Chapter 15
From Mechanisms to Mathematical Models and Back to Mechanisms: Quantitative Mechanistic Explanations

Tudor M. Baetu

Abstract Despite the philosophical clash between deductive-nomological and mechanistic accounts of explanation, in scientific practice, both approaches are required in order to achieve more complete explanations and guide the discovery process. I defend this thesis by discussing the case of mathematical models in systems biology. Not only such models complement the mechanistic explanations of molecular biology by accounting for poorly understood aspects of biological phenomena, they can also reveal unsuspected 'black boxes' in mechanistic explanations, thus prompting their revision while providing new insights about the causal-mechanistic structure of the world.

Keywords Scientific explanation • Quantitative-dynamic explanation • Mechanism • Mathematical model • Systems biology

1 Introduction

Inspired by the deductive-nomological tradition, some philosophers suggested that at least some mathematical models explain biological phenomena by applying the laws and theories of physics and chemistry to biological systems (Smart 1963; Weber 2005, 2008). Knowledge of the peculiarities of a given biological system supplies a list of parameters corresponding to the various parts of the system, their interactions and organization, as well as the initial and boundary conditions to which the system is subjected under experimental or physiological conditions. The laws and theories associated with the model provide a set of rules describing the relationships between parameters and how these relationships change with time. These rules, suitably expressed in mathematical language, play a key role in explanation by making possible the derivation of descriptions of phenomena by

T.M. Baetu (✉)
Programa de Filosofia, Universidade do Vale do Rio dos Sinos, Av. Unisinos, 950 Bairro Cristo Rei, 93022-000 São Leopoldo, RS, Brasil
e-mail: tudormb@unisinos.br

means of analytic and, more recently, computational methods. For example, using the Hodgkin and Huxley (1952) model as a case study, Marcel Weber (2005) argues that the mechanistic descriptions of molecular and cellular entities, activities, and organizational features (e.g., cell membranes, ions) specify how a physicochemical theory should be applied, while the explanatory burden falls on the regularities that describe the behavior of the system (in this case, a physicochemical law known as Nernst equation), and on how a description of the phenomenon of interest can be derived from these regularities.

Drawing on the highly influential view that biological phenomena are explained by showing how they are produced by mechanisms[1] (Bechtel 2006; Craver 2007; Darden 2006; Wimsatt 1972), Carl Craver (2006, 367) defends a very different view of the explanatory role of mathematical models: "[m]odels are explanatory when they describe mechanisms." In response to Weber, Craver points out that the Hodgkin and Huxley model also includes hypothetical assumptions and uninterpreted parameters whose purpose is to render the model empirically adequate. Empirical adequacy could have been obtained by appealing to a different set of assumptions and parameters (e.g., a different number of ion currents), while a different physical interpretation of certain parameters would have yielded a different mechanistic explanation of the phenomenon under investigation. Craver concludes that the Hodgkin and Huxley model is an incomplete 'how-possibly' account providing some preliminary insights about the possible mechanisms responsible for generating and propagating action potentials along axons, but should not be confused with the explanation of the phenomenon. In order to be explanatory, a model should provide a complete description of the mechanism actually responsible for a phenomenon. This description must "include all of the relevant features of the mechanism, its component entities and activities, their properties, and their organization" (Craver 2006, 367),[2] and "exhibit productive continuity without gaps from the set up to termination conditions" (Machamer et al. 2000, 3). In turn, a complete inventory of the explanatorily relevant mechanistic components and features, along with a specification of their causal-role and productive continuity provide an intuitive understanding of how phenomena are produced.[3]

---

[1]Machamer, Darden, and Craver define mechanisms as "entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (2000, 3). Alternatively, a mechanism is "a complex system that produces that behavior by the interaction of a number of parts, where the interactions among parts can be characterized by direct, invariant, change relating generalization" (Glennan 2002), or "a structure performing a function in virtue of its component parts, component operations, and their organization [. . .] responsible for one or more phenomena" (Bechtel and Abrahamsen 2005). McKay and Williamson (2011) propose a more generally applicable characterization, according to which a "mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon."

[2]Explanatory relevance is equated to causal relevance and demonstrated by means of experimental interventions (Baetu 2012a; Craver 2007; Woodward 2003).

[3]"Intelligibility [. . .] is provided by descriptions of mechanisms, that is, through the elaboration of constituent entities and activities that, by an extension of sensory experience with ways of working, provide an understanding of how some phenomenon is produced" (Machamer et al. 2000, 22).

Both Weber and Craver agree that mathematical models can be explanatory, but they attribute explanatory value to very different features of these models. For Weber, the explanation lies specifically in the derivation of a description of phenomena from mathematically formulated law-like regularities, while for Craver a mathematical model is explanatory only to the extent it identifies the physical structures actually responsible for causing the phenomena.[4] Ultimately, propositions can be derived from other propositions whether or not these derivations reflect the causal structure of the world, while mechanistic structures can be identified experimentally without relying on any kind of conceptual derivation. This divergence about what counts as an explanation often translates into a direct contradiction: the same model may be deemed explanatory under a deductive-nomological approach, but not under a mechanistic one, as demonstrated in the case of the Hodgkin and Huxley model; conversely, qualitative descriptions of biological mechanisms count as explanations under a mechanistic approach, but have no explanatory value under a deductive-nomological approach.

I argue that despite this philosophical clash between deductive-nomological and mechanistic accounts, scientific practice can rely on an explanatory pluralism in which the two approaches are not only complementary, as recently argued by William Bechtel and Adele Abrahamsen (2010, 2011), but also dynamically integrated in a process of reciprocal validation. The view defended in this paper is that mathematical models play an explanatory role by attempting to provide an answer to the question "Can the proposed mechanism generate the phenomenon of interest in all its minute quantitative/dynamic details?" Using examples from the recent scientific literature, I show how the answer to this research question reveals both agreements and disagreements between mechanistic explanations and mathematical models of mechanisms involving derivations of descriptions of phenomena. Agreements are used to infer or reinforce the completeness of mechanistic explanations; conversely, disagreements between models and mechanisms prompt revisions of either or both models and mechanisms. I argue that, at least in some cases, attempts to reach an agreement between models and mechanisms generate progressive research programs, in the sense that these cycles of modifications and revisions reveal novel ways of thinking about the ontology of mechanisms, as well as surprising explanations to seemingly unrelated scientific puzzles. For additional discussions on the relationship between mechanistic and mathematical models see chapters in the present volume by William Bechtel (2015, Chap. 9), Pierre-Alain Braillard (2015, Chap. 14), Ingo Brigandt (2015, Chap. 7), Tobias Breidenmoser and Olaf Wolkenhauer (2015, Chap. 11), Fridolin Gross (2015, Chap. 8), Tarik Issad and Christophe Malaterre (2015, Chap. 12), and Frédérique Théry (2015, Chap. 6).

