of these allosteric molecules, they can be frequently be reduced to simple cartoons where the details of conformational changes, substrate binding affinities, and more can be massaged into a small set of key details. Fig. 1.3 (A) shows the molecular structures of a variety of allosteric transcriptional repressors (top). While each has their own fascinating structure and continuum of conformational states, all can be coarse grained into a simple cartoon representation (bottom) with an active (red) and inactive (purple) state, both of which possess binding pockets (semicircular notches) for an inducer molecule (orange).

Much as we can reduce the complexity of allosteric molecules schematically, we can enumerate simple mathematical models that describe their behavior. Thermodynamic models built on an assumption of quasi-equilibrium are routinely used to describe complex biological phenomena despite the reality that being in thermodynamic equilibrium is synonymous with being dead. Even with this glaring

assumption, such models have been shown to be exceptionally predictive for a variety of complex systems, especially in modeling molecular binding reactions (Dill and Bromberg, 2010) and allostery writ large (Einav and Phillips, 2017; Einav et al., 2016; Keymer et al., 2006; Phillips, 2015; Swem et al., 2008). As the timescales of binding and unbinding reactions are orders of magnitude smaller than that of many other processes in the cell, it is fair to make the approximation that molecular binding is in equilibrium. Under this assumption, we are granted the powerful mathematical privilege to say that the probability of a given state of the system P_{state} follows a Boltzmann distribution,

$$P_{\text{state}} = \frac{e^{-\frac{\epsilon_{\text{state}}}{k_B T}}}{Z}, \quad (1.1)$$

where ϵ_{state} is the energy of that state, k_B is the Boltzmann constant, and T is the system temperature. The denominator Z is the partition function of the system and is the sum

$$Z = \sum_{i \in \text{states}} e^{\frac{\epsilon_i}{k_B T}}, \quad (1.2)$$

ensuring that the distribution is normalized. Therefore, if we are interested in computing the probability of a given allosteric protein being in the active state, we merely have to enumerate all of the Boltzmann weights (given by the numerator in Eq. 1.1) and compute

$$P_{\text{active}} = \frac{\text{sum over all possible active states}}{\text{sum over all possible states}}. \quad (1.3)$$

This probability, defined as a function of the inducer concentration, is shown schematically in Fig. 1.3 (B). While we have passed over some of the more subtle details of this calculation, the plot in Fig. 1.3 (B) presents a *quantitative* prediction of how the activity of an allosteric molecule should scale as a function of the inducer, in this case becoming less active as more inducer is present).

In **Chapter 2** and the associated supplementary **Chapter 6** of this dissertation, we use the Monod-Wyman-Changeux model of allostery (Monod et al., 1965) to build a predictive model of transcriptional regulation where the level of gene expression changes in response to changing activity of an allosteric transcriptional

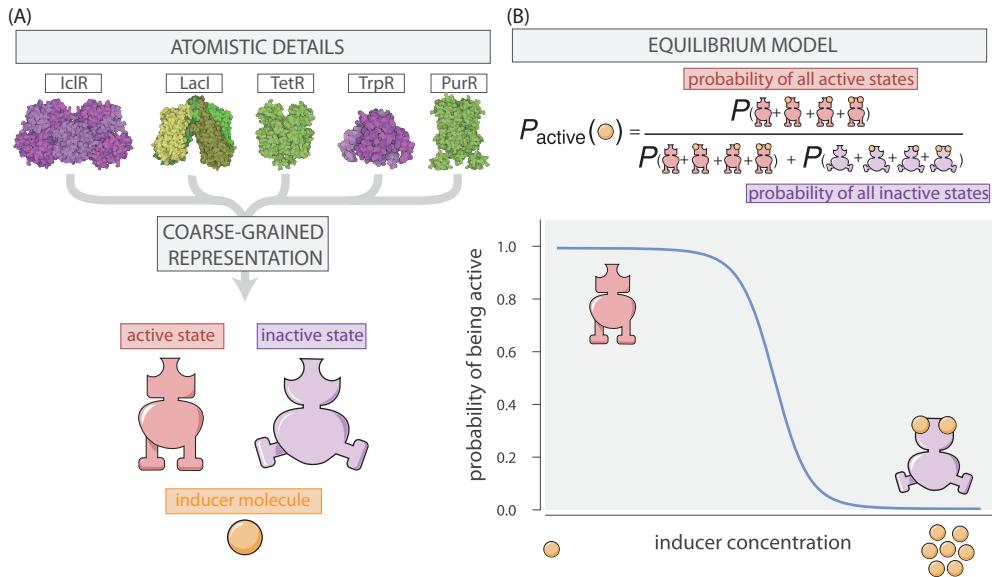


Figure 1.3: A Coarse grained representation of an allosteric molecule. (A) Crystal structures of a variety of allosteric transcription factors are shown at the top. In this thesis, we coarse grain away many of the details to a minimal model (bottom) where the protein can be represent as being either active (red) or inactive (purple), both of which can bind an inducer molecule (orange). (B) By making an assumption of quasi-equilibrium, we can trivially compute a mathematical description of the active probability of an allosteric protein as a function of the inducer concentration (top). In this particular case, the inactive state becomes more probable relative to the active state at higher concentrations of inducer molecule.

repressor. Using the same tricks given by Eq. 1.1 and Eq. 1.3, we expand upon a previously characterized thermodynamic model of the simple repression motif. This motif, schematized in Fig. 1.4 (A) is not just a convenient abstraction of a regulatory architecture. Rather, this motif is the most ubiquitous regulatory scheme in *E. coli* (???; Gama-Castro et al., 2016) and has been the target of much theoretical and experimental dissection (Bintu et al., 2005a; Brewster et al., 2014; Buchler et al., 2003; Garcia and Phillips, 2011; Phillips et al., 2019; Vilar and Leibler, 2003). However inclusion of allostery in a mathematical sense had yet to be experimentally dissected.

At the beginning of 2016, Manuel Razo-Mejia, Stephanie L. Barnes, Nathan M. Belliveau, Tal Einav, and I joined forces and set out to build a complete theoretical model for allosteric transcriptional regulation coupled with a thorough experimen-

tal dissection. This was no small task and would have likely taken a full Ph.D.’s worth of effort for a single person to do. Yet, within a year of project inception we had submitted a manuscript to preprint servers where all of us were annotated as equal contributors. This experience defined how I view collaboration in scientific research and serves as a shining example of scientific socialism.

Together, we enumerated a complete thermodynamic model for the inducible simple repression motif and defined a succinct input-output function for the fold-change in gene expression (schematized in Fig. 1.4 (B)). This model, which is explored in depth in Chapter 2, is defined by a minimal set of biophysical parameters, many of which can be directly measured using standard tricks of molecular biology and biochemistry. With a model in hand, we turned to a collection of 17 unique *E. coli* strains, each with different copy numbers of the repressor protein and different regulatory DNA sequences. Using our theoretical model, we inferred the lone two biophysical parameters which we did not know *a priori* from a single experimental strain (white points in middle panel of Fig. 1.4 (C)), and tested our predictions on all other experimental strains. We found the model to be remarkably predictive, suggesting that our “toy” model of an allosteric repressor captured the underlying physics of the system.

A key feature of this work is a derivation of thermodynamic state variable of this regulatory architecture which we term the *free energy*. This parameter provides an intuition for the effective free energy difference between states of the promoter in which the repressor is bound relative those states in which the repressor is not bound to the promoter. This parameter accounts for all of the ways in which one can tune the parameter values and still achieve the same fold-change in gene expression, as is diagrammed in Fig. 1.5 (A). While we leave the details of this derivation to Chapter 2, we emphasize that this formalism provides a means by which all of the experimental measurements plotted in Fig. 1.4 (C) can be collapsed onto a master curve defined *only* by the free energy, which is illustrated in Fig. 1.5 (B). This scaling, often referred to as “data collapse” in physics, concretely shows that

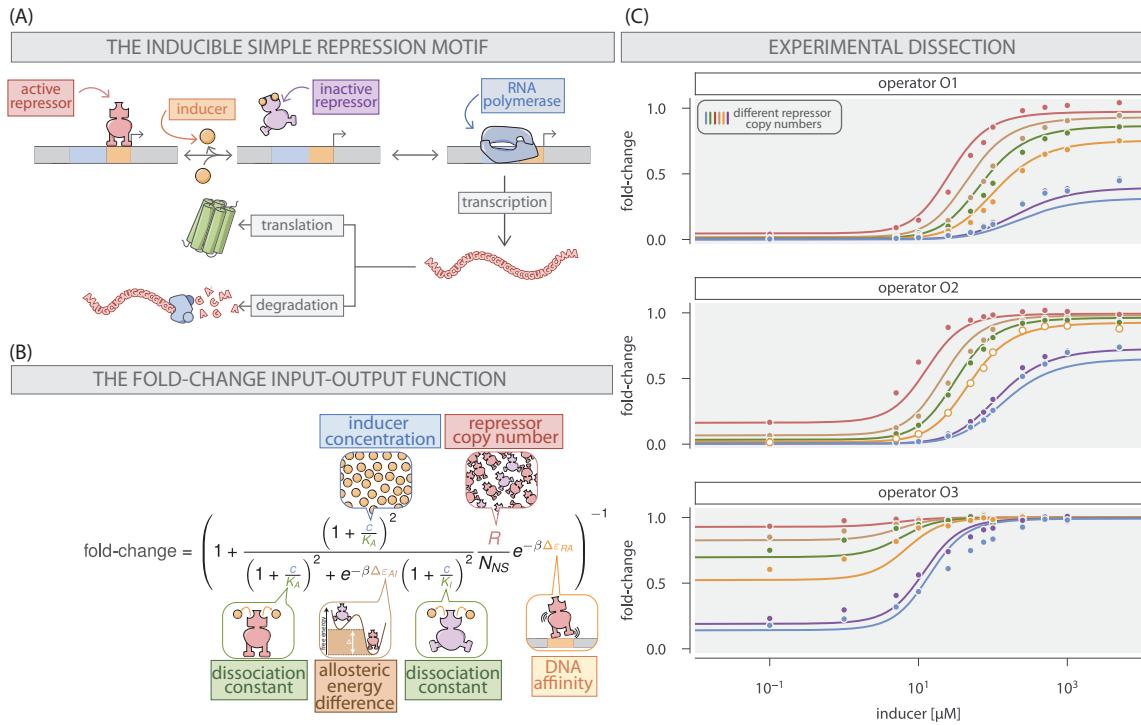


Figure 1.4: Experimental dissection of the inducible simple repression input-output function. (A) Schematic diagram of the inducible simple repression motif. (B) Schematic diagram of the input-output function as is derived in Chapter 2. (C) Experimental measurements of the fold-change in gene expression using the *lac* repressor from *E. coli*. Different rows correspond to different operator sequences and therefore different values for $\Delta \varepsilon_{RA}$. Different colors correspond to different values for the average repressor copy number R . While filled points in the middle panel represent the experimental strain used to infer the values of the inducer dissociation constants. All points correspond to the mean of at least 10 biological replicates.

one has identified the *natural variable* of the system. With this scaling function in hand, we are able to make a measurement of the biophysical parameters, compute the free energy, and make a concrete prediction of what the fold-change in gene expression will be. Or, as we will see in the following section, details of the biophysical parameters can be determined directly from an empirical measurement of the free energy.

