

Enriching the Strategies for Creating Mechanistic Explanations in Biology

William Bechtel
Department of Philosophy and Center for Circadian Biology
University of California, San Diego

Abstract

To demonstrate that a proposed mechanism could explain a phenomenon, biologists must recombine the mechanism. Traditionally they have relied on mentally rehearsing the operations, often aided by a mechanism diagram. Such a strategy has reached its limits in contemporary biology. For example, through mental rehearsal alone researchers cannot determine whether a feedback mechanism will generate sustained oscillation. Accordingly, mechanistic inclined biologists are enriching their strategies, relying on computational simulation and graph-theoretical analyses of networks. The limitations of traditional approaches and the additional strategies biologists are employing are illustrated in research on glycolytic and circadian oscillations.

Keywords: Mechanistic explanation; decomposition; recombination; computational simulation; network analysis

1. Introduction

The approach of explaining phenomena by identifying and characterizing responsible mechanisms has a long history in biology (for historical reviews, see Coleman, 1971; Allen, 1979). After Theodor Schwann (1839) identified cells as the basic units in which metabolic processes such as fermentation occur, a host of researchers developed and deployed a very impressive variety of strategies for structurally decomposing cells and functionally characterizing what their components did. The discovery of chromosomes and the characterization of the operations of mitosis and meiosis was a prominent 19th century success. Vitalist critics persisted throughout the century in arguing that phenomena such as fermentation could not be explained mechanistically, but Eduard Buchner's (1897) demonstration of fermentation in a cell-free extract inspired the quest for explanations of cell activities in terms of chemical reactions catalyzed by enzymes. As reflected in Edmund Cowdry's (1924) *General cytology*, mechanistically inspired biologists in the first decades of the 20th century were developing new techniques for decomposing mechanisms into their parts (e.g., staining cell preparations to identify organelles) and operations (e.g., inhibiting enzymes to identify steps in reactions). As reflected in Cowdry's own chapter on the mitochondrion and the Golgi apparatus, a major aspiration was to localize different cellular activities in specific organelles. The development of new techniques continued in the decades after Cowdry's book. Cell fractionation and electron microscopy played central roles, facilitating what George Palade (1987) described as a bridge between morphology (providing increased detail about cell structure) and biochemistry (characterizing cell reactions). Within the newly constituted discipline of cell biology, researchers combined

these techniques to offer new, mechanistic explanations of phenomena such as oxidative metabolism and protein synthesis (Bechtel, 2006).

Only much more recently did philosophers of science develop detailed accounts of mechanistic explanations (and Machamer, Darden, & Craver, 2000; Bechtel & Abrahamsen, 2005) but they have now become a central focus of philosophical analysis. Much of the philosophical inquiry focuses on the processes by which scientists develop mechanistic explanations by (1) linking phenomena to be explained to mechanisms and (2) decomposing these mechanisms into responsible parts and operations (Bechtel & Richardson, 1993/2010; Craver & Darden, 2013). All accounts of mechanisms also acknowledge the importance of how these components are organized in generating the phenomena. Organization becomes a focus for biologists as they attempt to recompose a mechanism, at least conceptually, to show that it can generate the phenomenon. But philosophers have paid less attention to how biologists recompose mechanisms. In contemporary research papers, in which most of the text and figures are devoted to presenting new experimental findings about parts or operations, the researchers may conclude with a description of a mechanism, often accompanied by a diagram. The diagram typically does much of the work of conveying how the researchers conceive of the parts and operations fitting together. The emphasis on recomposing mechanisms, often relying on mechanism diagrams, is more frequent in commentaries on research papers and in review articles.

Despite playing a central role in biologists' attempt to recompose mechanisms, philosophers, including philosophers discussing mechanistic explanations, have offered little analysis of how diagrams play that role (but see Abrahamsen, Sheredos, & Bechtel, in press). Here I note two features of diagrams that make diagrams particularly well suited for representing the organization of a mechanism. First, abstract shapes (glyphs, Tversky, 2011) or icons are used to represent parts and arrows (often of varied formats) are employed to show how an operation performed by one part affects other parts. Second researchers can use the two spatial dimensions of a diagram¹ to represent the relations between parts (sometimes how they are spatially related, as in Figure 1 below) and sometimes how they are functionally related (as in Figure 2). A diagram can better present the often very complicated relations between parts and operations of a mechanism than text (often, if no diagram is presented, readers construct diagrams for themselves). One important advantage of diagrams is that they allow viewers to direct their attention to different components of the mechanism as desired. A diagram, however, is static and itself doesn't show how the parts, working together, are able to produce the phenomenon. To understand this, the viewer must mentally animate the diagram (Hegarty, 1992) by rehearsing the operations specified by the arrows in his or her imagination. An explanatory text often includes a narration of the operations of the mechanism and serves to guide the viewer in mentally animating the diagram.

¹ Drawn or printed diagrams are limited to two dimensions (although on occasion overlays will be used to reflect additional dimensions); sometimes researchers find it important to represent a mechanism in three dimensions and build physical models.

Mental animation of a diagram, guided by narration, usually suffices for establishing understanding when the operations of the proposed mechanism are construed as executed sequentially, “from start or set-up to finish or termination conditions” (to borrow the language used in Machamer, Darden, and Craver’s definition of mechanism) whenever the starting conditions are present. Assuming regular sequential operation is a good heuristic starting point and the resulting explanations, which I refer to as *basic mechanistic explanations* (Bechtel, 2011), have often seemed to suffice to explain biological phenomena. Researchers, for example, typically view biochemical pathways as sequences of individual reactions.

However, the very strategies of decomposing mechanisms to identify parts and operations have continued unabated. In the case of almost all biological mechanisms, including those thought to be well understood, researchers are discovering more and more parts. Researchers often confront three problems in recomposing these parts into a mechanistic explanation that they could mentally simulate. First, the parts do not operate just in a sequence but in many parallel streams that proceed at different rates yet interact with each other at numerous points. It is challenging to rehearse mentally multiple interacting sequences of activity at the same time. Second, researchers frequently discover that operations they view as later in the sequence feed back on those they view as earlier. In such circumstances, their mental simulations have to take into account the effects of later operations when rehearsing the earlier one. Third, many of the additional parts researchers identify for one mechanism turn out to be parts that also belong to other mechanisms. Mechanisms cease to have well-delineated boundaries but are embedded in large interactive networks in which long-range connections modulate the behavior of what were taken to be independent mechanisms (Bechtel, 2015).

Diagrams can represent multiple pathways, non-sequential organization and the embeddedness of mechanisms in larger networks. But the challenge of understanding them—understanding how the mechanism portrayed generates the phenomenon one is trying to explain—begins to stress human cognitive capacities. To appreciate the problems, one can look ahead to Figures 2 and 3 or one can think about the challenges in understanding human agents that are engaged in complex systems—for example, an orchestra player who not only must coordinate with her instrument but also with those around her, whose behavior is in part affected by her own behavior. Humans cannot mentally simulate all the operations and interactions proposed in the account of the mechanism to determine whether the proposed mechanism would generate the phenomenon. To make progress, researchers have had to enrich their strategies for investigating mechanisms beyond those that worked in developing basic mechanistic explanations.

