CONTENTS

1 IT'S AN ALLOSTERIC WORLD

> What's in a name? That which we call a rose by any other word would smell as sweet.
>
> —William Shakespeare

1.1 The Second Secret of Life

The 1953 discovery of the structure of DNA ushered in the molecular era in biology with a vengeance. As with many other great discoveries, the determination of the molecular basis of heredity spawned a host of new questions. One of the dominant mysteries was the nature of regulation. How are the many molecules of the cell (including DNA itself) regulated so that they carry out their functions both when and where they are needed and not otherwise? Such questions arise in all corners of biology. In the context of metabolism, it was clear that bacterial cells rank-order their preferences for different carbon sources, raising the question of how the cell acts on these preferences. The study of enzymes revealed that some enzymes are active in the absence of some inhibitor and are shut down in its presence. Animal body plans are set up by particular spatiotemporal patterns of gene expression, making it clear that whole batteries of genes are switching between different states. These examples and many others reveal that regulation is one of the most widespread molecular processes in all of biology.

Ten short years after the secrets of the great molecule of heredity were uncovered, a second molecular discovery of central importance to biology was announced. That discovery, the formulation of the allostery concept, is relevant to thinking about the function of molecules across the entire domain of biological inquiry. Stated simply, the allostery concept harkens back to the Roman deity Janus, shown in Figure 1.1(A), symbolized by his two faces and noted for presiding over all kinds of transitions. In this book, we will take a broad view of allostery as the phenomenon in which a molecule has more than one state of activity, as shown in Figure 1.1(B), with the relative probabilities of those different states controlled by some effector(s).

One of the discoverers of the allostery concept, Jacques Monod, referred to this discovery as the "second secret of life." But what exactly was this secret? We will refine our answer to that question through the various case studies that make up the chapters that follow, though here we give a brief qualitative sketch

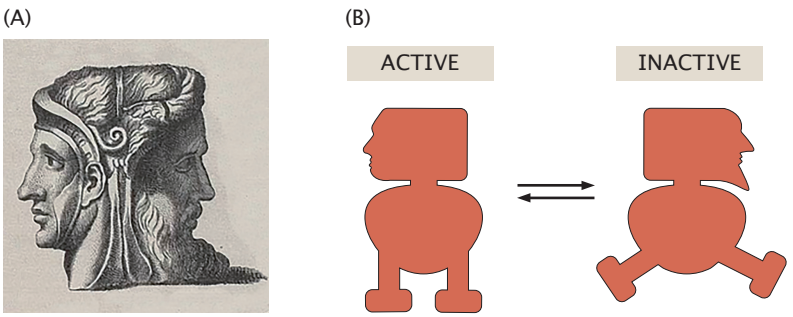


Figure 1.1

Allostery defined. (A) The Roman god Janus. (B) Molecules such as transcriptional repressors have a Janus-like existence as they switch between active and inactive conformations.

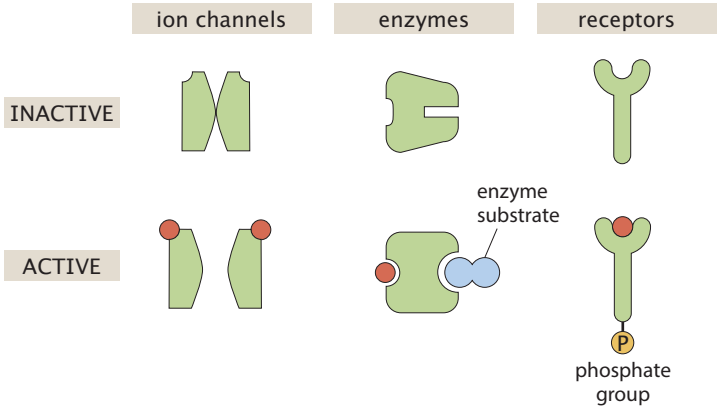


Figure 1.2

The molecular switch. Different classes of molecules exemplify the allostery phenomenon, highlighting the inactive and active conformational states of these molecules. Ion channels can be either closed or open, with the binding of a ligand favoring the open state. An enzyme can be in an inactive state, in which it is unable to cleave a substrate, or in an active state, in which it is competent to perform such a cleavage reaction. The presence of an effector (red) favors the active state. A membrane-bound receptor can be in either an inactive state or an active state when it is bound to a ligand, in which it can perform a phosphorylation reaction leading to a subsequent signaling cascade. For all the examples shown here, the ligands shown in red tip the balance in favor of the active conformation over the inactive conformation.

of the key elements of the concept. Stated simply, many biological molecules behave as molecular switches. In its most basic incarnation, the idea is that many biological molecules have two distinct conformations, which we will often think of as inactive and active, as shown in Figure 1.2. As a result of the exchange of energy with the molecules of the surrounding solution (i.e., thermal energy), these molecules are constantly flipping back and forth between these two conformational states. In equilibrium, the relative proportion of these two states is fixed by the energy difference between them. However, the interesting regulatory behavior of these molecules is that the binding of a ligand can change the relative probabilities of the inactive and active states. Specifically, the binding affinity of the ligand for each state is different, resulting in a shift in the relative probabilities of the inactive and the active conformations when the ligand concentration is changed. The outcome is that a ligand can serve to regulate when molecules like those shown in Figure 1.2 are active. Our task is to explore different biological phenomena that are controlled by such molecules and to examine what physical models of these molecules have to say about their function.

1.2 The Broad Reach of the Allostery Concept

Biology is a science full of beautiful and fascinating exceptions. But the ease with which we can find such exceptions is not a proof of the absence of broad and overarching ideas. One such motif that has captured my imagination and

which serves as the basis of this book is that there is a unifying mathematical description of the way that many of the macromolecules of life can exist in several distinct states (see Figure 1.2, p. 4). For example, as we have already remarked, ion channels can be open or closed. Proteins can be phosphorylated or not. Receptors can be active or not. Often, which of these two different states is more likely depends in turn upon whether or not a particular ligand is bound to the molecule of interest. In such cases, by titrating the amount of ligand competing for the attentions of our molecule of interest, we can shift the balance between these two states. But only when viewed using equations rather than words and cartoons is the full impact of the allostery idea made clear. To foreshadow the kinds of phenomena that fall within the purview of the allostery concept, here we consider several illustrative examples.

Figure 1.3 shows how allosteric processes are central to many signaling pathways. As seen in the figure, both the membrane-bound receptors that receive signals and the soluble proteins that mediate the processes at the end of the signaling cascade are often themselves allosteric. For example, in the context of metabolism, as shown in the first panel of the figure, many enzymatic reactions are catalyzed by proteins that are subject to feedback such that the reaction occurs only when it is needed. This case illustrates a metabolic enzyme that is activated only in the presence of some effector molecule (shown in red). We will explore these processes more deeply in the next section.

Further, the processes of the central dogma of molecular biology such as transcription involve regulatory proteins. Transcription factors can be localized to the nucleus and bind DNA depending upon the presence or absence of some ligand, as shown in the middle panel of Figure 1.3. In the presence of inducer for the regulatory architecture shown here, the repressor protein is released from its binding site on the DNA, allowing for the expression of the regulated gene.

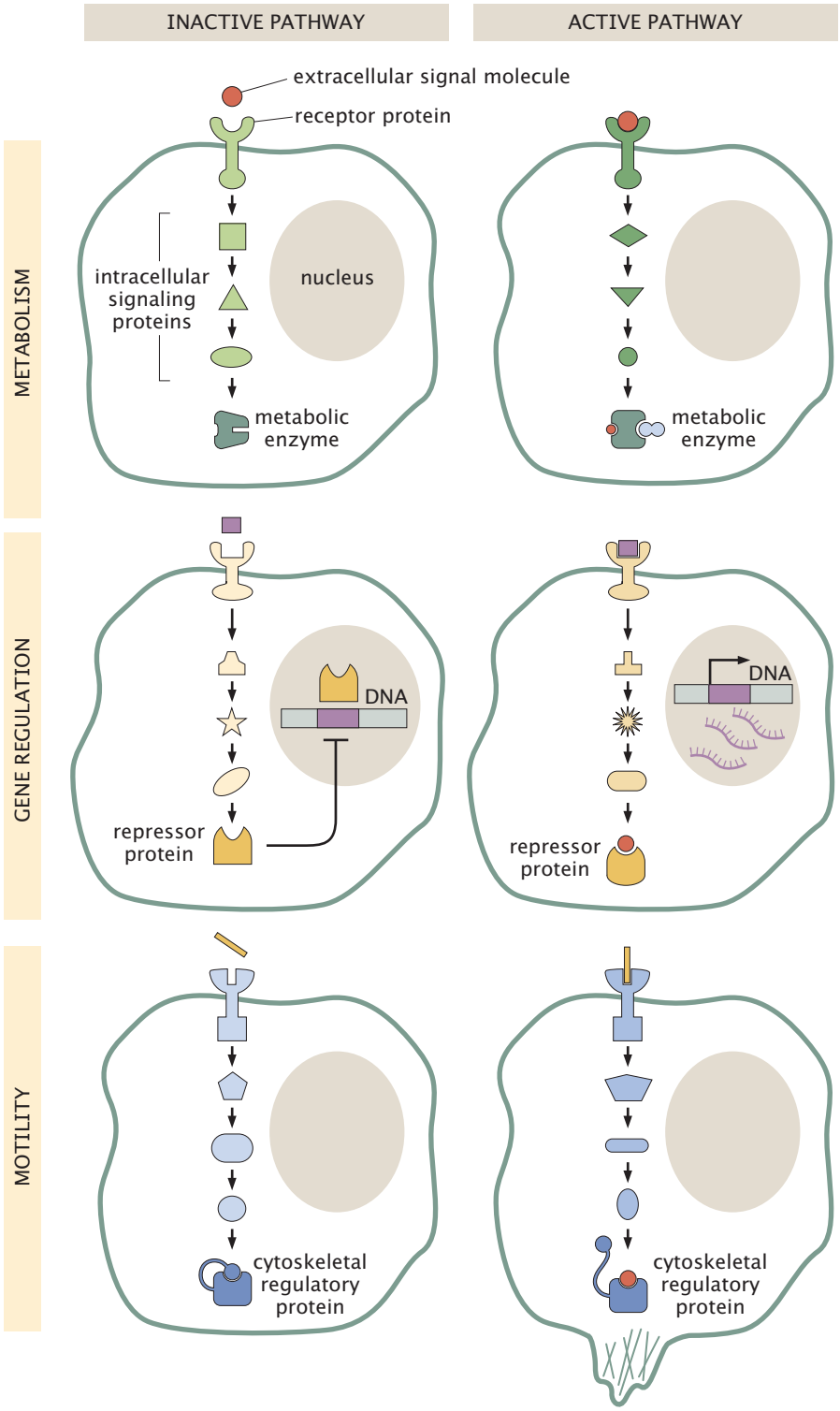
As indicated schematically in the final panel of the figure, activation of cytoskeletal growth and remodeling makes processes such as eukaryotic chemotaxis possible. Indeed, one of the most inspiring microscopy time-lapse sequences of all time shows a neutrophil engaged in the process of tracking down a bacterium. Several snapshots from this process are shown in Figure 1.4. In order for the cell to change direction, the leading edge has to be remodeled with new actin filaments synthesized in the correct direction of motion. To that end, these motile cells have an impressive signaling pathway that allows them to detect extracellular ligands, resulting in a subsequent molecular cascade within the cell that ends with the construction of actin filaments at the leading edge. A schematic representation of this pathway is given in Figure 1.5 and is mediated by molecules that can exist in several conformational states (i.e., allosteric) with different abilities to catalyze reactions.

1.2.1 Sculpting Biochemistry via Allostery

We can get a higher-resolution view of the ubiquitous nature of allosteric regulation by turning to one of the best-understood biochemical pathways, namely, that associated with carbon metabolism. Figure 1.6 gives a depiction of the key enzymes that mediate glycolysis, as well as their various substrates. What is not at all evident from this figure is that many of the enzymes in this pathway are

Figure 1.3

Signaling pathways and allostery. Each panel shows a schematic representation of inactive and active signaling pathways. In each case, an extracellular ligand binds to a receptor resulting in a cascade involving intracellular signaling proteins. These proteins in turn influence a variety of other proteins that can carry out specific biological processes, including activating metabolic enzymes (top panel), turning on the transcription of key genes (middle panel), or turning on cytoskeletal polymerization in particular regions within the cell (bottom panel).



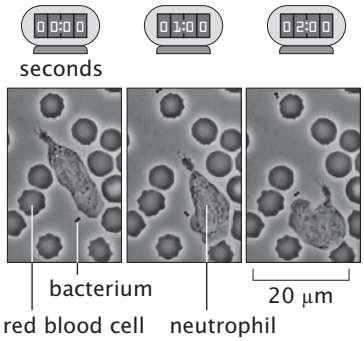


Figure 1.4

Eukaryotic chemotaxis. Snapshots from a video taken by David Rogers of the dynamics of a neutrophil hunting down a bacterium. This classic video raises myriad questions about the mechanisms of signaling and motility in cell biology. We use snapshots from this video as a reminder of the ubiquitous nature of cell signaling, exemplified here by the way the neutrophil is "aware" of its environment. From Phillips, R., J. Kondev, J. Theriot, and H. Garcia (2013), *Physical Biology of the Cell*, 2nd ed. Reproduced by permission of Taylor & Francis LLC, a division of Informa plc. Adapted from from a video by David Rogers.

