Chemical Arbitrariness and the Causal Role of Molecular Adapters

Oliver M. Lean

Published in *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* (2019, issue 78, 101180)
Preprint - Please cite published version

Abstract

Jacques Monod (1971) argued that certain molecular processes rely critically on the property of *chemical arbitrariness*, which he claimed allows those processes to "transcend the laws of chemistry". It seems natural, as some philosophers have done, to interpret this in modal terms: a biological relationship is chemically arbitrary if it is possible, within the constraints of chemical "law", for that relationship to have been otherwise than it is. But while modality is certainly important for understanding chemical arbitrariness, understanding its biological role also requires an account of the concrete causal-functional features that distinguish arbitrary from non-arbitrary phenomena. In this paper I elaborate on this under-emphasised aspect by offering a general account of these features: arbitrary relations are instantiated by mechanisms that involve molecular *adapters*, which causally couple two properties or processes which would otherwise be uncorrelated. More specifically, adapters work by acting as *intermediate* rather than *cooperating* causes.

1. Introduction

It's common to hear that certain biological phenomena at the molecular level involve relations that are "arbitrary" or "chemically arbitrary". This is most commonly attributed to the genetic code—specifically, the relationship between the three-base triplets of mRNA and the amino acids they specify. By contrast, other molecular relations are cited as chemically "necessary" or "non-arbitrary", including those between DNA and complementary RNA bases in copying. This difference has been taken by some to be deeply important to biological functioning and evolution, and yet it has proven difficult to spell out just what this difference amounts to.

The intuition behind this distinction is that non-arbitrary relations are in some sense determined, or at least constrained, by the chemical properties of the molecules involved in a way that arbitrary ones are not. Transcription, for example, occurs by complementary pairing between the bases: a G on the parent strand produces a C on the daughter strand, for example, because one binds the other stably through their complementary stereochemical structures.

This reliance on the "intrinsic" chemical properties of the bases, the thinking goes, constrains the pairings that are possible, suggesting a degree of chemical *necessity* to the pairings that actually occur.

In contrast, the translation of an RNA transcript into sequences of amino acids seems markedly different: The fact that "AUG" codes for methionine isn't due to any direct contact between the codon and its corresponding amino acid. Instead, this is because AUG binds to specific tRNA molecules that carry methionine at their other end. If the tRNAs were different, then, the genetic code would be different (indeed, in some organisms it *is* different). Hence, the RNA-amino acid pairings seem more contingent—less determined by the dictates of chemistry—than the DNA-RNA pairings in transcription. This is the reason typically given for calling the genetic code chemically arbitrary.

Here it is critical to distinguish this notion of chemical arbitrariness from what we might call *functional* arbitrariness. Francis Crick (1958) famously suggested that the particular assignments of codons to amino acids may be a "frozen accident"—one that became entrenched due to the severe consequences of deviating from it once it was established. This claim is now widely questioned: there is now a mass of evidence suggesting that the standard genetic code is optimal or near-optimal in several respects relative to other possible codes (see e.g. Maynard Smith and Szathmary, 1995; Butler et al, 2009; Kumar and Saini, 2016). For example, point mutations in DNA that lead to amino acid changes are likely to substitute amino acids that are chemically similar to the original, which minimises the chances of loss of function. This suggests that the standard code is not a frozen accident at all, but that it was reached after a period of variation and selection; understanding how the code was reached is a fascinating scientific puzzle (Koonin and Novozhilov 2008). Nevertheless, this matter is quite independent of the claim that alternative codes are at least *chemically* possible. In fact, the claim that the genetic code has been subject to selection *implies* that variations existed to be selected from in the first place, and chemical arbitrariness is said to be the property that makes this variation possible (Bergstrom and Rosvall 2009). So the concept of chemical arbitrariness—the theme of the present paper—is quite independent of claims about functional arbitrariness and frozen accidents. (However, as I will revisit in Section 4, the two are connected in important ways.)