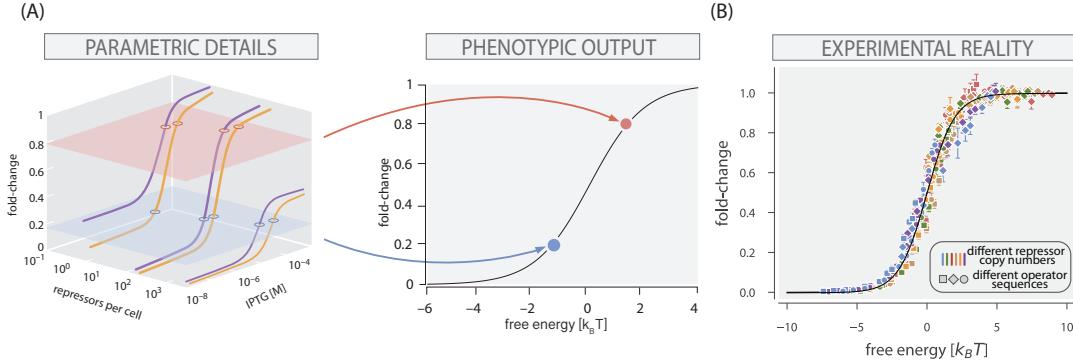


Figure 1.5: Collapse of individual induction profiles onto a simple scaling function. (A) Many different combinations of parameter values yield the same value of the fold-change in gene expression, shown as red and blue horizontal planes. Any point on those planes corresponds to a single value of the free energy (middle) and will appear on the master curve. (B) Data presented in Fig. 1.4 (C) collapsed onto the master curve defined by the predicted value of the free energy.

1.3 Using Free Energy to Examine Evolutionary Adaptation

Allow us to briefly return to Monod and his biphasic growth curves in the mid 1940's. At this point in scientific history, the French vision of biology had taken a strongly finalistic and vitalistic turn (Loison, 2013). In particular, a neo-Lamarckian view had been employed to explain the phenomenon of enzymatic adaptation where the enzymes appropriate for digesting the substrate could be spontaneously formed out the bacterial cytoplasm and inherited by the cell's descendants, completely independent of genes. In general, this approach to biology deeply frustrated Monod and strongly influenced his desire to "physicalize" the science (Loison, 2013). One tool he knew was critical to this mission was the burgeoning field of genetics. In the mid 1930's Monod undertook a short retreat to Thomas Hunt Morgan's lab at Caltech where he was introduced to genetics which he later remarked to as "biology's first discipline" (Loison, 2013). This visit had a profound impact on Monod, who reflected upon it some three decades later:

"Upon my return to France, I had again taken up the study of bacterial growth.

But my mind remained full of the concepts of genetics and I was confident of its ability to analyze and convinced that one day these ideas would be applied

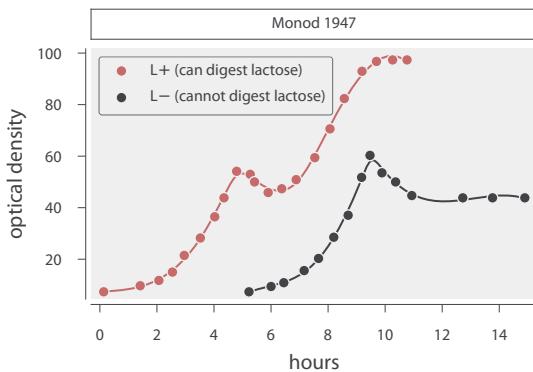


Figure 1.6: Growth curves of lactose-positive and lactose-negative *E. coli* strains on a glucose/lactose mixture. Black curve shows the growth curve of an *E. coli* strain unable to digest lactose grown on a glucose/lactose mixed medium. Red curve shows a mutant of the same *E. coli* strain which is able to consume lactose. The latter displays a diauxic growth cycle with an adaptive period (highlighted in white), illustrating that enzymatic adaptation is a truly genetic property. Figure adapted from Monod (1947).

to bacteria.” (Monod, 1966)

Once he returned to his study of bacterial growth and enzymatic adaptation, he was confronted with incorporating the role of genetic inheritance into his mechanistic explanations. In the mid 1940’s, Monod and his coworkers had begun experimenting with a strain of *E. coli* which was unable to digest lactose ,termed *L-*. When grown on a mixture of glucose and lactose, this strain would not display a diauxic shift and would only be able to consume the glucose in the medium (Fig. 1.6, black). However, Monod and his coworker Alice Audureau discovered a mutation in this strain which *enabled* the digestion of lactose, termed *L+* (Monod, 1947). The growth curve of this strain had the striking feature of diauxic growth. Rather than this mutation merely enabling the digestion of lactose, it did so in a non-constitutive manner and preserved the phenomenon of adaptation. This was an important step forward in Monod’s understanding of enzymatic adaptation (Loison, 2013), revealing that it was a “truly genetic property” (Monod, 1966).

This finding illustrates the level of evolutionary adaptation operating at the level of molecules. While it is difficult to find any literature dissecting this particular

L+ mutation, it is not difficult to imagine several different mechanisms by how it could be manifest. One such explanation is that this *L+* mutation is within a transcriptional regulator itself where a deficiency in the ability to respond to the presence of lactose (and decreasing glucose concentration) had been restored. Such mutations are the crux of **Chapter 3** and the corresponding supplemental **Chapter 7** of this dissertation.

As summarized in the previous section and discussed in depth in Chapter 2, Chapter 3 and the associated supplemental Chapter 7 of this dissertation focus on the influence of mutation within the allosteric transcription factor. Furthermore, Chapter 3 presents a generic mechanism by which shifts in the free energy can be mapped directly to changes in values of the biophysical parameters. This chapter, much like Chapter 2, was borne out of a wonderful collaboration with Manuel Razo-Mejia, Stephanie L. Barnes, Nathan M. Belliveau, Tal Einav, and Zofii A. Kaczmarek. Being able to launch another collaborative effort afforded us the opportunity to both develop a new theoretical interpretation for how mutations influence the free energy and acquire enough experimental data to thoroughly test it.

The primary conceptual development of Chapter 3 is illustrated in Fig. 1.7. Theoretically, we consider a bacterial strain with an allosteric repressor (which we term the “wild-type” repressor) that has been sufficiently characterized. Given enough parametric knowledge of the system, we can easily compute both its predicted average fold-change in gene expression along with the corresponding free energy. However, once a mutation has been introduced *into the repressor protein* (resulting in a non-synonymous amino acid change), we are once again ignorant *a priori* of what changes, if any, that mutation may have imparted on the system. In Fig. 1.7, we examine two separate hypothetical mutations, shown in purple and orange, which significantly change the character of the system by either increasing or decreasing the fold-change in gene expression, respectively. If we assume that these mutations do not change the underlying physics of the system, we are

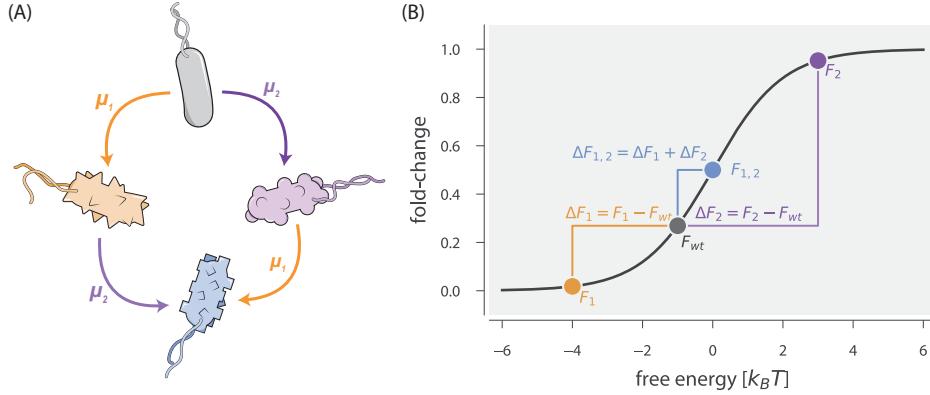


Figure 1.7: Mutations lead to shifts in free energy, permitting prediction of double mutant phenotypes. Consider a wild-type bacterium which on, on average, exhibits a fold-change of ≈ 0.3 and a free energy of $-1 k_B T$ (grey point in (B)). We can consider that a single mutation (either orange or purple) changes the mean fold-change and therefore the free energy, translating the measurement about the master curve (black line in (B)). Assuming there are no epistatic interactions between the two single mutations, a null hypothesis predicts that for the double mutant (blue bacterium in (A) and point in (B)) 0 , the net free energy is simply the sum of the individual free energy shifts.

permitted to use the theoretical framework outlined in Chapter 2 and in Fig. 1.4 to characterize each mutation and determine what biophysical parameters have been changed. This permits us to calculate the new free energy of the system ($F_{\text{mutation } 1}$) as well as the shift in free energy from the wild-type value,

$$\Delta F_{\text{mutation } 1} = F_{\text{mutation } 1} - F_{\text{wt}}. \quad (1.4)$$

As will be described in detail in Chapter 3 and the supplemental Chapter 7, the precise value of this free energy shift ΔF can be directly computed given sufficient parametric knowledge.

This formalism provides a mathematical hypothesis for how double mutants may behave. Given known values for ΔF of each mutation in isolation, can we compute the shift in free energy of the pairwise double mutant $\Delta F_{\text{mutations } 1 \& 2}$? Eq. 1.4 presents a mathematical null hypothesis that the net shift in the free energy is simply the sum of the individual shifts in free energy,

$$\Delta F_{\text{mutations } 1 \& 2} = \Delta F_{\text{mutation } 1} + \Delta F_{\text{mutation } 2}, \quad (1.5)$$

assuming there are no epistatic interactions between the mutations. Given the fact that we can compute the fold-change in gene expression given knowledge of the free energy, we can therefore predict the double mutant phenotype *a priori*, a prediction not possible prior to this work.

Over the course of two years (while this theory was in the works), the experimental cast of characters (Stephanie L. Barnes, Nathan M. Belliveau, Manuel Razo-Mejia, and Zofii A. Kaczmarek) and I made a series of mutations in the LacI repressor that we had characterized in the work presented in Chapter 2. These mutations included three point mutations in the DNA binding domain of the repressor, four mutations in the inducer binding domain, nine double mutants (one inducer binding and one DNA binding each), across four repressor copy numbers and three operator sequences. While this process of strain generation and data collection is not the primary focus of the work, it took $\approx 80\%$ of the effort. Without them, this work would have remained an untested theoretical novelty. While we leave many of the rich details of this prediction to the reader in Chapter 3, we showcase our experimental success in Fig. 1.8 (B) where the predicted induction profiles of nine double mutants (light blue shaded regions) are overlaid with their experimental measurements (points). The near 100% agreement between theory and experiment illustrates the utility of using free energy shifts as a means to predict new phenotypes.