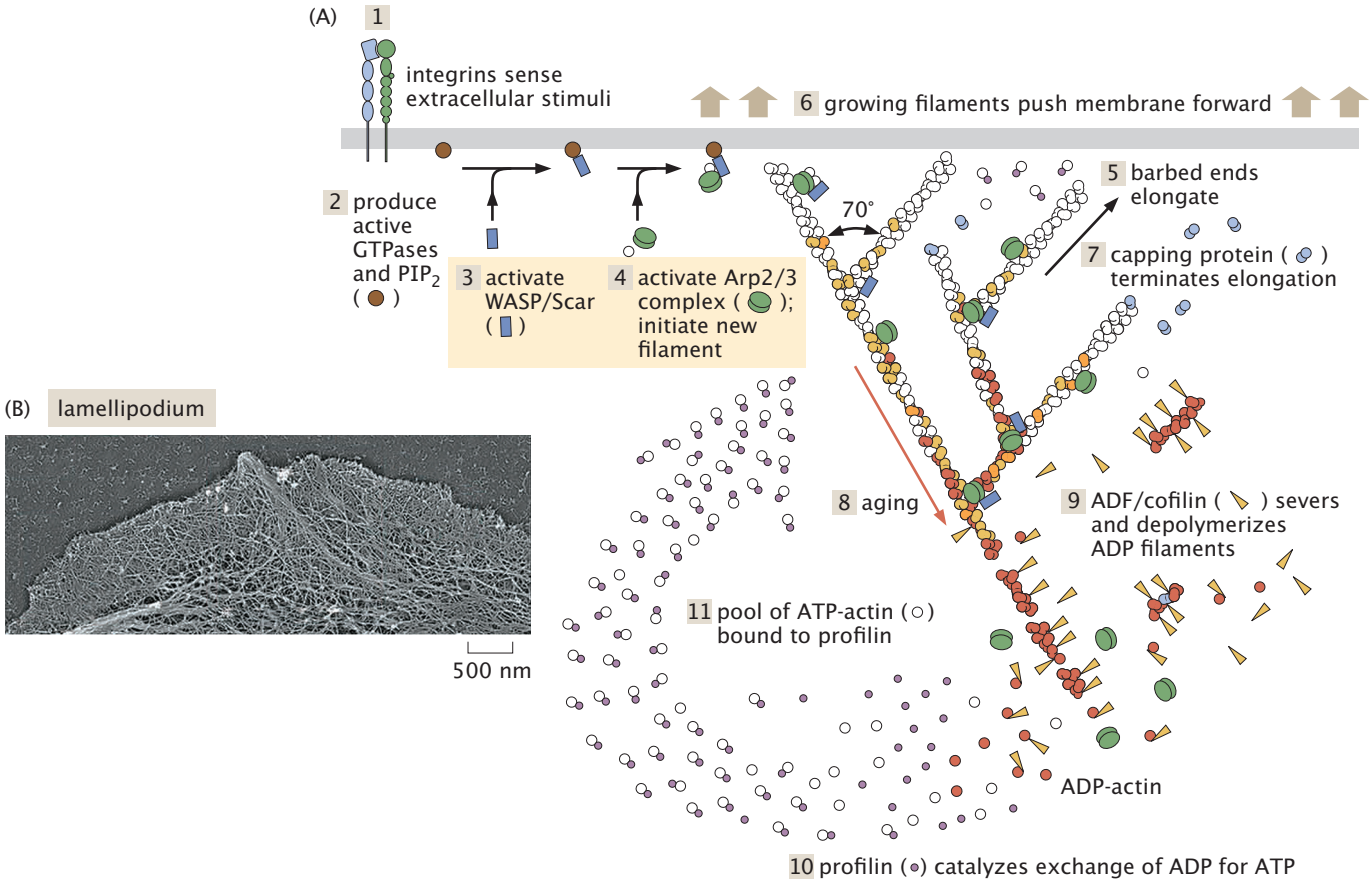


Figure 1.5

Cell signaling and the dendritic nucleation model. (A) The nucleation of new actin filaments such as those involved in the famous video of a neutrophil chasing down a bacterium (see Figure 1.4) require the activation of proteins such as N-WASP and Arp2/3 (highlighted in yellow). Protein activation can be viewed through the prism of the allosteric models this book discusses. (B) Leading edge of a motile cell as viewed using an electron microscope. The positioning of the actin network shown here is controlled by the signaling cascade shown in part (A). (A) Adapted from Pollard, T. D., and G. G. Borisy (2003) "Cellular motility driven by assembly and disassembly of actin filaments," *Cell* 112:456. with permission of Elsevier. (B) Adapted from Phillips, R., R. Milo (2016) *Cell Biology by the Numbers* (Garland Science), Estimate 3–7. Adapted from Svitkina, T. M., A. B. Verkhovsky, K. M. McQuade, and G. G. Borisy (1997) *J. Cell Biol* 139:397–415.

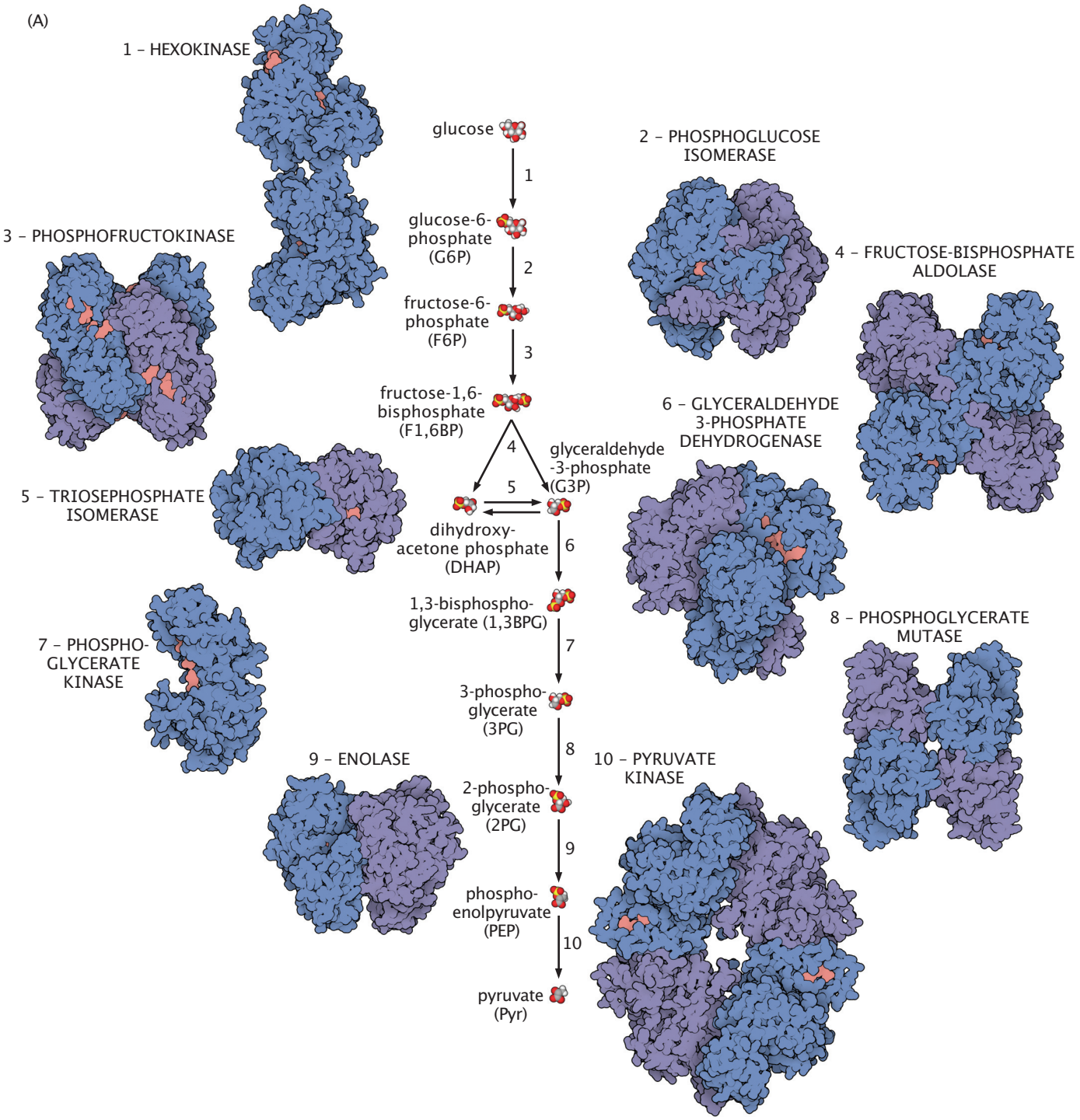


Figure 1.6

Enzymes in the glycolysis pathway. The molecules running down the center of the figure reveal the life history of a glucose molecule once it enters the glycolysis pathway. The large protein structures mediate these molecular transformations as glucose is turned into pyruvate. As shown in Figure 1.7, many of these proteins are modulated by the binding of small molecules. Courtesy of David Goodsell.

activated and inhibited by a suite of small molecules. Figure 1.7 gives an impression of this small-molecule regulatory landscape by presenting the complement of known inhibitors and activators. For example, if we look at phosphofructokinase, we see that it is inhibited by four distinct molecules and is activated by three others, giving the cell a suite of regulatory knobs with which to regulate this one step of this complex and important pathway.

As seen in Figure 1.7, there are a number of small molecules that play a role in repeated signaling and regulation. That qualitative impression has been made more concrete by counting up the number of regulatory interactions these small molecules participate in. A systematic analysis of the panoply of small molecules that preside over the regulation of proteins is shown in Figure 1.8 which counts the number of times that a given small molecule is known to participate in either inhibitory or activating interactions. It is clear that potassium ions are one of the most critical players in regulating the activity of proteins. Perhaps even more interesting is the role that ATP plays both as inhibitor and activator, including in the glycolysis pathway.

Sometimes biochemical reactions are mediated by ligands that are tethered to their receptor, giving rise to biochemistry on a leash. An example of this generalized allostery is offered in Figure 1.9. For example, as seen in Figure 1.9(A), when the tethered ligand is bound to the tethered receptor, the protein is in an inactive state. As the concentration of the free ligand is increased, at some point that concentration exceeds the "effective concentration" (estimated as $c_{eff} = 1/\frac{4}{3}\pi R^3$, where R is approximately the tether length), and hence the soluble ligands bind the tethered receptor, thus opening up the protein to its active conformation. The second example (N-WASP) shown in Figure 1.9(B) features a multidomain protein with a tethered linker. The relative equilibrium of active and inactive states is modulated by several binding partners (Cdc42 and PIP_2).

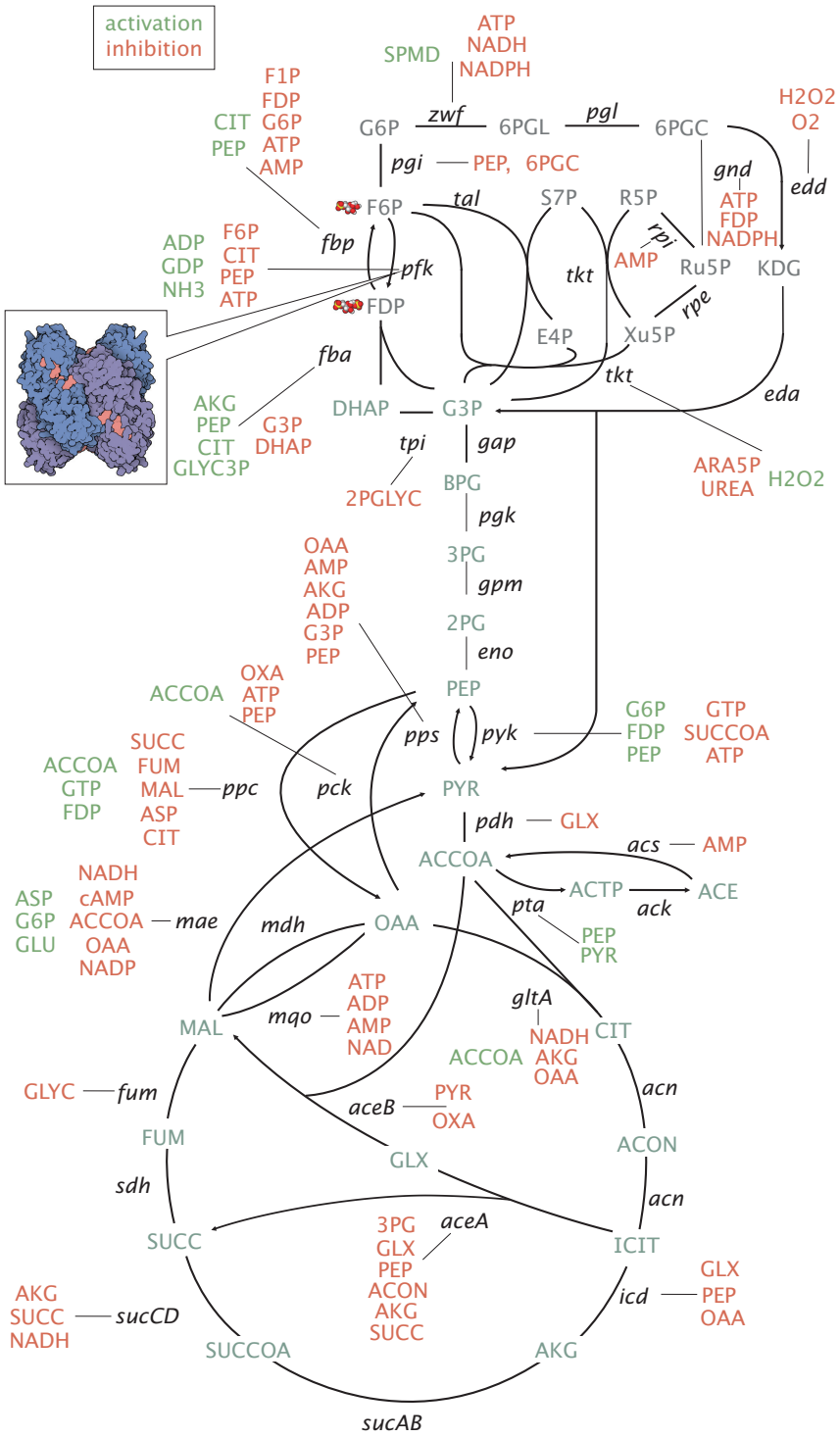
The example in Figure 1.9(B) teaches us another lesson as well. In particular, this is our first encounter with the concept of combinatorial control, as illustrated schematically in Figure 1.10. Here the idea is that a given cellular action, whether the activation of transcription or the activity of an enzyme, is dependent upon the status of several inputs simultaneously. For example, in the genetic network shown in Figure 1.10(A), the gene at the bottom of the diagram is regulated by two distinct activators. For this particular construct, the genetic circuit behaves as an AND gate, with high expression occurring only when both activators are present. As we will explore in detail in chapter 9, (p. 303), allosteric molecules can themselves behave as logical elements, as indicated schematically in Figure 1.10(B). Only in the presence of both inputs will the allosteric molecule be active, as already shown in the case of Cdc42 and PIP_2 in the context of Figure 1.9(B).

