**Biological robustness: design, organization and mechanisms**

*Maria Serban and Sara Green*

**Abstract**. Recent engineering-based modelling efforts in systems biology suggest that some forms of biological robustness depend on relational features of the organization of living systems. To explain such features, abstract representations of patterns of organization or *design principles* within the framework of control theory have been proposed. We characterize the results of such model-based practices as *structural-causal explanations* and situate our account in the broader philosophical debate about explanation in biology.

*Keywords*: robustness, systems biology, design, structural explanation, biological explanation, mechanism

1. **Introduction**

Systems biology makes extensive use of formal tools from engineering sciences and applied mathematics to explain dynamic features of living systems (Stelling et al. 2004; Wagner 2005; Alon 2007). The success of these epistemic practices has sparked philosophical debates concerning the ability of the mechanistic framework to account for the abstract features of the resulting models and explanations (Brigandt, Green & O’Malley 2018). Many arguments have centred on the adequacy of the mechanistic model of explanation to accommodate strategies drawing on network science and control theory to account for certain types of biological phenomena (e.g., Huneman 2010; Woodward 2013; Skillings 2015; Craver 2016; Chirimuuta 2017; Halina 2018). Focusing on how systems biologists model the phenomena of biological robustness, we aim to clarify the explanatory role of design principles identified via control theoretic analysis.[[1]](#footnote-0)

To illustrate our account, we shall refer to the modelling efforts involved in the investigation of the robustness of bacterial chemotaxis. This case study has been discussed by several philosophers of science without reaching a consensus on whether it entails a causal-mechanistic or a non-causal model of explanation (e.g., Braillard 2010; Woodward 2013; Matthiessen 2017; Green and Jones 2016). This paper aims to clarify this debate through a more detailed analysis of the mathematical modelling involved in what we describe as an engineering approach to bacterial chemotaxis. Our account offers an alternative to the dichotomy between interpretations of design principles as schemas for more detailed mechanistic explanations (Matthiessen 2017) and non-causal explanations citing *necessary* conditions for the target phenomena (Braillard 2010).

The notion of *design principles* discussed in this paper does not refer to an evolutionary “design” process. Instead, it relies on a “thin” notion of design, referring to “generalizable patterns of organization which play a role-functional part in present-day biological systems.” (Green, Levy and Bechtel 2015, 16). Modelling efforts in systems biology often describe the dynamics of target systems in terms of abstract patterns of organization that the system implements or realizes, such as negative feedback control, network motifs, or bi-stable switching. An important feature of design principles is that they are represented abstractly with the help of formal mathematical tools.

The application of design principles in the explanatory projects of systems biology could be broadly analysed in terms of Pincock’s recent account of *abstract explanation* in science (Pincock 2015). However, our analysis of a control theoretic model of biological robustness will show that the abstract character of design principles is not sufficient to capture the distinctive character of the explanation associated with this type of model. While the abstractness of design principles allows for generalizations of biological mechanisms and for the detection of organizational patterns across different biological systems (Levy and Bechtel 2013), we claim that their explanatory value should be analyzed as a separate epistemic contribution. We argue that engineering models of biological robustness draw their explanatory force from a formal mathematical model which can be causally interpreted. The mathematical model offers a derivational warrant for the formal representation of a certain type of design principle instantiated in the target biological system. Such design principles have an explanatory force in virtue of the fact that they can be interpreted as representing the *causal structures* upon which the robustness of certain biological properties depends. In short, we show how design principles can be discovered via control theoretic engineering analyses and used in *structural-causal explanations* of biological robustness.

The strategy is this. Section 2 explains why robustness is a joint topic of interest to both engineering and biology and introduces the case of bacterial chemotaxis. Section 3 summarizes the modelling efforts developed to understand protein network robustness in the case of bacterial chemotaxis. Here we draw a contrast between a model that interprets robustness as a result of a fine-tuning process and two related models showing that the perfect adaptation to constant stimuli by the *E. coli* chemotaxis network depends on certain structural features of this network. Section 4 argues that the design principle introduced by the engineering model, viz. the principle of integral feedback control, figures in a *structural-causal* explanation of the robustness of the bacterium’s adaptive chemotactic behaviour. Section 5 compares these engineering-based explanations to non-causal design explanations (Braillard 2010) and mechanistic explanations offered in systems biology. In section 6 we summarize our conclusions and clarify how structural-causal explanations differ from canonical mechanistic explanations which represent both causal and constitutive factors responsible for some target phenomenon.

1. **Modelling robustness – from engineering to biology**

Robustness is defined as the maintenance of functional stability in the face of external or internal perturbations (Kitano 2004).[[2]](#footnote-1) Stability, homeostasis and canalization are forms of biological robustness that have been intensively studied, but they do not exhaust the wide range of robust biological capacities and properties (e.g., Slepchenko and Terasaki 2004; Barkai and Shilo 2007). Alongside studies which focus on different types of biological robustness, there is currently a growing interest in identifying similarities and differences in the ways in which biological and engineered systems achieve robust properties. Since robustness is an important capacity in both engineered and biological systems, an interesting question is whether robustness in both kinds of systems relies on similar design principles (Stelling et al. 2004).

Here we focus on how modelling and conceptual tools from engineering sciences impact the explanation and understanding of cases of biological robustness. Our case study revolves around the theoretical and experimental efforts involved in understanding chemotactic behaviour in the *E.coli*. We start with a brief presentation of the phenomenon of bacterial chemotaxis, narrowing our focus to the feature of perfect/exact adaptation. We then contrast a fine-tuned model of the cell’s adaptive behaviour to two related, but distinct, models – what we term a *dynamical* and an *engineering model* of the robust perfect adaptation of *E. coli’s* chemotactic movement. By clarifying the relations between these three models, we aim to pin down the role of design principles in understanding of the robustness of a specific type of cellular behaviour.

##### *2.1. Bacterial chemotaxis and robust perfect adaptation*

Chemotaxis is the process that enables motile bacteria to approach beneficial chemical environments and escape hostile ones. *E. coli* swims via movement of external flagella that alternate between two kinds of motions called runs and tumbles (see Figure 1). The bacterium executes a ‘run’ by rotating its flagellar motor counterclockwise (CCW). This aligns all of its flagella into a bundle, resulting in a straight line movement for cca. 1 sec. For a ‘tumble’, the bacterium rotates its flagellar motor clockwise (CW). During a tumble the bundle breaks and the asynchronized flagella produce stationary changes of direction (cca. 0.1 sec). Thus, *E. coli* is randomly reoriented after each tumble.



*Figure 1: Motile behaviour of E. coli (adapted from Eisenbach 2001)*

The motion of the bacterium in a uniform external environment resembles a random walk since it does not have any ability to control or select its direction of motion. Its straight runs are subject to Brownian motion. In the presence of a chemical attractant, the bacterium moves in the direction of the attractant. Its motile behaviour involves less frequent tumbles, which results in longer runs and so gradual motion towards the attractant (the opposite behaviour is observed for repellents such as metal ions or leucine). Due to their small size most bacteria are unable to sense chemical gradients along the length of their body. Instead, their motion is determined by a *temporal gradient* (dX/dt) capturing temporal changes in the chemical concentration of nutrients or toxins in their environment.