As the above suggests, the distinction between arbitrary and non-arbitrary phenomena is not thought to be an idle observation: Not only does such a distinction *exist* in the biological world, but this distinction has been claimed to be extremely important for understanding many aspects of life processes and their evolution. This importance was emphasised by the molecular biologist Jacques Monod in his book *Chance and Necessity* (1971). Monod argued that chemical arbitrariness (or *gratuité*) was a critical property of biological phenomena that permitted living things a near-limitless plasticity of function. His central example is the *lac* operon system that he famously co-discovered—a segment of the *E. coli* genome containing genes for a number of proteins involved in the uptake and metabolism of lactose. Transcription of all these genes is controlled by a single *repressor* protein, LacI. In the absence of lactose, LacI binds to an "operator" region upstream of the *lac* genes, which prevents them from being read by the transcription machinery. When lactose is

present, however, it binds to LacI and causes a conformational change that releases LacI from the operator, allowing the genes to be transcribed. The upshot is that the *lac* genes are switched on when there is lactose for them to process, and switched off when there isn't.

The key feature of this system that Monod highlights is the particular way that LacI couples the expression of the *lac* genes with the presence of lactose. While there is no "chemically necessary relationship" (*ibid.* p. 76) between these two things, LacI achieves this regulation through an *allosteric* interaction: the binding of lactose at one of its sites distorts its shape at another, making it unable to bind to the operator. Allosteric interactions like these are what Monod calls "gratuitous" (or elsewhere "chemically arbitrary"—I will use this term), a property he views as critical to understanding the plasticity of biological systems:

> In a word, the very gratuitousness of these systems, giving molecular evolution a practically limitless field for exploration and experiment, enabled it to elaborate the huge network of cybernetic interconnections which makes each organism an autonomous functional unity, whose performances appear to *transcend the laws of chemistry if not ignore them altogether*." (Monod, 1971, p. 78, emphasis added)

Aside from Monod's claims about the importance of arbitrariness itself, the notion has also been used in philosophical work to explicate other biological concepts. In particular, John Maynard Smith (2000a,b) draws on Monod's ideas to elaborate his account of biological information. In his view, biological phenomena that we recognize as "informational", such as DNA and hormones, differ from non-informational ones in key ways. In particular, DNA and hormones can be said to carry information or "meaning" because what they do—their function—depends on their "interpretation" by other "evolved receivers" such as tRNAs or cell surface receptors. In contrast, what an enzyme does is "directly determined by its structure" (Maynard Smith 2000a, p. 193), and hence there is no need for anything resembling *interpretation* of a signal that the enzyme carries.

Yet despite the importance they attribute to chemical arbitrariness for biological systems and our understanding of them, Monod and Maynard Smith offer little by way of sustained philosophical analysis of the concept itself—of exactly what the implied lack of a "necessary connection" amounts to, for instance. When we try to clarify the difference between arbitrary and non-arbitrary, as some have done, we run into some difficulties. For example, one suggestion is that the difference is a matter of directness of interaction: the necessary connection between DNA and mRNA bases, or between enzyme and substrate, arises from their direct stereochemical contact, whereas there is no such direct interaction between codons and amino acids in translation. However, Godfrey-Smith (2000) questions whether directness of interaction is able to ground the distinctions typically made between arbitrary and non-arbitrary phenomena:

> "Perhaps there is a difference in degree here; an enzyme's catalytic action is strongly constrained by its physical structure, whereas a hormone could have a huge variety of effects depending on the location and structure of the receptors with which it

> interacts. On the other hand, the hormone's interaction *with those receptors* is certainly a matter of its physical structure; why look further "downstream" with the hormone but not with the enzyme?" (Godfrey-Smith 2000, p. 203, emphasis in the original)

In other words, it is trivially true that all "indirect" causal interactions are mediated by a series of steps each of which is directly related to the next. So the idea that arbitrariness is just a matter of indirectness is hard to square with the crucial biological importance that Monod and others have afforded it. In particular, it is hard to interpret it the way Monod does, as allowing living things to "transcend the laws of chemistry". In summary, it has proven difficult to make explicit why certain phenomena in molecular biology have been described as chemically arbitrary and others not, without deflating the concept to a trivial matter of indirectness and losing sight of why so much has been made of it.