1.2.2 One- and Two-Component Signal Transduction and the Two-State Philosophy

So far, we have focused on specific molecular pathways that feature allosteric molecules. These case studies naturally lead us to wonder about the broader reach of the allostery concept. Despite the amazing advances of the high-throughput era, it remains a daunting challenge to identify genome-wide which

Figure 1.7

Small-molecule interactions in *E. coli* central carbon metabolism. The enzymes that mediate various steps in the pathway are denoted in black. Activators of those enzymes are denoted in green, and inhibitors are denoted in red. The structure of the enzyme phosphofructokinase (pfk) is shown, as well as its substrate F6P and product FDP. Adapted from Reznick et al. (2017).



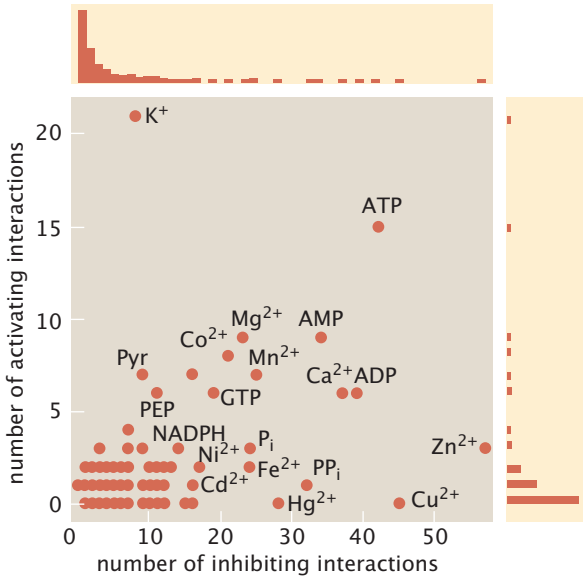


Figure 1.8

Small molecules with broad regulatory reach. These plots show the number of activating and inhibiting interactions that different small molecules engage in. Clearly, ATP is especially important both as an activator and inhibitor. Adapted from Reznick et al. (2017).

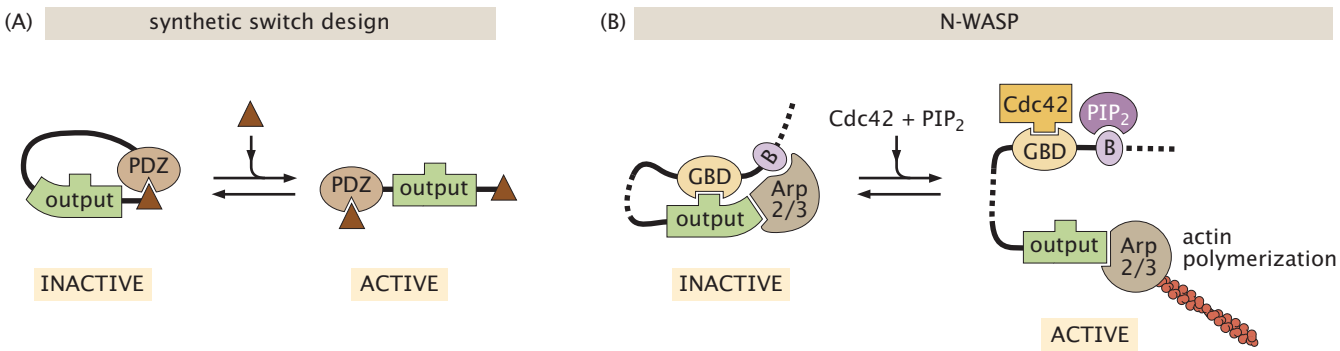


Figure 1.9

Biochemistry on a leash. Allostery can be based upon tethering motifs. (A) Tethering a receptor to a protein allows the state of activity of the protein to be controlled by the binding of soluble ligands which can outcompete the tethered ligand. The active state is characterized by exposure of the output domain. (B) Tether motif in the context of activation of actin polymerization. In the presence of Cdc42 and PIP_2, N-WASP is activated, which in turn activates Arp2/3. Adapted from Dueber et al. (2003).

proteins are allosterically regulated. Even more tricky is identifying what small molecules regulate them.

One context in which attempts have been made to survey the allosteric landscape is signaling in bacteria, specifically in the context of the two-component signal transduction systems in bacteria. The idea broadly is that the cell membrane is occupied by a wide variety of different receptors which flip between inactive and active states, as indicated schematically in Figure 1.11. In particular, as a result of the presence or absence of some external ligand, these receptors then switch between states in which they are either active or inactive for phosphorylating their cytoplasmic response regulator. These soluble proteins are now able to perform cellular functions such as changing the frequency

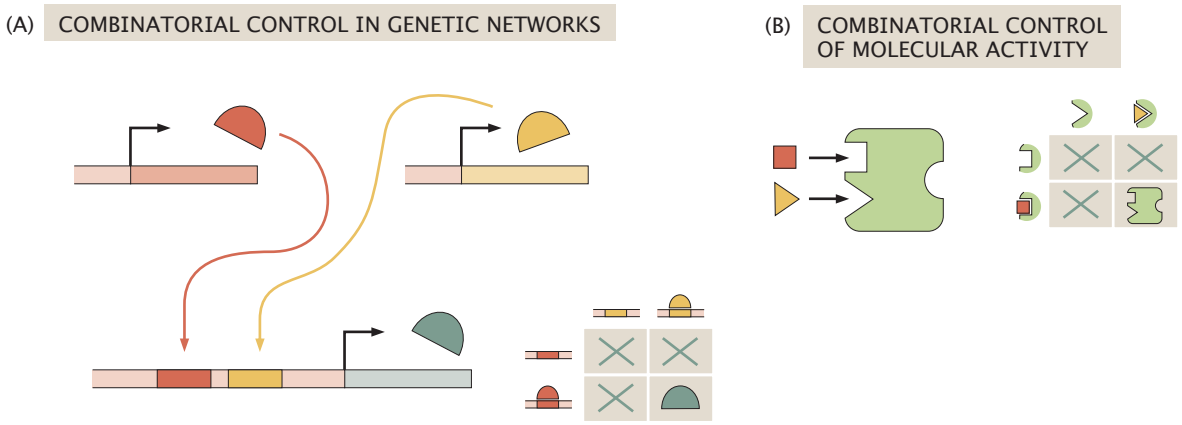
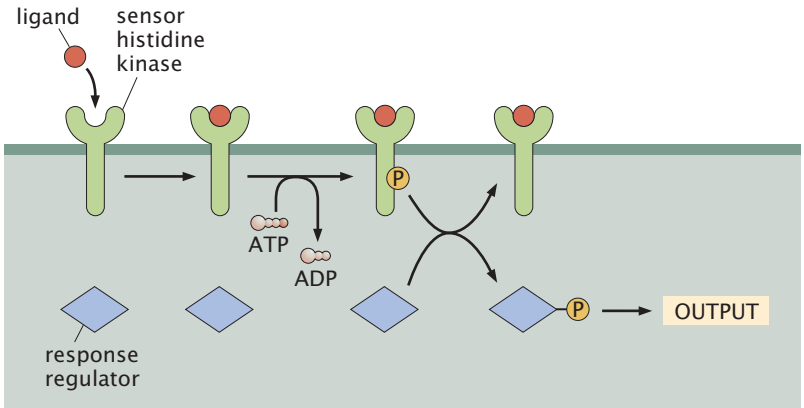


Figure 1.10

Introducing combinatorial control. (A) Combinatorial control in genetic networks. The gene in blue at the bottom of the schematic has binding sites for two distinct transcription factors. The logical truth table shows that the gene is on only when both transcription factors are bound. (B) Combinatorial control of allosteric molecules. The enzyme or signaling molecule (green) requires the presence of both ligands to be activated.

Figure 1.11

Schematic showing reactions of the sensor histidine kinase and response regulator in a typical bacterial two-component system. In the presence of ligand, the kinase has a higher probability of being in the active state where it is competent to carry out the phosphorylation reaction.



of tumble motions during bacterial chemotaxis, as will be taken up in detail in chapter 4 (p. 124).

To give a sense of the diversity of such two-component signaling systems in *E. coli*, Figure 1.12 shows the sensor histidine kinases and their corresponding response regulators. We see that there are a wide variety of inputs into these signaling systems that then lead to changes in cellular physiology and behavior. Figure 1.13 goes further by providing some insight into the distribution of one- and two-component signaling systems in bacteria by examining the databases of sequenced bacterial and archaeal genomes as of 2005 (it would be great to see these studies modernized). Specifically, the deeply interesting question that was examined in that work was the nature of the signaling systems in bacteria with special reference to whether those are one-component or two-component signaling systems. In the one-component signaling systems, the input and output domains are present on the same protein. Two classic

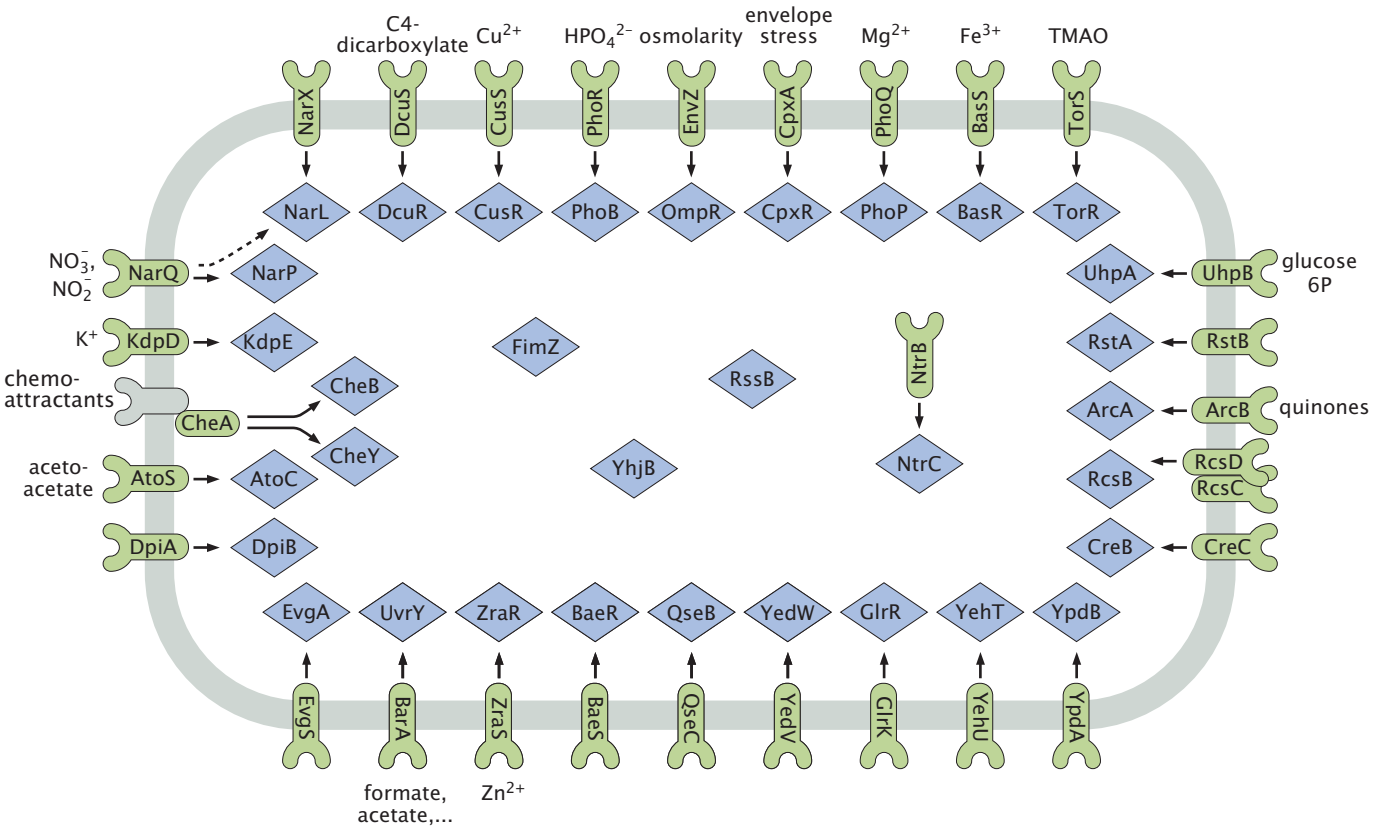


Figure 1.12

The vast array of sensor histidine kinases and response regulators found in *E. coli*. Schematic showing all sensor histidine kinases (green) and response regulators (blue) as identified by sequence in the *E. coli* genome. Both the sensor and the response regulator can be allosteric. Courtesy of Mark Goulian and Michael Laub.