1.4 Topic IV: The Physiological Adaptability of Transient Molecular Interactions

In **Chapters 4 and 5** (and the associated supplementary **Chapters 8 and 9**), we explore the final level of adaptation in Fig. 1.1 – physiological adaptation. We do so in two distinctly different systems. The first (Chapter 4) builds upon our discussion of transcriptional regulation, but now examines how robust the biophysical parameters of the thermodynamic model are to changes in physiology, either by changing the available carbon source or by changing the temperature. Secondly, we examine physiological adaptation in the context of osmoregulation – a true

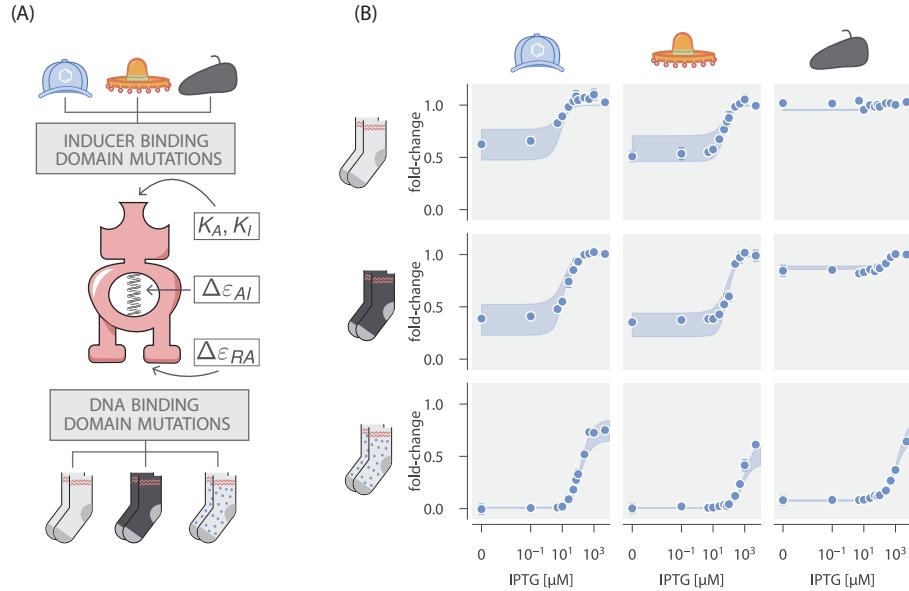


Figure 1.8: Theoretical prediction and experimental validation of double mutant phenotypes. (A) Cartoon representation of the LacI repressor with mutations in the inducer binding domain and DNA binding domain represented by hats and socks, respectively. While the mutations have known chemical features, we characterize each mutation as potentially modifying four biophysical parameters, K_A , K_I , $\Delta\epsilon_{AI}$ for inducer binding mutants, or $\Delta\epsilon_{RA}$ for DNA binding mutants. (B) Predicted induction profiles for pairwise double mutants are shown as blue shaded regions representing the uncertainty in our predictions. Experimental measurements are shown as blue points (means of at least 10 biological replicates). Each row corresponds to a single DNA binding domain mutation and each column to a single inducer binding domain mutation.

matter of life and death in the single-celled world.

Up to this point in our travels through scientific history, we have examined Monod's growth curves in various pairwise combinations of sugars. A feature of note is that the presence of diauxic shifts can be seen in various organisms and for many different types of sugars such as sucrose/arabinose, glucose/sorbitol, and glucose/lactose pairings (Monod, 1947). These combinations reveal that cells are able to juggle dual-input logic systems where the "decision" to digest one carbon source or another relies on monitoring changes in concentrations of either sugar. In his 1947 treatise, Monod showed that this phenomenon was not limited to dual-carbon mixtures and presented a "triauxic" growth curve of *E. coli* grown on a glu-

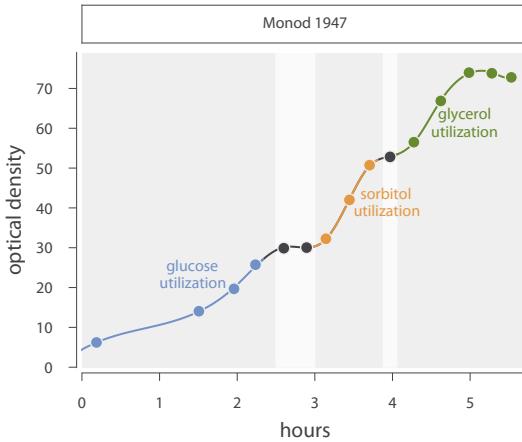


Figure 1.9: A metabolic hierarchy in a three-component growth mixture of glucose, sorbitol, and glycerol. A “triauxic” growth curve illustrating a hierarchy of carbon source metabolism. An *E. coli* culture was grown in a medium with equal parts glucose, sorbitol and glycerol with utilization in that order. Auxic transitions are shown as black points and white vertical lines. Regions of the growth curve where glucose, sorbitol, and glycrol are primarily consumed are colored in blue, orange, and green, respectively. Data digitized from Monod (1947).

cose/sorbitol/glycerol mixture, shown in Fig. 1.9. This result illustrated to Monod that the mechanisms underlying enzymatic adaptation “have the character of *competitive interactions* between different specific enzyme forming systems” (Monod, 1947).

These competing interactions must be resilient to a variety of physiological states. Despite the fact that the carbon atoms in glucose, sorbitol, and glycerol are all ultimately incorporated into the same biomolecules, their pathways to utilization are all distinct and include a variety of different metabolic intermediates. Furthermore, the exponential growth phases in Fig. 1.9 for each carbon source have different growth rates which itself results in large changes in cell volume (Jun et al., 2018; Taheri-Araghi et al., 2015a), genome copy number (Nordström and Dasgupta, 2006), and global gene expression patterns (Hui et al., 2015; Li et al., 2014; Schmidt et al., 2016). Despite these changes in cellular physiology, the regulatory systems underlying enzymatic adaptation still function with binding of transcription factors being ignorant of the majority of possible metabolic states of the cell.

Despite this empirical observation, it has been commonly assumed that the utility of thermodynamic models of gene expression are limited and that the precise values of the biophysical parameters are directly tied to the physiological state in which they were determined. It has even been said that thermodynamic models of gene expression have been a “tactical success, yet strategic failure” in building an understanding of how genomes operate (Phillips et al., 2019).

In **Chapter 4** of this dissertation, we quantitatively assess these assumptions in the context of gene expression by considering the theoretical models built in Chapters 2 and 3 and directly measuring the adaptability of the inducible simple repression regulatory architecture across different physiological states. Namely, we explore how predictive our thermodynamic model can be when modulating either the quality of the carbon source (glucose, glycerol, or acetate, Fig. 1.10 (A)) or by changing the temperature of the growth medium (32° C , 37° C , or 42° C , Fig. 1.10 (B)). The culture doubling time varies by nearly a factor of four across the different conditions, illustrating the diversity in physiological states.

How could this variation in cellular physiology be incorporated into our thermodynamic model? Up to this point in this thesis, experiments have been conducted in a growth medium supplemented with glucose held at a balmy 37° C . However, nowhere in the thermodynamic model schematized in Fig. 1.4 (B) is it specified *which* carbon source must be present whereas the temperature of the system is explicitly included as a multiplicative factor $\beta = (k_B T)^{-1}$ in front of the exponentiated terms. These features of the model allow us to make explicit predictions of how these perturbations should influence the observed fold-change in gene expression, if at all.

The parameter that we can say *a priori* is very likely to change is the repressor copy number R . In Chapters 2 and 3, we knew the total repressor copy number from previous work where the copy numbers were directly measured via quantitative Western blotting in a particular physiological state (Garcia and Phillips, 2011). However, it has been known for nearly three-quarters of a century that the

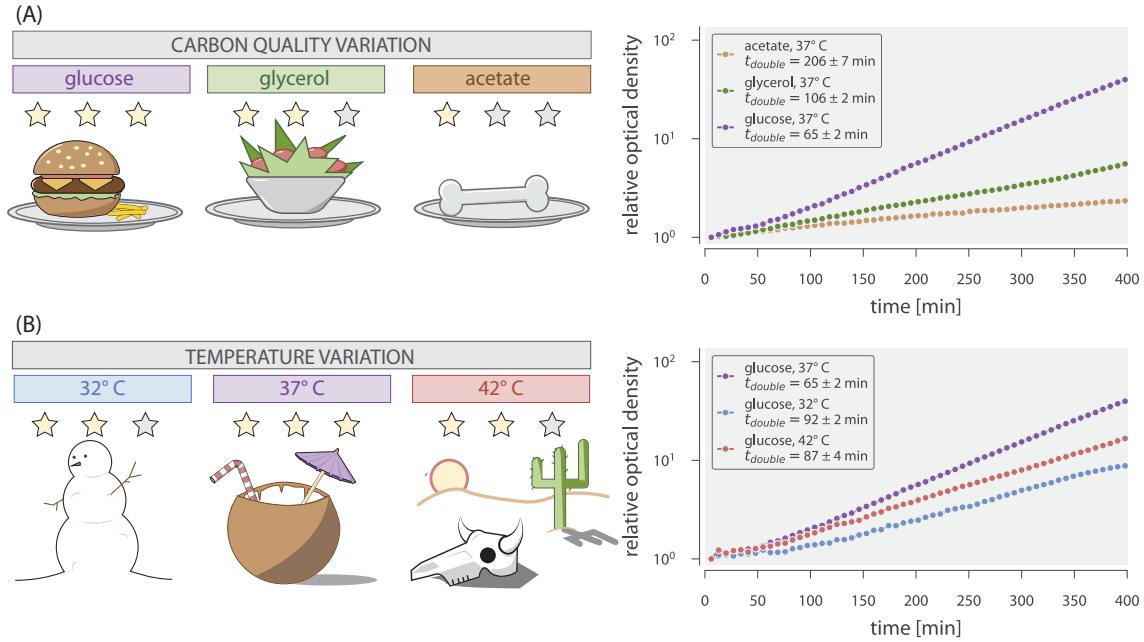


Figure 1.10: Control of cellular physiology via carbon source and temperature variation. (A) Carbon sources used in work presented in Chapter 4 (left) with “star rating” indicating quality of the carbon source. Growth curves for the three carbon sources, all at 37° C are shown on right-hand panel. (B) Growth temperatures explored in Chapter 4 (left) with “star rating” indicating fastest growth rate. Growth curves (right) are shown for the three temperatures, all of which use glucose as the sole carbon source. For right-hand panels in (A) and (B), optical density is computed relative to the initial optical density of the culture.

total protein content of the cell scales linearly with the growth rate (Jun et al., 2018; Schaechter et al., 1958), a phenomenon that has recently been queried at the single-protein level through proteomic methods (Li et al., 2014; Peebo et al., 2015; Schmidt et al., 2016; Valgepea et al., 2013). Thus, we cannot assume that the protein copy number of the strains used in Chapters 2 and 3 will not be perturbed. To account for this fact, we used a fluorescence-based method to directly count the number of LacI repressors per cell in each growth condition, a method which is discussed in extensive detail in Chapter 9. This experimental approach, while necessary, is extremely laborious. I am indebted to the work of Zofia A. Kaczmarek for her heroic efforts in conducting a large number of the experiments presented in Chapter 4.

Our work revealed two key features of this thermodynamic model of gene expression. First, we found that the values of the biophysical parameters inferred

from a single physiological state were remarkably predictive when the quality of the carbon source was decreased (Fig. 1.11 (A, left)). This indicates that this genetic circuit is largely insulated from the metabolic state of the cell. This is exemplified in our ability to collapse the measurements across different carbon sources as a function of the free energy, shown in Fig. 1.11 (A, right). For the carbon sources studied in this Chapter, we conclude that this simple thermodynamic model can be considered both tactically and strategically successful.

Yet when it comes to temperature, we find that a simple rescaling of the thermal energy of the system is *not* sufficient to predict the output of this genetic circuit when the temperature is varied (dashed lines in left-hand side of Fig. 1.11, (B)). This is not necessarily a surprising result as binding of transcription factors is not strictly an enthalpic process. Temperature is known to have a strong influence on many material properties of DNA ,such as persistence length and salt release (Goethe et al., 2015), excluded volume effects (Driessens et al., 2014), and repressor-DNA solubility (Elf et al., 2007), to name a few of many effects. To phenomenologically characterize the influence of temperature on the fold-change in gene expression, we considered that there was a constant entropic penalty (though inclusion of a temperature-dependent entropic cost is discussed in Chapter 9). We found that inclusion of this parameter markedly improved the description of the data (solid lines in left-hand side of Fig. 1.11 (B)) and permitted data collapse within experimental noise of data collected at 37° C (right-hand side of Fig. 1.11 (B)).