The aim of this paper is to develop an account of chemical arbitrariness that clarifies what it is and why it is thought to be so important for our understanding of molecular biological systems. As we will see, a number of philosophers have aimed to interpret the "non-necessity" of chemically arbitrary phenomena in *modal* terms; that is, they offer a way of understanding chemical arbitrariness by clarifying the sense in which the dictates of chemistry permit certain alternatives and forbid others. I will argue that, while modal considerations are an important piece of the puzzle, understanding the importance Monod and others have afforded this concept requires some further elaboration, which I offer here.

The paper proceeds as follows: Section 2 outlines the aforementioned attempts to clarify the concept of chemical arbitrariness in modal terms, and explain why they are in need of elaboration. It is this elaboration that I offer in Section 3, which involves an account of the causal-functional features that characterise phenomena that we recognize as arbitrary. Specifically, I argue that arbitrary causal relations rely on molecular *adapters*. I define adapters as molecules that couple two other properties or processes by acting as *intermediate*, rather than *cooperating* causes—a distinction I clarify in terms of the interventionist theory of causation. In Section 4, I discuss how this helps make sense of what is biologically important about chemical arbitrariness; that is, it highlights that the causal relations we typically call arbitrary are problems to which adapters are solutions. Section 5 concludes.

2. Modal Accounts of Chemical Arbitrariness

This section outlines two existing attempts to characterise the notion of arbitrariness. I call both of these "modal" accounts because, while they differ in significant ways, they both tie arbitrariness to the possibility of alternatives within certain constraints imposed by theory. I will argue that while these accounts are not incorrect, they are in need of elaboration with causal-functional detail that I offer in Section 3.

The first and most dedicated discussion of chemical arbitrariness comes from Ulrich Stegmann (2004), who offers a general set of conditions that aim to explicate Monod's use of

the term. His account includes a pair of definitions—first, a set of conditions for chemical necessity, and then for chemical arbitrariness as the failure to meet those conditions:

> "**Chemical necessity**. A relation ***R*** between two molecules is chemically necessary with respect to ***R′*** if and only if (i) ***R*** is a necessary condition for another relation ***R′*** to hold between them, and (ii) the ***R*** *is a necessary condition for* ***R′***-relation holds in virtue of a chemical principle." (Stegmann, 2004, p. 210)

> "**Chemical arbitrariness**. The relation ***R*** between molecules M_1 and M_2 is chemically arbitrary with respect to ***R′*** if and only if either (i) ***R*** is not required for ***R′*** or (ii) if ***R*** is required for ***R′***, then it is not a chemical principle that ***R*** is required for ***R′***. *"* (ibid. p. 214)

To illustrate how these conditions apply to specific cases, consider again the examples of transcription and translation. In transcription (*ibid.* p. 218*)*, the relation between the corresponding DNA and mRNA nucleotides is chemically necessary because (i) the two must be chemically complementary (***R***) for one to specify (***R′***) the other, and (ii) this is due to one or more chemical principles; namely, those that pertain to specific binding interactions. In translation (*ibid.* p. 215), by contrast, the relation between codon and amino acid—say, CAC and histidine—is chemically arbitrary, for the following reason: While it is true that (i) a codon must have certain chemical properties in order to specify histidine (i.e. properties possessed by CAC and its "synonyms") there is (ii) no *chemical principle* dictating that whatever codes for histidine must have these particular properties. This is a formalisation of the intuition that there is no "chemical reason" for the genetic code to be as it is (again, this doesn't imply that there are no *functional* reasons).