examples of one-component signaling systems in bacterial transcription will be considered in chapter 8 (p. 272) when we discuss repression by the Lac repressor and activation by CRP. By way of contrast (see Figure 1.11), two-component signaling systems are characterized by having the input domain on one protein (usually a membrane-bound receptor) and the output domain (usually a cytoplasmic protein) on another protein. As we see in Figure 1.13, one-component signaling systems far outnumber their two-component counterparts, suggesting a rich proteomic reservoir of possible allosteric signaling molecules for further investigation.

Though data like those described here are not a proof that those molecules are allosteric, it is at least a tantalizing hint that the proteins that serve as one-component signaling systems may well be allosteric. Several key observations were made on the basis of this bioinformatic analysis of sequenced genomes and their putative signaling systems. First, that work found that roughly 85% of the output domains on these one-component signaling molecules are DNA-binding helix-turn-helix domains, indicating that often these signaling pathways appear

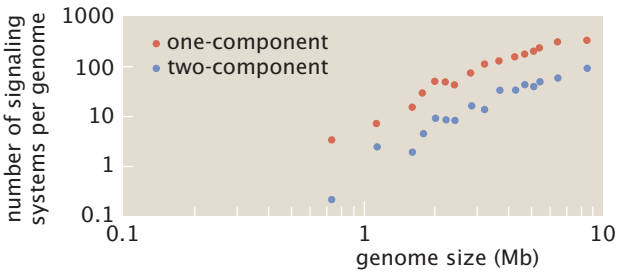


Figure 1.13

Number of one-component and two-component signal transduction systems as a function of genome size. Adapted from Ulrich, Koonin, and Zhulin (2005).

to be tied to transcriptional regulation, the example of which we will take up in detail in chapter 8 (p. 272). Similarly, this work found systematic trends in the input domains of these one-component systems, with more than 90% of them involving small-molecule binding domains, again providing a tantalizing hint of the possible allosteric control of these proteins.

1.3 Reasoning about Feedback: The Rise of Allostery

1.3.1 The Puzzle

Feedback is one of the greatest of ideas. A visit to a science museum such as the Musée des Arts et Métiers in Paris reveals century-old machines with their spinning "governors" that served to prevent them from running out of control, as shown in Figure 1.14. As hinted at in the figure, as the system rotates ever more quickly, the two balls will lift higher and in so doing will let some of the pressure bleed off, thus reducing the driving force that increases the rate of rotation.

In a fascinating reminder of the way that science and technology have always gone hand in hand, Figure 1.15 shows how the very same James Clerk Maxwell of Maxwell's equations fame worked on governors, noting that they are machines "by means of which the velocity of the machine is kept nearly uniform, notwithstanding variations in the driving power or the resistance." The modern world depends upon mechanical governors as well. In fact, we need look no farther than our toilets to see feedback in action, as also shown in Figure 1.14.

Like their macroscopic counterparts, the macromolecules of life are replete with examples of molecular governors which inhibit or enhance key biochemical reactions in response to either too much or too little of some substrate of interest. The bottom panel of Figure 1.14 shows regulatory feedback in the context of transcriptional autorepression, where the gene product of the gene of interest "governs" its own production. This will be the subject of chapter 8 (p. 272), where we will consider transcriptional regulation and its connection to allostery in detail.

But more generally, how do molecular governors work? One of the original ideas in the context of enzymes was that there are inhibitory molecules which compete for the attentions of the active site of some enzyme (as a concrete molecular example), thereby slowing down the reaction of interest. To be specific, people envisioned that the inhibitor molecule could actually bind to

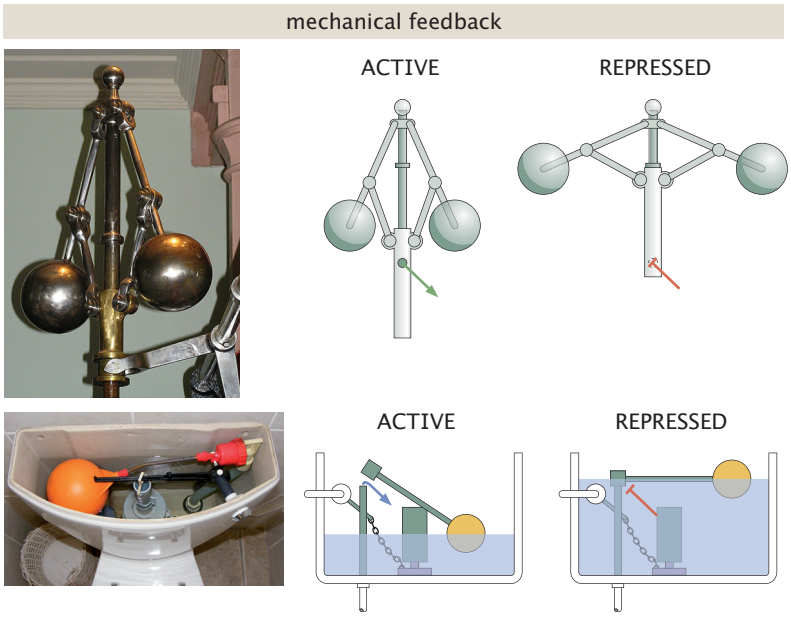
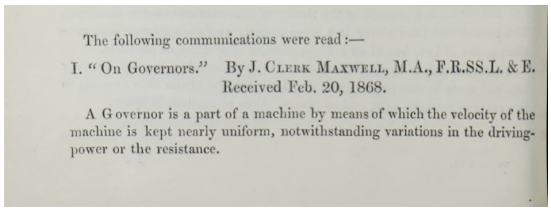


Figure 1.14

Control by feedback. (Top) Mechanical governors for feedback on a steam engine and in a toilet tank. (Bottom) Negative feedback in a gene regulatory circuit. The gene produces a protein that represses itself.



The following communications were read:—

I. "On Governors." By J. Clerk Maxwell, M.A., F.R.SS.L. & E. Received Feb. 20, 1868.

A Governor is a part of a machine by means of which the velocity of the machine is kept nearly uniform, notwithstanding variations in the driving-power or the resistance.

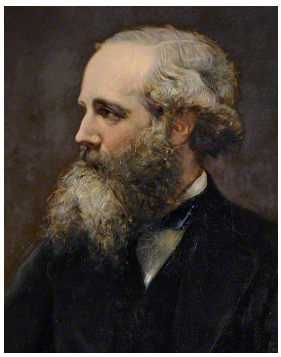


Figure 1.15

James Clerk Maxwell and the mechanics of governors. (*Lower left*) Lowes Cato Dickinson, portrait of James Clerk-Maxwell (1891). Photo courtesy of Master and Fellows of Trinity College, Cambridge. (*Lower right*) Courtesy of Ray Tomes of Auckland, New Zealand. Taken at MOTAT (Museum of Transport and Technology) in Auckland.

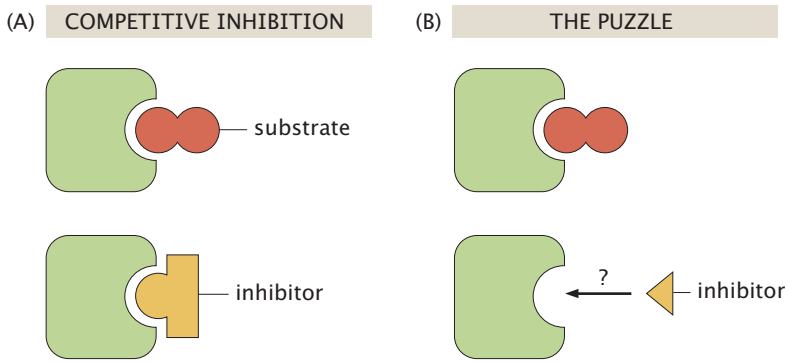


Figure 1.16

Regulation of an enzyme by an inhibitor. (A) The inhibitor fits in the active site occluding the site from possible binding by the correct substrate. (B) Not all inhibitors have the correct shape to fit into the active site of the enzyme. In the early 1960s, this phenomenon led to the question of how such inhibitors regulate their target enzymes.

the enzyme in such as way as to literally block access of the real substrate to the active site, as indicated schematically in Figure 1.16(A). However, for such a mechanism to work, it would seem that the inhibitory molecule would need to have the same shape and size as the substrate molecule whose enzymatic modification was being inhibited. The question of how molecules would activate an enzyme was even more puzzling to contemplate.

One of the early molecules that focused the attention of scientists on questions of enzymatic regulation was aspartate transcarbamoylase, one of the key enzymes in pyrimidine biosynthesis. Recall that the DNA double helix is made up of the repeated pairing of pyrimidine-purine pairs, with cytosine, thymine, and uracil making up the pyrimidine derivatives. The *E. coli* version of aspartate transcarbamoylase is made up of 12 subunits with half coming from two trimers and the other half coming from three dimers, the two classes serving as the enzymatic and regulatory parts of the complex (see Figure 6.1 for more details). This enzyme functions in pyrimidine biosynthesis by mediating the interaction of aspartate and carbamyl phosphate to form *N*-carbamyl-L-aspartate and inorganic phosphate. For our purposes, the reason this enzyme was and is so interesting is because the rate with which it carries out the reaction is modulated by the levels of both pyrimidines and purines. Specifically, the final product of the pyrimidine pathway, namely, CTP, feeds back into the reaction and slows it down. By way of contrast, ATP, the final product in the purine pathway, speeds up the pyrimidine synthesis reactions. Thus, the system is subject to both negative and positive feedback in order to tune the quantities of the pyrimidine substrate. The kinds of questions that arose in light of these observations centered on how molecules such as CTP and ATP could interact with the enzyme itself in such a way as to tune the reaction rate.

The puzzle faced by early investigators of proteins that were subject to inhibition and activation was how a battery of regulatory molecules could be fine-tuned to fit into the active sites of their binding partners, as shown in Figure 1.16(B). The simple answer is that often they don't. Rather, groups in Paris and Berkeley realized that a different regulatory strategy could be "action at a distance," in which the binding of a regulatory ligand in one part of a macromolecule could lead to a conformational change elsewhere in the molecule such that the activity of the enzyme was changed. This thinking has been codified in the so-called Monod-Wyman-Changeux (MWC) model. We now turn to this allosteric resolution of the regulatory puzzle.

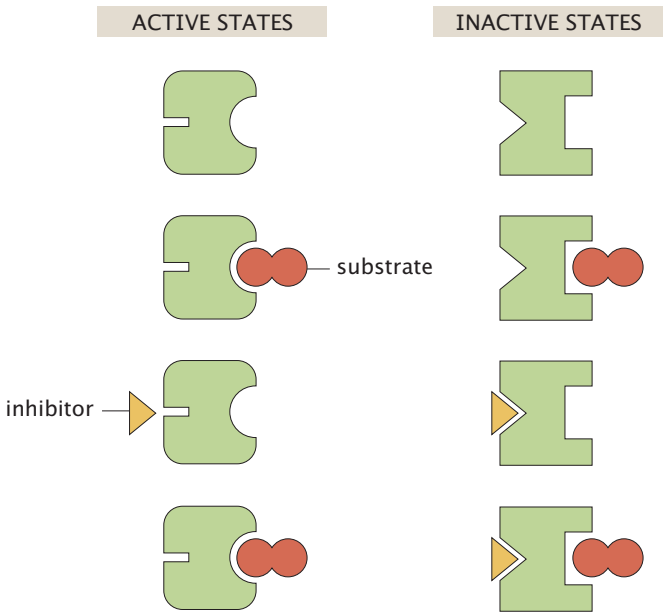


Figure 1.17

Allosteric regulation of an enzyme by an inhibitor. The enzyme can exist in both active (left column) and inactive (right column) states. For each conformation, there are four states of occupancy—empty, bound by substrate, bound by inhibitor, and bound by inhibitor and substrate. The binding affinities for both inhibitor and substrate are different in the two states, with the binding of the inhibitor favored in the inactive state.