*E.coli* display a remarkable sensory adaptation which involves a return to pre-stimulus levels despite continuous presence of attractants or repellents. In the presence of a chemical attractant mixed uniformly into the environment at a constant concentration, *E.coli*’s chemotactic behaviour is *perfectly adaptive*. This means that after a period of decreased tumbling frequency (called adaptation time, τ), *E.coli*’s tumbling frequency (*f*) increases toward and returns to the *exact* same frequency (*f0*) that it had prior to detecting the attractant in the cell’s environment. The *functional stability* of this characteristic adaptive behaviour is the *robustness* property that has been of high interest to biologists and engineers alike and which we shall analyse further.

##### *2.2. The chemotactic network*

As outlined in 2.1, the ability of *E.coli* to migrate in a particular direction depends on the control of the direction of rotation of its flagella. This control is in turn the result of a series of intracellular signals transmitted through a network of interacting proteins and receptors that sense concentration changes. The biomolecular network for *E.coli* chemotaxis is currently one of the best understood signal transduction systems. It comprises transmembrane receptors, the adaptor protein CheW, a two-component signaling system (proteins CheA and CheY), whose output controls the motor’s tumbling frequency (i.e., the probability that the cell will tumble). Moreover, it involves a protein controlling the output of the signaling system (CheZ) and two proteins, CheB and CheR, that modify CheA’s activity via the methylation state of the receptor complex.

The *perfect* adaptation response in *E.coli* is the result of interactions between two biochemical pathways which are schematically presented in Figure 2. When an environmental attractant attaches to a receptor, the receptor lowers the activity of the CheW-CheA complex. Less activity from this complex reduces the rate of CheY phosphorylation, which results in less CheY-P diffusing to the flagella. CheY-P induces clockwise rotation of the flagellar motor, so decreased levels of CheY-P result in less frequent tumbling (or longer straight runs). After the new attractant has been detected by receptors, the lower activity of the CheW-CheA complex also induces less CheB activity which means reducing the rate at which methyl groups are removed from the CheW-CheA complex. This, together with the continuous methylation of the CheR receptor leads to the increased methylation of the CheW-CheA complex. However, more methylation of this complex implies more overall activity which in turn leads to more phosphorylation of CheY that will diffuse to the flagellar motor, increasing the clockwise motor rotation and raising the tumbling frequency.

 

*Figure 2: Chemotactic network of E. coli (adapted from Registry of Standard Biological Parts)*

Not only is the chemotactic behaviour of *E.coli* perfectly adaptive, but the same molecular mechanism (the phosphorylation pathway whose activity is under the control of the methylation pathway) leads to the chemotactic network returning to the same steady state tumbling frequency value (*f0*) irrespective of the concentration value of the protein CheR. That is, the perfect adaptation of the chemotactic behaviour is *robust* across a wide range of value of CheR concentrations. This capacity is commonly referred to as *robust perfect adaptation* (RPA), and we shall in the following discuss efforts to model this capacity mathematically.

##### **Modelling the robustness of perfect adaptation**

The three models we discuss next capture the efforts of elucidating and explicating *why* the perfect adaptation of *E.coli* chemotaxis is robust to CheR protein concentrations (*R*) in the presence of a well-distributed chemical attractant. The first model sketched below aims to determine whether the perfect adaptation of bacterial chemotaxis is sensitive to the concentrations of the specific proteins that constitute *E. coli’*s chemotactic network. As we will briefly show, this *fine-tuned* model entails the wrong predictions concerning the behaviour of the bacterium. So while it might initially seem the most obvious choice for an explanatory model of the target behaviour, the fine-tuned model gets things wrong (cf. Alon 2007). The dynamical and the engineering models presented next show that the perfect adaptation of chemotactic behaviour is a robust phenomenon, i.e., it remains stable under large variations of chemotactic proteins. After introducing the main elements of the two modelling strategies, we focus on delineating the role that design principles play in explaining this case of biological robustness.

###### 3.1. A fine-tuned model

The fine-tuned modelling strategy relies on a direct, albeit simplified description of the interactions taking place in *E. coli*’s chemotactic network. It also assumes that the exact adaptation response depends on the realization of a precise balance between the different chemotactic parameters. For instance, Albert Goldbeter, Lee Segel and colleagues (Knox et al. 1986) developed a fine-tuned model in which adaptation is explained in terms of receptor modification which acts as a direct response to changed external conditions. According to their model only methylated receptor molecules trigger activity in a sensory system which is proportional to a weighted combination of the fractions of molecules that are in each of the four possible states. Modelling the dynamics of the receptors’ methylation requires tracking the activities of the methylating enzyme CheR and demethylating enzyme CheB. Further assuming that CheR works at saturation with rate VR while CheB works with Michaelis-Menten kinetics, the model requires very precise tweaking of the concentration rates of the chemotactic proteins in order to obtain exact adaptation to step increases in stimuli.

In other words, exact adaptation in this model depends on a strict relation between chemotactic proteins so that changing the value of the parameters leads to loss of the perfect adaptation of chemotactic behaviour. For example, changing the value of CheR concentration by a factor of 20% entails a loss of the precise adaptation response, contrary to what has been observed in the actual behaviour of *E. coli* (cf. Alon 2007: 145-6)*.*

Limitations of the fine-tuned model to capture the actual adaptation response observed in bacterial chemotaxis led to the development of an alternative dynamical model. Barkai and Leibler’s (1997) modelling strategy aimed to capture the stability of *E.coli’*s adaptive behaviour in the face of wide variations of chemotactic parameters. For this purpose the researchers abandoned the idea that the adaptive response is the result of a fine-tuning internal mechanism and focused on explaining the robustness (or functional stability) of this cellular behaviour.

###### 3.2. A dynamical model of robustness

The first step of the dynamical modelling strategy is to specify when and how chemotactic proteins (CheW-CheA, Che R, CheB, CheY and CheZ) affect each other. Using the same simplifying assumptions as the fine-tuned model, Barkai and Leibler (1997) built a very simple network with a single receptor species *E* (standing for the complex MCPs+CheA+CheW), which could be in either one of the two states: *active* or *inactive*. The active state is characterized by an increased CheA activity, which phosphorylates CheY, therefore inducing tumbling. The output of the model is the overall activity of the complex, calculated as a weighted average of all the individual forms of the receptor complex and their activity probabilities (i.e. the average number of receptors in the active state). This quantity is functionally dependent on the CheY-P level and on the kinetic rates of CheY dynamics. The activity probabilities are determined by the input value and the methylation level of the complex. The receptor complex can exist in either attracted–bound or attracted–free form and it can be successively methylated. The rates of change of all possibilities resulting from combining the states described above entail the 14 ordinary differential equations (ODEs) system proposed by Barkai and Leibler (1997).