Evidently, Stegmann's account leans heavily on the notion of a chemical principle, which he defines as "what 'must' be the case for chemical reasons" (p. 211). While he doesn't require them to be precise or exceptionless (*ibid*.), their modal character gives chemical principles a status closely akin to laws: they support counterfactual reasoning about whether the chemical conditions that apply in actual cases would apply to all *possible* cases, such as alternative evolutionary outcomes in which the milieu of biomolecules were different. Stegmann takes Monod's "principle of associative stereospecificity" to be an example of a chemical principle: according to this principle, the pairings between nucleotides in translation are "sterically predestined" (Monod 1971, p. 106): there could not have been a biological world in which A was paired with G and T with C, for instance.

Sahotra Sarkar (2000; 2005) also offers an account of chemical arbitrariness, albeit by a different name. Like Maynard Smith, his account aims to analyze what it means for DNA to contain information for proteins, or, in other words, for one kind of molecule to be a "sign" for another. More specifically, he aims to explicate the notion of information in genes without Maynard Smith's reliance on a notion of intentionality, opting instead for a purely causal or mechanistic account. In Sarkar's view, some entity *s* is a sign for some other entity σ only if two conditions are met 1) specificity, and 2) arbitrariness (which he refers to as "semioticity",

or elsewhere as "template assignment freedom"; I will retain the term "arbitrariness" to avoid confusion). The arbitrariness condition is met if "the theory which provides the mechanisms by which *s* produces σ allows that an *s′* different from *s* could have been the sign for σ" (Sarkar, 2000, p. 210). In his (2005, p. 274) he takes this to mean that alternative template assignments are "evolutionarily possible".

Despite differences in focus, these two accounts are relevantly similar. Firstly, both predicate arbitrariness of *relations* of some kind between two chemical entities. In Stegmann's case, these relations take various forms such as "is chemically complementary to" or "specifies in transcription". In Sarkar's case, the relation between sign and signified is effectively causal, evidenced by talk of how one "produces" the other. Secondly, both are appealing in some way to the dictates of chemistry to support counterfactual reasoning about evolutionary alternatives and the possibility thereof: Where Stegmann gets this counterfactual force from "chemical principles", Sarkar appeals to "the theory which provides the mechanisms" underlying the causal relationship between sign and signified. In any case, both in effect aim to define a space of possibilities whose boundaries are set by chemical theory. Subsequently, the relation in question is arbitrary if and only if that space of chemical possibilities contains alternatives to the relations in question.

My claim is not that these accounts are wrong-headed: The concept of chemical arbitrariness is certainly tied to the possibility of alternatives, and we should certainly look to chemistry (and neighbouring disciplines) to furnish us with the constraints on what is possible. However, even if we define chemical arbitrariness in something like the modal terms outlined above, more is needed if the concept is to do the work it is claimed to do. Specifically, if we are to understand chemical arbitrariness as a property of biological systems, and understand what that property does for those that possess it, we need a clear means of identifying it in the biological systems we study, and in a way that connects it to certain other features of those systems. As they stand, the modal accounts above do not yet meet this requirement.

To clarify the nature of this omission and, why it matters, consider again the case of transcription—a paradigm example of a *non*-arbitrary phenomenon. According to Stegmann's framework, there are chemical principles dictating that corresponding bases in transcription (***R′***) in must be chemically complementary (***R***). And because chemical principles hold under alternative evolutionary pathways, the requirement of complementarity is not simply because of how evolution happened to unfold. However, imagine a world in which DNA is transcribed into mRNA in a similar way to translation—using tRNA-like intermediates. In that world, the complementarity condition would not apply: an A in the DNA could specify an A in the mRNA, or indeed anything else. Cumbersome as this would be (remember that functional equivalence is not at issue here), it at least appears to be within the realm of chemical possibility. If this is so, the claim that the nucleotide relations in transcription are "chemically necessary" looks to be false: they are only due to how DNA happens to be transcribed. More generally, this may lead to the conclusion that *all* relations in biology are arbitrary: For any case in which two molecules are related in a chemically necessary way, we could imagine a version of the process in which they are not. If so, we no longer have a property that is

possessed by only *some* biological phenomena, one that makes those phenomena stand out as distinct or important in some way.