1.3.2 The Resolution of the Molecular Feedback Puzzle

Figure 1.17 gives a schematic view of the extremely clever hypothesis that was formulated as the allosteric alternative to the kind of direct regulation posited originally and schematized in Figure 1.16. The essence of the cartoon is that there is a regulatory site where a ligand binds that tunes the relative probability of the active and inactive states. Note however that there is a nuance that is captured in our cartoon. Specifically, the regulatory ligand does *not* have the same binding energy when bound to the active and the inactive states, as indicated by the difference in the shape of the binding site in the two conformations. This critical mechanistic feature is the entire basis of the allostery framework, as we will show in equation 1.2 (p. 26). Note further that the allosteric strategy is noncommittal with respect to the question of whether the regulatory ligand leads to inhibition or activation of its binding partner. If the effector molecule favors binding the active state, then it will serve as an activator. If the effector molecule favors binding the inactive state, then it will serve as an inhibitor. By way of contrast, for the strategy highlighted in Figure 1.16, there is no obvious mechanism for activation.

Though we provided a caricature of some of the classes of MWC molecules we will consider here, in Figure 1.2 (p. 4), in fact, the detailed atomic structures of some of these molecules are known for a variety of different conformational states both in the absence and presence of their substrates and effectors. Structure has become one of the most powerful tools in modern biology. The conceptual argument associated with the great push for structural insights into biological problems is a deep confidence in the structure-function paradigm that holds that function follows structure. Several examples of structures of key allosteric molecules that will occupy our attention throughout the book are shown in Figure 1.18. For example, our first concrete case study in quantitative allosteric thinking will focus in chapter 3 (p. 77) on the ligand-gated ion

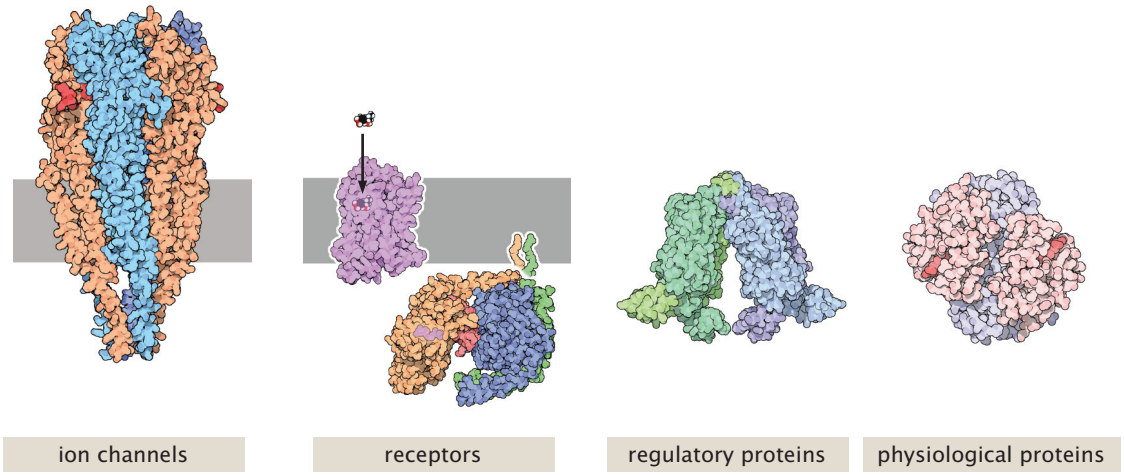


Figure 1.18

Structures of MWC molecules. Many different macromolecules can exist in distinct conformational states including ion channels, such as the nicotinic acetylcholine receptor shown here, G protein–coupled receptors such as the adrenergic receptor shown here, transcription factors such as the Lac repressor shown here, and proteins relevant to human physiology such as hemoglobin. Illustrations courtesy of David Goodsell.

channel known as the nicotinic acetylcholine receptor and shown in the left of Figure 1.18.

Chapters 4 (p. 124) and 5 (p. 170) take up the topic of allosteric membrane receptors like those shown in Figure 1.18. First, we tackle the behavior of the bacterial receptors responsible for chemotaxis and quorum sensing, followed by an in-depth examination of G protein–coupled receptors.

Another classic example that dates all the way back to the inception of the allostery concept itself is the Lac repressor molecule (third structure shown in Figure 1.18) that binds DNA, thus shutting down transcription of genes associated with lactose usage. This molecule can be thought of as an MWC molecule because in the presence of allolactose it undergoes a conformational change that reduces its binding affinity for DNA, thus permitting the transcription of the genes for β-galactosidase that make it possible to metabolize this alternative carbon source. There are also numerous examples of transcriptional activators.

The final example shown in Figure 1.18 is hemoglobin, the critical oxygen carrier. We devote chapter 7 (p. 231) to the fascinating allosteric mechanisms of hemoglobin and the basis for its physiological and evolutionary adaptation.

A zoomed-out view of the secondary structure of some representative allosteric proteins is shown in Figure 1.19. For example, in our discussion of gene regulation in chapter 8 (p. 272), we will consider both repression and activation, using classic examples from bacteria as our critical case studies. One of the most beloved and well-studied examples of activation is offered by the bacterial protein CRP, shown in the upper left panel of Figure 1.19, bound to its effector cAMP. The remaining structures in Figure 1.19 give other examples of allosteric proteins, revealing transcription factors, macromolecular assemblies, and enzymes.

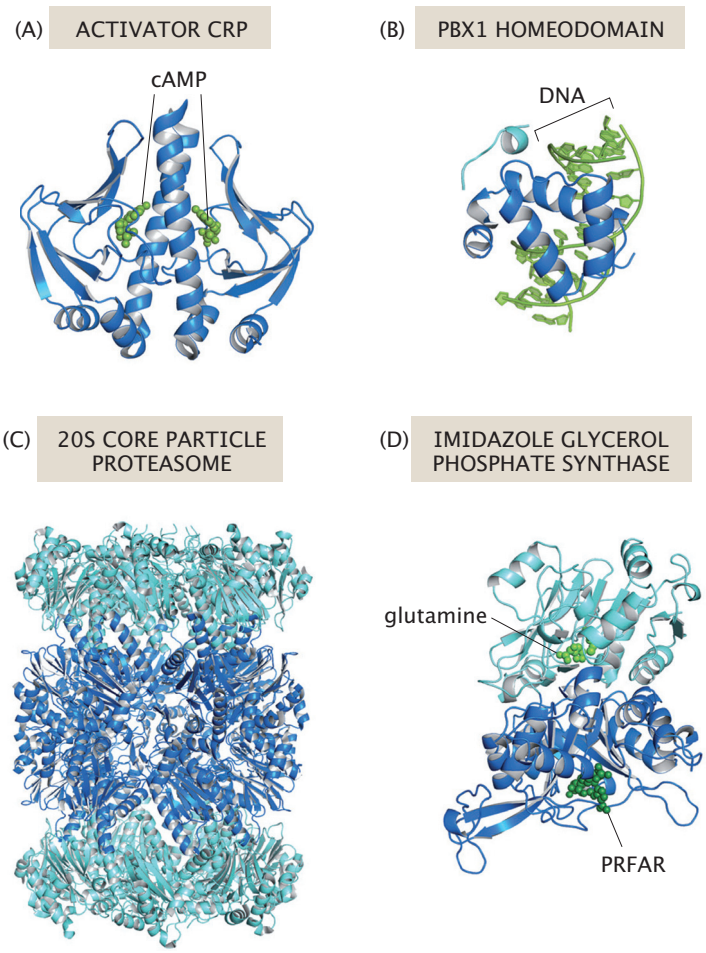


Figure 1.19

Structures of MWC molecules in complex with effectors and substrates. (A) The bacterial transcriptional activator CRP bound to cAMP (green). (B) PBX1 homeodomain bound to DNA. (C) Structure of the protein degradation apparatus from the archaeon *Thermoplasma acidophilum*. (D) Key metabolic enzyme relevant to both amino acid (histidine) and nucleotide (purine) biosynthesis. The structure shows both the substrate (glutamine) and the allosteric effector (PRFAR). Adapted from Grutsch, Bruschweiler, and Tollinger (2016), CC BY-NC 4.0.

Though it is clear that structural insights are a critical cornerstone of modern biology, the approach advocated here is quite different. Indeed, in many ways the entire goal of the kinds of models to be described throughout the book is to telescope out from the atomic-level mindset to more coarse-grained perspectives which make predictions about how MWC molecules will respond quantitatively in new situations. In some sense, the way that structure is internalized at the level of the models described here is through a very small set of parameters. As will be shown in the next section, there is one overarching conceptual framework for describing MWC molecules that in its simplest incarnation features three key parameters, namely, the difference in energy between the active and inactive states in the absence of ligand, and the dissociation constants (K_A and K_I) for ligand binding in the active and inactive states, respectively. In the context of the MWC model, the structural details for MWC molecules found in the Protein Data Bank influence only these three parameters. Our aim is to show how one can talk about the huge swaths of biology reflected in the case studies shown in Figure 1.18 in terms of abstract models without appealing to detailed atomic-level positions.

A fundamental pillar upon which the entire book is constructed is the idea of the power, the rigor, and the intuition that can be developed by self-consciously suppressing features of a system. As will be highlighted repeatedly throughout the book, statistical mechanics teaches us how to "integrate out" degrees of freedom. This does not mean that we approximate the system by ignoring some feature (such as the existence of intrinsically disordered domains in a protein) but rather that we formally and mathematically compute the implications of those hidden degrees of freedom for the rest of the system.

1.3.3 Finding the Allosterome

A puzzle that remains in the field of allostery in this high-throughput era is that we have had very limited tools that allow us to answer the general question of which proteins in the proteome are allosteric and who their binding partners are. Despite Monod's characterization of the allostery phenomenon as the second secret of life, because of this important knowledge gap, as a field we are often flying blind because of our ignorance of how the key molecular players in signaling pathways have their activity modified by other chemical agents, and because of our ignorance of the identity of those chemical agents themselves.

To that end, the emergence of mass spectrometry has provided an exciting opportunity to query not only the posttranslational modifications suffered by a given signaling molecule but also, because of recent innovations, when signaling molecules have bound a given small molecule. The idea of one such method is shown in Figure 1.20. We see that by lysing cells in the absence and in the presence of some small-molecule allosteric effector candidate, some proteins will bind that small molecule and, as a result, be resistant to limited proteolysis by proteinase K. This means that when the proteins are denatured and trypsin digested, the pattern of cuts in the polypeptide chain will be different for any protein that was bound to the candidate small molecule, as indicated in Figure 1.20(B). Approaches such as this hold the promise of systematic identification of the allosterome for any organism and will be a critical part of our resolution of the puzzle of how the macromolecules of the cell are controlled by a battery of small molecules.

1.4 Mathematicizing the Two-State Paradigm

By peering at allostery through a mathematical lens we learn there are many common biophysical features shared by these molecules, as shown in Figure 1.21. This figure focuses our attention on the function of these molecules rather than their structure. One important feature is how much activity they exhibit even in the absence of ligand, a quantity we will call the *leakiness*. Just as we interest ourselves in the activity of allosteric molecules in the absence of ligand, their behavior in "saturating" concentrations is also critical to their function. Another parameter of great physiological and evolutionary significance is the critical concentration at which the activity reaches the midpoint between the inactive and active states, sometimes denoted as the EC_{50}. We will also be deeply interested in how sharp the switching events are between the two

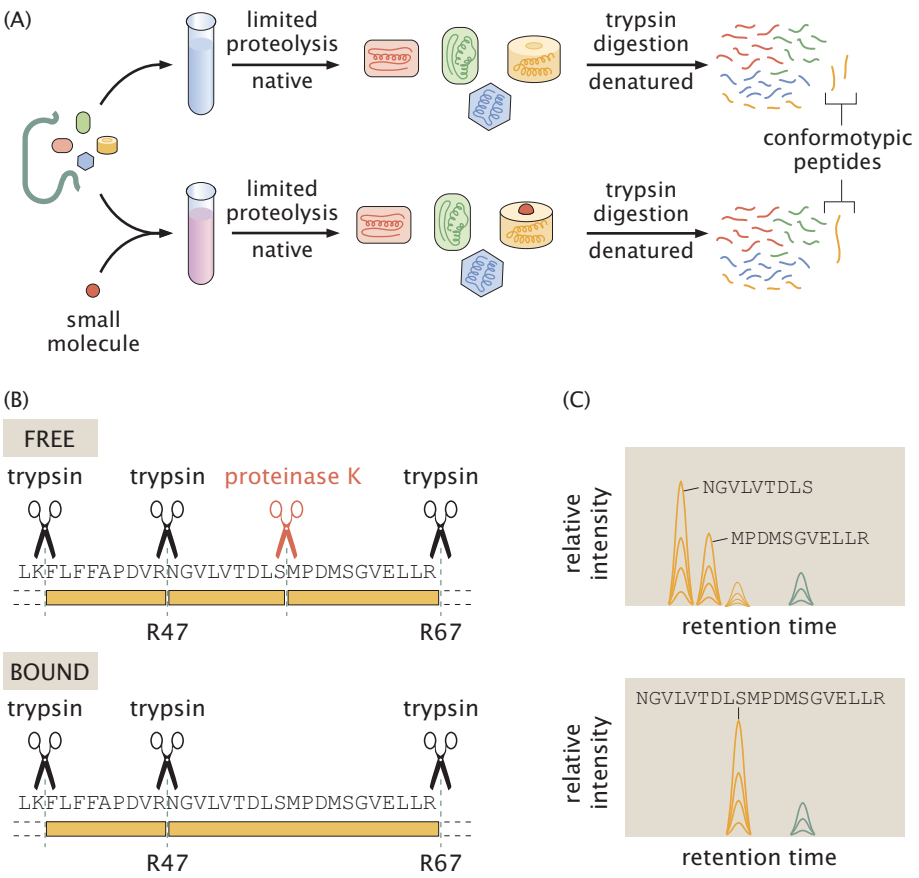


Figure 1.20

Finding allosteric proteins and the molecules that regulate them. (A) Incubation of cell lysate with and without some candidate small molecule leads some proteins to have a different pattern of peptide fragmentation because the binding of the small molecule protects some parts of the polypeptide. (B) Partial proteolysis is performed with proteinase K, followed by a trypsin digestion of denatured protein. This leads to peptide fragments that can be measured using mass spectrometry. In the example shown here, the measurement is made on the protein FixJ bound to aspartyl phosphate (PDB: 1DBW) and its ligand-free form (PDB: 1D5W) (C) Differences in spectrum resulting from mass spectrometry for the two different situations, revealing which parts of the molecule has been protected. Adapted from Piazza et al. (2018).