Seeking to address of these challenges is a major motivation of biologists who have adopted the name *systems biology* (see, for example, Alon, 2007; Ideker, Galitski, & Hood, 2001; Kitano, 2002).² I focus here on two major strategies systems biologists have adopted

² Systems biology has intellectual roots in the cybernetics and general systems theory movements of the 20th century, and these traditions have informed some of the modeling

to address these problems and how these complement the more traditional strategies of mechanistic research. The first involves mathematically characterizing multiple interacting components and simulating their collective behavior computationally while the second involves representing components as networks of nodes and edges and deploying graph theoretic and other tools for analyzing these graphs. I will introduce these explanatory strategies and show how they differ from yet complement those that gave rise to the development of basic mechanistic explanations.

In the next section I introduce the strategy of basic mechanistic explanation by describing its application to cells and show how it contributed to the growth of biological knowledge. In section 3 I explore how this strategy approached its limits as researchers encountered feedback loops that engender oscillatory behavior such as the daily raise and fall in human bodily temperature or in the concentrations of various proteins. Body temperature and protein synthesis are circadian phenomenon in that they are controlled by an internal mechanism that generates oscillations of approximately 24 hours. The strategies of basic mechanistic explanation could not determine whether the proposed mechanism would produce the observed sustained oscillations or would dampen. In the following two sections I will show how circadian researchers are supplementing basic mechanistic approaches with the two additional explanatory strategies introduced above.

As I will briefly discuss in the concluding section, the introduction of new strategies for advancing explanations is not novel. The strategies for developing basic mechanistic explanations of cell phenomena themselves developed over the 19th and 20th century and the mechanistic explanations of cell behavior that figure prominently in contemporary textbooks are a product of those strategies. Researchers have turned to new strategies of computational modeling and network analysis to cope with the success of those strategies in discovering more and more components organized and interacting in more complex ways. Almost certainly as biological research continues, additional strategies will be developed. But computational modeling and network analyses are already generating explanations that are more dynamic and integrated than the basic mechanistic explanations advanced in the 20th century.

2. The Quest for Basic Mechanistic Explanations

The idea of putting different types of components together to accomplish what individually they could not do has deep roots in engineering design. Ancient Greeks combined simple machines such as the wheel and axle and the pulley into compound machines such as a crane. Descartes vigorously advanced the idea that phenomena in the natural world, including those associated with living organisms, also result from machines. As was true of the mechanistic explanations espoused by Descartes, many of the early proposals of biological mechanisms were highly speculative. By the 19th century, however, researchers began to develop techniques that enabled them to pursue mechanistic accounts grounded in empirical knowledge of the parts and operations. These techniques involve identifying candidate mechanisms and then decomposing them into their parts and operations. For

approaches adopted in system biology. But far more fundamental was the development of techniques for collecting massive data about genes, proteins, and metabolites.

example, in the wake of the chemical revolution, chemists analyzed the constitution of different molecules in living organisms, identified reactions that transformed one into another, and discovered that biological tissues often contained catalysts (later termed *enzymes*) that facilitated these chemical reactions (Berzelius, 1836). It was in this same period that Schwann (1839), utilizing newly improved microscopes, identified cells as the basic units of living systems and went on to associate them with processes that transform food stuffs into new tissue and energy, which he named *metabolism*.

Schwann could go no further in explaining the metabolic power of cells than to appeal to their distinctive chemical composition, which he proposed resulted from a process of cell formation through an iterative depositing of different materials around a core as in crystal formation. Further advances awaited the development of new tools of biochemistry and cell biology. Starting around the beginning of the 20th century, biochemists identified biochemical groups that are transferred between substrates in reactions and by the 1930s Gustav Embden, Otto Meyerhof, and numerous other investigators had pieced together an account of the glycolytic pathway, a sequence of reactions transforming glucose to pyruvate that yielded modest production of ATP, the molecule that provides energy for other cellular functions. The introduction of the new techniques of cell fractionation and electron microscopy, beginning in the 1940s, was pivotal in generating a basic understanding of the mechanisms responsible for the subsequent oxidation of pyruvate to carbon dioxide and water coupled with much more synthesis of ATP than glycolysis (Bechtel, 2006). Cell fractionation enabled researchers to differentiate fractions with different enzymatic composition that originate in different organelles of the cell while electron microscopy permitted visual identification of these organelles and their structure. In particular, cell fractionation allowed the localization of the citric acid (Krebs) cycle and electron transport to the mitochondrion while electron microscopy revealed the distinctive organization involving the inner membrane protruding into the cell interior. As shown in Figure 1, the citric acid cycle was localized to the inner matrix while electron transport and the coordinated phosphorylation reactions to the cristae.

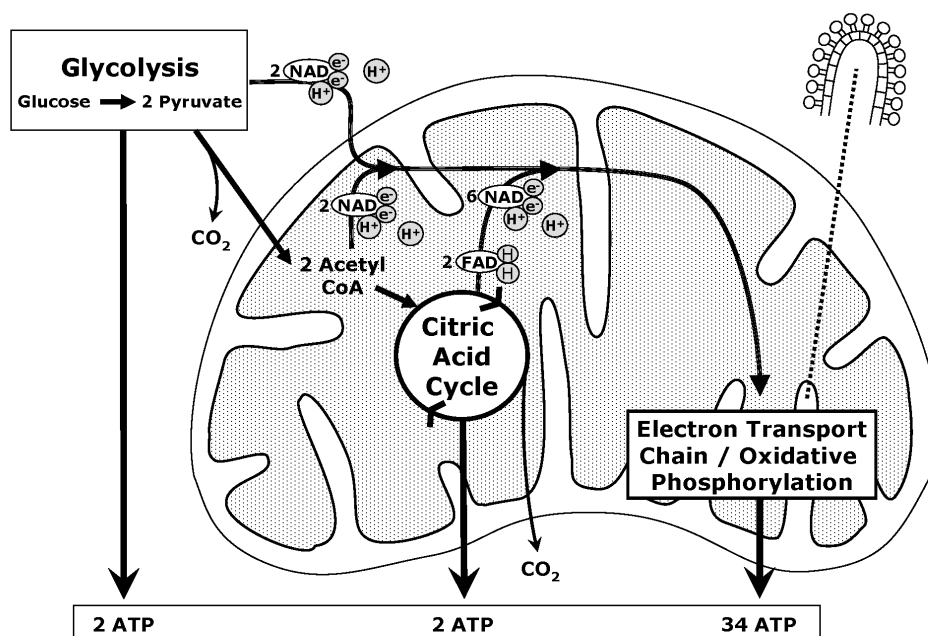


Figure 1. Conception of the mechanisms involved in basic energy metabolism as arrived at by the 1960s using the tools of cell fractionation and electron microscopy. As discussed above, decomposition is only part of the strategy for constructing mechanistic explanations. The behavior of a mechanism depends on how components are organized and how operations performed by one part affect others. Important graphical components of Figure 1 are the arrows that link the different operations, establishing what Machamer et al. refer to as *productive continuity*. To understand the activity of the mechanism, researchers mentally animate the operations. A researcher familiar with the different operations in Figure 1 can envisage a molecule of glucose being oxidized to yield two molecules of pyruvate, which, after being transformed into acetyl-CoA, enter the citric acid cycle, etc. Once research has succeeded in identifying a mechanism, decomposing it into its parts and operations, and recomposing these into an organized set of productively continuous operations that researchers can mentally animate, the search for a mechanistic explanation seems to have reached a successful conclusion.³

2. Pushing Basic Mechanistic Explanation to Its Limits

The account of metabolism in the previous section followed the strategy of basic mechanistic explanation in offering a largely feed-forward account from start to termination conditions. The one exception is that the citric acid cycle involves a feedback loop in which the initial component of the cycle is regenerated from a product and new incoming acetyl CoA. As typically approached, however, this does not present any special challenges as someone animating the diagram needs only follow the Krebs cycle far enough to where it outputs to other operations. In this section I will consider how further pursuit of mechanistic research did push the basic mechanistic strategy to its limits when research on glycolysis revealed the type of complex dynamics (the periodic increase and decrease in concentrations of intermediary metabolites) and linked these to feedback loops. I then turn to a different phenomenon, circadian rhythms, where the complex dynamics was known from the outset. In both cases, mechanistic strategies could generate part but not all of what was needed for an explanation.