There is an obvious way to block this conclusion, which is to claim that the alternative version of transcription we have imagined (using tRNA-like intermediates) is too radically different to count as a relevant alternative. In Stegmann's terms, we could argue that, while such a case is imaginable, the transcription relations between nucleotides in this world are not instances of ***R′*** at all. If so, this counterexample does not show that transcription is non-arbitrary; instead we have just substituted a relation that isn't arbitrary for one that is.

This may be so. However, for this to be a convincing response we must justify this move as more than just an *ad hoc* exclusion. This will require giving reasons why this radical alternative should not be admitted as an instance of the transcription relations denoted by ***R′***. So what exactly characterises this relation we're interested in? It won't do, it seems, to understand transcription just as the production of mRNA molecules whose sequence is determined by DNA sequences, since this would admit the radical alternative we want to exclude. Instead, then, we must appeal to a difference not just in what transcription does but how it does it. In short, the non-arbitrariness of transcription relations (for example) is essentially tied to the underlying *mechanism* by which those relations are realized.

Sarkar's account gets closer to acknowledging the relevance of these details, since he relates the possibility of alternatives to "the theory which provides the mechanisms" underlying the relation. However, to decide whether a given relation is arbitrary—whether the theory providing the mechanisms allows for alternatives—we must of course spell out what characterises the mechanisms in question. In turn, this requires an explicit account of which alternatives would count as an instance of the same mechanism and which don't—in other words, of that mechanism's essential or at least characteristic features.

Again, this does not mean that we should not characterise arbitrariness in much the way that Stegmann and Sarkar have done, i.e. as the possibility of alternatives as set out by the dictates of chemistry. The issue with these accounts is not that they define arbitrariness in terms of the chemical possibility of alternatives. Instead, the point is that to put this concept to work in picking out certain phenomena in biology we also need to clarify which chemically possible "alternatives" are relevant. That they do not provide this is quite understandable if their aim is simply to clarify the modal aspect of the notion of arbitrariness. However, *using* the term to understand and explain certain features of biological phenomena requires a means of determining which phenomena it applies to; in other words, how to recognise arbitrariness when we see it.

In summary even if we understand chemical arbitrariness in terms of modal conditions, we still need an account of the concrete, empirical features of the mechanisms that meet these conditions[1]. It is an account of these mechanistic features that I offer in Section 4. In Section 5, I address the question of how this account connects chemical arbitrariness to its functional importance to biological systems.

[1] Thank you to Ulrich Stegmann for pressing me for clarity on this point.

3. Molecular Adapters and Intermediate Causes

The previous section outlined two existing analyses of the concept of chemical arbitrariness—both modal in that they appeal in some way to the chemical possibility of alternatives. I argued that while tying arbitrariness to modality *per se* is not misguided, more is needed for an adequate understanding of which phenomena are arbitrary and why it matters. What is needed, in short, is an account of the essential or characteristic mechanistic features that make certain relations arbitrary and others not.

In this section, I offer an account of what these features might be. This account may not provide *necessary* conditions for a mechanism to be recognizable as a case of arbitrariness; to argue for that would require a radical exploration of biological possibility that I aim to avoid. I will therefore allow that there may be other kinds of causal or mechanistic structures that would similarly satisfy intuitions about the possibility of alternatives; that arbitrariness is multiply realizable, as it were. Nevertheless, I propose that the conditions I provide are at least *sufficient* for arbitrariness, and that they correctly sort the paradigm cases of arbitrary and non-arbitrary phenomena discussed so far.

My account will follow Sarkar, and depart from Stegmann, in attributing arbitrariness to *causal* relations. More specifically, I will characterise arbitrariness as pertaining to causal relations as understood on the interventionist theory of causation expounded by Woodward (2003). I will not undertake the task of comparing this framework with rivals, either in general or in the way it handles the issue of arbitrariness at hand. Instead, I will simply show that this framework contains useful tools for understanding the difference between arbitrary and non-arbitrary causal relations and the importance of this difference.