The inclusion of a phenomenological entropic parameter is not by any means meant to shut the book on temperature effects in this model. Rather, it serves as a representation of what *may* help explain these effects and demands more focused theoretical and experimental work. To say that the current disagreement between theory and experiments embodies the “strategic failure” of thermodynamic models is, in my view, disingenuous. To say so would be to deem the initial failures of elasticity theory to properly predict the influence of impurities and temperature on the elastic constants of materials as a “strategic failure” in material science.

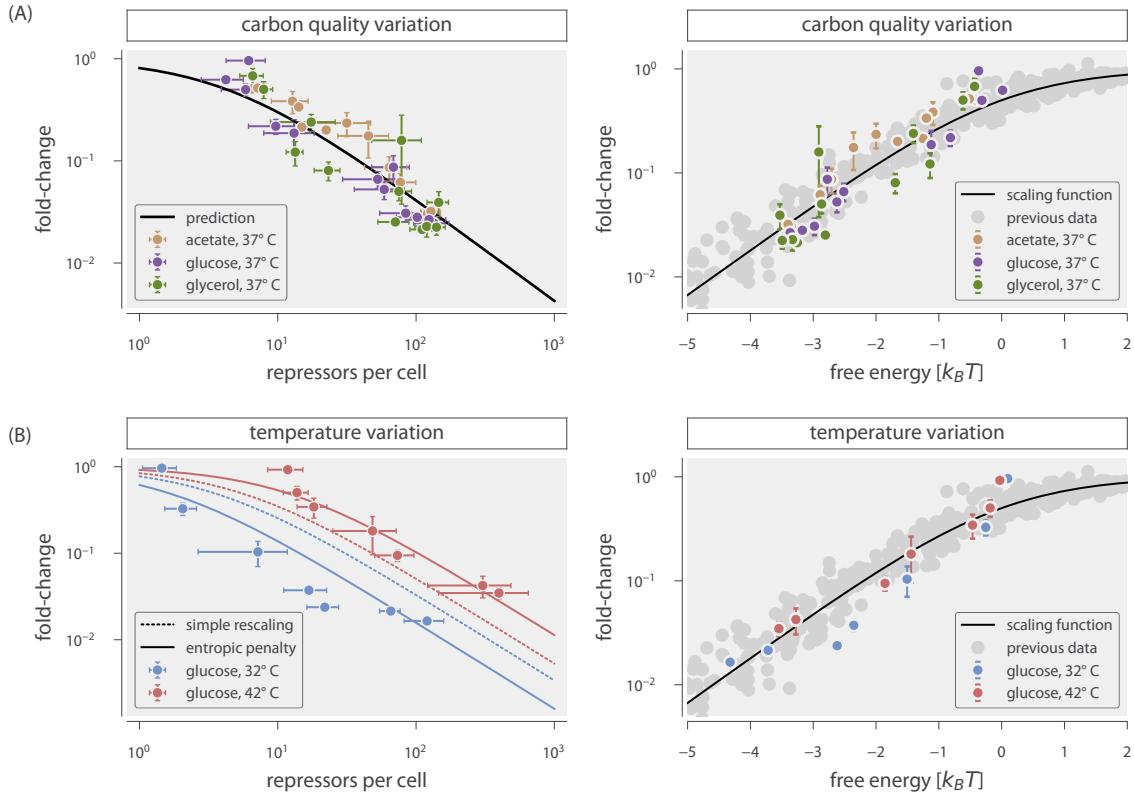


Figure 1.11: Performance of a simple thermodynamic model of simple repression in diverse physiological states. (A) Fold-change in gene expression measurements in different carbon sources plotted against the average repressor copy number (left) and free energy (right). Black line in the left-hand panel is the predicted fold-change assuming no parameters are modified. (B) Fold-change measurements at different temperatures plotted as a function of the repressor copy number (left) and free energy (right). Dashed-lines in left-hand plot show the predicted fold-change with a simple rescaling of the thermal energy. Solid lines are predicted fold-change upon inclusion of an entropic penalty. Points on right-hand plot were computed using parameters with an entropic penalty. All measurements and errors displayed are the mean and standard error of three to eight biological replicates. Light-grey points in right hand panels are data from Garcia and Phillips (2011), Brewster et al. (2014), Razo-Mejia et al. (2018), and Chure et al. (2019), all of which were measured in glucose-supplemented media at 37° C.

Initial phenomenological models of the effects of impurities and temperature on elastic properties of solids (Friedel, 1974) led to several decades of focused theoretical and experimental work that resulted in a complete predictive and mechanistic description (Phillips, 2001). Now is the time for a similar approach to biology in the context of temperature and the regulation of gene expression.

1.5 On Facing the Elements

The first four chapters of this work encompass myriad perspectives of adaptive processes at the level of transcription regulation. However, just as important as the regulation is the action of the gene that is ultimately expressed. While Monod's work described in the preceding sections was focused on the expression of enzymes, we now turn to yet another level of physiological adaptation in bacteria – the regulation of turgor pressure.

In the wild, microbes are constantly faced with an array of environmental insults ranging from changes in temperature, availability of oxygen for aerobic respiration, and even chemical warfare from neighboring microbial communities (Czaran et al., 2002). One such environmental challenge microbes often face is the variation in the osmolarity of their surroundings. Changes in ion concentrations can result in large volumes of water rushing across the cell membrane, leading to rupture of the membrane and ultimately cell death (Fig. 1.12 (A)). Unsurprisingly, all domains of life have evolved clever mechanisms to combat these osmotic shocks and regulate their internal turgor pressure.

One such mechanism for osmoregulation in *E. coli* is through the action of mechanosensitive ion channels – large, transmembrane structures which sense tension in the cell membrane. Exposure to a hypo-osmotic shock (where water rushes across the cell membrane *into* the cell), a change in membrane tension is sensed by these mechanosensitive channels, triggering a conformational change which opens a pore in the membrane without rupture (Fig. 1.12 (B)). This acts as a pressure release valve, providing a means for turgor pressure to be relieved without a potentially fatal burst. This phenomenon represents yet another system in which

adaptation can be found at the molecular, evolutionary, and physiological levels. In **Chapters 5 and 9** of this dissertation, we explore a fundamental question – how many mechanosensitive channels does a cell need to have an appreciable chance at surviving an osmotic shock?

To approach this question, Heun Jin Lee and I collaborated on the experimental and data analysis components, respectively. This is a project that had been in preparation for several years before I had the privilege of joining the team in the summer of 2017. While the experimental techniques used to probe transcriptional regulation were far from simple, they pale in comparison to those employed by Heun Jin. This project required an enormous amount of molecular biology to generate the necessary strains in which the number of mechanosensitive channels could be tuned across three orders of magnitude and measured with precision. This process involved reworking classic techniques in molecular biology to remove the presence of osmotic shocks which would prove fatal for strains with few or no mechanosensitive channels. On top of the complex biochemistry, Heun Jin developed a clever microfluidic system where osmotic shocks could be imaged in real time at the single cell level. While the majority of Chapter 5 focuses on the analysis and interpretation of the data, none of it would have been possible without Heun Jin's Herculean efforts.

While we leave the details of the inference to Chapter 5 and the supplemental Chapter 9, the survival probability as a function of the total mechanosensitive channel number is given for “slow” and “fast” osmotic shocks in Fig. 1.12 (C) and (D), respectively. The credible regions in this plot illustrate that for an $\approx 80\%$ chance of surviving either a slow or a fast osmotic shock, at least ≈ 500 channels are needed. This number is in agreement with recent proteomic measurements in *E. coli* (Li et al., 2014; Schmidt et al., 2016; Soufi et al., 2015), but are at odds with current theoretical models. While it is difficult to theoretically define a survival/death criterion, current physical models predict only a few mechanosensitive channels (specifically, MscL) are needed to relieve even large increases in membrane ten-

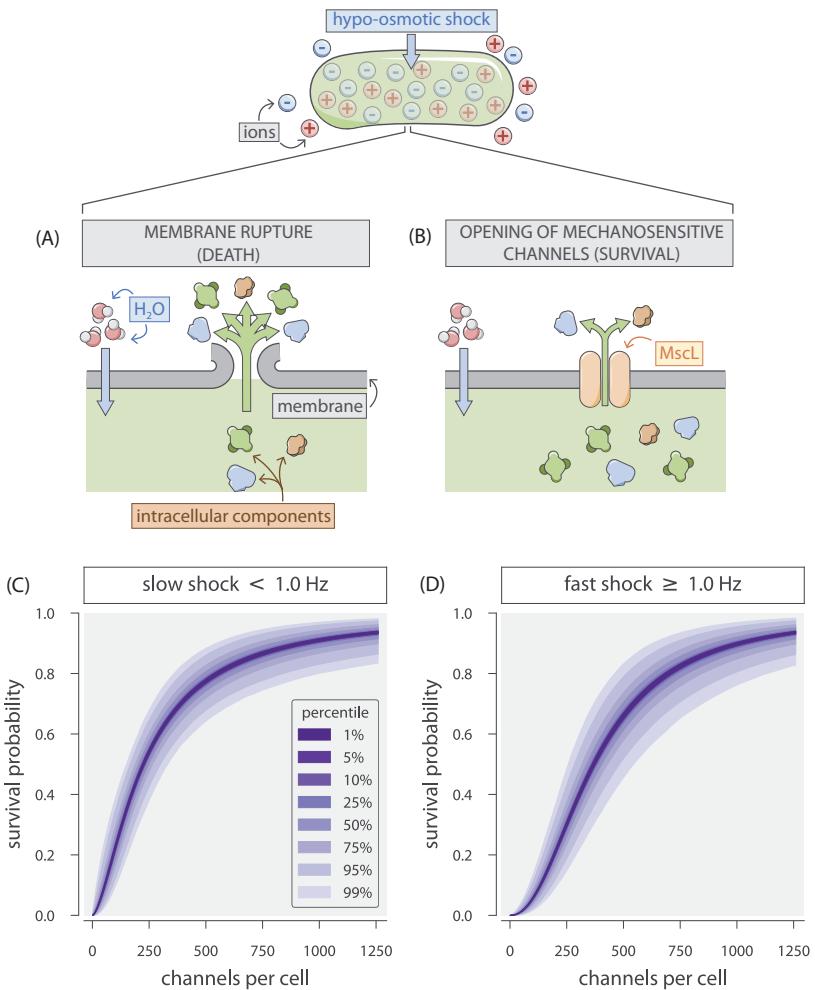


Figure 1.12: The connection between mechanosensitive channel copy number and probability of survival. (A) In the absence of mechanosensitive channels, water rushing across the membrane during a hypo-osmotic shock can lead to membrane rupture and large-scale release of intracellular components into the extracellular space, resulting in cell death. (B) In the presence of mechanosensitive channels (specifically, the major *E. coli* mechanosensitive channel MscL as shown in yellow), increased membrane tension results in a conformational change of the channel, resulting in the expulsion of water and some small constituents of the intracellular milieu. The inferred survival probability curves for slow and fast shock exchange rates are shown in (C) and (D), respectively. Different shaded purple regions correspond to different credible regions of the estimates.

sion. These findings illustrate another avenue in which the disagreement between theory and careful, quantitative experiments reveal gaps in our understanding of fundamental biological phenomena.