The reason the discovery of complex dynamics posed a problem for basic mechanistic explanation is that in pursuing such accounts researchers assume that mechanisms function in a regular manner. Any change in how the mechanism responds is assumed to be due to the external input, not processes endogenous to the mechanism. Whenever glucose is available to the first reaction in the glycolytic pathway, the subsequent reactions are thought to occur seriatim. To a first approximation, many biological mechanisms do operate in this manner, presenting ideal conditions for the successful application of mechanistic research strategies. However, the data biologists record often manifests substantial variability. Researchers often attribute this to measurement errors or dismiss it as noise. In some cases, however, a pattern is found that reveals underlying dynamical

³ Biochemists since the beginning of the 20th century have invoked mathematical representations to characterize operations within a mechanism. For example, the Michaelis-Menten equation is employed to determine how the concentration of the substrate affects the rate of production of the product. But this mathematical analysis is not required for understanding the behavior of the mechanism, characterized qualitatively.

behavior that is fundamental to the phenomenon. This happened in research on glycolysis. While measuring concentrations of glycolytic intermediates in yeast using spectrophotometric techniques, Amal Ghosh and Britton Chance (1964) discovered that the concentration of NADH oscillated with a period of about a minute. Benno Hess, Arnold Boiteur and J. Krüger (1969) subsequently demonstrated periodic oscillations in the concentrations of other reactants. Moreover, neighboring reactants in the glycolytic pathway oscillated in phase with each other, whereas those on opposite sides of two major reactions were phase reversed (i.e., 180° out of phase). One of these is the reaction in which fructose-6-phosphate (F6P) is phosphorylated to fructose-diphosphate (FDP) at the expense of transforming ATP to ADP through the action of the enzyme phosphofructokinase (PFK).

By focusing on PFK, researchers were able to extend the mechanistic explanation to partly account for this oscillation. PFK is an allosteric enzyme that contains binding sites for multiple molecules and the binding at some sites causes conformation changes at other sites, altering reactivity at those sites. In particular, PFK binds with three of the products of the main reaction, namely FDP and ADP, as well as AMP (a product generated by removing another phosphate group from ADP). When PFK is bound to FDP, ADP, and AMP the reaction from F6P to FDP runs faster. The downstream effect of this positive feedback loop is to increase the production of NADH and, even further downstream, the synthesis of ATP from ADP or AMP. This further effect realizes a long-range negative feedback loop that counters the short-term positive feedback loop; their joint action is first to speed up NADH synthesis and then to slow it down. This verbal narrative suggests how the parts of the glycolytic system combine to produce oscillations, but it is important to note that there is an alternative possibility—that the system reaches an equilibrium at which NADH concentrations stop oscillating. This presents a limit to the basic mechanistic strategy: it is not able to determine which outcome will be realized.

Glycolytic oscillation was discovered in the context of an already worked out mechanism and it remains unclear whether it plays a functional role in the metabolism of yeast. But in the case of many other physiological functions, oscillations clearly play a functional role. For example, rhythmic contraction of muscles is crucial for the circulation of blood. Increasingly neuroscientists are discovering that sub-threshold oscillations of ions in neurons, detectable by EEG or in resting state fMRI, figure in the coordination of processing in different regions of the brain. From here out I will focus on a system whose primary function is to maintain an oscillation, circadian rhythms. These are oscillations with a period of approximately 24 hours that are generated endogenously in many living organisms, can be entrained to the light-dark cycle in the environment, and regulate many physiological and behavioral activities.

One of the first challenges circadian researchers faced was to establish that these daily rhythms, observed in physiological measures such as body temperature or in physical behaviors such as running, are generated within the organism and are not simply responses to cues from the environment. The crucial evidence was provided by the fact that the period found through studies of behavior in conditions in which cues have been removed (referred to as free-running conditions) varied slightly from 24 hours. After

establishing that these rhythms are endogenous, research turned to figuring out the mechanism (which, early on, was referred to as a *clock*).

Initially, progress in identifying and decomposing the clock was slow. Research on fruit flies resulted in one of the first breakthroughs occurred. Fruit flies exhibit circadian oscillations in locomotive behavior as well as in timing of their eclosion from their pupae. Ronald Konopka and Seymour Benzer (1971) identified a gene, *period* (*per*), which, when mutated in different ways, resulted in slow or fast rhythms or arrhythmic behavior. Once cloning became available, Paul Hardin, Jeffrey Hall, and Michael Rosbash (1990) demonstrated that both *per* mRNA and the protein PER oscillate with a period of 24 hours, with the concentration of the mRNA peaking about four hours before the concentration of the protein peaks. Knowing that negative feedback is a design principle that can generate oscillations, Hardin et al. proposed that the mechanism had the form of a transcription-translation feedback loop (TTFL) according to which PER feeds back to inhibit the transcription of *per*.

To see how such a mechanism could generate oscillations, one can try to simulate its operation mentally. Start in the state in which the concentration of PER is low. Since there is little inhibition on the rate of *per* transcription and translation, the concentration of PER gradually increases. But as it does so, it increasingly inhibits *per* transcription, stopping the increase in its concentration. Since PER gradually degrades, the concentration will now start to decline. As it declines, the inhibition is reduced, and the concentration of PER begins to rise again. As with glycolytic oscillation, basic mechanistic explanation reaches a limit in that one could also tell a narrative in which the mechanism approaches a steady state and stopped oscillating. I return to this in the next section, but first will follow the history a bit more to see how the conception of the circadian clock expanded.

One shortcoming of the TTFL proposed by Hardin et al. was that researchers could not find a DNA-binding site on the PER protein. Such a site is required if PER is to bind to the promoter of its own gene and block its own transcription. This gap in the account was only filled when Joseph Takahashi's group (Vitaterna, King, Chang, Kornhauser, Lowrey, McDonald, Dove, Pinto, Turek, & Takahashi, 1994) undertook a comparable search for mutants in mice and identified a gene they named *Clock* (Circadian Locomotor Output Cycles Kaput).⁴ Mutations to *Clock* resulted in altered circadian rhythms but, more importantly, CLOCK possessed the needed region to bind to the promoter (known as an E-box) on *per*. A homolog of *Clock* was soon identified in fruit flies, and within a few years, three homologs of *per* were found in mammals. Together these genes and proteins constitute the major negative feedback loop shown in the top portion of Figure 2 (the figure shows PERs as dimerizing with CRYs, and CLOCK with BMAL1). Figure 2 presents many but not all the parts and operations that had been identified and fit into an account of the circadian mechanism by 2005. As the figure makes clear, these additional parts figure in various feedback loops, both positive and negative.

⁴ The naming convention for genes in fruit flies is to use lower case italics. In mammals, gene names are also in italics but begin with an initial capital. Protein names conventionally are all in roman capital letters.