A key feature of interventionism that will become relevant is that it ties causality to manipulability: two properties, or *variables* representing properties, are causally related, as opposed to merely correlated, if one can be changed by intervening on the other. For example, to say that there is a causal relationship between the sequence of the DNA gene (*D*) and the mRNA sequence it produces (*R*) is to say that some ideal interventions[2] changing the DNA sequence would change the mRNA sequence. (Here, *D* and *R* represent sequences of the four DNA bases {A,T,G,C} and the four RNA bases {A,U,G,C}, respectively.) Similarly, to say that there is a causal relationship between mRNA sequence and protein sequence *P* is to say that some interventions changing mRNA sequence would change the sequence of the resulting amino acid chain. (Since the genetic code is *redundant*, not all of these interventions would change the protein, yet the fact of a causal relation between them holds.) These causal relations can be represented in a *directed graph*, as follows:

[2] An ideal intervention is "a causal process that changes the value of X in an appropriately exogenous way, so that if a change in the value of Y occurs, it occurs only in virtue of the change in the value of X and not through some other causal route" (Woodward 2003, p. 94).

Clearly, the facts established so far do not reveal why the causal relation $R \rightarrow P$ is arbitrary while $D \rightarrow R$ is not. To see the difference, then, we must add more detail to the causal structure than is included above. In other words, it is not just *that* these relations obtain; rather it has something to do with *how* they obtain. This requires us to open the black box and provide a description of the molecular machinery—the proteins and other components—that underwrite or explain that causal relationship. With this in mind, the purpose of the rest of this section is to express the difference between arbitrary and non-arbitrary causal relationships in terms of certain differences in the mechanisms—the underlying causal processes[3]—that underwrite those relationships.

My approach will elaborate on the following idea: that arbitrary mechanisms (read: "mechanisms underwriting arbitrary causal relationships") involve entities that play a particular kind of functional role[4]. This idea has been in the background of the discussion since the beginning: Maynard Smith suggests that arbitrariness, which he takes to be a signature of informational processes in biology, connects to the role of "evolved receptors" in carrying out the function of those processes (2000b, p. 215). Similarly, Monod suggests that allosteric proteins, his central example in discussing chemical arbitrariness, are "transducers" of chemical signals (1971, p. 76). Earlier still, the mechanism of translation was hypothesised by Crick (1958) to involve an "adapter", a role that was attributed to tRNA "intermediates" by Zamcnik and colleagues around the same time (Hoagland et al. 1958). I therefore should not overstate the novelty of my proposal; the idea that arbitrariness has something to do with these functional roles is not itself new. My contribution will be to put the entities playing these functional roles—which I will collectively call *adapters*—under the spotlight, characterise their role explicitly in causal terms, and explain how this functional role connects the modal aspect of arbitrariness to the key biological importance that Monod and others have granted it.

To that end, this section first characterises "adapter" as a functional class. By "functional class" I mean a way of categorising parts of a system based on their functional role—on what they do and how. Other functional classes in this sense include "pump", "catalyst", "hinge", "gated channel", "energy source", and so on—the kind of categories that are deemed useful for understanding functional systems in fields that study them, including engineering and the biological sciences.

With this in mind, molecular entities that can be called adapters, I propose, exhibit two key features: 1) they establish a *causal coupling* between two properties or processes that would be uncoupled without that adapter; 2) they create this coupling by acting as

[3] In this paper, this is what I mean when I refer to "mechanisms": they are explanations for causal relationships in molecular biology that cite molecular entities and their causal interactions. If this happens to differ from the way the term is used by Craver (2001) or Machamer et al. (2001), for instance, I do not take these differences to be important for present purposes.