---

[4]Both views admit gradations in explanatory value. Models incorporating fundamental laws provide deeper explanations than models relying on more superficial regularities describing the behavior of a certain type of systems. Likewise, under a mechanistic approach, a model incorporating a more complete description of a mechanism is better than a model relying on a sketchier description. For an account of the completeness of mechanistic explanations, see (Baetu 2015)

The paper is organized as follows: I begin with a brief introduction of mathematical models of molecular networks in systems biology (Sect. 2), followed by a discussion of how mathematical models combine the application of laws, modeling and analysis strategies from chemistry, cybernetics, and systems theory (Sect. 3) with knowledge of mechanisms (Sect. 4). In Sect. 5, I show how mechanistic explanations and mathematical models complement each other. I then proceed in Sect. 6 to show that mechanisms and models don't always coexist in a state of static complementarily, but can also contradict each other. Based on an analysis of recent examples from science, I argue that mathematical models can provide criteria for assessing the completeness of mechanistic explanations and I show how the disagreements between models and mechanisms prompt important revisions of the current understanding of the molecular-mechanistic basis of biological activity. Finally, in Sect. 7, I summarize my findings and arguments.

2 A Brief Introduction to Mathematical Models of Molecular Networks

Molecular biology is one of the most important scientific achievements of the twentieth century. In conjunction with biochemistry and cell biology, it succeeded in explaining biological phenomena in mechanistic terms. Nevertheless, the approach pioneered by molecular biology has one important shortcoming: it yields primarily a qualitative understanding of mechanisms acting in isolation. The challenge for the twenty-first century biology is to integrate current knowledge of mechanisms into a conceptual framework that is "holistic, quantitative and predictive" (Kritikou et al. 2006, 801). This challenge was answered by an emerging field of investigation, dubbed 'systems biology.'

One strategy by means of which systems biologists hope to achieve their goals relies on mathematical modeling of molecular networks. Networks are abstract representations of physical systems consisting of parts connected by a web of relationships. Molecular networks represent molecular mechanisms broadly construed, including signal transduction pathways, metabolic pathways, as wells as more comprehensive systems composed of several mechanisms. The nodes of the network represent mechanistic entities (proteins, genes), while the connections between nodes represent mechanistic activities (chemical reactions rates, activation/inhibition). Other information about the modeled mechanisms (e.g., structural details of proteins, their tridimensional configuration, chemical mechanisms of reaction) is usually ignored, hence the abstract representational nature of molecular networks.

Mathematical modeling quantifies qualitative descriptions of networks, with most models falling in two distinct categories: discrete and continuous. Discrete quantification works on the assumption that the behavior of the network is determined by thresholds (e.g., a gene can be either expressed or not expressed). The same network can also be modeled in such a way that its nodes can take continuous values (e.g., the concentration of gene products can take any values within given intervals). Once the nodes are quantified, one can mathematically represent all the

possible states of the network; that is, all the possible combinations of the values of the nodes. The next step is to determine how the state of the network changes over time because of its internal wiring and external inputs, something which is achieved by introducing transition rules for each node. In discrete networks, the rules by means of which nodes act on each other are represented by logical functions. In continuous networks, systems of differential equations represent the rates of change of the value of any given node in terms of the values of other nodes and external inputs.[5]

It is interesting to note that while Craver allows for mathematical models of mechanisms, we are never told what, if anything, quantification and mathematical formalism add to the explanatory value, the completeness, or the intelligibility of the model. We are now in a better position to answer this question. Since mathematical models incorporate very precise assumptions about how a system changes from one state to another, we can conclude that they allow for a more detailed description of productive continuity. The fact that this description deals with quantitative changes further suggests that mathematical models may be necessary in order to achieve precise quantitative manipulations of mechanistic components (as opposed to semi-quantitative ones, such as knockout or over-expression) and to design synthetic mechanisms that behave in quantitatively precise ways [e.g., the repressilator (Baetu 2015; Elowitz and Leibler 2000; Morange 2009)]. Finally, mathematical models reveal that networks can exhibit a number of unsuspected and perhaps physiologically significant properties such as attractors and steady states,[6] robustness and sensitivity,[7] adaptability,[8] and hysteresis[9] (for the evolutionary significance of these novel properties, see Brigandt, Chap. 7).

---

[5]For a more technical description, consult Shmulevich and Aitchison (2009). Additional assumptions are required in order to construct a model of a network. One has to choose between a synchronic and an asynchronous updating scheme, between a binary, multi-value, or stochastic logic, between different kinetic laws, between ordinary and partial differential equations, etc. Without these assumptions, it is impossible to model the dynamic behavior of the network.

[6]These states amount to long-term behaviors of networks and can be experimentally measured, thus allowing for precise quantitative predictions, as well as an assessment of the empirical adequacy of the model.

[7]Robustness is insensitivity to the precise values of biochemical parameters (changes in reaction rates, concentrations of substrates), thus allowing a system to function in a wide range of conditions and resist certain perturbations. Sensitivity denotes the contrary, namely a situation where a mechanism is operational only if the values of its parameters are fine-tuned to specific values. Robustness and sensitivity allow for optimization analysis, which is especially useful for identifying which mechanistic components should be targeted in order to achieve a desired result with maximal efficiency and minimal side effects.

[8]Ability to adapt to 'background noise': the smallest change in stimulus intensity that can be sensed (ΔS) increases with the background stimulus intensity (S), such that $\Delta S/S$ remains constant (Weber-Fechner law).

[9]A network may display more than one 'stable state', and it is possible that a change in the system's state caused by a transient stimulus (e.g., external input, temporary change in gene expression) is not followed by a return to the initial state when the stimulus is withdrawn. It has been hypothesized that such states may underlie developmentally differentiated cell types (Kauffman 2004) or physiological cell states [e.g., proliferating vs. apoptotic cells; (Huang 1999)].