As noted above, although this diagram shows many of the crucial parts and operations, it is static. The viewer must supply the dynamics as she tries to mentally simulate the operations. The additional feedback loops make this task even more difficult, and they have not alleviated the problem that one could equally simulate sustained oscillations or ones that dampen to a steady state.

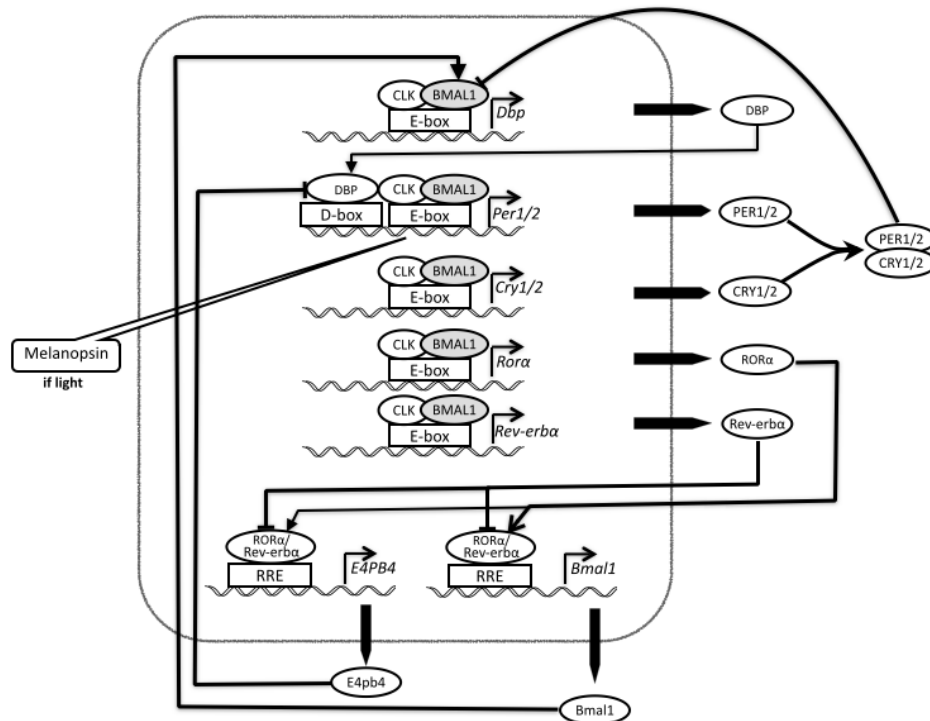


Figure 2. A mechanism diagram of the mammalian circadian clock showing the principal genes and proteins that figure in feedback loops as understood circa 2005.

In this section I have considered two cases in which the strategies of basic mechanistic research were pushed to their limits. With the discovery of the occurrence of oscillation in glycolysis, researchers were able to localize the responsible enzyme and provide a qualitative narrative of how it operated. In the case of circadian rhythms, researchers likewise began with a single feedback loop and could offer a narrative as to how it generated oscillations. But in neither case could they differentiate this narrative from one in which the system approached a steady state at which oscillations ceased. A further limit was reached when researchers identified multiple feedback loops in a diagram such as Figure 2. Mentally simulating more than one feedback loop operating at once is challenging. Moving forward to show how these mechanisms generated the phenomena required supplementing the strategies for developing mechanistic explanations with new explanatory strategies.

3. Explaining Dynamic Behavior through Computational Simulation

The previous section revealed one respect in which oscillatory phenomena stretches basic mechanistic explanation to its limits. Feedback systems generate oscillations, but many dampen over time as the system reaches a steady state. From mentally simulating a mechanism such as than shown in Figure 2, or even the initial proposal for a feedback loop

involving *per* alone, one cannot tell whether the oscillation will be maintained or whether it will reach a point at which the increase of PER from transcription and translation is perfectly compensated by the rate at which it degrades.

The main alternative to which scientists turn when mental simulation fails is computational simulation. When applied in cell and molecular biology, modelers create differential equations to describe how the concentrations of individual components are affected by the concentrations of one or more other components. What one realizes quickly with feedback systems is that if one assumes that all the equations are linear, oscillations resulting from feedback will quickly dampen. One or more non-linear equation is required. To determine how the TTFL proposed by Hardin et al. would behave, Albert Goldbeter (1995) constructed a computational model⁵ consisting of five differential equation. The first equation represents how *per* mRNA (M) changes as a result of being increased by the rate of transcription of the gene *per* (first term after the equal sign) and decreased by the rate of decay of the protein PER (second term):

$$\frac{dM}{dT} = v_s \frac{K_l^n}{K_l^n + P_N^n} - v_m \frac{M}{K_m + M}$$

v_s , v_m , K_l , and K_m , as well as n , are parameters, and the choice of values for parameters has a significant effect on how the simulation behaves. I will, however, only comment on n , which appears as the coefficient of the variable P_N in the denominator of the first term. P_N represents the concentration of PER in the nucleus. Putting it in the denominator has the effect of reducing the increase of M as P_N increases, thereby capturing the role of PER in the nucleus in inhibiting its own transcription. How much of an inhibitory effect P_N has is critically affected by the coefficient n , known as the Hill coefficient, which reflects how many molecules of PER are required to inhibit *per* transcription. In the second term M appears in the numerator as well as the denominator, where a parameter is added to it. This has the effect of increasing the rate of decay in a non-linear manner as the amount of M increases. Using what he claimed were biologically plausible values for parameters and applying each of the five equations iteratively, Goldbeter showed they generated stable oscillations in variables such as M .

The introduction of more components into the mechanistic account raised the possibility that the new system would yield dampened oscillations. Accordingly, Jean-Christophe Leloup and Goldbeter (2008), added additional equations and used simulations to show that the expanded computational model would, with what again were assumed to be biologically plausible parameters, generate sustained oscillations. Leloup and Goldbeter also incorporated components in this model enabling them to account for the entrainment of the oscillator by light and for known circadian pathologies such as delayed sleep-phase syndrome.

⁵ The term *model* is used in a wide variety of ways. Explanatory accounts, such as those presented in a mechanism diagram, are often referred to as (mechanistic or explanatory) *models*. Sets of equations that describe a mechanistic model and are used to generate a simulation are referred to as a *computational model*, which is often abbreviated as just *model*. It is usually clear from context what is meant by the term. In this section, *model* refers to a computational model.

Goldbeter deployed these computational models to understand how an already proposed mechanism would behave. They provide what Adele Abrahamsen and I have referred to as *dynamic mechanistic explanations* (Bechtel & Abrahamsen, 2010). By using that term, we sought to emphasize the complementary roles of computational modeling in ascertaining the dynamics of a system and mechanistic analysis of its composition. Although the roles are complementary, computational models contribute to explanations in a very different manner than does the identification of the parts and operations of the mechanism. Mechanistic explanations have often been presented as an alternative to the type of deductive-nomological (DN) explanations that figured in earlier philosophy of science. According to the DN model, explanation involved deriving a description of a phenomenon from statements of one or more laws and initial conditions. Basic mechanistic explanations do not invoke laws or perform derivations. Computational simulations are more closely aligned with DN explanations than mechanistic explanations (Krakauer, Collins, Erwin, Flack, Fontana, Laubichler, Prohaska, West, & Stadler, 2011). First, while the equations used in computational simulations are typically not what one would characterize as laws, but rather mathematically characterized regularities, they are a basis for deriving results. Second, equations, like laws, are general. They specify relations between values of variables without specifying the parts whose properties are varying. Different parts with varying properties may be characterized by the same generalization. Third, invoking Craver's distinction between *how-possibly* and *how-actually* accounts, computational models provide *how-possibly* accounts—whether they characterize a putative or an actual mechanism, what they show is that such a mechanism *could* generate the phenomenon.