*Figure 3: Representation of the simplified network used in Barkai and Leibler’s (1997) model; where E represents the receptor complex in its activated state (non-methylated and methylated), f is the tumbling frequency, and R and B are the concentrations of the methylation proteins CheR and CheB.*

The second step of Barkai and Leibler’s modelling consists in writing the relevant equations and implementing them in a computer simulation program. Running the simulations yields a series of predictions concerning the adaptation of the network in the presence of a well-distributed chemical attractant for CheR concentrations varying over several orders of magnitude. That is, the output of the simulation model (*f*) is checked against variations in other network parameters (in particular CheR concentrations). This *sensitivity analysis* shows that the ratio *f/f0 → 1* despite changes in relevant network protein concentrations. This in turn establishes the *robustness* of perfect adaptation since it shows that adaptation does not depend on fine-tuning the values of the various network proteins.

Alon and colleagues (1999) tested experimentally the predictions of Barkai and Leibler’s (1997) dynamical model. The experimental manipulation involved controlling the changing concentrations of each of the chemotaxis proteins (which are the target disturbances for the robustness of perfect adaptation). Using genetically engineered *E.coli* strains, Alon et al. first deleted the gene for CheR from the chromosome, and then introduced into the cell a copy of the gene under control of an inducible promoter. Addition of the inducer led to higher CheR concentration in the cells. The *E. coli* population response to a saturating step of attractant was monitored using videomicroscopy. Subsequent runs of the experiment involved changes in the expression levels of all the other chemotaxis proteins. These experimental manipulations confirmed the predictions of the simulations done as part of Barkai and Leibler’s modelling strategy, showing the functional stability of the adaptation response to varying CheR concentrations.

###### 3.3. An engineering model

Further research showed that there is a formal analogy between *E.coli*’s chemotactic network robust perfect adaptation and the capacity of some engineered systems to maintain some desired function despite significant external perturbations (e.g., how a car’s automatic pilot maintains a constant speed). The model developed by Doyle’s group (and published in Yi et al. 2001) identifies a structural similarity between the robustness of the two types of systems, captured by the *integral control feedback principle* (IFC), which is a principle well-known from control theory.

Control theory is a mathematical framework which describes the behaviour of continuously operating dynamic systems typically in engineered processes and machines. Applied to engineering disciplines, the framework offers algorithms and methods for controlling such systems under specific constraints such as to avoid delay and overshots and to maintain stability of the control action. Doyle’s group relied on these mathematical techniques

to reduce the set of ODEs which characterize the main processes in the chemotactic network to an equation which describes formally the action of a specific type of controller used in several engineering systems.

As illustrated on Figure 4, IFC establishes that the integral of the difference (or error) between the actual output and the desired (steady-state) output is fed back into the system, thus controlling its activity.



*Figure 4. A block diagram illustrating the IFC-principle which ensured that responses to an input u (here: chemoattractant concentration) is normalized to provide an output y that calibrates the difference between the actual output (y1) and the desired output (y0: steady-state receptor activity) - modelled as the integral of the system error on x. Source: Yi TM, Huang Y, Simon MI, Doyle J (2000) Robust perfect adaptation in bacterial chemotaxis through integral feedback control. Proc Natl Acad Sci USA 97:4649-4653. Copyright (2000) National Academy of Sciences, USA.*

In the following, we explain how this analogy has been established and how it can be used to explore the robustness of some biological properties, like perfect adaptation. The starting point of the engineering model (Yi et al. 2000) was Barkai and Leibler’s set of ODEs that was reduced by a series of analytic techniques to a single equation, which corresponds in the control theoretic framework to one of the formulas for computing IFC. [[3]](#footnote-2) Yi et al. (2000) argued that the chemotactic network realizes the type of feedback -- abstractly represented by the IFC principle -- via the methylation state of the receptor. They also claimed that in both types of systems (biological and engineered), IFC is a *necessary and sufficient* condition for obtaining robust control of some target steady-state property (Yi et al. 2000, 4652). We return to this stronger claim below.

The conclusion that the robustness of perfect adaptation should be characterized and understood in terms of the IFC principle is supported by the mathematical equivalence derived between the initial set of ODEs and the IFC characteristic equation. This equivalence guarantees that there is a formal analogy between the internal organization of engineered and biological systems, which in turn helps explain why both systems achieve robust properties in virtue of instantiating a specific type of organization described by the IFC design principle.

The engineering model also guided the researchers in their exploration of the empirical assumptions underlying Barkai and Leibler’s original model. In deriving the IFC equation from the original set of 13 ODEs, Doyle’s group argued that only four assumptions are necessary to obtaining the perfect adaptation response (1. CheB demethylates only active receptors; 2. The kinetic rate constants of CheR and CheB are relatively independent of methylation state and ligand occupancy of receptor complex; 3. The activity of unmethylated receptor is negligible relative to methylated receptor forms; and 4. The concentration of bound CheR is independent of ligand level). These assumptions need to be in place to obtain the precise adaptation output either via the simulations or in solving the control engineering equation. Derivation from these (e.g,, through mathematical manipulation of the parameters) will not give the desired output. Therefore, Yi et al. argue that variations in the robustness of perfect adaptation behaviour can be accounted for in terms of departures from these assumptions.

In our analysis so far we have insisted on the continuity between the dynamic and the control theoretic characterization of the robustness of perfect adaptation. However, it can be argued that the control theoretic model is relatively independent from the specific dynamic assumptions of Barkai and Leibler’s model. On the one hand, the formula for IFC can be derived from a different set of ODEs which would correctly approximate the dynamics of the target network. On the other hand, the precise IFC formula depends on the complexity of the system. In other words, there is no unique formula for IFC. The integral of the error between the actual output and the desired output can be fed back into the system in a direct way or via a series of nested feedback loops. Implementing different types of nested feedback loops will yield different formulas for IFC. However, the IFC is sufficiently abstract and general to apply to a variety of cases in both biology and engineering. Therefore, changing some of the empirical or mechanistic assumptions about the target chemotactic network might still lead to a dynamical description which can be reduced to an IFC formula, understood as a more general design principle. As discussed in the following, this has some important implications for the explanatory characteristics of the model.

1. **Structural-causal explanations**

A first striking feature of the account associated with the control theoretic model is that it characterizes the robustness of perfect adaptation in terms of the *abstract* category of a design principle (the IFC principle). Following the terminology used by the scientists themselves (e.g., Alon 2007), we will sometimes refer to these explanations as *design explanations*, although they differ from another class of contrastive functional explanations that have been discussed in the philosophy of biology under the same name (Wouters 2007; van Eck and Mennes 2016, 2018). Section 5 discusses the relevant differences but for now we stress that our aim is to clarify the sense in which design principles from engineering function as explanatory abstractions[[4]](#footnote-3) in accounts of biological robustness. For this purpose we can rely on a thin notion of a design principle (Green, Levy and Bechtel 2015; Levy and Bechtel 2013) which refers to typically formal representations of ways in which complex systems can be internally organized or relationally linked.