1.6 On Molecular Biophysics and Evolutionary Dynamics

This thesis as a whole presents an attempt to understand how adaptive processes operate in biological systems at a mechanistic level beyond qualitative description. The thermodynamic model derived and explored in Chapter 2 presents a concrete theoretical framework through which we can understand how mutations and environmental perturbations influence the output of a simple genetic circuit with quantitative precision. While the work here specifically explores the *mean* level of gene expression of a population, I've had the privilege to be involved in several projects which explore the complete distribution of gene expression of various regulatory motifs using non-equilibrium models(Laxhuber et al., 2020; Razo-Mejia et al., 2020). Both equilibrium and non-equilibrium approaches, while differing in their fundamental assumptions of the system, can be used to understand how the regulation of gene expression occurs *in vivo* and should be viewed as complementary rather than adversarial approaches.

A combination of these types of approaches will be necessary to attack what I believe is the next great frontier of biological physics – predicting evolution. While this thesis is primarily focused on a single type of regulatory architecture regulating a single promoter via a single species of transcription factor, it is worth remembering that systems-level phenotypes are often complex resulting from the concerted action of an array of biological processes. As was mentioned in our discussion on physiological adaptation, it has been known for nearly a century that the bacterial growth rate is directly correlated to the total protein content of the cell, with recent works illustrating rich phenomenology in the structure of the bacterial proteome as a whole (Hui et al., 2015; Klumpp and Hwa, 2014; Li et al., 2014; Schmidt et al., 2016; Scott et al., 2010).

In collaboration again with Nathan M. Belliveau, we have begun to explore how

the composition of the bacterial proteome is structured at the single-protein level. Fig. 1.13(A) shows a compiled data from a variety of different proteomic data sets (using either quantitative mass spectrometry (Peebo et al., 2015; Schmidt et al., 2016; Valgepea et al., 2013) or ribosomal profiling (Li et al., 2014)) where the abundance of different molecular constituents of the bacterial proteome are plotted as a function of the growth rate. These components, broken down by their functional designation according to their Cluster of Orthologous Groups (COG) annotation (Galperin et al., 2015), reveal varied dependencies on the growth rate. Of note are the COG classes “cellular processes and signaling”, “metabolism”, and “information storage and processing” which all appear to have a correlation between the cellular growth rate and the total mass of that proteome sector. However, when plotted as the total *mass fraction* of the proteome instead of the total mass, a striking result is observed. Fig. 1.13 (B) reveals a very strong, negative correlation between the mass fraction of the proteome dedicated to information storage and processing (including ribosomal and transcriptional machinery) and the proteome fraction dedicated to metabolism. This direct competition for resources between the proteins involved in translation (ribosomes, elongation factors, etc.) and metabolic networks has been shown previously (Hui et al., 2015; Klumpp and Hwa, 2008; Scott et al., 2010) and suggests a strong evolutionary constraint in how resources can be optimally partitioned.

As of this writing, our understanding of the cellular resource allocation visible in Fig. 1.13 remains largely phenomenological (Scott et al., 2014). This is in part due to the tremendously high-dimensional nature of systems-level organization. Our understanding of systems with such huge degrees-of-freedom have classically benefited enormously from the application of statistical mechanics as this thesis shows in the context of transcriptional regulation. The quantitative framework derived and carefully dissected in this thesis, I believe, lays the groundwork to understand how phenomena such as that shown in Fig. 1.13 (B) arise, and perhaps more importantly, evolve. A recent work from Michael Lässig, Ville Mustonen, and Aleksandra Walczak entitled *Predicting evolution* (Lässig et al., 2017) describes

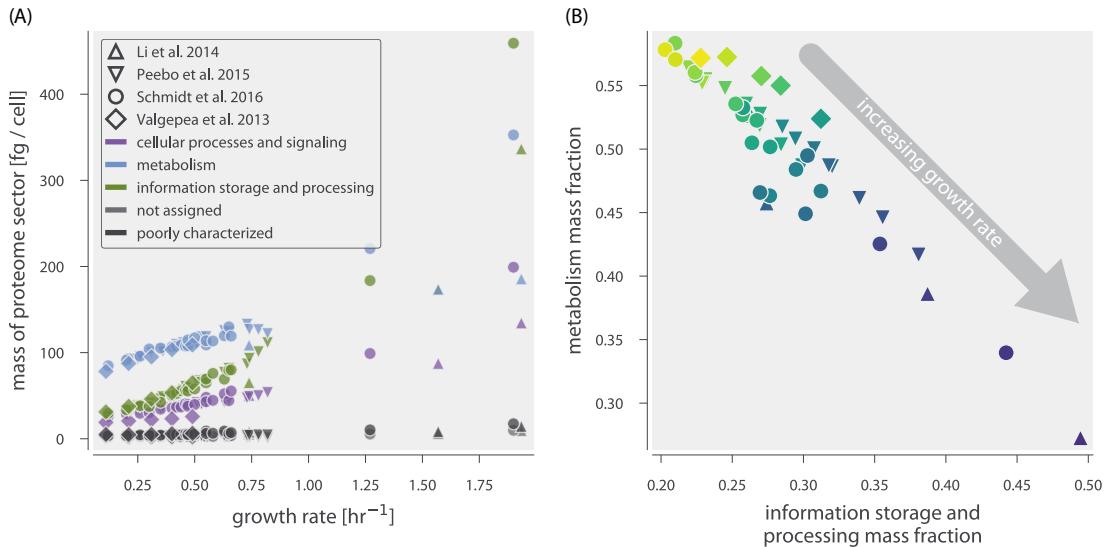


Figure 1.13: Allocation of cellular resources induces compositional structure in the *E. coli* proteome. (A) The total proteome mass of the five major annotated COG categories is shown as a function of the experimental growth rate. Different marker shapes represent different data sets. (B) The total fraction of the proteome dedicated to metabolic machinery plotted as a function of the total proteome mass dedicated to the processes of the central dogma. Different shapes correspond to the different data sets shown in (A). Color indicates increasing growth rate from yellow to dark blue. Data shown in this figure come from Peebo et al. (2015) (inverted triangles), Li et al. (2014) (triangles) Schmidt et al. (2016) (circles) and Valgepea et al. (2013) (diamonds)

what the future of evolutionary theory may look like given these types of models. Recent technological advancements in sequencing, microscopy, and computation coupled with theoretical advancements in the biophysics of gene regulation present an opportunity for a rich theoretical dialogue between molecular biophysics and evolutionary dynamics coupled with experimental dissection (Fig. 1.14).

The somewhat recent paradigm shift in our understanding of noise in biological networks illustrates how cross-disciplinary approaches to scientific discovery can solve (and more-often) create new fields of biological inquiry. I can only hope that some of the material described in the coming chapters can help contribute to a systems-biology approach to evolution.

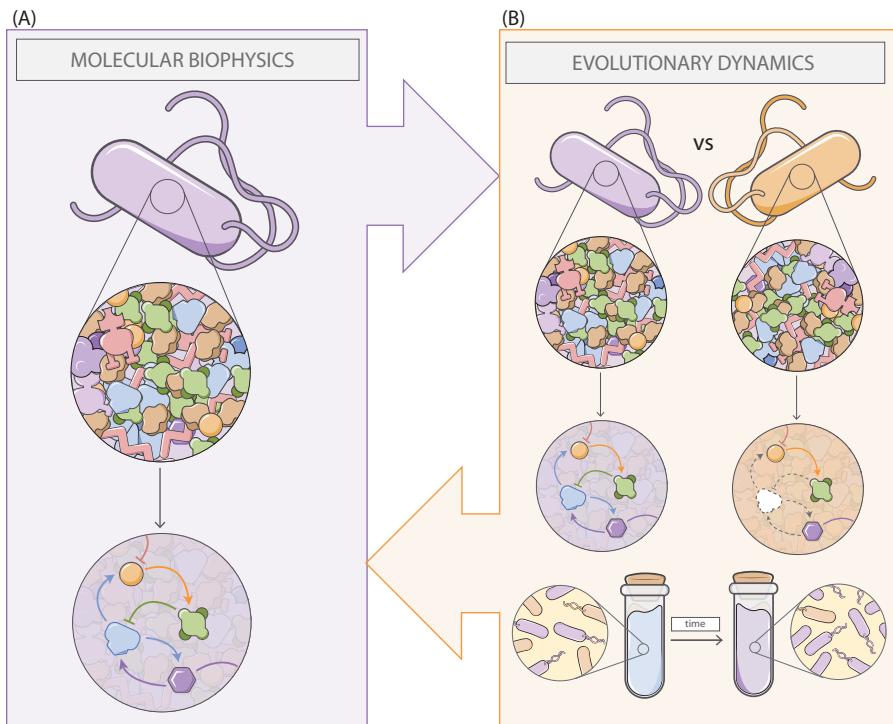


Figure 1.14: The coming interplay between molecular biophysics and evolutionary dynamics. (A) Recent progress in our understanding the structure and function of biological networks has resulted in many examples where high-dimensional biological phenomena can be boiled down to effective phenomena. Future work will draw from our understanding of these networks to place them in an evolutionary perspective (B) where the connection between perturbations at the level of nodes in biological networks can be drawn to fitness and evolutionary trajectories can be predicted.

REFERENCES

- Ackers, G.K., and Johnson, A.D. (1982). Quantitative model for gene regulation by A phage repressor. *Proc. Natl. Acad. Sci. USA* 5.
- Aghaeepour, N., Finak, G., Hoos, H., Mosmann, T.R., Brinkman, R., Gottardo, R., and Scheuermann, R.H. (2013). Critical assessment of automated flow cytometry data analysis techniques. *Nature Methods* 10, 228–238.
- Allen, R.J., and Waclaw, B. (2018). Bacterial growth: A statistical physicist's guide. *Reports on Progress in Physics* 82, 016601.
- Anderson, R.P., Jin, R., and Grunkemeier, G.L. (2003). Understanding logistic regression analysis in clinical reports: An introduction. *The Annals of Thoracic Surgery* 75, 753–757.
- Auerbach, A. (2012). Thinking in cycles: MWC is a good model for acetylcholine receptor-channels. *The Journal of Physiology* 590, 93–98.
- Balleza, E., Kim, J.M., and Cluzel, P. (2018). Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature Methods* 15, 47–51.
- Barnes, S.L., Belliveau, N.M., Ireland, W.T., Kinney, J.B., and Phillips, R. (2019). Mapping DNA sequence to transcription factor binding energy in vivo. *PLOS Computational Biology* 15, e1006226.
- Bavi, N., Cortes, D.M., Cox, C.D., Rohde, P.R., Liu, W., Deitmer, J.W., Bavi, O., Strop, P., Hill, A.P., Rees, D., et al. (2016). The role of MscL amphipathic N terminus indicates a blueprint for bilayer-mediated gating of mechanosensitive channels. *Nature Communications* 7, 11984.
- Berg, H.C., and Purcell, E.M. (1977). Physics of chemoreception. *Biophysical Journal* 20, 193–219.
- Berg, J., Willmann, S., and Lässig, M. (2004). Adaptive evolution of transcription factor binding sites. *BMC Evolutionary Biology* 4, 42.