3 Mathematical Models as Applications of Chemistry, Cybernetics, and Systems Theory

Early system theorists aimed to formulate a set of general principles governing the behavior of systems in all fields of scientific investigation. Most famously, von Bertalanffy (1976, 32) argued that "there exist models, principles, and laws that apply to generalized systems or their subclasses, irrespective of their particular kind, the nature of their component elements, and the relationships or 'forces' between them." This view is sometimes echoed by contemporary systems biologists. For example, in their discussion of a previously proposed continuous model of the bacterial chemotaxis signal transduction pathway revealing robustness, Baker et al. (2006, 190) claim that "[o]ne of the primary objectives of systems biology is to formulate biological laws that are akin to the laws of physics."[10] In turn, a generally applicable understanding of the relationship between a system's structure and its properties can yield 'design explanations' showing "why a given structure or design is necessary or highly preferable in order to perform a function or to have an important property like robustness" (Braillard 2010, 55); for a discussion of generalizable mechanistic explanations based on graph-theoretic considerations, see Bechtel, Chap. 9.

Under this approach, mathematical models in systems biology can be treated as applications of the modeling, analytic and computational methods of analysis pioneered by systems theory. This clearly amounts to a form of explanatory heteronomy of systems biology on systems theory,[11] although it is important to keep in mind that explanatory tools are indirectly borrowed from many fields of investigation, systems theory being more an interdisciplinary rendezvous point than a thoroughly unified science. For example, discrete modeling has its roots in cybernetics, computer science and engineering, while continuous models incorporate laws and mathematical models from chemical kinetics, some of which are themselves applications of statistical mechanics and thermodynamics to the field of chemistry (see also Braillard, Chap. 14).

It is also interesting to note that any given network may be mathematically modeled in more than one way, and, depending on the modeling strategy and associated assumptions, the same network may or may not be shown to possess

---

[10]The 'laws' to which they refer have a variable degree of generality, ranging from common features of networks displaying specific properties, such as robustness or adaptability, to more general properties of generic networks. Examples of the latter are found in Kauffman's (1993) seminal work on Random Boolean Networks (networks in which the connections between nodes are wired randomly). By investigating the behavior of such networks, some general principles emerged; for instance, networks become chaotic as the number of connections per node increases.

[11]Weber (2005) argues for an 'explanatory heteronomy' of biology on physics and chemistry, and spells out the sense in which the former reduces to the latter: biology relies on the laws of physics and chemistry in order to generate explanations, and, in this respect, can be viewed as applied physics and chemistry.

certain properties. In order to claim that the model explains certain aspects of biological phenomena, one needs to know to what extent a given set of modeling assumptions is true or approximately true of a given biological system; this determines the degree to with which the biological system is expected to behave as described by its corresponding model. As it turns out, mathematical models tend to incorporate several unverified, and potentially idealizing assumptions. Idealizations are problematic in many fields of investigation, yet the problem is more acute in systems biology. For example, in physics, it is usually clear what exactly is being idealized and to what extent; think of Newton's idealization of Earth as a homogenous sphere or as a material point. By contrast, in systems biology, very little is known about real time concentrations of substrates and real time kinetics of reaction, to the point that it is seldom clear to what extent a model idealizes these features, what are the potential drawbacks of these idealizations, and what can be done to correct the situation.[12]

This incertitude creates a paradoxical situation: mathematical models have the potential to explain, in Weber's deductive-nomological sense, yet it is often not clear to what extent they actually explain. Another way to frame this situation is to think of the explanatory value of a mathematical model as being contingent on additional evidence demonstrating that the model is a suitable surrogate of the physical system being modeled, such that knowledge generated by studying the model can be safely extrapolated to the target physical system and its ability to generate the phenomenon of interest (Baetu 2014).

Finally, mathematical modeling involves a certain degree of abstraction. Since a lot is known about the molecular basis of biological activity, the starting point of current modeling efforts is, more often than not, prior knowledge of molecular mechanisms. This prior knowledge determines the wiring of molecular networks, as well as the physical and causal interpretation of nodes and connections. However, not all that is known about mechanisms is incorporated in molecular networks. As a general rule, specific modes of action, the structure of mechanistic entities and other high-resolution biochemical details are not represented in molecular networks, which tend to amount to simplified schemas consisting of unstructured entities (proteins, genes) and generic activities (activation/inhibition, increase/decrease in concentration). Thus, simplified representations of mechanisms, known as 'mechanism schemas',[13] act as bridges mediating the transfer of knowledge from molecular

---

[12]For instance, current discrete modeling strategies assume that a network is either updated synchronously (the values of all its nodes are updated at the same time) or asynchronously (no two nodes are updated at the same time). Set aside the difficulty of finding out which of the two assumptions holds true of the particular system under investigation – a situation that makes it such that investigators simply test several models until they find one that simulates well characterized features of the system –, it is also possible that no real biological system will perfectly fall into one or the other of these two categories.

[13]"A mechanism schema is a truncated abstract description of a mechanism that can be easily instantiated by filling it with more specific descriptions of component entities and activities" (Darden 2006, 111–12).

biology to other fields of investigation, such as bioinformatics and systems biology (Baetu 2011b, 2012b).

On the one hand, abstraction generates networks simple enough to allow for computable solutions, while highlighting those features of biological systems most amenable to mathematical modeling (Baetu 2011a, 2012b). For example, a Gene Regulatory Network (GRN)[14] representation of signal transduction pathways does not tell us how a specific repressor protein regulates gene expression; the repressor may compete with a transcriptional activator for the same DNA binding site, it may bind the activator and cause a conformational change affecting the activity of the latter, or it may trap the activator in the cytoplasm. Removing such details makes it possible to reduce the total number of parameters in the model, thus increasing computability. At the same time, the network highlights specific features of the physical system it represents, in this case, the inputs and outputs of genome expression, showing how each gene module in the GRN behaves like a logic gate integrating several regulatory inputs in order to yield a single output, namely the presence or absence of its corresponding gene product.

On the other hand, however, models built from mechanism schemas cannot account for all the features of the phenomena produced by these mechanisms. Because specific modes of action and structural details of entities are abstracted, a mathematical model may tell us which mechanistic components should be targeted in order to achieve a desired result, but doesn't tell us how to operate changes in a physical system.[15] For this reason, it seems highly unlikely that mathematical models can replace current mechanistic explanations.