The explanatory import of computational modeling in biology extends beyond contexts in which it is used to understand the behavior of a hypothesized mechanism. I will briefly highlight four additional roles. Sometimes they are the object of experiments that are designed to better understand how the modeled mechanisms functions. This involves intervening on the computational model by, for example, investigating the effects of other parameter values or of removing or adding components to the model. As experimental researchers discovered more components of the mammalian circadian clock, modelers tried strategies such as fixing the values of variables for some of the components in their models to see if that affected the ability of the model to generate sustained oscillations. It is much easier to intervene on a computational model than on the biological tissue. The results of such experiments on computational models only provide information about what would happen in a real biological preparation to the extent that the model correctly describes the actual biological system. Yet, by drawing attention to what is possible under the current hypothesis, they can play an important role in the interactive engagement of modeling and experimentation.

A second additional use of computational models is to identify and characterize design principles (Green, Levy, & Bechtel, 2014). Increasingly biologists are approaching biological systems with the mindset of engineers. When engineers design systems, they often put together modules that are themselves composed of components, but organized according to principles from which their behavior in various contexts can be determined. Biologists can make use of such principles to understand how mechanisms they encounter operate. Uri Alon (2007) pioneered the investigation of motifs: small networks of two, three, or four nodes that are organized in a particular way (e.g., two units negatively feeding back on each

other). Through computational simulations (typically models using either Boolean or differential equations), he and others have determined how motifs will behave in any system in which parameter values fall within a specified range. When researchers identify an instance of such a motif in a biological system, they can immediately infer its behavior. Design principles are not tied to any particular mechanism but abstract from them; nonetheless, they can be applied in understanding the behavior of actual mechanisms in which components are organized in the manner specified.

A third use of computational models is closely related to the search for design principles. While one can construct detailed computational models that adhere closely to the details of a particular mechanism, one can also relax those constraints to develop models that generalize across a broader range of phenomena. This involves abstracting (Levy & Bechtel, 2013) or coarse graining (Krakauer, Collins, Erwin, Flack, Fontana, Laubichler, Prohaska, West, & Stadler, 2011) by, for example relaxing constraints of the range of variables or considering parameters in an extended range. When successful, this approach can reveal general principles. Biology is often contrasted with physics insofar as there do not seem to be a small set of basic principles (laws) that can be applied universally. But that does not mean one cannot achieve varying degrees of generality and computational models provide one vehicle for doing so.

A final use of computational models is to provide an understanding of global states of systems and how systems might evolve. A useful way to represent the behavior of a complex system with many components each changing their state over time is in terms of a state space in which each dimension corresponds to a variable characterizing the system. The current state of the system will correspond to a point in the space and change in the system will correspond to a trajectory through the possible states of the system. By studying trajectories in state space, investigators can identify the structure in the state space—for example, discovering that it contains attractor states to which the system will evolve from a variety of other states (the basin of the attractor). This structure can be productively represented as a landscape in which the points at which the system will stabilize located at the bottom of valleys. Although one can develop a state-space representation of an empirically-studied system, it is much easier to run multiple simulations with a computational model of a dynamical system. Using the simple computational model of the TFFL that Goldbeter advanced in 1995, he was able to map a landscape that contained a limit cycle attractor—a closed loop of states in state space corresponding to the oscillation of *per* mRNA and PER concentrations. The cycle is called a limit cycle since from a variety of starting points not on the cycle, the system will evolve towards the cycle. In more complex dynamical systems, there are multiple attractors in the landscape. Moreover, one can represent alterations to the system as changes in the identity and location of attractors. By constructing computational models of hypothetical complex systems, and studying the resulting landscapes and how they can change, one can acquire intuitive ideas about what is happening in natural systems (e.g., that the perturbation that leads to cancerous growth creates a new attractor) (Huang & Kauffman, 2012).

Computation modeling invokes a very different explanatory strategy than the mental animation of mechanism diagrams that figured in basic mechanistic explanation. It emphasizes the abstract, possible system, not the concrete actual mechanism. When

directed at particular mechanisms, it can provide information about how the mechanism will behave that cannot be generated from the mechanistic account itself. Through its extended uses it can facilitate the discovery of generalized principles and enable researchers to address additional questions about possible systems that lie beyond the scope of basic mechanistic accounts.

4. Explaining the Integration of Mechanisms through Graph-Theoretic Analyses

In this section I address a second limitation that often confronts basic mechanistic explanations—that the same strategies used for initially identifying parts of a mechanism end up identifying a large-holistic system implicated in multiple phenomena, not just the one under investigation. Mechanistic research often assumes that biological mechanisms are independent entities that, when provided the right inputs, depend only on their inner workings to generate the phenomenon they are invoked to explain. That is, it assumes there is a natural boundary to a mechanism and that the parts one identifies reside within it. In some cases, research on a mechanism begins with a delineated structure (e.g., a cell organelle or a brain region) and a characterization of what it does. But in many cases, as in the investigations of circadian rhythms discussed above, research begins with a part of the mechanism (e.g., a gene), and the mechanism is identified by determining which other parts (1) interact with that part in generating the phenomenon and (2) change the phenomenon when they are altered. When researchers turn to recomposing the mechanism, they seek to identify the place of these entities within the mechanism.

The problem with the strategy just outlined is that it is extremely sensitive to the techniques available at a time to identify components that have an effect on the phenomenon being explained and to measure those effects. Traditional strategies such as inhibiting or stimulating parts were limited to investigating a small number of potential parts. Often this would yield on the order of ten parts, as seen in Figure 2 in the case of circadian rhythms. With the development of new techniques in the 2000s, however, researchers have identified many more genes that have effects on circadian oscillations. For example, using small interfering RNAs to knock-down 17,631 known and 4,837 predicted human genes in U2OS (human osteosarcoma) cells containing a luciferase reporter attached to the known clock gene *Bmal1*, John Hogenesch, Steve Kay, and their collaborators identified nearly 1000 genes that resulted in low amplitude circadian oscillations (Zhang, Liu, Hirota, Miraglia, Welch, Pongsawakul, Liu, Atwood, Huss, Janes, Su, Hogenesch, & Kay, 2009). Due to challenges in analyzing period in these cases, they were not further analyzed. They focused instead on 343 genes that clearly increased the amplitude or altered the period of circadian rhythms (they only counted the gene if it produced deviations more than three standard deviations from the mean).

The researchers selected 17 genes on which to perform a dose-dependent knockdown and in 16 cases established dose-dependent effects comparable to those that previous research had found with genes already regarded as clock genes. In addition, the researchers analyzed protein interactions and determined that some of the proteins synthesized from these genes interacted directly with known clock genes whereas others are further removed. Many of them are part of pathways such as those for insulin and hedgehog signaling, cell cycle, and folate metabolism. For example, down regulating several

components of the insulin pathway (*JNK*, *IKK*, *MTOR*, *APKC*, and *PYK*) results in longer period oscillations while down regulating another, *PFK*, results in shorter period oscillations. Since these are also pathways that previously had been shown to be regulated by the circadian clock, Zhang et al. conclude: “Collectively, we conclude the clock is massively interconnected and functionally intertwined with many biological pathways” (p. 207).