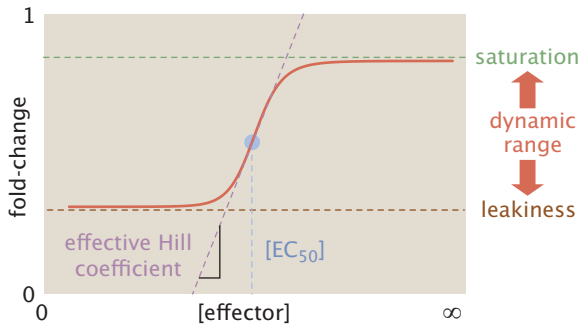


Figure 1.21

Macromolecular activity as a function of ligand concentration. This curve shows key phenotypic properties of allosteric molecules including leakiness, dynamic range, the EC_{50}, and the effective Hill coefficient, which gives a measure of the sharpness of the regulatory response.

conformational states of interest as a function of the ligand concentration that drives this shift. The examination of such sharpness in molecular responses led to one of the most preeminent ideas in biology, namely, cooperativity, an idea that falls very naturally within the purview of the statistical mechanical models of allostery to be described throughout the book.

The MWC model is the mathematical framework that was introduced to describe molecules with two states of activity and modulated by the binding of regulatory ligands. Monod, Wyman, and Changeux saw how to enumerate the various microscopic states of the system and to compute their relative probabilities. As any lover of statistical mechanics knows, the two-state paradigm is one of the centerpieces of statistical physics and has had enormous reach in the form of the Ising model and its generalizations. In the physics setting, the two states referred to in the statistical mechanical setting of Ising models can refer to the orientations of magnetic "spins," for example. The kinds of questions that people were interested in addressing with such two-state models centered on phase transitions between the low-temperature magnetic state and a high-temperature nonmagnetic state of some materials. Monod, Wyman, and Changeux, without knowing it at the time, were introducing another overwhelmingly important statistical mechanical model that could have impact in biology similar to that of the Ising model in physics. To get a better idea of how specific models can have such broad reach, we consider examples of such transcendent concepts in physics.

1.4.1 Transcendent Concepts in Physics

How can we describe the behavior of allosteric transitions mathematically? The answer to that question is the subject of this book! The goal will be to show in many different biological contexts like those described earlier in the chapter how the allostery phenomenon can be described in mathematical terms. In particular, we aim to reveal how to enumerate both the various microscopic states that are available to an MWC molecule and the probabilities of these different states as a function of the concentration of various ligands. The simplest version of these ideas will unfold here, and then in the remaining chapters, we will see how those ideas can be generalized to include features such as oligomerization, cooperativity, and applications to a variety of distinct biological situations.

Certain scientific concepts like the MWC model have very broad reach. To clarify what I mean by that, let's explore some examples of scientific broad reach in physics. Young scientists and engineers of all stripes are subjected to a first indoctrination in physics during their early years in university. Shortly after beginning a foray into mechanics, these students are exposed to the seemingly sterile world of masses and springs. After drawing a free-body diagram to reckon how all the forces act on the mass, they obtain an equation of motion that gives the position of that mass as a function of time. Little do they expect that in talking about the abstract behavior of blocks and springs, they have opened a vista onto one of the most far reaching of ideas: periodic motion around an equilibrium point. If they are lucky, these same students will later see that, in fact, the mass-spring problem lays the groundwork for thinking about very different problems such as the pendulum and electrical circuits built up of

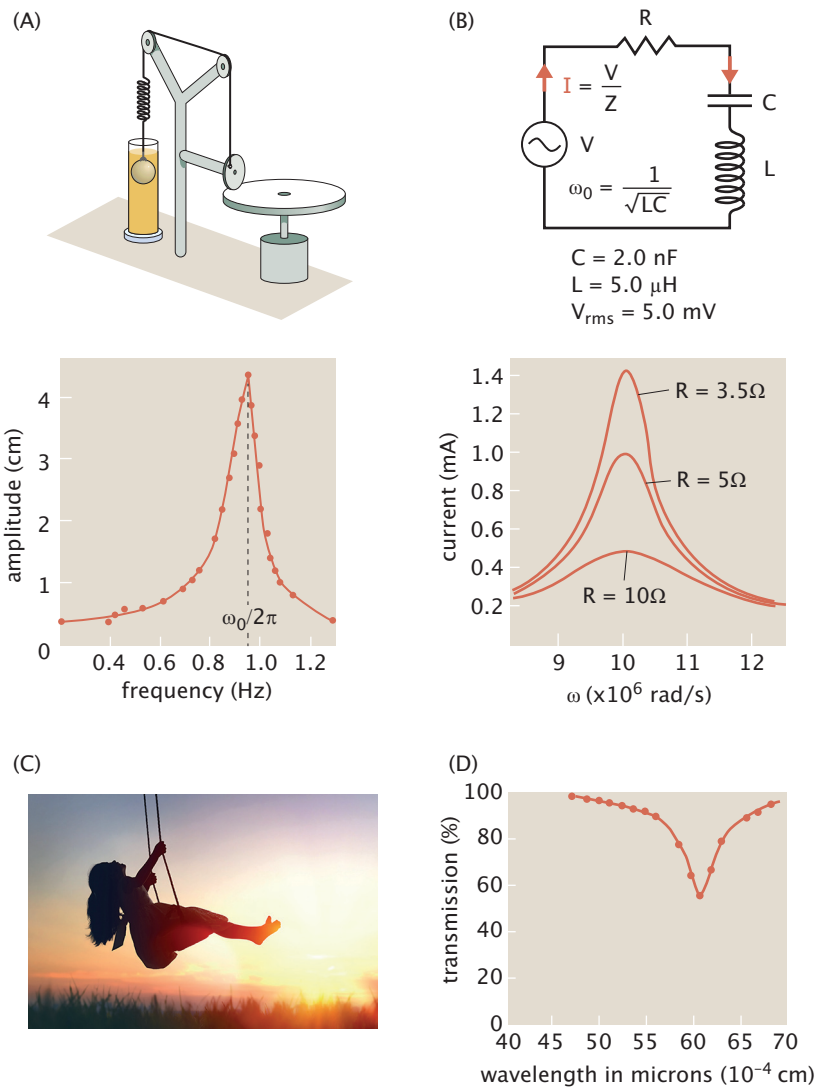


Figure 1.22

Resonance as a concept that transcends any particular example. (A) Mass-spring system. Forced oscillator with damping due to motion in a fluid. The graph shows the amplitude of the oscillations as a function of the frequency of the driving force. (B) Resonance in an *RLC* circuit. The circuit is composed of a resistor (R), a capacitor (C), and an inductor (L). Z is the impedance. The graph shows the current as a function of frequency of the voltage for several choices of resistor. (C) Child pushed on a swing. (D) Transmission of infrared radiation passing through a thin film of sodium chloride as a function of the wavelength of the incident radiation. (A) adapted from French (1965); (B) adapted from Hyperphysics website (C.R. Nave); (C) © Konstantin Yuganov, Dreamstime.com / ID 75499691; and (D) adapted from Feynman, Leighton, and Sands (1963), see Further Reading.

resistors, capacitors, and inductors. All are surprisingly described by the same equation,

$$\ddot{x} + \gamma\dot{x} + \omega^2 x = F(t), \tag{1.1}$$

where x is the displacement from equilibrium, $\dot{x} = dx/dt$ is the velocity, $\ddot{x} = d^2x/dt^2$ is the acceleration, γ provides a measure of the damping, and ω is the vibrational frequency. Furthermore, depending upon the behavior of the forcing function $F(t)$, the periodic motions can give rise to the general phenomenon of resonance, as shown in Figure 1.22. Here we see that for certain driving frequencies, the amplitude of the vibrations become very large—the phenomenon of resonance familiar to anyone who has pushed a child on a swing.

This resonance idea is so far-reaching as to be astonishing. The mechanics of a pendulum, represented by a child on a swing set in the figure, can be

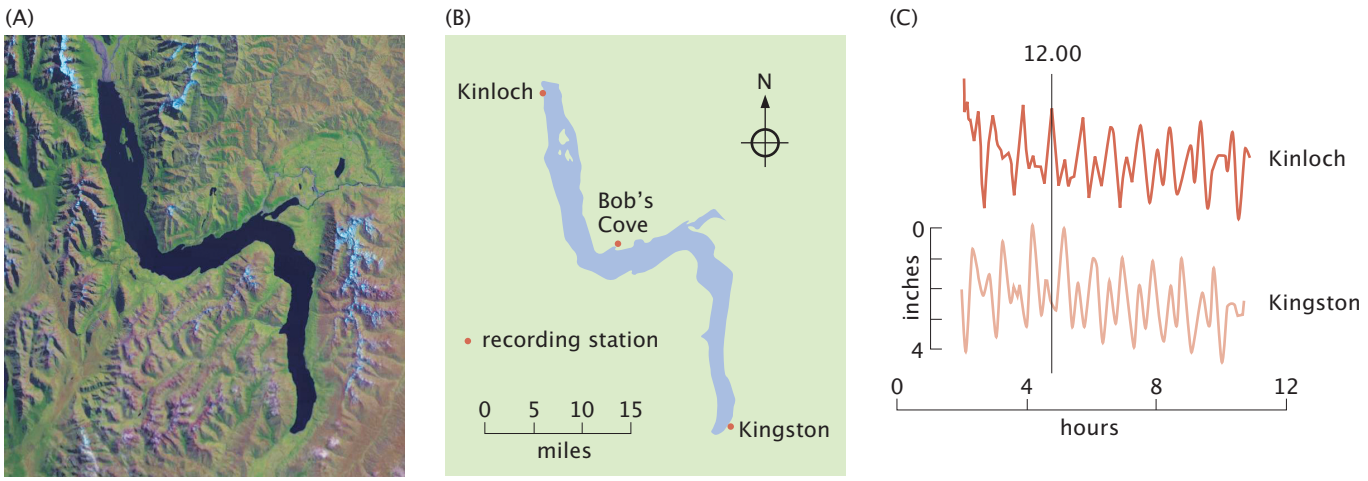


Figure 1.23

Seiche in Lake Wakatipu. A large-scale example of resonance in an unexpected place. (A) Satellite image of the lake. (B) Scale map showing the sites of measurement. (C) Lake height at different points in the lake as a function of time. Adapted from Bottomley (1956).

mapped onto the problem of a mass-spring system, and hence all the things we learned about resonance in the context of the mass-spring system apply just as well to the pendulum (as long as the amplitude of the swinging is not too large). Things become increasingly surprising when we learn that precisely the same mathematics describes the dynamics of charge flow in simple electrical circuits featuring capacitors and inductors. These insights become even more impressive when we find that the very same thinking helps us understand the sloshing motion of a giant lake such as Lake Wakatipu in New Zealand, home to a famed seiche, as shown in Figure 1.23. My point here is to demonstrate a fundamental principle in physics that is now ready for prime time in biology, too: the mathematical unity of apparently disparate phenomena.

Another example of this kind of surprising deep connection between apparently quite different phenomena is offered by the ubiquitous random-walk concept, shown in Figure 1.24. The key point here is that an idea so simple as rolls of a die or flips of a coin can be repurposed to help us understand phenomena as diverse as the diffusion of molecules in solutions or cells, or the statistical conformations of polymer molecules such as DNA. But how? The middle panel of the figure shows how we can think of a random walker as being able to march in any one of six directions each step: east, west, north, south, and up or down. The outcome of the roll of our die tells the walker which one of those steps to make. Importantly, the outcome of this analysis is a *statistical* description of the molecular configurations.