In summary, we can conclude that mathematical models in biology are not exclusively applications of what are traditionally deemed to be more fundamental sciences, such as physics and chemistry, but also of many other fields, such as systems theory and computer engineering. Second, as applications of other sciences and fields of investigation, mathematical models tend to yield putative explanations because it is usually not clear to what extent a given set of modeling assumptions apply to biological systems and to what extent they introduce idealizations. Finally, since details accounting for certain aspects of the phenomena under investigation are abstracted from the model, it seems unlikely that mathematical models can or are meant to replace mechanistic explanations indigenous to the biological sciences.

---

[14]GRNs are "hardwired genomic regulatory codes, the role of which is to specify the sets of genes that must be expressed in specific spatial and temporal patterns. [. . .] these control systems consist of many thousands of modular DNA sequences. Each such module receives and integrates multiple inputs, in the form of regulatory proteins (activators and repressors) that recognize specific sequences within them. The end result is the precise transcriptional control of the associated genes" (Davidson and Levine 2005, 4935).

[15]For example, in order to physically inhibit the activity of a repressor, detailed knowledge of its structure, such as a mapping of the amino acids responsible for DNA binding, is required; the repressor activity is tempered with by mutating specifically these DNA-binding amino acids.

4   The Mechanistic Content of Mathematical Models of Molecular Networks

Craver (2007) correctly points out that the mathematical model developed by Hodgkin and Huxley had only a partial causal interpretation, and that the model played primarily a role in discovery by guiding the subsequent elucidation of the mechanism responsible for generating and propagating action potentials along axons. Most notably, the model suggested the existence of transmembrane ion channels and provided some insights about their possible properties. In light of this kind of historical examples, it might be tempting to conclude that mathematical models are primarily instrumental constructs eventually superseded by mechanistic explanations. However, this is not always the case. The starting point of many modeling efforts, especially in today's practice, is prior knowledge of molecular mechanisms. This knowledge provides a straightforward physical interpretation of the nodes of the network (they correspond to genes and proteins) and the inter-nodal connections (they are activities of entities affecting other entities in the system/mechanism). Most modeling assumptions also refer to features of physical systems (e.g., reaction rates), and therefore have a clear physical interpretation; it is just that it cannot be easily determined to what extent the modeled systems actually possess these features, and how realistically these featured are modeled. Thus, many mathematical models incorporate substantial knowledge of the causal-mechanistic structure of the world and are not plagued by incomplete or problematic causal interpretations.

5   How Quantitative Models Complement Qualitative Descriptions of Molecular Mechanisms

The fact that many models are constructed in light of prior knowledge of molecular mechanisms raises an interesting question: Why would anyone bother designing and testing mathematical models when mechanistic explanations are already available? The simplest solution to the puzzle would be the claim that, despite their mechanistic content and unambiguous causal interpretation, mathematical models in systems biology are, after all, primarily meant to generate predictions. This view is certainly defendable: since it is often the case that realistic models outperform idealized models in terms of empirical adequacy, the fact that a mathematical model incorporates substantial knowledge about the causal-mechanistic structure of the world is not necessarily an indication that the model has explanatory ambitions. Nevertheless, a survey of the scientific literature suggests otherwise. In many cases, mathematical models are used not only to generate predictions, but also to account for certain anomalies and poorly understood aspects of phenomena known to be produced by mechanisms that have been already elucidated (Bechtel and Abrahamsen 2010, 2011). Building on the earlier suggestion that mathematical models complement

current mechanistic explanations, I propose that the answer to the puzzle lies in a refutation of the assumption that a model is complete when it includes all the relevant features of the mechanism, their causal role, and productive continuity.

In order to understand how and why models of mechanisms fail to explain certain features of phenomena, let us consider the following example. Leukocytes exposed to antigens, inflammatory agents, and pathogens express a variety of genes required for mounting an immune response. After a brief period of activation, gene expression shuts down. The phenomenon to be explained amounts to a black-box correlation between input (cells are exposed to pathogens) and output (spike of gene expression) conditions (Fig. 15.1, top panel). A mechanistic explanation tells us what happens inside the black box: the spike of gene expression following stimulation is explained by a negative feedback regulatory mechanism whereby a transcriptional factor (nuclear factor κB, or NF-κB) is initially activated, then subsequently inactivated by an inhibitory protein (inhibitor of κB, or IκB) coded by a gene under its transcriptional control (Fig. 15.1, middle panel).[16]

The above explanation amounts to a qualitatively complete description of the regulatory mechanism responsible for the peak of gene expression. The causal contribution of the mechanism to the target phenomenon is well established, the mechanistic function of its various components is well understood, and there is a large body of evidence supporting the conclusion that there are no gaps in the causal chain linking input and output conditions. In sum, it is a typical description of a well understood mechanism one may very well find among the diagrams of a recent immunology textbook. Furthermore, the mechanistic explanation is also highly satisfactory relative to the pragmatic goals of molecular biology, such as treating illnesses and developing techniques for intervention and manipulation (Craver 2007, 38); for example, it shows how to achieve a loss of NF-κB activity (e.g., mutations in the DNA binding or nuclear localization domains).

Nevertheless, in some respects, qualitatively complete descriptions remain unsatisfactory (see also Brigandt, Chap. 7; Breidenmoser and Wolkenhauer, Chap. 17; Issad and Malaterre, Chap. 18). For instance, the above explanation supports the inference that, if the inhibitory protein IκB is synthesized too fast, there is no spike of activation, and conversely, if the inhibitor is synthesized too slowly or in insufficient quantity, gene expression is never turned off. However, this does not tell us whether the mechanism is sensitive or robust relative to the exact amounts of the inhibitor. If sensitive, the mechanism will malfunction in response to mutations that affect the amount and stability of the inhibitor; in contrast, if robust, the mechanism can adapt and continue to operate irrespective of changes in the concentration of the inhibitor. Likewise, the model doesn't tell us why there is an increase in the