- Bialecka-Fornal, M., Lee, H.J., DeBerg, H.A., Gandhi, C.S., and Phillips, R. (2012). Single-Cell Census of Mechanosensitive Channels in Living Bacteria. *PLoS ONE* 7, e33077.
- Bialecka-Fornal, M., Lee, H.J., and Phillips, R. (2015). The Rate of Osmotic Down-shock Determines the Survival Probability of Bacterial Mechanosensitive Channel Mutants. *Journal of Bacteriology* 197, 231–237.
- Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. (2005a). Transcriptional regulation by the numbers: Models. *Current Opinion in Genetics & Development* 15, 116–124.
- Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., Kuhlman, T., and Phillips, R. (2005b). Transcriptional regulation by the numbers: Applications. *Current Opinion in Genetics & Development* 15, 125–135.
- Blount, P., Sukharev, S.I., Schroeder, M.J., Nagle, S.K., and Kung, C. (1996). Single residue substitutions that change the gating properties of a mechanosensitive channel in *Escherichia Coli*. *Proc Natl Acad Sci U S A* 93, 11652–11657.
- Blount, P., Sukharev, S.I., Moe, P.C., Martinac, B., and Kung, C. (1999). Mechanosensitive channels of bacteria. *Methods in Enzymology* 294, 458–482.
- Bochner, B.R., Huang, H.-C., Schieven, G.L., and Ames, B.N. (1980). Positive selection for loss of tetracycline resistance. *Journal of Bacteriology* 143, 926–933.
- Boedicker, J.Q., Garcia, H.G., Johnson, S., and Phillips, R. (2013a). DNA sequence-dependent mechanics and protein-assisted bending in repressor-mediated loop formation. *Physical Biology* 10, 066005.
- Boedicker, J.Q., Garcia, H.G., and Phillips, R. (2013b). Theoretical and Experimental Dissection of DNA Loop-Mediated Repression. *Physical Review Letters* 110, 018101.
- Booth, I.R. (2014). Bacterial mechanosensitive channels: Progress towards an understanding of their roles in cell physiology. *Current Opinion in Microbiology* 18,

16–22.

Booth, I.R., Edwards, M.D., Murray, E., and Miller, S. (2005). The role of bacterial ion channels in cell physiology. In *Bacterial Ion Channels and Their Eukaryotic Homologs*, A. Kubalsi, and B. Martinac, eds. (Washington DC: American Society for Microbiology), pp. 291–312.

Boulton, S., and Melacini, G. (2016). Advances in NMR Methods To Map Allosteric Sites: From Models to Translation. *Chemical Reviews* *116*, 6267–6304.

Brewster, R.C., Jones, D.L., and Phillips, R. (2012). Tuning Promoter Strength through RNA Polymerase Binding Site Design in *Escherichia coli*. *PLoS Computational Biology* *8*, e1002811.

Brewster, R.C., Weinert, F.M., Garcia, H.G., Song, D., Rydenfelt, M., and Phillips, R. (2014). The Transcription Factor Titration Effect Dictates Level of Gene Expression. *Cell* *156*, 1312–1323.

Brophy, J.A.N., and Voigt, C.A. (2014). Principles of genetic circuit design. *Nature Methods* *11*, 508–520.

Buchler, N.E., Gerland, U., and Hwa, T. (2003). On schemes of combinatorial transcription logic. *Proceedings of the National Academy of Sciences* *100*, 5136–5141.

Canals, M., Lane, J.R., Wen, A., Scammells, P.J., Sexton, P.M., and Christopoulos, A. (2012). A Monod-Wyman-Changeux Mechanism Can Explain G Protein-coupled Receptor (GPCR) Allosteric Modulation. *Journal of Biological Chemistry* *287*, 650–659.

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software* *76*, 1–32.

Cass, J.A., Stylianidou, S., Kuwada, N.J., Traxler, B., and Wiggins, P.A. (2017). Probing bacterial cell biology using image cytometry. *Molecular Microbiology* *103*, 818–828.

- Cheng, W., and Hüllermeier, E. (2009). Combining instance-based learning and logistic regression for multilabel classification. *Machine Learning* 76, 211–225.
- Chure, G., Lee, H.J., and Phillips, R. (2018). MCMC chains generated in "Connecting the Dots between Mechanosensitive Channel Abundance, Osmotic Shock, and Survival at Single-Cell Resolution" accessible through DOI 10.22002/D1.942.
- Chure, G., Razo-Mejia, M., Belliveau, N.M., Einav, T., Kaczmarek, Z.A., Barnes, S.L., and Phillips, R. (2019). Predictive Shifts in Free Energy Couple Mutations to Their Phenotypic Consequences. *Proceedings of the National Academy of Sciences* 116.
- Colin, R., and Sourjik, V. (2017). Emergent properties of bacterial chemotaxis pathway. *Current Opinion in Microbiology* 39, 24–33.
- Cruickshank, C.C., Minchin, R.F., Le Dain, A.C., and Martinac, B. (1997). Estimation of the pore size of the large-conductance mechanosensitive ion channel of *Escherichia Coli*. *Biophysical Journal* 73, 1925–1931.
- Czaran, T.L., Hoekstra, R.F., and Pagie, L. (2002). Chemical warfare between microbes promotes biodiversity. *Proceedings of the National Academy of Sciences* 99, 786–790.
- Daber, R., Sharp, K., and Lewis, M. (2009). One Is Not Enough. *Journal of Molecular Biology* 392, 1133–1144.
- Daber, R., Sochor, M.A., and Lewis, M. (2011). Thermodynamic Analysis of Mutant lac Repressors. *Journal of Molecular Biology* 409, 76–87.
- Diénert, F. (1900). Sur la fermentation du galactose et sur l'accoutumance des levures à ce sucre, (Sceaux).
- Dill, K.A., and Bromberg, S. (2010). *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology* (New York: Garland Science).
- Downey, A. (2014). Probably Overthinking It: Bayes's theorem and logistic regression. *Probably Overthinking It*.

- Dreiseitl, S., and Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics* 35, 352–359.
- Driessen, R.P.C., Sitters, G., Laurens, N., Moolenaar, G.F., Wuite, G.J.L., Goosen, N., and Dame, R.T. (2014). Effect of Temperature on the Intrinsic Flexibility of DNA and Its Interaction with Architectural Proteins. *Biochemistry* 53, 6430–6438.
- Edelstein, A.D., Tsuchida, M.A., Amodaj, N., Pinkard, H., Vale, R.D., and Stuurman, N. (2014). Advanced methods of microscope control using μ Manager software. *Journal of Biological Methods* 1, 10.
- Edwards, M.D., Black, S., Rasmussen, T., Rasmussen, A., Stokes, N.R., Stephen, T.L., Miller, S., and Booth, I.R. (2012). Characterization of three novel mechanosensitive channel activities in *Escherichia Coli*. *Channels (Austin)* 6, 272–281.
- Einav, T., and Phillips, R. (2017). Monod-Wyman-Changeux Analysis of Ligand-Gated Ion Channel Mutants. –.
- Einav, T., Mazutis, L., and Phillips, R. (2016). Statistical Mechanics of Allosteric Enzymes. *The Journal of Physical Chemistry B* 120, 6021–6037.
- Eldar, A., and Elowitz, M.B. (2010). Functional roles for noise in genetic circuits. *Nature* 467, 167–173.
- Elf, J., Li, G.-W., and Xie, X.S. (2007). Probing Transcription Factor Dynamics at the Single-Molecule Level in a Living Cell. *Science* 316, 1191–1194.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. *Science* 297, 1183–1186.
- Espah Borujeni, A., Channarasappa, A.S., and Salis, H.M. (2014). Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Research* 42, 2646–2659.
- Feeling-Taylor, A.R., Yau, S.-T., Petsev, D.N., Nagel, R.L., Hirsch, R.E., and Vekilov, P.G. (2004). Crystallization Mechanisms of Hemoglobin C in the R State. *Biophys-*

ical Journal 87, 2621–2629.

Fernández-Castané, A., Vine, C.E., Caminal, G., and López-Santín, J. (2012). Ev-
idencing the role of lactose permease in IPTG uptake by Escherichia coli in fed-
batch high cell density cultures. Journal of Biotechnology 157, 391–398.

Finch, J.T., Perutz, M.F., Bertles, J.F., and Dobler, J. (1973). Structure of Sickled
Erythrocytes and of Sickle-Cell Hemoglobin Fibers. Proceedings of the National
Academy of Sciences 70, 718–722.

Forsén, S., and Linse, S. (1995). Cooperativity: Over the Hill. Trends in Biochemical
Sciences 20, 495–497.

Friedel, J. (1974). On the stability of the body centred cubic phase in metals at high
temperatures. Journal de Physique Lettres 35, 59–63.

Frumkin, I., Lajoie, M.J., Gregg, C.J., Hornung, G., Church, G.M., and Pilpel, Y.
(2018). Codon usage of highly expressed genes affects proteome-wide translation
efficiency. Proceedings of the National Academy of Sciences 115, E4940–E4949.

Galperin, M.Y., Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2015). Expanded
microbial genome coverage and improved protein family annotation in the COG
database. Nucleic Acids Research 43, D261–269.

Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-
Rascado, L., García-Sotelo, J.S., Alquicira-Hernández, K., Martínez-Flores, I., Pan-
nier, L., Castro-Mondragón, J.A., et al. (2016). RegulonDB version 9.0: High-level
integration of gene regulation, coexpression, motif clustering and beyond. Nucleic
Acids Research 44, D133–D143.

Garcia, H.G., and Phillips, R. (2011). Quantitative dissection of the simple repres-
sion input-output function. Proceedings of the National Academy of Sciences 108,
12173–12178.

Garcia, H.G., Lee, H.J., Boedicker, J.Q., and Phillips, R. (2011a). Comparison and
Calibration of Different Reporters for Quantitative Analysis of Gene Expression.

- Biophysical Journal 101, 535–544.
- Garcia, H.G., Lee, H.J., Boedicker, J.Q., and Phillips, R. (2011b). Comparison and Calibration of Different Reporters for Quantitative Analysis of Gene Expression. Biophysical Journal 101, 535–544.
- Garcia, H.G., Sanchez, A., Boedicker, J.Q., Osborne, M., Gelles, J., Kondev, J., and Phillips, R. (2012). Operator Sequence Alters Gene Expression Independently of Transcription Factor Occupancy in Bacteria. Cell Reports 2, 150–161.
- Gardino, A.K., Volkman, B.F., Cho, H.S., Lee, S.-Y., Wemmer, D.E., and Kern, D. (2003). The NMR Solution Structure of BeF₃-Activated Spo0F Reveals the Conformational Switch in a Phosphorelay System. Journal of Molecular Biology 331, 245–254.
- Gerland, U., Moroz, J.D., and Hwa, T. (2002). Physical constraints and functional characteristics of transcription factor-DNA interaction. Proceedings of the National Academy of Sciences 99, 12015–12020.
- Gilbert, W., and Müller-Hill, B. (1966). ISOLATION OF THE LAC REPRESSOR. Proceedings of the National Academy of Sciences of the United States of America 56, 1891–1898.
- Goethe, M., Fita, I., and Rubi, J.M. (2015). Vibrational Entropy of a Protein: Large Differences between Distinct Conformations. Journal of Chemical Theory and Computation 11, 351–359.
- Good, I.J. (1950). Probability and the Weighting of Evidence (New York City).
- Harman, J.G. (2001). Allosteric regulation of the cAMP receptor protein. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1547, 1–17.
- Hase, C.C., Minchin, R.F., Kloda, A., and Martinac, B. (1997). Cross-linking studies and membrane localization and assembly of radiolabelled large mechanosensitive ion channel (MscL) of *Escherichia Coli*. Biochem Biophys Res Commun 232, 777–

782.