These results present a new challenge in conceptualizing the clock mechanism. It would not be productive to simply add all of these genes and proteins to what is regarded as the clock mechanism. As noted, many are part of other mechanisms; such a move would quickly lead to treating the whole cell or organism as the mechanism for all phenomena. The great success of mechanistic research has stemmed from its ability to decompose systems into relevant mechanisms and their parts and to show how these contribute to the phenomena under investigation. If biology is to continue to build upon this success, researchers need ways to draw boundaries around mechanisms in order to generate recognizably if more complex mechanistic explanations. How can they do so?

As in many other fields, biologists are increasingly invoking tools to analyze networks to understand biological mechanisms (Barabasi & Oltvai, 2004; Mitra, Carvunis, Ramesh, & Ideker, 2013; Prokop & Csukás, 2013). Most fundamentally, network approaches provide new tools for representing biological organization. They also provide new tools to reason about organization and its consequences for the behavior of mechanisms. To illustrate the potential of this approach, I start with the network diagram (reproduced in Figure 3) that Zhang et al. used to present their results. Shown in light and dark blue are the proteins that are normally construed as constituting the circadian clock (*ARNTL* and *ARNTL2* are alternate names for *BMAL1* and *BMAL2*; *NR1D2* and *NR1D2* are alternate names for *Rev-Erb α* and *Rev-Erb β*). In purple, red, and green they show proteins that when knocked down increased the amplitude or altered the period of the clock. In pink are proteins that link those proteins that affect the clock when knocked down and the core components of the clock.

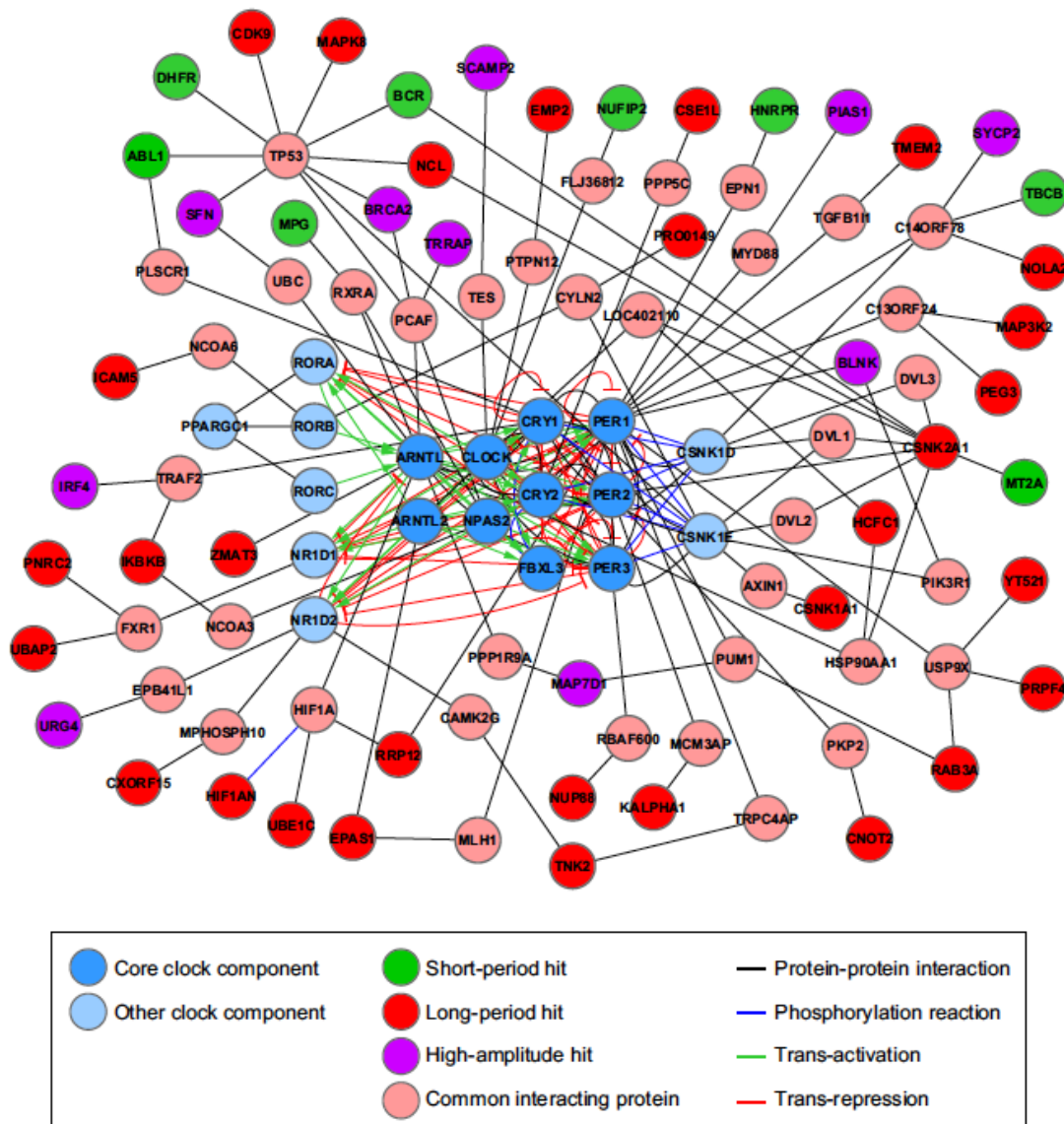


Figure 3. Zhang et al.'s (2009) representation of the various proteins that when knocked-down have effects on circadian rhythms. See text and legend for details.

The first important role of network representations in biology is to provide new perspectives on organization both within and between mechanisms. These perspectives draw from graph theory, which provides a number of measures for analyzing network organization. In the vocabulary of graph theory, networks consist of nodes (the circles in the above diagram) and edges (the lines connecting the circles). One graph-theoretic measure, cluster analysis, identifies as modules nodes that are highly interconnected (clustered). As a result of these connections, modules represent candidate mechanisms. Sometimes the mechanisms identified in this manner correspond roughly to those identified by the classical procedure of starting with the phenomenon and finding parts that affect it. The highly connected nodes in the center, colored in light or dark blue, correspond to the traditionally construed clock mechanism. Network analysis, however, is

most useful when it offers accounts that differ with those advanced directly from mechanistic research. One such role is to identify additional components beyond those differentiated by classical mechanistic research (see, for example, Ravasz, Somera, Mongru, Oltvai, & Barabasi, 2002; Kelley & Ideker, 2005). Other times modules may point to the existence of functional organization within and between mechanisms that may not have been discovered by traditional approaches (Costanzo, Baryshnikova, Bellay, Kim, Spear, Sevier, Ding, Koh, Toufighi, Mostafavi, Prinz, St Onge, VanderSluis, Makhnevych, Vizeacoumar, Alizadeh, Bahr, Brost, Chen, Cokol, Deshpande, Li, Lin, Liang, Marback, Paw, San Luis, Shuteriqi, Tong, van Dyk, Wallace, Whitney, Weirauch, Zhong, Zhu, Houry, Brudno, Ragibizadeh, Papp, Pal, Roth, Giaever, Nislow, Troyanskaya, Bussey, Bader, Gingras, Morris, Kim, Kaiser, Myers, Andrews, & Boone, 2010).