---

[16]In resting cells, NF-κB is held in the cytoplasm by IκB (Huxford et al. 1998). When cells are stimulated (Fig. 15.1, middle panel, A), a chain of protein-protein interactions leads to the degradation of IκB (B); NF-κB is freed (C), translocates to the nucleus (D) where it binds κB sequences in the promoter regions of target genes drastically enhancing their transcription (Pahl 1999). NF-κB also binds the promoter of the IκB gene (E), and the newly synthesized IκB binds NF-κB, trapping it back in the cytoplasm (Sun et al. 1993).
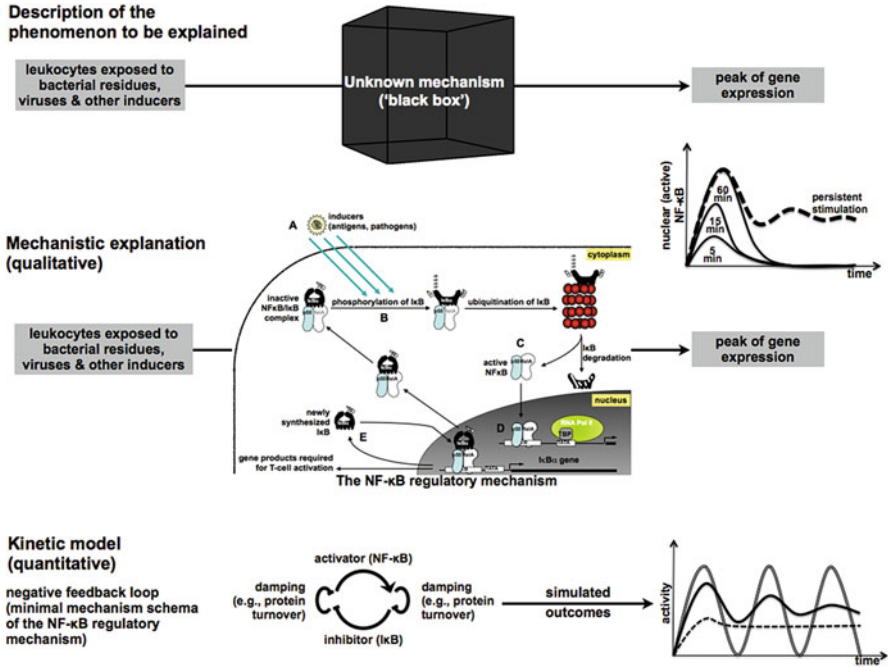
Fig. 15.1 The NF-κB regulatory mechanism (Adapted from Baetu and Hiscott 2002; Hoffmann et al. 2002)

intensity (but not the duration) of gene expression when cells are exposed to longer pulses of stimulation (Fig. 15.1, top right graph, solid curves), and only an increase in the duration (but not the intensity) of gene expression in the case of persistent stimulation (dashed curve). The failure to account for these minute features of the phenomenon under investigation stems primarily from a poor understanding of how the various parts of the mechanism change quantitatively over time.

In more general terms, the experimental data supporting the qualitative description of a mechanism demonstrates the necessary causal contribution of the proposed mechanism to the phenomenon of interest. This does not prove, however, that the mechanism can and does generate the phenomenon, exactly as it is measured in all its quantitative-dynamic details, thus leaving open the question whether the mechanism is also sufficient to generate the phenomenon. This constitutes a serious problem for a mechanistic account of explanation: if the mechanism doesn't generate the phenomenon, exactly as it is measured, then the mechanistic explanation fails. The immediate corollary is that quantitative parameters (e.g., concentrations of components), as well as their temporal dynamics (e.g., rates of reactions) must be included in an ideally complete mechanistic explanation, along with a complete list of mechanistic components, their organization and causal roles, and evidence for productive continuity. However, the inclusion of this additional

information in a computationally useful way – that is, in a format that allows the derivation of the quantitative-dynamic features of the target phenomenon –, requires mathematical modeling, which has its drawbacks, namely the introduction of unverified or difficult to verify assumptions, potential idealizations, and the abstraction of some biochemical and molecular details. Thus, it is reasonable to conclude that qualitatively complete descriptions of mechanisms and their mathematical model counterparts stand in a relationship of explanatory complementarity rather than direct competition.

By the same token, this means that a quantitative model, such as the Hodgkin and Huxley model, cannot be replaced by a qualitatively more detailed description of a molecular mechanism. The Hodgkin and Huxley model, or rather a version of it benefiting from the retrospective hindsight of a more detailed causal interpretation, continues to play a very important explanatory role in contemporary neuroscience: it shows that the molecular mechanism qualitatively described by currently accepted mechanistic explanations can generate action potentials that closely fit quantitative experimental measurements. In other words, 'how-possibly' mathematical models are not only useful in guiding the subsequent elucidation of mechanisms or generating predictions, but also provide putative explanations for precise quantitative-dynamic aspects of biological phenomena.

Recent trends in biology support this interpretation. More and more studies published in leading journals complement qualitative descriptions of mechanisms supported by the experimental practice of molecular biology with quantitative models aiming to demonstrate that the proposed mechanisms can generate the quantitative-dynamic aspects of the phenomena of interest. For example, recent studies investigating the development of tubular organs rely on mathematical models in order to show that the proposed mechanisms can produce the phenomena under investigation in the right amount/intensity [e.g., precise allometric growth ratio of bronchioles in the lung (Tang et al. 2011)]. Furthermore, in some cases mathematical models are required in order to show that no additional mechanisms or mechanistic components are likely to be needed in order to generate the target phenomenon. For example, Taniguchi et al. (2011) propose that probabilistic biases in the distribution of adhesion proteins suffice to generate the right amount of twisting in the developing gut. Given the stochastic nature of the proposed mechanism, it is not possible to rely on commonsense mechanistic intuitions to establish that the mechanism can generate (even approximately) the target phenomenon (Horne-Badovinac and Munro 2011). The authors of the study reasoned that since the predictions of the mathematical model match the experimental data, a more complex model including additional parameters is not needed; since the model does not include physically uninterpreted parameters, it can be established that no corresponding mechanistic components are missing; hence a more complex mechanism, including additional entities and activities, or additional mechanisms are not likely to be needed to produce the target phenomenon (i.e., the proposed mechanism is necessary and sufficient to produce the phenomenon).

Due to uncertainties and potential idealizations associated with modeling assumptions, the above studies offer only putative explanations. The results they

yield can be safely extrapolated to biological mechanisms and their ability to generate phenomena on condition that these models accurately describe their target biological mechanisms. Nonetheless, these putative explanations cover ground beyond the reach of qualitative descriptions of mechanisms. In order to determine whether a proposed mechanism can generate a phenomenon exactly as it is measured, down to minute quantitative-dynamic details, we usually cannot rely on qualitative descriptions of mechanisms and commonsense intuitions about how mechanisms work. Numerical computations are required, and this is precisely what mathematical models make possible.