As shown in section 3, characterizing the robustness of the perfect adaptation of bacterial chemotaxis in terms of the principle of integral feedback control was the result of a modelling process which involved “leaving out” various details about *E. coli*’s chemotactic network. The result of this process of abstraction is a mathematical formula which characterizes a certain type of internal organization or a pattern of organization that can be implemented either in an engineering or a biological circuit. This representation of the pattern of organization implemented by the chemotactic network during perfect adaptation is *more abstract* than the representation proposed in Barkai and Leibler’s dynamical model of the same target phenomenon. By making this comparison explicit we want to highlight that the abstractness of the design principle used in the engineering-based explanation of biological robustness is a relative feature of the representation being used. One representation or model is more abstract than the other (Levy forthcoming). Further we argue that the more abstract model, despite containing less causal details, is explanatory because it conveys the *right* type of(i.e., relevant, sufficient) information concerning the explanatory questions, i.e., about the generic factors on which the perfect adaptation response depends.

The abstract representation identifies that the target property (that of robust perfect adaptation) depends on a certain form of internal organization of the system (as illustrated in Figure 4). To capture the explanatory work that engineering design principles can do in contexts such as that of bacterial chemotaxis research we can repurpose one recent philosophical account of *abstract explanation* in science. Pincock (2015) claims that:

we have an explanation when we have found (1) a classification of systems using (2) a more abstract entity that is (3) appropriately linked to the phenomenon being explained. Whenever an explanation has these three features I will say that we have an *abstract explanation*. (Pincock 2015, 867, *m.e*.)

Pincock’s (2015) account of abstract explanation dovetails nicely with our discussion of the key features of the engineering-based account of robust perfect adaptation. We showed that the engineering model builds upon the *abstract* representation of the chemotactic network in the dynamical model. The formal derivation methods allow researchers to explain the property of robustness in terms of the model incorporating the abstract IFC design principle. The abstract description is “appropriately linked to the phenomenon being explained” because it is derived from the ODEs which describe the behaviour of the chemotactic network. In addition, the design principle plays a systematization (or classification) function by grouping together the causal systems (biological and engineered) which exhibit robust behaviours in virtue of the causal structures in which their internal parts (whatever they might be) are embedded (Green and Jones 2016: 355).[[5]](#footnote-4)

Since the systematization function is typically associated with the generality of an explanatory account, it is important to clarify that although engineering explanations using design principles tend to have a wider scope than, for instance, particular mechanistic explanations, their generality is not a mere consequence of their abstract character. That is, we do not take either abstractness nor generality to be sufficient to distinguish this class of explanations of biological robustness. As argued by Matthiessen (2017), the mechanistic framework can also accommodate certain types of abstract explanations.[[6]](#footnote-5) Moreover, cataloguing the result of the engineering model as an abstract explanation *tout court* fails to be informative about the role that engineering design principles play in this and other similar cases of biological explanation. For this reason we need to take our analysis of engineering explanations of biological robustness beyond Pincock’s account of abstract explanations. We propose that such explanations should be further qualified as a subtype of abstract explanations in science, namely as *structural-causal explanations*.

This qualification might strike some as oxymoronic given that current philosophical debates tend to use examples of structural explanations in science to articulate and defend a view of non-causal explanations (e.g., Ladyman and Ross 2007; French 2014; Feline 2015; Glennan 2017). However, there are exceptions to this rule. Some philosophical accounts of structural explanation do not build in a strong contrast between the causal and structural character of an explanation (e.g., Bokulich 2009; Leuridan 2013; Haslanger 2016). To avoid a purely terminological dispute, we proceed next to clarify the two labels that we use in characterising these explanations of biological robustness.

Firstly, the term *structural* can be directly linked to the thin notion of design introduced in the opening of this section. Thus we take the notion of “structure" to designate the same thing as “patterns of organization" did in our definition of the thin notion of design, namely a set of interrelationships holding between the objects that constitute a complex system. Stuart Shapiro offers such a definition of the generic notion of structure:

A structure is the abstract form of a system, highlighting the interrelationships among the objects, and ignoring any features of them that do not affect how they relate to other objects in the system. (Shapiro 1997, 73)

We explicitly cite his definition in order to avoid a potential confusion with another sense of the term “structure" which in biology is often taken to refer to the concrete or material constituents of a biological system. To keep the two senses apart, we will talk about relational structure when the “patterns of organization" meaning is intended. Following this line of thought, abstract explanations such as those that appeal to engineering design principles are structural explanations because they emphasize the link between the relational structure of a system (biological or engineered) and the robustness of some specific functional property or behaviour that the system exhibits. An important aspect of the case discussed here is that the model provides a mathematical warrant for the realization of a certain type of organizational pattern, and thus for the explanation of the target robustness phenomenon.

Secondly, we also claim that this class of design explanations have a *causal* character. The formal model articulated in terms of the IFC principle is linked to the original explanandum (the robustness of perfect adaptation) via a causal interpretation. However, this causal interpretation does not suffice to justify the explanatory force of the model. The formal mathematical derivation of the IFC equation from the set of ODEs describing the interactions in the chemotactic network is also a key element in securing the relevance of the explanation. Therefore, we claim that there are two features that make the control theoretic model explanatory relative to its target phenomenon (the robustness of perfect adaptation): (1) the mathematical derivation of the IFC equation which provides a mathematical warrant for the robustness of the perfect adaptation, regardless of many variations of causal details, and (2) the causal interpretation of the variables in the IFC equation which ensures the empirical relevance of the design principle for the case of bacterial chemotaxis. An important aspect to notice is that the causal interpretation of the IFC design principle does not amount to a causal story about how the robust perfect adaptation response is achieved in particular systems. Rather, its role is to guarantee that the variables in the formal model correspond to causal relations in *E.coli’*s chemotactic network. In addition, the mathematical model can be causally interpreted in the context of other systems (biological or engineered) which exhibit similar robust behaviours. As we will show in section 5, this distinction is important for differentiating structural causal explanations from mechanistic explanations in biology.

In summary, we maintain that the engineering modelling strategy identifies a particular type of *causal structure*, viz. integral feedback control, upon which the robust adaptive behaviour of the bacterium *E. coli* depends. The causal structure represented formally via the characteristic equation of IFC explains why the adaptive behaviour is robust rather than fine-tuned. That is, it explains why the adaptation occurs despite wide variations in biochemical parameters rather than being sensitive to their precise values.

In more general terms, engineering-based design explanations of robustness work by showing that sometimes the functional stability of some property or behaviour of a biological system depends on certain causal structural features of its internal organization. In addition to identifying this type of structural dependence relation, design explanations help to systematize or group together those causal systems (biological and/or engineered) whose robustness depends on the relational structure in which their components are organized. The fact that such relational structures are typically characterized in formal terms, e.g., in our case study by deriving the characteristic equation of the IFC principle from the equations describing the core dynamics of the chemotactic system, is what makes the proposed explanatory account a sub-type of abstract explanations in science.[[7]](#footnote-6) By appealing to design principles which represent the causal structures on which the robustness of certain biological properties or behaviours depends, structural-causal explanations are “insensitive” to details about the causal story of the target robust phenomenon. Since they bypass detailed causal descriptions as well as constitutive features of the chemotactic system’s adaptive behaviour, they differ from canonical causal-mechanistic explanations in biology. We return to this point in the following section.