Another important measure used in graph analysis that provides insights to the organization within and between mechanisms is degree distribution—the distribution of the number of edges from a given node. Early graph theorists assumed that degree would be distributed normally, but in many networks it is not normally distributed. Rather, a few nodes have unusually high degree. These nodes are referred to as hubs and they may serve either to integrate a module/mechanism (TP53 in the upper left in Figure 3) or to facilitate integration between modules/mechanisms (CNSK2A1 on the right). Some components of the core clock, such as Per1, have both extensive connections to other components of the clock and to units elsewhere, suggesting an important role in integrating clock components with components of other mechanisms that are both regulated by and regulate clock function.

Second, networks are not only the product of numerous experimental inquiries, but they can also serve as a guide both to further experimentation and modeling. A network representation reveals many unsuspected indirect connections between nodes and so can guide inferences about how the effects of perturbing one node will spread to others. In many cases, modeling of the network will provide a guide to what sorts of effects one should expect. Researchers often annotate network representations with information from Gene Ontology about where in the cell genes are expressed, the functions the proteins perform, and the larger biological processes in which they figure. Particularly valuable from the point of view of understanding mechanisms is that Gene Ontology represents parts and functions hierarchically, which then supports predictions as to the effects of perturbing individual genes (Yu, Kramer, Dutkowski, Srivas, Licon, Kreisberg, Ng, Krogan, Sharan, & Ideker, 2016). Given that Gene Ontology involves curated information about the entities and their functions that have been identified in experimental research, it is not surprising that the hierarchical units more or less correspond to traditional mechanisms. What is less to be expected that new data driven approaches to generating ontologies, such as NeXO (Dutkowski, Kramer, Surma, Balakrishnan, Cherry, Krogan, & Ideker, 2013), also yield hierarchical structures corresponding to mechanisms. When combined with information from these ontologies, network representations enable researchers to make new inferences about the effects of perturbations (e.g., knocking out of two genes) not only within traditionally characterized mechanisms but also across mechanisms.

The challenge I raised at the beginning of this section was where to draw the boundaries around a mechanism once we recognize how the components of a mechanism are

interconnected with other entities in large networks. In the above paragraphs I noted that modules in graph representations often correspond to mechanisms initially identified in more traditional ways. But given that there are a number of entities outside the modules that affect activity within, researchers must often exercise discretion as to where to draw the boundaries. Recent research, for example, has identified several instances in which an operation within the circadian mechanism depends directly on a component usually treated as part of the metabolic mechanism. One involves the binding of the dimer of Clock and Bmal1 to the E-box on *per* and other genes, as shown in Figure 2. Clock has been identified as a histone acetyl transferase—by adding acetyl groups to the chromatin, it affects how tightly the chromatin is bound and hence whether genes such as *Per* can be transcribed. This function of Clock is regulated by SIRT1 binding to it. The concentration of SIRT1 is itself dependent on levels of NAD⁺, a central component in glycolysis and other metabolic pathways (Bellet, Orozco-Solis, Sahar, Eckel-Mahan, & Sassone-Corsi, 2011). If one focuses on the way in which NAD⁺ modulates SIRT1 activity, and thereby modulates circadian rhythms, then one might include it in the circadian mechanism. If the research question involves how transcription of circadian genes is modulated, a researcher might include not just SIRT1 but even the metabolic processes such as glycolysis that oxidize and reduce NAD⁺ in the relevant mechanism. Decisions as to where to locate the boundaries of mechanisms are constrained by both the interconnectivity of components in larger networks and the questions researcher are investigating.

The characterization and analysis of network organization constitutes a different strategy than traditionally employed in mechanistic research. As in the case of computational modeling, in network analysis, researchers abstract from the specific composition of the mechanism and focus on the ways edges connect nodes. The goal of network analysis is to identify organizational principles at an abstract or coarse-grained levels and determine the behaviors they make possible. The application of these results to specific actual networks that realize a pattern of organization involves a reasoning process much like derivation. The resulting abstract analysis plays a different explanatory role than the specification of the parts and operations. By identifying a mechanism with a module in a network, one can use the network representation to identify ways in which the module is affected by other modules. To investigate how such modulation is actually achieved, researchers need to engage further with mechanistic strategies that identify specific parts and operations. Network analysis performs a function complementary to traditional mechanistic inquiry, revealing multiple ways in which a mechanism is situated among other mechanisms and creating conditions in which researchers can select which entities to treat as a mechanism and investigate further.

5. Expanding mechanism's explanatory strategies

The strategies scientists pursue to explain the phenomena of interest to them change over time. Cytology and cell biology through the 19th and 20th centuries witnessed the cultivation of new techniques to identify mechanisms and decompose them into their parts and operations. This is evident in *General Cytology* and again with the introduction of electron microscopy and cell fractionation. These techniques enabled cell biologists to identify organelles and localize specific chemical reactions in each. Based on these results, researchers set about recomposing mechanisms, often using diagrams and mentally

rehearsing the operations depicted (Figure 1). They offered accounts of how cells perform a wide range of activities that fit the pattern of basic mechanistic explanations. Even as contemporary cell researchers move beyond these accounts, they remain success stories.

But as a result of developing new techniques for decomposing cells to identify genes and proteins, especially those that operate on a mass scale, cell biologists found that the parts and operations could no longer be recomposed into basic mechanistic explanations. The discovery of multiple operations occurring in parallel, feedback relations between operations regarded as later onto those regarded as earlier, and interactions with components regarded as parts of other mechanisms rendered the project of invoking mental simulation to animate static diagrams insufficient. One could not settle whether the mechanism would generate, for example, sustained circadian rhythms or dampened oscillations.

To address these questions, biologists availed themselves of new strategies that could complement those of basic mechanistic explanation. I have focused on two, computational simulation and network analysis, that are playing important roles in contemporary systems biological approaches to explaining cellular phenomena such as circadian rhythms. I have emphasized the differences between these strategies and mechanistic strategies of the past. Both computational simulations and network analyses abstract from the details of the composition of mechanisms and appeal to general principles. Researchers apply the results of such abstraction to understand the operation of specific mechanisms through processes like derivation. From a computational analysis, researchers can find out how a proposed mechanism will behave and explore other possibilities. From a network analysis, researchers can make inferences based on characteristic modes of organization and propose plausible boundaries for mechanisms while recognizing that they are situated in an environment that affects their operation.

Deploying new strategies to advance explanatory objectives is a recurring theme in biology. What is important to recognize is that the conception of explanation is also being extended. Recomposition achieved through computational modeling and network analysis differs from that achieved by mentally animating a mechanism. Nonetheless, contemporary cell biology is still mechanistic. In particular, biologists are still decomposing biological systems into parts and operations (with high-throughput procedures, this process is accelerated). And the resulting explanation still appeals to these parts and operations. And when possible, biologists still attempt to mentally rehearse the operations in mechanism diagrams. The new techniques complement existing mechanistic strategies. The resulting perspective is a pluralistic one in which different explanatory strategies each make a complementary contribution to the pursuit of mechanistic explanation in contemporary cell biology.

Acknowledgment

I thank Garland Allen, the audiences at the two Updating Cowdry Workshops at Woods Hole, and the editors of this volume, Jane Maienschein, Karl Matlin, and Manfred Laublicher, for their very helpful and constructive comments on earlier versions of this chapter.