6   The Integration of Mathematical Modeling and Mechanistic Explanations

In the previous section, I argued that mathematical models, viewed as applications of chemistry, cybernetics, and systems theory to biological phenomena, and mechanistic explanations, viewed as descriptions of causal-mechanistic structures, complement each other. As a rule of thumb, qualitative descriptions embody experimental evidence of the causal contribution of mechanisms (entities, activities, and their organization) to the target phenomenon, while mathematical models provide further 'proof or principle' demonstrations that the mechanisms in question can generate phenomena down to minute quantitative-dynamic details. The former rely on actual experimental control over the mechanism and its components, while the latter rely on the ability to mathematically/computationally derive close approximations of quantitative descriptions of phenomena.

I will now go a step further and argue that mathematical models and qualitative mechanistic descriptions don't always stand in a state of passive complementarity whereby each explains aspects of a phenomenon inaccessible to the other. Rather, there is a constant interaction between the two. Since molecular networks are abstract representations of molecular mechanisms, revisions of the latter may entail revisions of the former; in such situations, mathematical models of molecular networks are revised as well. Conversely, mathematical models can reveal anomalies and unsuspected explanatory holes in previously accepted mechanistic models, thus prompting their revision. This interaction is progressive,[17] in the sense that it addresses not only the immediate research aims – in this case, accounting for quantitative-dynamic aspects of phenomena produced by molecular mechanisms –, but provides unexpected explanations to seemingly unrelated and thus far unexplained phenomena.

Mathematical models can support current mechanistic explanations by showing that the proposed mechanisms can generate phenomena down to minute

---

[17] By analogy with Lakatos' (1978, 33) notion of 'progressive research programme' in which "each new theory [. . .] predicts some novel, hitherto unexpected fact."

quantitative-dynamic details. They can also reveal discrepancies and anomalies. Let us consider again the spike of gene expression in leukocytes. According to the mechanistic explanation depicted in the middle panel of Fig. 15.1, the spike is generated by a negative regulatory feedback loop mechanism (NF-κB activates IκB, and IκB inhibits NF-κB). However, a mathematical model of a generic negative feedback loop mechanism [e.g., (Goodwin 1963)] has three possible outputs, namely perpetual oscillation, damped oscillation, or a plateau of activation (Fig. 15.1, bottom right graph); none matches the observed spikes of gene expression (Fig. 15.1, top right graph). Hoffmann et al. (2002) interpreted this mismatch between the simulated results of the quantitative model and the experimental data as an indication that the currently accepted mechanistic explanation is inadequate.

It is not question of doubting the causal relevance of the NF-κB regulatory mechanism, or that this mechanism consists of a negative feedback loop. Experimental evidence clearly shows that the mechanism, as described in Fig. 15.1, is necessary for the generation of normal immune responses in vitro (in cell models) and in vivo (in animal models). In this particular case, it is also hard to doubt the results of the mathematical model or that the model employed is a suitable model of a negative feedback loop. Models of negative feedback loop systems have been extensively validated in physics, engineering and cybernetics. Thus, there is a direct explanatory conflict here. The explanandum is always the same, namely the characteristic spike of gene expression associated with an immune response, yet the mechanistic explanation shows that the NF-κB feedback mechanism is a necessary cause of this spike, while the mathematical model of the mechanism fails to accurately predict the quantitative-dynamic description of the spike, thus pointing to the opposite conclusion, that the phenomenon is not explained.

In order to solve the puzzle, researchers had to resort to a seemingly irrelevant piece of information. It was known for some time that there are three isoforms of IκB present in mammalian cells, and that only the gene coding for the α isoform is under the transcriptional control of NF-κB. The other two isoforms, β and ε, are expressed constitutively. There are, therefore, two overlapping versions of the NF-κB regulatory mechanism, one with a negative feedback loop, and one without. The β and ε isoforms were largely ignored because the results of knockout experiments suggested that they may play a different physiological role,[18] and, most importantly, the version of the mechanism without a feedback loop doesn't explain in any way how a spike of activation could be produced (one can only expect a plateau, like the one depicted at the bottom of Fig. 15.2).

Hoffmann et al. (2002) observed that cells that express only the α isoform generate a damped oscillation pattern of gene expression, closely matching the one predicted by the mathematical model. This suggested that when the β and ε isoforms

---

[18]In mice, IκBα$^{-/-}$ is associated with exacerbated inflammation and embryonic lethality, while IκBβ/ε$^{-/-}$ females have a shorter fertility span. Nevertheless, other experiments suggest that the three forms are partially redundant. For a review of the original scientific literature, consult Hoffmann (2002, 1241–42).
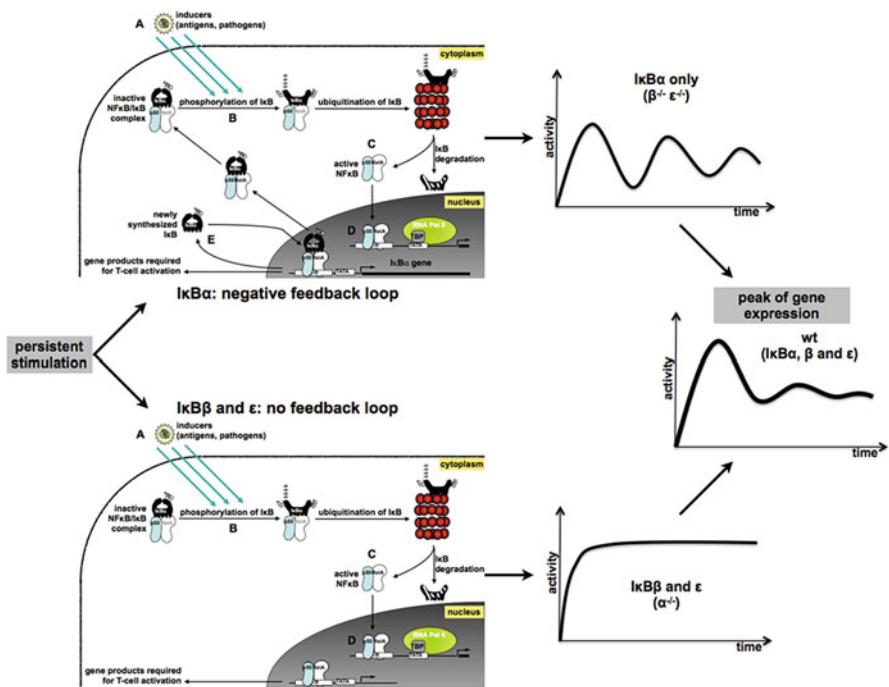
Fig. 15.2 The revised model of the NF-κB regulatory mechanism (Adapted from Hoffmann et al. 2002)

are also present, the damped oscillation pattern is somehow transformed into the characteristic wild-type spike of gene expression (Fig. 15.2). In order to test the combined contribution of two overlapping mechanisms, it is not possible to rely on a qualitative understanding of mechanisms. Precise numerical computation is necessary, and this requires mathematical modeling. Explorations with mathematical models showed that the observed spike, as well as a differential temporal dynamics of gene expression following pulse and persistent stimulation, can be derived from a mathematical model of a molecular network combining the two versions of the NF-κB regulatory mechanism, one involving a negative regulatory loop and the other a constitutive expression of the inhibitor IκB.