1. **Causal but non-mechanistic explanations**

The robustness of bacterial chemotaxis has been used before as a case study for philosophical accounts of biological explanation. Here we briefly review two such proposals which take the engineering modelling to support non-causal and mechanistic approaches to explanation, respectively. We start with Braillard’s account of design explanations, highlighting the problems of a non-causal interpretation of the engineering modelling strategy. Then we turn to Matthiessen’s mechanistic proposal of viewing design principles as mechanism schemas. We argue that the latter proposal downplays the distinctive character of the modelling strategy outlined in section 3.

*5.1. A non-causal model of design explanations*

The engineering model of robust perfect adaptation discussed in Section 3.3 is presented by Braillard (2010) as a clear-cut example of a systems biology strategy yielding non-causal, non-mechanistic explanations of biological robustness. Braillard’s approach extends Arno Wouters’ (2007) account, according to which *design explanations* are a special type of functional explanations that target questions such as: “Why do fish respire with gills rather than lungs? Why do tetrapods respire with lungs rather than gills? Why don’t these organisms respire through their skin?" (Wouters 2007). Wouters argues that design explanations differ both from etiological and causal role functional explanations since:

they [design explanations] explain neither by showing that during evolution these structures have been selected nor by showing how the mechanism contributes to some function of the system. Rather they point to synchronic and non-causal functional dependencies between the system’s structure and its environment. The aim is to understand why the actual design is better than some alternatives. (cf. Braillard 2010, 51).

So, the explanatory value of descriptions invoking design principles depends, according to Wouters, on the fact that they invoke physical laws or constraints. The explanatory answers to the physiological questions cited above connect the organism’s need for oxygen, laws of gas diffusion, and the physical properties of the environmental medium, the respiratory surface and the size of the animal. For instance, it follows from Fick’s law of gas diffusion that an organism with a high demand for oxygen needs a large respiratory surface. Design explanations that invoke such general physical constraints or laws are further qualified as a type of *abstract contrastive explanations*. They establish that the occurrence of some target feature of a system depends on some physical law or constraint.[[8]](#footnote-7)

Extending Wouters’ account from comparative physiology to systems biology, Braillard claims that the explanation of robust perfect adaptation in terms of the IFC design principle is a *non-causal design explanation*. The core of Braillard’s argument for the non-causality of design explanations is that they provide necessary conditions for the functionality of a biological system. He writes: “[d]esign explanations […] show 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).

Braillard does not provide a criterion for delimiting causal from non-causal model-based explanations. It is therefore not clear what blocks the causal interpretation of the explanation associated with the Doyle’s group engineering modelling efforts. Braillard suggests that causal explanations must have temporality built into them. Since design explanations using engineering principles like the IFC do not make explicit the temporal distinction between cause and effect (what precedes what) he concludes that they are non-causal. However, a closer look at the relational structure represented by the IFC equation shows that the temporality of the cause-effect relation is present in the type of feedback corresponding to the IFC principle. According to this design principle, the system is organized so that it uses information about the difference between two states (actual and steady-state). This implies that at some time t2, the system’s activity (in the case of the bacterial chemotactic system, the level of methylated protein concentrations) depends on something that happened at t1 when the activity level was different from the steady-state (t0) activity levels.

To strengthen the non-causal interpretation, Braillard emphasises the link between the abstract character and the generality of design explanations. Although abstractness and generality are often thought to go hand in hand, we argue that they are best understood as distinct properties of scientific models. Abstractness concerns details, or more precisely, lack thereof, while generality is about the scope of a model (Levy forthcoming). So, while we agree with Braillard that model-based explanations of robustness which invoke design principles are both abstract and general, we don’t think this characterization is incompatible with their causal character. That is, the variables in the formally derived engineering model can be interpreted causally as representing relations in the target chemotactic system.

Moreover, whether the generality of design explanations is a virtue in the context of explaining biological robustness needs to be empirically established via additional scientific modelling and experimentation, rather than derived conceptually from the assessment of the abstract character of mathematical models. For this reason, we disagree that the generality of design explanations of biological robustness can be inferred conclusively from the abstract (mathematical) character of the engineering model. What is required is systematic empirical investigation of the scope of applicability of design principles in framing explanations of different instances of biological robustness.

Braillard’s interpretation is motivated by Yi et al.’s claim that IFC is a *necessary and sufficient condition* for the robustness of perfect adaptation (and similar biological properties). Such claims are clearly different from mechanistic explanations, and possibly instantiate a higher-level description of mechanistic explanations, in the sense that IFC provides “a formal constraint on possible mechanistic models capturing the causal operation of chemotaxis in different systems” (Green 2015, 641). Yet, one could be cautions to take such claims about the scope of models at face value. IFC might be both necessary and sufficient for achieving the robust control of certain steady-state properties in simple systems which operate in a noisy environment. But this is different from showing that the robustness of perfect adaptation *always* depends on the IFC-like organization of biological organisms. Indeed, other researchers commenting on the paper by Yi et al. have cautioned against the idea that RPA necessarily entail the modelled form of integral control (Briat et al. 2016).

Still, the discussion about RPA and the IFC principle shows that an important epistemic aim in this research context is to explore the genetic features of biological (and engineering) systems with the same capacities. The engineering modelling efforts support the search for nested types of control feedback organization on which the robustness of other biological properties might depend (Stelling et al. 2004; Rao, Kirby, and Arkin 2004; Rao, Glekas, and Ordal 2008; Typas and Sourjik 2015). In these contexts, the focus shifts from that of explaining how specific systems (or mechanisms) work to that of investigating the relational structures that can support a specified type of function (Green 2015). While we believe that the non-causal character of Braillard’s alternative account of design explanations is disputable, we agree with Braillard that a mechanistic interpretation does not adequately capture the explanatory virtues of the engineering model. We now turn to such an example of a mechanistic interpretation of the same case study.

###### 5.2. Design principles and mechanism schemas

As a foil we consider a philosophical analysis which argues for the mechanistic character of design explanations of biological robustness. Matthiessen (2017) argues that the abstractness of design explanations does not suffice to demonstrate that they constitute a distinctive non-mechanistic type of explanations in biology. His argument is in line with other mechanist philosophers who emphasize the role of abstraction and idealization in the construction and validation of mechanistic explanations (e.g., Levy and Bechtel 2013; Craver and Darden 2013: 32-4; Baetu 2015). We agree with this perspective and with much of Matthiessen criticism of non-causal interpretations (see Section 4.1). Yet, we take issue with his central contention that the explanatory features of control theoretic analyses of biological robustness can be understood only on the backdrop of a mechanistic framework.

In particular, Matthiessen (2017) argues that the dynamical and engineering models of the robustness of perfect adaptation are explanatory only insofar as they count as mechanism *schemas* for the target phenomenon. More generally, he implies that all explanatory projects pursued in systems biology must satisfy mechanistic standards. To have a mechanistic explanation of a phenomenon implies, minimally, showing how lower-level entities are organized and interact with each other in order to produce a higher-level target phenomenon (e.g., Illari and Williamson 2012; Skillings 2015). According to a widespread consensus among philosophers of biology, mechanistic explanations differ from canonical causal explanations in virtue of their citing constitutive or compositional explanatory facts (e.g., Fagan 2015; Glennan 2017). In other words, mechanistic explanations are best understood as hybrid causal-constitutive accounts of target phenomena or regularities, such as the robustness of perfect adaptation. The absence of reference to constitutive relevance factors in the control theoretic explanation of biological robustness is one of the challenges facing a straightforward mechanistic rendition of this type of modelling.