- Haswell, E.S., Phillips, R., and Rees, D.C. (2011). Mechanosensitive Channels: What Can They Do and How Do They Do It? *Structure* *19*, 1356–1369.
- Herbig, U., Jobling, W.A., Chen, B.P.C., Chen, D.J., and Sedivy, J.M. (2004). Telomere Shortening Triggers Senescence of Human Cells through a Pathway Involving ATM, p53, and p21CIP1, but Not p16INK4a. *Molecular Cell* *14*, 501–513.
- Huang, M., Song, K., Liu, X., Lu, S., Shen, Q., Wang, R., Gao, J., Hong, Y., Li, Q., Ni, D., et al. (2018). AlloFinder: A strategy for allosteric modulator discovery and allosterome analyses. *Nucleic Acids Research* *46*, W451–W458.
- Hui, S., Silverman, J.M., Chen, S.S., Erickson, D.W., Basan, M., Wang, J., Hwa, T., and Williamson, J.R. (2015). Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology* *11*.
- Jones, D.L., Brewster, R.C., and Phillips, R. (2014). Promoter architecture dictates cell-to-cell variability in gene expression. *Science* *346*, 1533–1536.
- Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiologyHistory, recent progress, and the future with focus on cell size control: A review. *Reports on Progress in Physics* *81*, 056601.
- Kao-Huang, Y., Revzin, A., Butler, A.P., O'Conner, P., Noble, D.W., and Hippel, P.H.V. (1977). Nonspecific DNA binding of genome-regulating proteins as a biological control mechanism: Measurement of DNA-bound *Escherichia coli* lac repressor in vivo. *Proceedings of the National Academy of Sciences* *74*, 4228–4232.
- Kepler, T.B., and Elston, T.C. (2001). Stochasticity in Transcriptional Regulation: Origins, Consequences, and Mathematical Representations. *Biophysical Journal* *81*, 3116–3136.
- Keymer, J.E., Endres, R.G., Skoge, M., Meir, Y., and Wingreen, N.S. (2006). Chemosensing in *Escherichia coli*: Two regimes of two-state receptors. *Proceedings of the National Academy of Sciences* *103*, 1786–1791.

- Kim, N.H., Lee, G., Sherer, N.A., Martini, K.M., Goldenfeld, N., and Kuhlman, T.E. (2016). Real-time transposable element activity in individual live cells. *Proceedings of the National Academy of Sciences* *113*, 7278–7283.
- Kim, P.-J., Lee, D.-Y., Kim, T.Y., Lee, K.H., Jeong, H., Lee, S.Y., and Park, S. (2007). Metabolite essentiality elucidates robustness of *Escherichia coli* metabolism. *Proceedings of the National Academy of Sciences of the United States of America* *104*, 13638–13642.
- Klumpp, S., and Hwa, T. (2008). Growth-rate-dependent partitioning of RNA polymerases in bacteria. *Proceedings of the National Academy of Sciences* *105*, 20245–20250.
- Klumpp, S., and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition. *Current Opinion in Biotechnology* *28*, 96–102.
- Ko, M.S.H. (1991). A stochastic model for gene induction. *Journal of Theoretical Biology* *153*, 181–194.
- Krembel, A., Colin, R., and Sourjik, V. (2015a). Importance of Multiple Methylation Sites in *Escherichia Coli* Chemotaxis. *PLoS ONE* *10*.
- Krembel, A.K., Neumann, S., and Sourjik, V. (2015b). Universal Response-Adaptation Relation in Bacterial Chemotaxis. *Journal of Bacteriology* *197*, 307–313.
- Kuhlman, T., Zhang, Z., Saier, M.H., and Hwa, T. (2007). Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. *Proceedings of the National Academy of Sciences* *104*, 6043–6048.
- Lanfranco, M.F., Gárate, F., Engdahl, A.J., and Maillard, R.A. (2017). Asymmetric configurations in a reengineered homodimer reveal multiple subunit communication pathways in protein allostery. *Journal of Biological Chemistry* *292*, 6086–6093.
- Laxhuber, K.S., Morrison, M., Chure, G., Belliveau, N.M., Strandkvist, C., Naughton, K., and Phillips, R. (2020). Theoretical Investigation of a Genetic Switch for Metabolic

Adaptation. PLOS ONE.

Lässig, M., Mustonen, V., and Walczak, A.M. (2017). Predicting evolution. *Nature Ecology & Evolution* 1, 0077.

Levantino, M., Spilotros, A., Cammarata, M., Schirò, G., Ardiccioni, C., Vallone, B., Brunori, M., and Cupane, A. (2012). The Monod-Wyman-Changeux allosteric model accounts for the quaternary transition dynamics in wild type and a recombinant mutant human hemoglobin. *Proceedings of the National Academy of Sciences* 109, 14894–14899.

Levina, N., Totemeyer, S., Stokes, N.R., Louis, P., Jones, M.A., and Booth, I.R. (1999). Protection of *Escherichia Coli* cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: Identification of genes required for MscS activity. *EMBO J* 18, 1730–1737.

Lewis, M., Chang, G., Horton, N.C., Kercher, M.A., Pace, H.C., Schumacher, M.A., Brennan, R.G., and Lu, P. (1996). Crystal Structure of the Lactose Operon Repressor and Its Complexes with DNA and Inducer. *Science* 271, 1247–1254.

Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. *Cell* 157, 624–635.

Li, X.-t., Thomason, L.C., Sawitzke, J.A., Costantino, N., and Court, D.L. (2013). Positive and negative selection using the tetA-sacB cassette: Recombineering and P1 transduction in *Escherichia Coli*. *Nucleic Acids Research* 41, e204–e204.

Lindsley, J.E., and Rutter, J. (2006). Whence cometh the allosterome? *Proceedings of the National Academy of Sciences* 103, 10533–10535.

Liu, F., Morrison, A.H., and Gregor, T. (2013). Dynamic interpretation of maternal inputs by the *Drosophila* segmentation gene network. *Proceedings of the National Academy of Sciences of the United States of America* 110, 6724–6729.

Loison, L. (2013). Monod before Monod: Enzymatic Adaptation, Lwoff, and the

Legacy of General Biology. History and Philosophy of the Life Sciences 35, 167–192.

Louhivuori, M., Risselada, H.J., van der Giessen, E., and Marrink, S.J. (2010). Release of content through mechano-sensitive gates in pressurized liposomes. Proc Natl Acad Sci U S A 107, 19856–19860.

Lovely, G.A., Brewster, R.C., Schatz, D.G., Baltimore, D., and Phillips, R. (2015). Single-molecule analysis of RAG-Mediated V(D)J DNA cleavage. Proceedings of the National Academy of Sciences 112, E1715–E1723.

Lutz, R., and Bujard, H. (1997). Independent and Tight Regulation of Transcriptional Units in Escherichia Coli Via the LacR/O, the TetR/O and AraC/I1-I2 Regulatory Elements. Nucleic Acids Research 25, 1203–1210.

Martinac, B., Buechner, M., Delcour, A.H., Adler, J., and Kung, C. (1987). Pressure-sensitive ion channel in *Escherichia Coli*. Proc Natl Acad Sci U S A 84, 2297–2301.

Martins, B.M.C., and Swain, P.S. (2011). Trade-Offs and Constraints in Allosteric Sensing. PLOS Computational Biology 7, e1002261.

Martínez-Gómez, K., Flores, N., Castañeda, H.M., Martínez-Batallar, G., Hernández-Chávez, G., Ramírez, O.T., Gosset, G., Encarnación, S., and Bolívar, F. (2012). New insights into *Escherichia coli* metabolism: Carbon scavenging, acetate metabolism and carbon recycling responses during growth on glycerol. Microbial Cell Factories 11, 46.

Marzen, S., Garcia, H.G., and Phillips, R. (2013). Statistical Mechanics of Monod-Wyman-Changeux (MWC) Models. Journal of Molecular Biology 425, 1433–1460.

Michel, D. (2010). How transcription factors can adjust the gene expression flood-gates. Progress in Biophysics and Molecular Biology 102, 16–37.

Milo, R., Hou, J.H., Springer, M., Brenner, M.P., and Kirschner, M.W. (2007). The relationship between evolutionary and physiological variation in hemoglobin. Proceedings of the National Academy of Sciences 104, 16998–17003.

- Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Research* *38*, D750–D753.
- Mirny, L.A. (2010). Nucleosome-mediated cooperativity between transcription factors. *Proceedings of the National Academy of Sciences* *107*, 22534–22539.
- Mishra, V., Skotak, M., Schuetz, H., Heller, A., Haorah, J., and Chandra, N. (2016). Primary blast causes mild, moderate, severe and lethal TBI with increasing blast overpressures: Experimental rat injury model. *Scientific Reports* *6*, 26992.
- Mondal, J., Bratton, B.P., Li, Y., Yethiraj, A., and Weisshaar, J.C. (2011). Entropy-Based Mechanism of Ribosome-Nucleoid Segregation in *E. coli* Cells. *Biophysical Journal* *100*, 2605–2613.
- Monod, J. (1941). Sur un phénomène nouveau de croissance complexe dans les cultures bactériennes. *Comptes Rendus Des Séances de L'Académie Des Sciences* *212*, 934–939.
- Monod, J. (1947). The Phenomenon of Enzymatic Adaptation And Its bearings on Problems of Genetics and Cellular Differentiation. *Growth Symposium* *9*, 223–289.
- Monod, J. (1966). From Enzymatic Adaptation to Allosteric Transitions. *Science* *154*, 475–483.
- Monod, J., Changeux, J.-P., and Jacob, F. (1963). Allosteric proteins and cellular control systems. *Journal of Molecular Biology* *6*, 306–329.
- Monod, J., Wyman, J., and Changeux, J.-P. (1965). On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology* *12*, 88–118.
- Moon, T.S., Lou, C., Tamsir, A., Stanton, B.C., and Voigt, C.A. (2012). Genetic programs constructed from layered logic gates in single cells. *Nature* *491*, 249–253.
- Motlagh, H.N., Wrabl, J.O., Li, J., and Hilser, V.J. (2014). The ensemble nature of allostery. *Nature* *508*, 331–339.

- Murphy, K.F., Balázsi, G., and Collins, J.J. (2007). Combinatorial promoter design for engineering noisy gene expression. *Proceedings of the National Academy of Sciences* *104*, 12726–12731.
- Nagai, T., Ibata, K., Park, E.S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature Biotechnology* *20*, 87–90.
- Naismith, J.H., and Booth, I.R. (2012). Bacterial mechanosensitive channelsMscS: Evolution's solution to creating sensitivity in function. *Annu Rev Biophys* *41*, 157–177.
- Nayak, C.R., and Rutenberg, A.D. (2011). Quantification of Fluorophore Copy Number from Intrinsic Fluctuations during Fluorescence Photobleaching. *Bioophysical Journal* *101*, 2284–2293.
- Nordström, K., and Dasgupta, S. (2006). Copy-number control of the Escherichia coli chromosome: A plasmidologist's view. *EMBO Reports* *7*, 484–489.
- Oehler, S., Amouyal, M., Kolkhof, P., von Wilcken-Bergmann, B., and Müller-Hill, B. (1994). Quality and position of the three lac operators of E. Coli define efficiency of repression. *The EMBO Journal* *13*, 3348–3355.
- O'Gorman, R.B., Rosenberg, J.M., Kallai, O.B., Dickerson, R.E., Itakura, K., Riggs, A.D., and Matthews, K.S. (1980). Equilibrium binding of inducer to lac repressor operator DNA complex. *Journal of Biological Chemistry* *255*, 10107–10114.
- Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in Escherichia coli with increasing specific growth rate. *Molecular BioSystems* *11*, 1184–1193.
- Perutz, M.F., and Mitchison, J.M. (1950). State of Hæmoglobin in Sickle-Cell Anæmia. *Nature* *166*, 677–679.
- Phillips, R. (2001). Crystals, Defects and Microstructures by Rob Phillips ([/core/books/crystals-defects-and-microstructures/85B8CE3A333532ABC0C21CC7E2096B50](http://core/books/crystals-defects-and-microstructures/85B8CE3A333532ABC0C21CC7E2096B50)).