From a philosophical point of view, this is an extremely interesting result. It shows that mathematical models can prompt revisions of mechanistic explanations; in this case, the initial negative feedback loop mechanism was augmented to include a parallel pathway of activation not subjected to negative feedback. I take this two-way interaction whereby mathematical models are revised as mechanisms are elucidated in more detail, while mechanistic explanations are revised as a result of mathematical modeling to be a strong indication that a mixed, mechanistic and deductive-nomological approach is necessary in order to achieve more complete explanations, as well as to guide the discovery process in biology.

More important, the back and forth interaction between mechanistic and deductive-nomological approaches is progressive, in the sense that it generates new insights about the ontological status of mechanisms while offering unexpected explanations to seemingly unrelated and thus far unexplained phenomena. In the above example, the overlapping mechanisms explanation challenges our metaphysical intuitions about the nature of mechanisms. While we might be tempted to treat molecular mechanisms as neatly individuated objects, such as clocks and other man-made mechanical devices, they are in fact populations consisting of a large number of identical mechanisms collectively generating a biological phenomenon. As a general rule, each 'individual mechanism' in the population has a physiologically insignificant contribution to the target phenomenon; is ephemeral (Glennan 2010), not only because molecular components have a relatively short life-time, but also because, after having fulfilled its causal contribution, a component is randomly replaced with another copy of the same kind; and doesn't necessarily operate in perfect synchrony with other 'individual mechanisms.' The study by Hoffmann et al. further suggests that some mechanistic components come in several variants, with each variant acting along a partially distinct causal pathway, such that molecular mechanisms may in fact amount to heterogeneous populations of partially overlapping mechanisms.[19]

In turn, this renewed appreciation of the internal variability of molecular mechanisms suggests an unexpected explanation for a seemingly unrelated phenomenon. Many genes and gene products come in several copies displaying very small differences in terms of sequence, structure, and biochemical function. The physiological role of these seemingly redundant molecular components is poorly understood. The overlapping mechanisms explanation suggests that they are not mere spare parts or evolutionary accidents, but may in fact play a very important physiological role: generate partially overlapping versions of the same mechanism, collectively required in order to generate highly complex patterns of gene expression and associated phenotypes, adaptive responses to external stimuli, and other biological phenomena. Heterogeneous populations of partially overlapping mechanisms may

---

[19]Furthermore, there are cases when significantly distinct mechanisms responsible for distinct phenomena nevertheless share mechanistic components. Thus, in addition to a modular mode of organization whereby systems of mechanisms are organized serially or in parallel, the output of a mechanism serving as input for one or more other mechanisms (Bechtel 2006; Craver 2007; Darden 2006), significantly distinct mechanisms may also be firmly interlocked in the same manner as the partially overlapping mechanisms described above. The lessons learned form the NF-κB regulatory mechanism raise the possibility that non-modular sharing of mechanistic components plays a physiologically relevant role in adjusting quantitative-dynamic aspects of the phenomena produced by these mechanisms. If this turns out to be the case, then molecular mechanisms are unlike any man-made mechanisms, first because they are heterogeneous populations rather than individual objects, and second because they operate both in a modular and a non-modular fashion. In other words, there is a sense in which a cell or organism cannot be decomposed into a set of mechanism-modules, but is one integrated mechanism consisting of heterogeneous populations of mechanisms overlapping to various degrees.

also account for fine grained phenotypic variability, thus bridging the gap between variability required for evolution, and the seemingly monolithic explanations in molecular biology.

7   Conclusion

The emergence of systems biology is marked by a renewed interest in mathematical modeling. From a philosophical standpoint, this 'mathematical turn' in biology constitutes an excellent opportunity to investigate the relationship between deductive-nomological and mechanistic accounts of scientific explanation. I argue that mathematical models in systems biology combine substantial knowledge of molecular mechanisms with the application of laws, modeling and analysis strategies borrowed from chemistry, cybernetics and systems theory. Mechanism schemas obtained by abstracting high-resolution biochemical details act as bridges between molecular mechanistic explanations and mathematical models of networks. In turn, mathematical models account for poorly understood aspects of biological phenomena, most notably minute quantitative-dynamic features. Thus, in scientific practice, deductive-nomological and mechanistic approaches to explanation are not mutually exclusive, but complementary. Furthermore, mathematical models can reveal unsuspected 'black boxes' and motivate revisions of mechanistic explanations. This interplay between mechanistic explanations and their mathematical counterparts constitutes a progressive research approach that generates explanations of novel phenomena, and reveal strange properties of molecular mechanisms that have thus far escaped our attention.

References

Baetu, T. M. (2011a). A defense of syntax-based gene concepts in postgenomics: 'Genes as modular subroutines in the master genomic program'. *Philosophy of Science, 78*(5), 712–723.

Baetu, T. M. (2011b). Mechanism schemas and the relationship between biological theories. In P. McKay, J. Williamson, & F. Russo (Eds.), *Causality in the sciences*. Oxford: Oxford University Press.

Baetu, T. M. (2012a). Filling in the mechanistic details: Two-variable experiments as tests for constitutive relevance. *European Journal for Philosophy of Science, 2*(3), 337–353.

Baetu, T. M. (2012b). Genomic programs as mechanism schemas: A non-reductionist interpretation. *British Journal for the Philosophy of Science, 63*(3), 649–671.

Baetu, T. M. (2014). Models and the mosaic of scientific knowledge. The case of immunology. *Studies in History and Philosophy of Biological and Biomedical Sciences, 45*, 49–56.

Baetu, T. M. (2015). When is a mechanistic explanation satisfactory? Reductionism and antireductionism in the context of mechanistic explanations. In G. Sandu, I. Parvu, & I. Toader (Eds.), *Romanian studies in the history and philosophy of science*. Dordrecht: Springer.