Matthiessen’s strategy is to show that abstract representations of engineering design principles can be understood as mechanism *schemas* and that the IFC principle is explanatory only insofar as it is a stage towards a *more complete* mechanistic description of how *E. coli* achieves the robust property of perfect adaptation.

A first problem with this strategy is that it points to a contentious issue concerning the standards of mechanistic explanations, namely to the completeness ideal or the idea that the more concrete the explanation the better. This normative ideal implies that nothing short of a complete (highly detailed) description of the microlevel constituents, their interactions and modes of organization would satisfy the standards for a mechanistic explanation of a target phenomenon. But Matthiessen seems to take the completeness ideal to be compatible with the explanatory value of engineering design principles interpreted as examples of mechanism schemas. Given the current state of scientific research, and the nature of scientific inquiry, few if any mechanistic descriptions can be expected to meet the completeness standard. So, on primarily pragmatic grounds, mechanism schemas should get the honorific title of being explanatory.

We do not object to the pragmatist defence of the completeness ideal, but rather to the more implicit and problematic claim that the overall (or perhaps ultimate) aim of modelling efforts in fields like systems biology is to elaborate more concrete (or detailed) mechanistic descriptions of target phenomena such as the robustness of perfect adaptation in *E. coli* and in other specific bacterial species. This methodological assumption is captured in the statement that biological explanations are *primarily* concerned with particular phenomena (“It is important to retain this sense that explanations in biology are constantly oriented towards particular phenomena." cf. Matthiessen 2015, 18). Granting that more complete (or more concrete) mechanistic explanations are desirable in certain research contexts does not imply that they are always to be preferred. Crucially, the engineering modelling methodology in the case examined here seems to take quite the *opposite* direction – towards enlarging the scope of the applicability of the modelling results and exploring more general underlying design principles (cf. Green 2015).

Hence, one problem with narrowing the focus on particularist explanations in biology is that it fails to do justice to the variety of explanatory practices in modern biology. This challenge is particularly important when examining interdisciplinary practices like systems biology which may be inspired by explanatory ideals from the physical and engineering sciences. Another serious problem with analyzing the explanatory contributions of control theoretic analyses in terms of mechanistic schemas is that the application of the label “mechanistic” risks to become a purely stipulative matter (Huneman 2010). This goes against the idea that mechanistic models of explanation aim to capture the standards used in specific scientific practices to evaluate the outcomes of different modelling strategies.

A second related problem is that mechanistic standards require that explanations cite both causal and constitutive factors for some target phenomenon. Or, as we showed in our analysis of the engineering modelling strategy, constitutive or compositional claims do not figure in control theoretic approaches which analyse robust properties in terms of engineering design principles. This suggests that even on a minimal construal of the mechanistic standards for explanation, one should avoid to hastily classify design principles as mechanism schemas and pin their explanatory value to their being a part on a continuum of mechanistic descriptions.

One mechanistic interpretation that comes close to our account is formulated by Levy and Bechtel (2013). They grant that design principles can inform about generic features of system organization by specifying how mechanisms with a specific pattern of connectivity will behave. Unlike Matthiessen (2017), they thus argue that sometimes less (causal detail) is more, and that the virtues of design principles cannot be reduced to mechanism schemas. While we agree with this interpretation, we caution against interpreting design principles merely as abstract mechanisms. To show *why* the same behavior occurs in all or most systems in which a specified type of function is realized is a different explanatory project, compared to traditional mechanistic research. Importantly, research on RPA also goes beyond the causal structures of existing systems in exploring potential design principles for artificial systems within systems biology (Briat el al. 2016). We believe that our suggestion of structural-causal explanations better captures the specific explanatory aim of research aiming to identify generic principles and patterns among causally distinct systems.

5.3. Summary

In summary, our primary agreement with Braillard (2010) lies in the acceptance of the explanatory irrelevance of many causal details in abstract engineering models. One strength of the engineering model lies in the applicability of the same principle to systems of varying causal details that share characteristic dynamical and structural features. In fact, if explanation is about the identification of dependence relations, then omitting these details is an *epistemic requirement* for selecting and identifying the right factors on which the target robust property depends. On this point we differ from Matthiessen’s (2017) interpretation of the same case study. Where we agree with Matthiessen is that the engineering model is not an instance of derivation of necessary and sufficient dependence-relations that are insensitive to empirical demonstration of applicability. For instance, an argument that the IFC is also responsible for the robust adaptive behaviour s in other bacterial species, e.g., in *B. subtilis*, requires an empirical demonstration showing that this target behaviour is similarly insensitive to details about the specific constitution of its chemotactic network.

Finally, we have argued that pointing out the abstract or mathematical character of the engineering model (or of the design principle which figures in the explanans) does not suffice to establish the non-causal character of the explanation associated with it. Many models cite mathematical dependencies between the parameters or variables characterizing a given phenomenon, and some of dependencies represent causal relations whereas others capture non-causal relations (Glennan 2017: 237). Whether the mathematical content of a model determines the causal or non-causal character of an explanation depends on the kind of dependency relation that the mathematical descriptions captures in the world, or on the type of question being asked. In the case of the engineering model, the characteristic equation of the IFC principle represents the causal structural features of the organization of *E. coli*’s chemotactic network. The explanatory force of the model depends *both* on the mathematical warrants provided by the derivation of the IFC characteristic equation and on the causal interpretation of its variables. The *structural-causal* account of design explanations tries to do justice to both these sources of the explanatory value of the engineering model.

1. **Concluding remarks**

Bacterial chemotaxis has been the topic of philosophical as well as scientific debates. The case has received great scientific interest because of the remarkable behaviour of *E. coli* to retain sensitivity to concentration changes, “disregarding” remaining high concentrations of attractant in its environment. Due to this characteristic behaviour, the cell can continuously explore and detect novel changes in its environment. To explain the feature of robust perfect adaptation, a model emphasizing fine-tuning has turned out to have severe limitations in failing to obtain robustness even for small variations in the concentration of the key enzyme CheR. The combined efforts of a dynamic and an engineering model have demonstrated that the robust capacity depends on structural features of the chemotaxis network. We take this case study to raise important philosophical questions about the epistemology of engineering-based approaches to modelling and explaining biological robustness.

In Section 3 we showed that the dynamic model (Barkai and Leibler 1997) offers a representation of perfect adaptation in *E. coli* chemotaxis that combines a mechanistic schema with a quantitative description of key dynamic features of the chemotactic network. The corresponding set of ODEs was the basis for a simulation study which investigated whether the tumbling frequency variable (*f*) returns to the same steady-state value (*f0*) despite internal and external changes. If, after a certain adaptation time (τ), the ratio p= *f/f0=1*, then the adaptation of the bacterium’s movement to its environment is said to be *precise*. The simulation-based sensitivity analysis established that changes in the chemotactic protein concentrations leave this ratio invariant.