- Phillips, R. (2015). Napoleon Is in Equilibrium. *Annual Review of Condensed Matter Physics* *6*, 85–111.
- Phillips, R., Belliveau, N.M., Chure, G., Garcia, H.G., Razo-Mejia, M., and Scholes, C. (2019). Figure 1 Theory Meets Figure 2 Experiments in the Study of Gene Expression. *Annual Review of Biophysics* *48*, 121–163.
- Pilizota, T., and Shaevitz, J.W. (2012). Fast, Multiphase Volume Adaptation to Hyperosmotic Shock by *Escherichia coli*. *PLOS ONE* *7*, e35205.
- Pilizota, T., and Shaevitz, J.W. (2014). Origins of *Escherichia coli* Growth Rate and Cell Shape Changes at High External Osmolality. *Biophysical Journal* *107*, 1962–1969.
- Poelwijk, F.J., Heyning, P.D., de Vos, M.G., Kiviet, D.J., and Tans, S.J. (2011). Optimality and evolution of transcriptionally regulated gene expression. *BMC Systems Biology* *5*, 128.
- Raman, A.S., White, K.I., and Ranganathan, R. (2016). Origins of Allostery and Evolvability in Proteins: A Case Study. *Cell* *166*, 468–480.
- Razo-Mejia, M., Boedicker, J.Q., Jones, D., DeLuna, A., Kinney, J.B., and Phillips, R. (2014). Comparison of the theoretical and real-world evolutionary potential of a genetic circuit. *Physical Biology* *11*, 026005.
- Razo-Mejia, M., Barnes, S.L., Belliveau, N.M., Chure, G., Einav, T., Lewis, M., and Phillips, R. (2018). Tuning Transcriptional Regulation through Signaling: A Predictive Theory of Allosteric Induction. *Cell Systems* *6*, 456–469.e10.
- Razo-Mejia, M., Marzen, S., Chure, G., Morrison, M., Taubman, R., and Phillips, R. (2020). First-principles prediction of the information processing capacity of a simple genetic circuit. In Preparation.
- Reynolds, K.A., McLaughlin, R.N., and Ranganathan, R. (2011). Hot Spots for Allosteric Regulation on Protein Surfaces. *Cell* *147*, 1564–1575.
- Rogers, J.K., Guzman, C.D., Taylor, N.D., Raman, S., Anderson, K., and Church,

- G.M. (2015). Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Research* *43*, 7648–7660.
- Rohlhill, J., Sandoval, N.R., and Papoutsakis, E.T. (2017). Sort-Seq Approach to Engineering a Formaldehyde-Inducible Promoter for Dynamically Regulated *Escherichia coli* Growth on Methanol. *ACS Synthetic Biology* *6*, 1584–1595.
- Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* *323*, 785–793.
- Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., and Elowitz, M.B. (2005). Gene Regulation at the Single-Cell Level. *Science* *307*, 1962–1965.
- Rosenfeld, N., Perkins, T.J., Alon, U., Elowitz, M.B., and Swain, P.S. (2006). A Fluctuation Method to Quantify In Vivo Fluorescence Data. *Biophysical Journal* *91*, 759–766.
- Rydenfelt, M., Cox, R.S., Garcia, H., and Phillips, R. (2014a). Statistical mechanical model of coupled transcription from multiple promoters due to transcription factor titration. *Physical Review E* *89*, 012702.
- Rydenfelt, M., Garcia, H.G., Cox, R.S., and Phillips, R. (2014b). The Influence of Promoter Architectures and Regulatory Motifs on Gene Expression in *Escherichia coli*. *PLoS ONE* *9*, e114347.
- Salis, H.M., Mirsky, E.A., and Voigt, C.A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology* *27*, 946–950.
- Schaechter, M., Maaløe, O., and Kjeldgaard, N.O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology* *19*, 592–606.
- Schatz, D.G., and Baltimore, D. (2004). Uncovering the V(D)J recombinase. *Cell* *116*, S103–S108.

- Schatz, D.G., and Ji, Y. (2011). Recombination centres and the orchestration of V(D)J recombination. *Nature Reviews Immunology* *11*, 251–263.
- Schlake, T., and Bode, J. (1994). Use of Mutated FLP Recognition Target (FRT) Sites for the Exchange of Expression Cassettes at Defined Chromosomal Loci. *Biochemistry* *33*, 12746–12751.
- Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann, M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology* *34*, 104–110.
- Schumann, U., Edwards, M.D., Rasmussen, T., Bartlett, W., van West, P., and Booth, I.R. (2010). YbdG in *Escherichia Coli* is a threshold-setting mechanosensitive channel with MscM activity. *Proc Natl Acad Sci U S A* *107*, 12664–12669.
- Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z., and Hwa, T. (2010). Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science* *330*, 1099–1102.
- Scott, M., Klumpp, S., Mateescu, E.M., and Hwa, T. (2014). Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Molecular Systems Biology* *10*, 747–747.
- Setty, Y., Mayo, A.E., Surette, M.G., and Alon, U. (2003). Detailed map of a cis-regulatory input function. *Proceedings of the National Academy of Sciences* *100*, 7702–7707.
- Sharan, S.K., Thomason, L.C., Kuznetsov, S.G., and Court, D.L. (2009). Recombiengineering: A homologous recombination-based method of genetic engineering. *Nat Protoc* *4*, 206–223.
- Shehata, T.E., and Marr, A.G. (1975). Effect of temperature on the size of *Escherichia coli* cells. *Journal of Bacteriology* *124*, 857–862.
- Shis, D.L., Hussain, F., Meinhardt, S., Swint-Kruse, L., and Bennett, M.R. (2014).

- Modular, Multi-Input Transcriptional Logic Gating with Orthogonal LacI/GalR Family Chimeras. *ACS Synthetic Biology* 3, 645–651.
- Sivia, D., and Skilling, J. (2006). *Data Analysis: A Bayesian Tutorial* (OUP Oxford).
- Sochor, M.A. (2014). In vitro transcription accurately predicts lac repressor phenotype in vivo in *Escherichia coli*. *PeerJ* 2, e498.
- Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth and ethanol stress. *Frontiers in Microbiology* 6.
- Sourjik, V., and Berg, H.C. (2002). Receptor sensitivity in bacterial chemotaxis. *Proceedings of the National Academy of Sciences* 99, 123–127.
- Stahler, G.J., Mennis, J., Belenko, S., Welsh, W.N., Hiller, M.L., and Zajac, G. (2013). Predicting Recidivism For Released State Prison Offenders. *Criminal Justice and Behavior* 40, 690–711.
- Stokes, N.R., Murray, H.D., Subramaniam, C., Gourse, R.L., Louis, P., Bartlett, W., Miller, S., and Booth, I.R. (2003). A role for mechanosensitive channels in survival of stationary phase: Regulation of channel expression by RpoS. *Proceedings of the National Academy of Sciences* 100, 15959–15964.
- Swain, P.S., Stevenson, K., Leary, A., Montano-Gutierrez, L.F., Clark, I.B.N., Vogel, J., and Pilizota, T. (2016). Inferring time derivatives including cell growth rates using Gaussian processes. *Nature Communications* 7, 13766.
- Swem, L.R., Swem, D.L., Wingreen, N.S., and Bassler, B.L. (2008). Deducing Receptor Signaling Parameters from In Vivo Analysis: LuxN/AI-1 Quorum Sensing in *Vibrio harveyi*. *Cell* 134, 461–473.
- Taheri-Araghi, S., Bradde, S., Sauls, J.T., Hill, N.S., Levin, P.A., Paulsson, J., Vergasola, M., and Jun, S. (2015a). Cell-size control and homeostasis in bacteria. *Current Biology : CB* 25, 385–391.
- Taheri-Araghi, S., Bradde, S., Sauls, J.T., Hill, N.S., Levin, P.A., Paulsson, J., Vergas-

sola, M., and Jun, S. (2015a). Cell-size control and homeostasis in bacteria. *Current Biology* : CB 25, 385–391.

Talts, S., Betancourt, M., Simpson, D., Vehtari, A., and Gelman, A. (2018). Validating Bayesian Inference Algorithms with Simulation-Based Calibration. arXiv:1804.06788 [Stat].

Thomason, L.C., Costantino, N., and Court, D.L. (2007). E. Coli Genome Manipulation by P1 Transduction. *Current Protocols in Molecular Biology* 79, 1.17.1–1.17.8.

Tungtur, S., Skinner, H., Zhan, H., Swint-Kruse, L., and Beckett, D. (2011). In vivo tests of thermodynamic models of transcription repressor function. *Biophysical Chemistry* 159, 142–151.

Ullmann, A. (2011). In Memoriam: Jacques Monod (19101976). *Genome Biology and Evolution* 3, 1025–1033.

Ursell, T., Phillips, R., Konddev, J., Reeves, D., and Wiggins, P.A. (2008). The role of lipid bilayer mechanics in mechanosensation. In *Mechanosensitivity in Cells and Tissues 1: Mechanosensitive Ion Channels*, A. Kamkin, and I. Kiseleva, eds. (Springer-Verlag), pp. 37–70.

Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). Escherichia coli achieves faster growth by increasing catalytic and translation rates of proteins. *Molecular BioSystems* 9, 2344.

van den Berg, J., Galbiati, H., Rasmussen, A., Miller, S., and Poolman, B. (2016). On the mobility, membrane location and functionality of mechanosensitive channels in *Escherichia Coli*. *Scientific Reports* 6.

Velyvis, A., Yang, Y.R., Schachman, H.K., and Kay, L.E. (2007). A solution NMR study showing that active site ligands and nucleotides directly perturb the allosteric equilibrium in aspartate transcarbamoylase. *Proceedings of the National Academy of Sciences* 104, 8815–8820.

Victorelli, S., and Passos, J.F. (2017). Telomeres and Cell Senescence - Size Matters

Not. EBioMedicine 21, 14–20.

Vilar, J.M.G., and Leibler, S. (2003). DNA Looping and Physical Constraints on Transcription Regulation. *Journal of Molecular Biology* 331, 981–989.

Vilar, J.M.G., and Saiz, L. (2013). Reliable Prediction of Complex Phenotypes from a Modular Design in Free Energy Space: An Extensive Exploration of the lac Operon. *ACS Synthetic Biology* 2, 576–586.

Weinert, F.M., Brewster, R.C., Rydenfelt, M., Phillips, R., and Kegel, W.K. (2014). Scaling of Gene Expression with Transcription-Factor Fugacity. *Physical Review Letters* 113.

Yakovchuk, P., Protozanova, E., and Frank-Kamenetskii, M.D. (2006). Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Research* 34, 564–574.

Zeldovich, K.B., and Shakhnovich, E.I. (2008). Understanding Protein Evolution: From Protein Physics to Darwinian Selection. *Annual Review of Physical Chemistry* 59, 105–127.

Zhu, X.-D., and Sadowski, P.D. (1995). Cleavage-dependent Ligation by the FLP Recombinase CHARACTERIZATION OF A MUTANT FLP PROTEIN WITH AN ALTERATION IN A CATALYTIC AMINO ACID. *Journal of Biological Chemistry* 270, 23044–23054.