References

- Abrahamsen, A., Sheredos, B., & Bechtel, W. (in press). Mechanism diagrams. In S. Glennan & P. Illari (Eds.), *Routledge handbook of mechanisms*. London: Routledge.
- Allen, G. E. (1979). *Life science in the twentieth century*. London ; New York: Cambridge University Press.
- Alon, U. (2007). *An introduction to systems biology: Design principles of biological circuits*. Boca Raton, FL: Chapman & Hall/CRC.
- Barabasi, A.-L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 5, 101-113.
- Bechtel, W. (2006). *Discovering cell mechanisms: The creation of modern cell biology*. Cambridge: Cambridge University Press.
- Bechtel, W. (2011). Mechanism and biological explanation. *Philosophy of Science*, 78, 533-557.
- Bechtel, W. (2015). Can mechanistic explanation be reconciled with scale-free constitution and dynamics? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*.
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 421-441.
- Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41, 321-333.
- Bechtel, W., & Richardson, R. C. (1993/2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Cambridge, MA: MIT Press. 1993 edition published by Princeton University Press.
- Bellet, M. M., Orozco-Solis, R., Sahar, S., Eckel-Mahan, K., & Sassone-Corsi, P. (2011). The time of metabolism: NAD⁺, SIRT1, and the circadian clock. *Cold Spring Harbor Symposia on Quantitative Biology*, 76, 31-38.
- Berzelius, J. J. (1836). Einige Ideen über bei der Bildung organischer Verbindungen in der lebenden Naturwirksame, aber bisher nicht bemerkte Kraft. *Jahres-Berkcht über die Fortschritte der Chemie*, 15, 237-245.
- Buchner, E. (1897). Alkoholische Gärung ohne Hefezellen (Vorläufige Mittheilung). *Berichte der deutschen chemischen Gesellschaft*, 30, 117-124.
- Coleman, W. (1971). *Biology in the nineteenth century : problems of form, function, and transformation*. New York: Wiley.
- Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L., Toufighi, K., Mostafavi, S., Prinz, J., St Onge, R. P., VanderSluis, B., Makhnevych, T., Vizeacoumar, F. J., Alizadeh, S., Bahr, S., Brost, R. L., Chen, Y., Cokol, M., Deshpande, R., Li, Z., Lin, Z. Y., Liang, W., Marback, M., Paw, J., San Luis, B. J., Shuteriqi, E., Tong, A. H., van Dyk, N., Wallace, I. M., Whitney, J. A., Weirauch, M. T., Zhong, G., Zhu, H., Houry, W. A., Brudno, M., Ragibizadeh, S., Papp, B., Pal, C., Roth, F. P., Giaever, G., Nislow, C., Troyanskaya, O. G., Bussey, H., Bader, G. D., Gingras, A. C., Morris, Q. D., Kim, P. M., Kaiser, C. A., Myers, C. L., Andrews, B. J., & Boone, C. (2010). The genetic landscape of a cell. *Science*, 327, 425-431.
- Cowdry, E. V. (1924). *General cytology*. Chicago: University of Chicago Press.
- Craver, C. F., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. Chicago: University of Chicago Press.

- Dutkowski, J., Kramer, M., Surma, M. A., Balakrishnan, R., Cherry, J. M., Krogan, N. J., & Ideker, T. (2013). A gene ontology inferred from molecular networks. *Nature Biotechnology*, *31*, 38-45.
- Ghosh, A. K., & Chance, B. (1964). Oscillations of glycolytic intermediates in yeast cells. *Biochemical and Biophysical Research Communications*, *16*, 174-181.
- Goldbeter, A. (1995). A model for circadian oscillations in the *Drosophila* Period protein (PER). *Proceedings of the Royal Society of London. B: Biological Sciences*, *261*, 319-324.
- Green, S., Levy, A., & Bechtel, W. (2014). Design sans adaptation. *European Journal for Philosophy of Science*, *5*, 15-29.
- Hardin, P. E., Hall, J. C., & Rosbash, M. (1990). Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature*, *343*, 536-540.
- Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1084-1102.
- Hess, B., Boiteux, A., & Krüger, J. (1969). Cooperation of glycolytic enzymes. *Advances in Enzyme Regulation*, *7*, 149-167.
- Huang, S., & Kauffman, S. A. (2012). Complex Gene Regulatory Networks – from Structure to Biological Observables: Cell Fate Determination. In R. A. Meyers (Ed.), *Computational Complexity* (pp. 527-560): Springer New York.
- Ideker, T., Galitski, T., & Hood, L. (2001). A new approach to decoding life: systems biology. *Annual Review of Genomics and Human Genetics*, *2*, 343-372.
- Kelley, R., & Ideker, T. (2005). Systematic interpretation of genetic interactions using protein networks. *Nature Biotechnology*, *23*, 561-566.
- Kitano, H. (2002). Systems biology: A brief overview. *Science*, *295*, 1662-1664.
- Konopka, R. J., & Benzer, S. (1971). Clock mutants of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences (USA)*, *89*, 2112-2116.
- Krakauer, D. C., Collins, J. P., Erwin, D., Flack, J. C., Fontana, W., Laubichler, M. D., Prohaska, S. J., West, G. B., & Stadler, P. F. (2011). The challenges and scope of theoretical biology. *Journal of Theoretical Biology*, *276*, 269-276.
- Leloup, J.-C., & Goldbeter, A. (2008). Modeling the circadian clock: From molecular mechanism to physiological disorders. *BioEssays*, *30*, 590-600.
- Levy, A., & Bechtel, W. (2013). Abstraction and the organization of mechanisms. *Philosophy of Science*, *80*, 241-261.
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*, 1-25.
- Mitra, K., Carvunis, A.-R., Ramesh, S. K., & Ideker, T. (2013). Integrative approaches for finding modular structure in biological networks. *Nature Reviews Genetics*, *14*, 719-732.
- Palade, G. E. (1987). Cell Fractionation. In J. E. Pauly (Ed.), *The American Association of Anatomists, 1888-1987. Essays on the History of Anatomy in America and a Report on the Membership--Past and Present*. Baltimore: Wilkin and Wilkins.
- Prokop, A., & Csukás, B. (2013). *Systems biology : integrative biology and simulation tools. Volume 1*. Dordrecht: Springer.
- Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., & Barabasi, A. L. (2002). Hierarchical organization of modularity in metabolic networks. *Science*, *297*, 1551-1555.

- Schwann, T. (1839). *Mikroskopische Untersuchungen über die Übereinstimmung in der Struktur und dem Wachsthum der Theire und Planzen*. Berlin: Sander.
- Tversky, B. (2011). Visualizing thought. *Topics in Cognitive Science*, 3, 499-535.
- Vitaterna, M. H., King, D. P., Chang, A.-M., Kornhauser, J. M., Lowrey, P. L., McDonald, J. D., Dove, W. F., Pinto, L. H., Turek, F. W., & Takahashi, J. S. (1994). Mutagenesis and mapping of a mouse gene, *Clock*, essential for circadian behavior. *Science*, 264, 719-725.
- Yu, Michael K., Kramer, M., Dutkowski, J., Srivas, R., Licon, K., Kreisberg, Jason F., Ng, Cherie T., Krogan, N., Sharan, R., & Ideker, T. (2016). Translation of Genotype to Phenotype by a Hierarchy of Cell Subsystems. *Cell Systems*, 2, 77-88.
- Zhang, E. E., Liu, A. C., Hirota, T., Miraglia, L. J., Welch, G., Pongsawakul, P. Y., Liu, X., Atwood, A., Huss, J. W., Janes, J., Su, A. I., Hogenesch, J. B., & Kay, S. A. (2009). A genome-wide RNAi screen for modifiers of the circadian clock in human cells. *Cell*, 139, 199-210.