Baetu, T. M., & Hiscott, J. (2002). On the TRAIL to apoptosis. *Cytokine & Growth Factors Reviews, 13*, 199–207.

Baker, M., Wolanin, P., & Stock, J. (2006). Systems biology of bacterial chemotaxis. *Current Opinion in Microbiology, 9*, 187–192.

Bechtel, W. (2006). *Discovering cell mechanisms: The creation of modern cell biology*. Cambridge: Cambridge University Press.

Bechtel, W. (2015). Generalizing mechanistic explanations using graph-theoretic representations. In P.-A. Braillard & C. Malaterre (Eds.), *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences* (pp. 199–225). Dordrecht: Springer.

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., & Abrahamsen, A. (2011). Complex biological mechanisms: Cyclic, oscillatory, and autonomous. In C. A. Hooker (Ed.), *Philosophy of complex systems* (pp. 257–285). New York: Elsevier.

Braillard, P.-A. (2010). Systems biology and the mechanistic framework. *History and Philosophy of Life Sciences, 32*, 43–62.

Braillard, P.-A. (2015). Prospect and limits of explaining biological systems in engineering terms. In P.-A. Braillard & C. Malaterre (Eds.), *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences* (pp. 319–344). Dordrecht: Springer.

Breidenmoser, T., & Wolkenhauer, O. (2015). Explanation and organizing principles in systems biology. In P. A. Braillard & C. Malaterre (Eds.), *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences* (pp. 249–264). Dordrecht: Springer.

Brigandt, I. (2015). Evolutionary developmental biology and the limits of philosophical accounts of mechanistic explanation. In P.-A. Braillard & C. Malaterre (Eds.), *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences* (pp. 135–173). Dordrecht: Springer.

Craver, C. (2006). When mechanistic models explain. *Synthese, 153*, 355–376.

Craver, C. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.

Darden, L. (2006). *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution*. Cambridge: Cambridge University Press.

Davidson, E., & Levine, M. (2005). Gene regulatory networks. *Proceedings of the National Academy of Science, 102*(14), 4935.

Elowitz, M., & Leibler, S. (2000). Synthetic gene oscillatory network of transcriptional regulators. *Nature, 403*, 335–338.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science, 69*, S342–S353.

Glennan, S. (2010). Ephemeral mechanisms and historical explanation. *Erkenntnis, 72*, 251–266.

Goodwin, B. (1963). *Temporal organization in cells: A dynamic theory of cellular control processes*. London: Academic.

Gross, F. (2015). The relevance of irrelevance: Explanation in systems biology. In P.-A. Braillard & C. Malaterre (Eds.), *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences* (pp. 175–198). Dordrecht: Springer.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology, 117*, 500–544.

Hoffmann, A., Levchenko, A., Scott, M., & Baltimore, D. (2002). The I κB–NF-κB signaling module: Temporal control and selective gene activation. *Science, 298*, 1241–1245.

Horne-Badovinac, S., & Munro, E. (2011). Tubular transformations. *Science, 333*, 294–295.

Huang, S. (1999). Gene expression profiling, genetic networks, and cellular states: An integrating concept for tumorigenesis and drug discovery. *Journal of Molecular Medicine, 77*(6), 469–480.

Huxford, T., Huang, D.-B., Malek, S., & Ghosh, G. (1998). The crystal structure of the I κB/NF-κB complex reveals mechanisms of NF- κB inactivation. *Cell, 95*, 759–770.

Issad, T., & Malaterre, C. (2015). Are dynamic mechanistic explanations still mechanistic? In P.-A. Braillard & C. Malaterre (Eds.), *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences* (pp. 265–292). Dordrecht: Springer.

Kauffman, S. (1993). *The origins of order: Self-organization and selection in evolution*. New York: Oxford University Press.

Kauffman, S. (2004). A proposal for using the ensemble approach to understand genetic regulatory networks. *Journal of Theoretical Biology, 230*(4), 581–590.

Kritikou, E., Pulverer, B., & Heinrichs, A. (2006). All systems go! *Nature Reviews Molecular Cell Biology, 7*, 801.

Lakatos, I. (1978). *Philosophical papers Vol I: The methodology of scientific research programmes*. Cambridge: Cambridge University Press.

Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.

McKay, P., & Williamson, J. (2011). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science, 2*, 119–135.

Morange, M. (2009). Synthetic biology: A bridge between functional and evolutionary biology. *Biological Theory, 4*(4), 368–377.

Pahl, H. L. (1999). Activators and target genes of Rel/NF-κB transcription factors. *Oncogene, 18*, 6853–6866.

Shmulevich, I., & Aitchison, J. (2009). Deterministic and stochastic models of genetic regulatory networks. *Methods in Enzymology, 467*, 335–356.

Smart, J. J. C. (1963). *Philosophy and scientific realism*. New York: Humanities Press.

Sun, S.-C., Ganchi, P. A., Ballard, D. W., & Greene, W. C. (1993). NF-κB controls expression of inhibitor I κB α: Evidence for an inducible autoregulatory pathway. *Science, 259*, 1912–1915.

Tang, N., Marshall, W., McMahon, M., Metzger, R., & Martin, G. (2011). Control of mitotic spindle angle by the RAS-regulated ERK1/2 pathway determines lung tube shape. *Science, 333*, 342–345.

Taniguchi, K., Maeda, R., Ando, T., Okumura, T., Nakazawa, N., Hatori, R., Nakamura, M., Hozumi, S., Fujiwara, H., & Matsuno, K. (2011). Chirality in planar cell shape contributes to left-right asymmetric epithelial morphogenesis. *Science, 333*, 339–341.

Théry, F. (2015). Explaining in contemporary molecular biology: Beyond mechanisms. In P.-A. Braillard & C. Malaterre (Eds.), *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences* (pp. 113–133). Dordrecht: Springer.

von Bertalanffy, K. (1976). *General system theory: Foundations, development, applications*. New York: George Braziller.

Weber, M. (2005). *Philosophy of experimental biology*. Cambridge: Cambridge University Press.

Weber, M. (2008). Causes without mechanisms: Experimental regularities, physical laws, and neuroscientific explanation. *Philosophy of Science, 75*(5), 995–1007.

Wimsatt, W. C. (1972). Complexity and organization. In K. F. Schaffner & R. S. Cohen (Eds.), *PSA 1972, Proceedings of the philosophy of science association* (pp. 67–86). Dordrecht: Reidel.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.