A final physical example of the transcendence of certain physical concepts is given by the all-important wave phenomenon of interference, one of the fruits of Thomas Young's interconnected thinking on physiology and physics (see Figure 1.25 and the article by Mollon (2002) referenced in the Further Reading section). Young was the first to see the phenomenon of interference in all of its sameness, applying it not only to the well-known example of light but also to auditory beats and to the seemingly obscure phenomenon of the tides in the Gulf of Tonkin, which don't exhibit the usual twice-daily tides we are

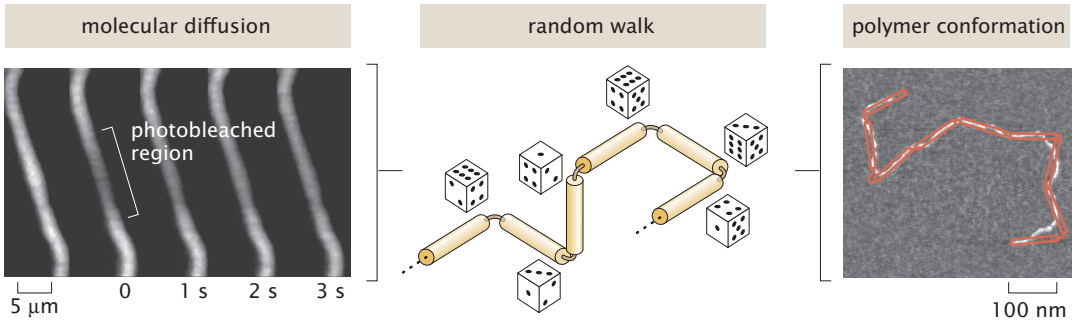


Figure 1.24

The broad reach of the random-walk concept. The center panel of the figure shows how successive rolls of a die determine the nature of the walk. Molecular diffusion, shown on the left, can be analyzed using nothing more than this simplified die rolling. In this case, a bacterium has had the fluorescence in its middle destroyed through photobleaching, and over time, because of diffusion, the fluorescence is restored in the photobleached region. Similarly, the conformations of a polymer such as DNA can be thought of using the same ideas. (*Left and right*) From Phillips, R., J. Kondev, J. Theriot and H. Garcia (2013), *Physical Biology of the Cell*, 2nd ed. Reproduced by permission of Taylor & Francis LLC. (*Left*) Adapted from Mullineaux, C. W., A. Nenninger, N. Ray, and C. Robinson (2006), *J. Bacteriol.* 188:3442, fig. 5. Amended with permission from American Society for Microbiology. (*Right*) Adapted from Wiggins, P. A., et al. (2006), *Nat. Nanotech* 1:37, fig. 1a.

accustomed to at most beaches. Most of us have experienced interference first hand in the form of beautifully colored oil slicks on our driveway after a rain. Isaac Newton mapped out how the color depends upon the thickness of the air layer between two glass plates (one of which was curved; see Figure 1.25(A)), and Thomas Young saw how to compute those colors on the basis of the simple idea of waves either reinforcing one another or canceling each other out, as shown in Figure 1.25(B). The critical idea of sameness is that in all these cases, the interference phenomenon can be simply expressed as the result of several waves adding up either constructively or destructively in a way that is relatively indifferent to whether those waves are sea waves, sound waves, or light waves. This idea was described by Young in what might be thought of as *The Feynman Lectures on Physics* of his time, and his explanation is shown in Figure 1.25(D). As Young responsed to critics of his idea, "I was so forcibly impressed with the resemblance of the phenomena that I saw, to those of the colours of thin plates, with which I was already acquainted, that I began to suspect the existence of a closer analogy between them than I could before have easily believed." It is just such a resemblance of phenomena that the allostery concept allows us to understand, as we will see in the pages that follow.

1.4.2 One Equation to Rule Them All

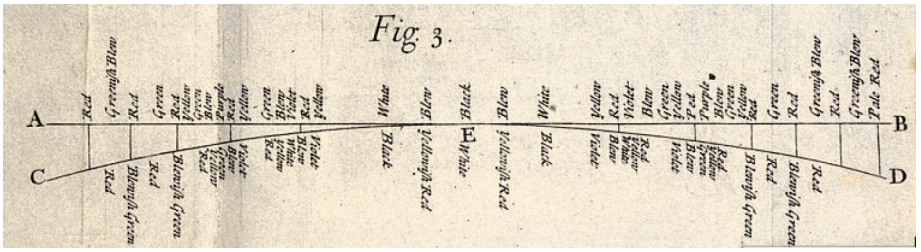
But what does this have to do with biology in general, and this book in particular? A superficial appearance that phenomena are different may mask a very subtle but deep connection between those phenomena or the theory used to describe them. The argument of this book is that the allostery phenomenon is such an example that has the same status as ideas such as resonance or random walks, but unlike those cases, the allostery universality is one that animates the subject of biology.

One of the deep appeals of the kind of universality offered by the allostery concept and its statistical mechanical implementation via the MWC model

Figure 1.25

The broad impact of the interference concept. (A) Newton examined the colors in a layer of air between two glass surfaces. (B) Using his theory of interference, Thomas Young predicted the observed wavelengths as a function of film thickness. The final column is the addition of Mollon (2002), who gave the wavelengths in nanometers. (C) Figure from Thomas Young's *Lectures on Natural Philosophy* (1807) that shows the interference in water "obtained by throwing two stones of equal size into a pond at the same instant." (D) Page from Thomas Young's *Lectures on Natural Philosophy* that show how waves interfere. The dark lines correspond to the waves being superposed, and the broken line shows their composition (though Young used a different scale.) Adapted from Mollon (2002), see Further Reading.

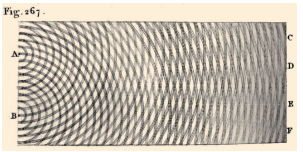
(A)



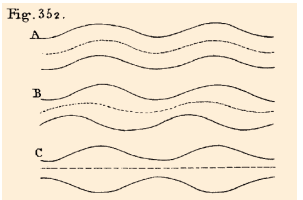
(B)

Colours.	Length of an Undulation in parts of an Inch, in Air.	Number of Undulations in an Inch.	Number of Undulations in a Second.	Wavelength (nm)
Extreme -	.0000266	37640	463 millions of millions	
Red - -	.0000256	39180	482	650
Intermediate	.0000246	40720	501	
Orange - -	.0000240	41610	512	609
Intermediate	.0000235	42510	523	
Yellow -	.0000227	44000	542	576
Intermediate	.0000219	45600	561 (= 2st nearly)	
Green - -	.0000211	47460	584	536
Intermediate	.0000203	49320	607	
Blue - -	.0000196	51110	629	497
Intermediate	.0000189	52910	652	
Indigo - -	.0000185	54070	665	469
Intermediate	.0000181	55240	680	
Violet - -	.0000174	57490	707	444
Extreme - -	.0000167	59750	735	

(C)



(D)



and its generalizations is that it provides a different way of connecting biological phenomena. Thus, we can imagine organizing biological phenomena on the basis of their biological proximity. For example, we might talk about ion channels and their role in muscle contraction in a physiology course and talk about transcription factors and their induction in a systems biology course. Alternatively, as suggested by Figure 1.26, we can organize biological phenomena according to their physical proximity. In this case, the ion channels and the transcription factors can be seen as the "same" phenomenon, despite how apparently different the biological phenomena they explain mechanistically may be. Figure 1.26 attempts to make this point by showing the MWC model as an intellectual node that links many disparate biological phenomena.

To be specific, Figure 1.26 puts forward the suggestion that phenomena as diverse as the packing of DNA in nucleosomes and the binding of oxygen to hemoglobin are related through physical proximity. But the relatedness of these different problems becomes really clear only when formulated mathematically, in precisely the same way that a mass-spring system and an *LC* circuit are the "same" thing is revealed by the underlying mathematics. The allostery phenomenon as embodied in the MWC model can be stated through the idea that there is one equation to rule them all, namely,

$$p_{active}(c) = \frac{\left(1+\frac{c}{K_A}\right)^n}{\left(1+\frac{c}{K_A}\right)^n + L\left(1+\frac{c}{K_I}\right)^n}. \tag{1.2}$$

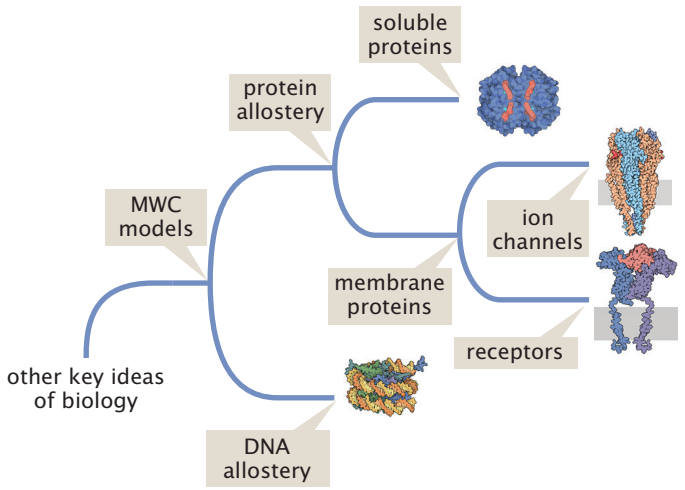


Figure 1.26

Physical proximity of diverse biological phenomena. The MWC model links hemoglobin, ligand-gated ion channels, G protein–coupled receptors, and nucleosomes all through one physical/mathematical framework. Courtesy of David Goodsell.

This equation tells us the probability that an MWC molecule with n ligand binding sites will be in the active state as a function of ligand concentration c. The model features three key parameters L, K_A and K_I. The parameter L is the equilibrium constant between the active and inactive states in the absence of ligand, K_A and K_I are the dissociation constants for ligand binding to the active and inactive states, respectively.

A recent exciting development in evolutionary cell biology is the realization that it is possible to explore the biophysical basis of the parameters that yield the different phenotypes shown in Figure 1.21 such as leakiness, EC_{50}, and effective Hill coefficient explored by evolution. In the context of the MWC model, the entire space of phenotypes is determined by the three molecular parameters introduced in the context of equation 1.2: L, K_A and K_I. That is to say, the only way molecular structure reflects on function in the context of the MWC model is through the value of these three parameters which set key characteristics such as the leakiness, the dynamic range, the EC_{50}, and the effective Hill coefficient which determine the sensitivity of the molecule to ligand concentration. Of course, this is a very naive view of molecules and their evolution, but it will serve as the jumping off point for our thinking.

In many ways, the task of the book is to show where equation 1.2 comes from, why it is the same for so many distinct biological problems, and what its implications are for thinking about biological phenomena ranging from quorum sensing to enzyme feedback. An array of examples of the way this equation can be used to describe different biological phenomena is highlighted in Figure 1.27. This figure is intended to whet the reader's appetite for statistical mechanical modeling of a host of different biological phenomena. There we see examples as diverse as the activity of enzymes of glycolysis such as phosphofructokinase (see chap. 6), oxygen binding to hemoglobin (see chap. 7), ligand-gated ion channels (see chap. 3), the activity of chemotaxis receptors (see chap. 4), and the activity of G protein–coupled receptors (see chap. 5). For now, we content ourselves with admiring these various activity curves for the generality of the allosteric phenomena that they reveal.

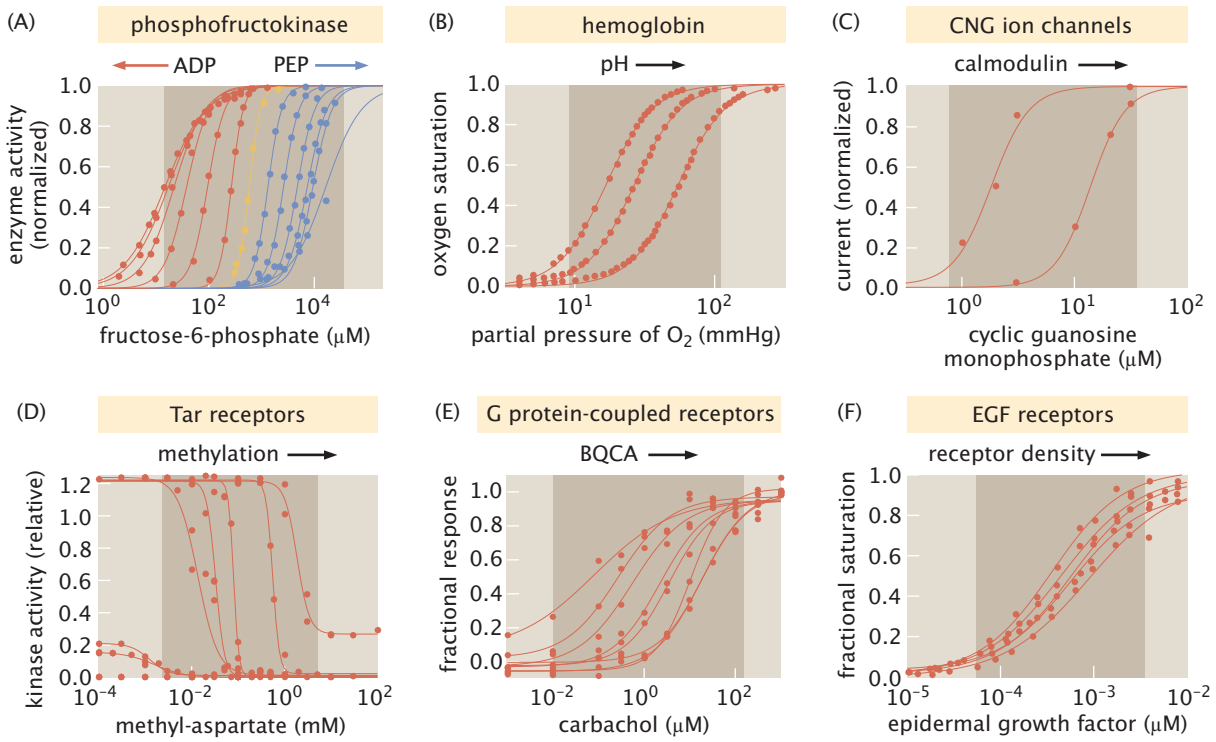


Figure 1.27

Diversity of activity curves of key allosteric molecules. Each dose-response curve shows the activity as a function of an effector molecule. In each case, there is a family of curves reflecting the fact that the activity curve can be tuned by other molecules (shown by the arrows above each graph) in the process of adaptation. Adapted from Olsman and Goentoro (2016).