The strategy adopted by Doyle’s group involves the mathematical derivation of the IFC characteristic equation from the original set of ODEs. Since the IFC is a design principle implemented in engineered systems that are built to exhibit a robust behaviour, a reverse formal analogy with biological systems suggests that the robustness of perfect adaptation in the *E. coli* case *depends* on the fact that its chemotactic network implements the IFC principle. That is, the difference between the current and steady-state output of the system is fed back into the system via the methylation pathway, and this leads to the robustness of perfect adaptation. Accordingly, Yi et al. (2000) argue for a broader inference to the effect that similar types of organization principles can explain the robustness of other cellular behaviours.

Research on the robustness of bacterial chemotaxis thus holds several interesting lessons for understanding the contributions of engineering approaches to biology. In our view, the diversity of scientific practice calls for philosophy of science to accommodate various models of scientific explanations, ranging from explanations of particular events to explanations of general types of behaviours or events observed in a wider variety of systems.

Our aim has been to clarify the *distinctive* character of the explanatory strategy associated with engineering-based modelling. The control theoretic strategy applied by Doyle’s group identified a specific design principle or abstract pattern of organization (the integral feedback control principle) as the main factor on which the robustness of perfect adaptation depends. We labelled the result of this modelling strategy a *structural-causal explanation* because the main explanatory category is an abstract representation of a causal relational structure or pattern of organization.

To clarify the explanatory features of the model, we have found reference to abstraction and generality insufficient. Similar examples of design principles have been interpreted as abstract models within a mechanistic framework (Levy & Bechtel 2013). While we generally agree with this account, we believe that interpreted design principles merely as abstract mechanisms risks downgrading epistemic virtues that Levy and Bechtel themselves highlight. Design principles show *why* the same behavior occurs in causally different systems that share a particular organizational structure. While this research question can inform mechanistic research, we regard this as an explanatory project distinct from the mechanistic aim of specifying specify how certain capacities are realized through interactions between specified causal entities.

We believe that our suggestion of structural-causal explanations better captures the virtues of the engineering model. More specifically, we have argued that the engineering model explains by representing a set of interrelationships within a given system. This interpretation highlights relational structures of a system (biological or engineered) which can be described in a formal framework. Our proposal of a structural causal account thus highlights the empirical connection to the experimental manipulations of the temporal concentration changes in the chemotactic network while highlighting that the engineering model increases understanding by identifying and ignoring the causal details that are not necessary for describing the dynamics as a result of the relational structure of the system.

### **References**

Alon, U. 2007. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC Press.

Alon, U., Surrette M.G., Barkai N., and S. Leibler. 1999. “Robustness in Bacterial Chemotaxis.” *Nature*, no. 397: 168–71.

Andersen, H. 2017. “Complements Not Competitors: Causal and Mathematical Explanations”. *British Journal for the Philosophy of Science.* [*https://doi.org/10.1093/bjps/axw023*](https://doi.org/10.1093/bjps/axw023)

Baetu, T. 2015. “The Completeness of Mechanistic Explanations” *Philosophy of Science* 82 (5): 775-786.

Baetu, T. 2016. “The ‘Big Picture’: The Problem of Extrapolation in Basic Research.” *British Journal for the Philosophy of Science* 67(4): 941-964.

Barkai, N., and S. Leibler. 1997. “Robustness in Simple Biochemical Networks.” *Nature* 387: 913–17.

Barkai, N., and B-Z. Shilo. 2007. “Variability and Robustness in Biomolecular Systems.” *Molecular Cell: Perspective* 28: 755–60.

Bechtel, W., and A. Abrahamsen. 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 (3): 321–33.

Bokulich, A. 2009. “How Scientific Models Can Explain.” *Synthese* 180 (1): 33–45.

Braillard, P. A. 2010. “Systems Biology and the Mechanistic Framework.” *History and Philosophy of the Life Sciences* 32 (1): 43–62.

Briat, C., Gupta, A., & Khammash, M. (2016). “Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks”. *Cell systems*, 2(1), 15-26.

Brigandt, I. 2013. “Systems Biology and the Integration of Mechanistic Explanation and Mathematical Explanation.” *Studies in History and Philosophy of Biological and Biomedical Sciences* 44 (4): 477–92.

Brigandt, I., Green, S. and O’Malley, M. 2018. “Systems Biology and Mechanistic Explanation” In S. Glennan and P. McKay Illari (eds.) *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. New York: Routledge, pp.362-374.

Chirimuuta, M. 2017. “Explanation in Computational Neuroscience: Causal and Non-causal” *The British Journal for the Philosophy of Science* [*https://doi.org/10.1093/bjps/axw034*](https://doi.org/10.1093/bjps/axw034)

Craver, C., and L. Darden. 2013. *In Search of Mechanisms: Discovery Across the Life Sciences*. University of Chicago Press.

Craver, C. 2016. “The Explanatory Power of Network Models”, *Philosophy of Science* 83 (5): 698-709.

Dretske, F. 1988. *Explaining Behaviour: Reasons in a World of Causes.* MIT Press.

Eisenbach, M. 2001. “Bacterial chemotaxis”. In *Encyclopedia of Life Sciences.* John Wiley & Sons, [www.els.net](http://www.els.net)

Eronen, M. 2015. “Robustness and Reality”. *Synthese* 192(12): 3961-3977.

Evans, M.R. et al. 2013. “Do Simple Models Lead to Generality in Ecology?” *Cell: Trends in Ecology and Evolution* 28 (10): 578–83.

Feline, L. 2017. “Mechanisms Meet Structural Explanation.” *Synthese*.

French, S. 2014. *The Structure of the World: Metaphysics and Representation.* Oxford University Press.

Glennan, S. 2017. *The New Mechanical Philosophy.* Oxford University Press.

Green, S. 2015. “Revisiting generality in biology: systems biology and the quest for design principles” *Biology and Philosophy* 30 (5):629-652.

Green, S., and N. Jones. 2016. “Constraint-Based Reasoning for Search and Explanation: Strategies for Understanding Variation and Patterns in Biology.” *Dialectica* 70 (3): 343–74.

Green, S., Levy A., and W. Bechtel. 2015. “Design Sans Adaptation.” *European Journal for Philosophy of Science* 5: 15–29.

Haslanger, S. 2016. “What Is a Structural Explanation?” *Philosophical Studies* 173 (1): 113–130.

Halina, M. 2018. “Mechanistic Explanation and Its Limits”. In S. Glennan and P. McKay Illari (eds.) *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. New York: Routledge, pp. 213-225.

Huneman, P. 2010. “Topological Explanations and Robustness in the Biological Sciences”. *Synthese* 177(2): 213-245.

Huneman, P. 2017. “Outlines of a Theory of Structural Explanations.” *Philosophical Studies*. doi:10.1007/s11098-017-0887-4.

Huneman, P. 2018. “Diversifying the picture of explanations in biological sciences: ways of combining topology with mechanisms”. *Synthese* 195(1): 115-146.

Illari, P. Williamson, J. (2012). “What is a mechanism? Thinking about mechanisms across the sciences”. *European Journal for Philosophy of Science* 2 (1):119-135.

Jansson, L. and Saatsi, J. 2016. “Explanatory Abstractions”. *British Journal for the Philosophy of Science*.