1.5 Beyond the MWC Two-State Concept

From the outset, I want to make clear that our subject has a long and rich history filled with subtle phenomena, deep and creative models, and colorful personalities. What this means to those that participated in the creation of the subject is that there are many nuances that represent years of work tied to fierce intellectual battles. Though it may seem to miss some of the nuance, my plan is largely to fly below the radar of these debates and to provide a warm statistical mechanical embrace to many different variations of the same basic theme. Specifically, we will argue that these different categories of models all fall within the same statistical mechanical fold because they are based upon discrete state spaces and because ligand binding tunes the relative free energies of inactive and active configurations. We provide a first view of these ideas now.

1.5.1 Molecular Agnosticism: MWC versus KNF versus Eigen

A central theme of the book is the power and beauty of coarse-grained descriptions that intentionally suppress reference to the microscopic degrees of freedom. A corollary of this point of view will be that sometimes we will have a measured indifference to the molecular particulars of a given problem. As such,

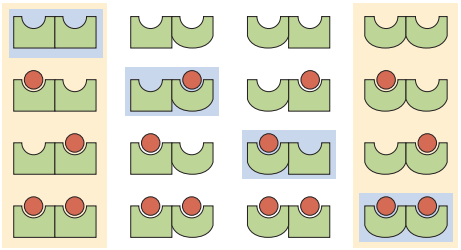


Figure 1.28

Compact representation of the MWC (yellow states), KNF (blue states), and Eigen models (all states) for a two-site allosteric molecule. The active state has the rectangular shape, and the inactive monomers have a rounded shape.

we will pass freely between descriptions based on either all-or-nothing (MWC) or sequential (KNF) or hybrid models of molecular conformations.

To get a flavor for the different conformational states and states of occupancy that allosteric molecules can sustain, Figure 1.28 shows some of the different ways that one can imagine assigning microscopic states to allosteric molecules. The MWC model will be our reference model, culminating in equation 1.2. However, an alternative picture is offered by the sequential model of Koshland, Némethy, and Filmer developed in 1966. This model imagines obligate conformational changes whenever a subunit is bound by a ligand, as seen in the blue states of Figure 1.28. An even broader generalization introduced by Manfred Eigen (1968) allows for the possibility that the different subunits independently change between the inactive and active states and that both sets of states allow for ligand binding, albeit with different dissociation constants. As already mentioned, the debates surrounding these different molecular possibilities have engendered passionate debates. For our purposes, we will view them much more liberally as comprising different collections of allowed states but with the key unifying property of the presence of inactive and active states that have different binding constants for ligands.

One concern justifiably put forth by those critical of the two-state concept is that real biological macromolecules sometimes have more than two dominant conformational states. Part of the reason that I find debates about the nature of allostery dated is that such arguments miss for me the statistical mechanical essences which are simpler and can be flexibly altered to account for the more complex situations, going all the way to the case of intrinsically disordered proteins in which there is a continuum of different states, as will be discussed in detail in chapter 11 (p. 347).

To be specific, we can easily accommodate generalizations to cases involving more substates, as shown in Figure 1.29, where we consider a molecule with three conformational states. In this case we show the "three-state MWC molecule" in which there are three conformational states, each of which is subject to ligand binding or not, resulting in a total of six possible states. The statistical mechanical protocol we will develop in the next chapter will allow us to write the energies of each of these states, as well as their statistical weights, resulting in expressions for their different probabilities, though we need to be alert to the proliferation of parameters as we accept more and more states. Similarly, we can also accommodate the even further generalizations of these multistate models to the full level of the Eigen model by allowing for all states of occupancy and all partial conformational states.

We can explore this proliferation of parameters more systematically by reflecting on what we have done thus far and by generalizing to cases involving N states and M binding sites. For a molecule with only one conformational state

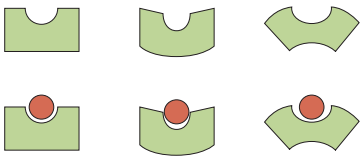


Figure 1.29

Generalizations of the two-state concept. In this example we consider a three-state, one-site model in which there are three conformational states, each of which can be empty or bound by ligand.

but M binding sites, the number of distinct states is 2^M, since each site can be either empty or occupied by ligand. The general rule for the generalized MWC model is that if we have N states and M binding sites, then the total number of distinct configurations that we need to consider in our states-and-weights diagram is

$$\sum_{i=1}^{N} 2^M = N2^M. \tag{1.3}$$

Each of the N conformational states will have its own ϵ_i characterizing the energy of that conformation. Further, for each of those states, there will be a distinct binding energy (or K_d).

The two-state paradigm can be generalized even further to the case in which there is a continuum of allowed states, as we will discuss in chapter 11 (p. 347). As early as the 1980s, Cooper and Dryden explored generalizations of the allostery concept in which there was not a strict conformational change, but, nevertheless, the inactive and active conformations had different free energies of binding because such binding altered the vibrational free energy of the system. An exciting recent development that builds on this kind of thinking has been the emergence of the paradigm of intrinsically disordered proteins, which forces us to think more broadly about the different states allowed to allosteric proteins.

1.6 On Being Wrong

As already highlighted in the preface, the aim of this book is to explore a highly idealized and abstracted view of models of allosteric molecules. There are many opinions on the value of models in the context of biology, some of which are negative and focus on how such models are wrong by "missing" some key element of the system. In that vein, it has become a cliché to quote George Box's refrain that "all models are wrong, but some are useful", at this point, however, I think that this quote itself has outlived its usefulness. In his great essay *Common Sense*, Thomas Paine notes "A long habit of not thinking a thing wrong gives it a superficial appearance of being right." I also worry that a long habit of not thinking a thing right gives it a superficial appearance of being wrong. In the context of the subject of this book, I have been amazed at the diversity of unsubstantiated opinions starting with the words "we all know that" that then go on to assert that either the MWC model has already been shown either to work or to fail, with both opinions held with equal conviction.

It is hard to escape the feeling after studying the literature that often the wrongness is not with the model but with a lack of a truly rigorous dialogue between theory and experiment or an incomplete generalization of the model

when faced with new circumstances. For example, there are some that continue to use equation 1.2 (p. 26) as "the MWC model" even when they are considering cases with more than one molecular species. However, in the case in which there is some competitor ligand, equation 1.2 needs to be generalized to

$$p_{active}(c_1, c_2) = \frac{(1 + \frac{c_1}{K_{A,1}} + \frac{c_2}{K_{A,2}})^n}{(1 + \frac{c_1}{K_{A,1}} + \frac{c_2}{K_{A,2}})^n + L(1 + \frac{c_1}{K_{I,1}} + \frac{c_2}{K_{I,2}})^n} \tag{1.4}$$

where c_1 is the concentration of species 1, and c_2 is the concentration of species 2, the competitor.

A second way in which the model of equation 1.2 must be generalized is in the case of important situations like those shown in Figure 1.7, (p. 10) in which there are multiple inputs to the same allosteric molecule. The one equation that rules them all (the traditional MWC model) must account for these multiple inputs. For example, if there are two ligands that interact with the MWC molecule, equation 1.2 must be generalized as

$$p_{active}(c_1, c_2) = \frac{(1 + \frac{c_1}{K_{A,1}})^n (1 + \frac{c_2}{K_{A,2}})^n}{(1 + \frac{c_1}{K_{A,1}})^n (1 + \frac{c_2}{K_{A,2}})^n + L(1 + \frac{c_1}{K_{I,1}})^n (1 + \frac{c_2}{K_{I,2}})^n}. \tag{1.5}$$

For the case in which there are M distinct sites subject to binding by different ligands labeled by the index i, the MWC equation introduced earlier as eqn. 1.2 must be generalized even further to

$$p_{active}(\{c_i\}) = \frac{\prod_i^M (1 + \frac{c_i}{K_{A,i}})^n}{\prod_i^M (1 + \frac{c_i}{K_{A,i}})^n + L \prod_i^M (1 + \frac{c_i}{K_{I,i}})^n}. \tag{1.6}$$

The point here is that these kinds of generalizations are really necessary in order to broaden the scope of the original MWC formulation. Each scenario will have a slightly different equation, and as far as this book is concerned, they all fall within the bailiwick of the MWC framework. Failure to use the right equation will invariably lead to a superficial appearance of wrongness that may not be justified at all.

1.7 Summary

Our first chapter set the stage for all that will follow. It began with one of the most important *facts* of biology (many of the macromolecules of life are allosteric) and one of the most important *concepts* in biology (the MWC model of allostery). I heard of a very distinguished and well-known biologist who refused to talk to physicists entering biology until they could properly define allostery, and the aim of this chapter was to give the reader enough background to pass that test. To that end, I showed how allosteric molecules are found in all corners of biology, whether neuroscience or physiology or evolution. The mathematical implementation of the allostery concept leads to a transcendent biological framework in many ways analogous to the way that concepts such as resonance or interference are transcendent in physics. The chapter ended by

openly acknowledging the many oversimplifications and caricatures inherent in this class of models but argued that despite these shortcomings, the framework is extremely potent.

1.8 Further Reading

Alon, U. (2007) *An Introduction to Systems Biology*. Boca Raton, FL: Chapman and Hall/CRC. This excellent book is visionary in showing the interplay between careful theoretical thinking and well-designed experiments to attempt to deeply understand biological problems.

Ben-Naim, A. (2001) *Cooperativity and Regulation in Biochemical Processes*. New York: Kluwer Academic/Plenum. This excellent book channels earlier work from Terrell Hill that demonstrates the naturalness of Gibbs's grand partition function for describing the binding and activity of allosteric molecules.

Cantor, C. R., and P. R. Schimmel (1980) *Biophysical Chemistry*. New York: W. H. Freeman. This series of books digs deeply into many aspects of allostery. The cover of my edition of the book pays homage to the concept of allostery.

Feynman, R. P., R. B. Leighton, and M. Sands (1963) *The Feynman Lectures on Physics*. Reading, MA: Addison-Wesley. Everything here is worth reading, but the chapter on resonance is particularly delightful.

Lim, W. B. Mayer, and T. Pawson (2014) *Cell Signaling*. New York: Garland Science. Over a happy and challenging six months I read every page of this excellent book. My book is an attempt to see what happens when one tries to mathematicize the ideas in this work.

Lindsley, J. E., and J. Rutter (2006) "Whence cometh the allosterome?" *Proc. Natl. Acad. Sci.* 103:10533–10535. This very important paper lays down the gauntlet by noting the challenge of figuring out which proteins are allosteric and if they are, what molecules control that allostery.

Martins, B. M. C., and P. S. Swain (2011) "Trade-offs and constraints in allosteric sensing." *PLoS Comput. Biol.* 7(11):e1002261. This article is a must-read for anyone truly interested in allostery. The authors explore many important facets of the statistical mechanics of models of allostery.

Mollon, J. D. (2002) "The origins of the concept of interference." *Phil. Trans. R. Soc. London* A360, 807–819. This fascinating article describes how Thomas Young brought unity to our understanding of the phenomenon of wave interference. For those interested in delving more deeply into the topic and especially how it touches on biology, see Nelson, P. (2017) *From Photon to Neuron: Light, Imaging, Vision*. Princeton, NJ: Princeton University Press. And any interested scientific reader should study Feynman, R. P. (2014) *QED: The Strange Theory of Light and Matter*. Princeton, NJ: Princeton University Press.

Motlagh, H. N., J. O. Wrab, J. Li, and V. J. Hilser (2014) "The ensemble nature of allostery." *Nature* 508:331. This paper reflects on the evolution of our understanding of the allostery phenomenon with the emergence of new experimental techniques such as NMR which have substantially generalized the mechanistic underpinnings of how molecules can behave allosterically.

(continued...)

INDEX