Jones, N. 2018. “Inference to the More Robust Explanation”. *British Journal for the* *Philosophy of Science* 69(1): 75-102.

Kitano, H. 2004. “Biological Robustness.” *Nature Reviews: Genetics* 5 (11): 826–37.

Knox, B.E. et al. 1986. “A molecular mechanism for sensory adaptation based on ligand-induced receptor modification”. Proc Natl Acad Sci USA.

Ladyman, J. and Ross, D. *Every Thing Must Go: Metaphysics Naturalized.* Oxford University Press.

Lange, M. 2017. *Because without Cause: Non-Causal Explanations in Science and Mathematics.* Oxford University Press.

Leuridan, B. 2013. “The Structure of Scientific Theories, Explanation, and Unification. a Causal-Structural Account.” *British Journal for the Philosophy of Science* 65 (4): 717–71.

Levy, A. forthcoming. “Idealization and Abstraction”. *Synthese* [*10.1007/s11229-018-1721-z*](https://philpapers.org/go.pl?id=LEVIAA-8&proxyId=&u=http%3A%2F%2Fdx.doi.org%2F10.1007%2Fs11229-018-1721-z)

Levy, A. 2017. “Causal Order and Kinds of Robustness.” In *Landscapes of Collectivity in the Life Sciences*, edited by Ehud Lamm Snait Gissis and Ayelet Shavit. MIT Press.

Levy, A., and W. Bechtel. 2013. “Abstraction and the Organization of Mechanisms.” *Philosophy of Science* 80 (2): 241–61.

Machamer, P., Darden L., and C. Craver. 2000. “Thinking About Mechanisms.” *Philosophy of Science* 67: 1–25.

Matthiessen, D. 2017. “Mechanistic Explanation in Systems Biology: Cellular Networks.” *The British Journal for the Philosophy of Science* 68(1): 1-25.

Pincock, C. 2015. “Abstract Explanations in Science.” *British Journal for the Philosophy of Science* 66: 857–82.

Rao, C. V., Glekas G. D., and G. W. Ordal. 2008. “The Three Adaptation Systems of Bacillus Subtilis Chemotaxis.” *Trends in Microbiology* 16 (10): 480–87.

Rao, C. V., Kirby J. R., and A. P. Arkin. 2004. “Design and Diversity in Bacterial Chemotaxis: A Comparative Study in Escherichia Coli and Bacillus Subtilis.” *PLOS Biology*.

Shapiro, S. 1997. *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press.

Slepchenko, B.M., and M. Terasaki. 2004. “Bioswitches: What Makes Them Robust?” *Current Opinioni in Genetics and Development* 14 (4): 428–34.

Skillings, D. 2015. “Mechanistic Explanations of Biological Processes”. *Philosophy of Science* 82 (5): 1139-1151.

Stelling, J., Sauer U., Szallasi Z., Doyle F. J., and J. Doyle. 2004. “Robustness of Cellular Functions.” *Cell Review* 118: 657–85.

Typas, A., and V. Sourjik. 2015. “Bacterial Protein Networks: Properties and Functions.” *Nature Reviews: Microbiology* 13: 559–72.

van Eck, D. and Mennes, J. 2016. “Design Explanation and Idealization”. *Erkenntnis* 81 (5):1051-1071.

van Eck, D. and Mennes, J. 2018. “Mechanism Discovery and Design Explanation: Where Role Function Meets Biological Advantage Function”. *Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie* 49 (3):413-434.

Wagner, A. 2005. “Distributed Robustness Versus Redundancy as Causes of Mutational Robustness.” *Bioessays* 27 (2): 176–88.

Weisberg, M. 2006. “Robustness Analysis” *Philosophy of Science* 73 (5): 730-742.

Wimsatt, W. C. 2007. *Re-engineering Philosophy for Limited Beings: Piecewise Approximations of Reality.* Harvard University Press.

Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation.* Oxford University Press.

Woodward, J. 2013. “Mechanistic Explanation: Its Scope and Limits.” *Aristotelian Society Supplementary* 87 (1): 39–65.

Wouters, A. 2007. “Design Explanation: Determining the Constraints on What Can Be Alive.” *Erkenntnis* 67 (1): 65–80.

Yi, T-M., Huang Y., Simon M.I., and J. Doyle. 2000. “Robust Perfect Adaptation in Bacterial Chemotaxis Through Integral Feedback Control.” *Proceedings of the National Academy of Sciences USA* 97 (9): 4649–53.

1. Notably, our focus on *biological robustness* differs from philosophical discussions of *robustness analysis* as a comparative method that takes several models of the same target phenomenon and seeks to determine how reliable or good knowledge of the target is (e.g., Wimsatt 2007; Weisberg 2006; Eronen 2015; Jones 2018). In this paper we are concerned with robustness as a property of biological systems, and how robustness is modelled and explained in biology and engineering (for a similar focus, see e.g., Wagner 2005; Huneman 2010, 2017, 2018; Levy 2017). [↑](#footnote-ref-0)
2. This working definition implies that robustness is a relative, not an absolute property. No system is robust with respect to all its behaviours (or properties) and under any kind of change. The system as a whole can have both robust and fragile properties or behaviours. In the biological as well as in the engineering sciences, the notions of robustness and functionality are tightly connected because what is taken to be robust is either some typical activity or behaviour performed by the system or some desired system characteristic. [↑](#footnote-ref-1)
3. The steps of the mathematical derivation are showed in the supplementary material of Yi et al. 2001. The derivation focuses on the activity of the receptor complex and applies the principle of mass action to the ODEs in the original Barkai and Leibler model. [↑](#footnote-ref-2)
4. Similar points about the explanatory value of abstract representations or categories have been made in Jansson and Saatsi (2016); Levy (forthcoming). [↑](#footnote-ref-3)
5. “Design principles categorize seemingly different biological processes into types through a demonstration of a general principle that they all instantiate, and this categorization makes many causal details or differences between biological processes irrelevant for understanding certain system behaviors. Importantly, unification in this context does not mean reduction of mechanistic explanations to principles. Rather unification is reached through higher-level reflections on types of system organization.” (Green and Jones 2016: 355) [↑](#footnote-ref-4)
6. See also (Levy and Bechtel 2013). [↑](#footnote-ref-5)
7. By emphasizing the structural elements of design explanations, our account has common features with so-called “explanations by constraints”, a model of scientific explanation recently defended by Huneman (2017, 2018), Green and Jones (2016) and Lange (2017). This family of accounts

describes explanations by constraints as *non-causal* and emphasizes the use of mathematical or formal tools in delineating the explanatory answers to the target questions. We have argued that design principles that figure in some structural explanations of robustness can be interpreted causally, despite their mathematical representational format. However, this difference does not rule out the possibility of there being non-causal explanations of other types of robustness (cf. Huneman 2017, 2018). [↑](#footnote-ref-6)
8. Discussing Wouter’s account, Van Eck and Mennes (2015) characterize the structure of design explanations as follows: “Design explanations address the following type of contrastive why-question: ‘why does organism o have trait t rather than trait t’? Or somewhat more fine-grained: ‘why does item i have characteristic c rather than c’?’ [↑](#footnote-ref-7)