[4] Here I adopt a causal role notion of function akin to Cummins (1975), as opposed to a teleological notion.

intermediate, rather than *cooperating*, causes. These properties are best illustrated using two real-world examples of systems that use molecular adapters:

Case 1: Synthetic. The first example is an instance of synthetic biology—a system derived from nature but artificially altered for human ends. The system in question, described by Bayley and colleagues (Gu et al. 1999), is a single-molecule sensor that can be used to detect a range of different organic molecules (or "analytes"—objects of analysis). This system involves two key components: the channel protein α-haemolysin, and ß-cyclodextrin which fits inside that channel and partially blocks it. This blockage causes a reduction in the passage of ions through the channel that can be measured as a reduction in conductance. It is the ß-cyclodextrin that the researchers refer to as an "adapter": While in place in the channel it can bind to a range of different analytes, which leads to a further measurable change in conductance. Importantly, different analytes block the channel by different amounts and remain there for different times. Hence, each analyte produces a characteristic signal pattern across two dimensions—the change in conductance it induces, and the duration of that change.

Case 2: Natural. The second example, from Medzhitov and colleagues (Horng et al. 2001), is natural rather than synthetic, observed to play a role in immune responses to bacterial infection in mice. Perhaps the most significant fact about this example, for present purposes, is the researchers' account of how it was discovered. It was already known that detection of telltale products of bacteria by Toll-like receptors (TLRs) could activate various immune responses via a signalling pathway that included the protein MyD88. However, the researchers found that at least one receptor, TLR4, could still activate those immune responses even when the MyD88 gene was knocked out. This suggested that TLR4 "is coupled to an additional signaling pathway that is independent of MyD88 and likely controlled by a distinct, and as yet unidentified, adapter" (p. 835). This prompted the search for a protein that connected the causal dots between TLR4 and the immune response. The adapter in question was found and named TIRAP.

What does it mean to say that ß-cyclodextrin and TIRAP each function as "adapters" in their respective systems? First, both achieve a *causal coupling*, a causal dependency between two properties that would otherwise not exist: in the first case, between the presence of analytes and measurable conductance; in the second, between TLR4 activation and immune response. In other words, adapters constitute what Dretske (1988) calls a *structuring cause*: they are not just causes of an effect, they underwrite or explain a cause-effect relationship[5].

However, merely facilitating another causal relationship, or being necessary for that causal relationship to obtain, is too liberal to qualify something as playing an adapter role. This is also true, for example, of enzymes with respect to the reactions they catalyze: Without RNA polymerase in transcription, for example, DNA sequences would not cause RNA sequences (at anywhere near a biologically salient rate). To see what is distinct about the role played by adapters, then, we need to look closer at the particular way in which they

[5] I'm grateful to Samir Okasha for reminding me of Dretske's notion of structuring causes.

underwrite the relationships in question—at the causal structure of the mechanism in which they play a part. This is where the distinction between *intermediate* and *cooperating* causes comes in.

This distinction is illustrated in Figure 1, which shows two different cases in which *X* and *Y* are causally related. In the first case (a_1), this relationship is indirect in the sense that there is an *intermediate* variable *A* between them: *X* causes *Y* because *X* causes *A,* which in turn causes *Y*. The three variables form what Pearl (2009) calls a "chain". This kind of structure carries an important implication: when *A* is intermediate between *X* and *Y*, intervention on *A* "breaks" the causal relationship *X*→*A* (a_2), thus eliminating *X*'s causal relevance to *Y* (see Woodward 2003, pp. 38-39). To put it another way, if we wanted to control *Y*, being able to intervene on *A* would make control over *X* unnecessary. Consider the elaborate chain of contraptions constituting a Rube Goldberg engine: one can intervene at a variety of points in the chain to either trigger the process or prevent its continuation, independently of what happens upstream.

In the second case (b_1), *X* directly causes *Y*, while the third variable *E* is a *cooperating* cause. (The three variables form a "collider" or "inverted fork", in Pearl's terminology.) Consider a flashlight with a battery (*X*) and a switch (*E*): To turn on the light bulb (*Y*), the battery must be in place *and* the switch must be in the "on" position. Control over the switch doesn't eliminate the need for a battery, and no amount of manipulation of the battery can turn on the light if the switch is off. More generally, this causal structure is one in which intervention on *E,* as seen in (b_2), *does not render X causally irrelevant to Y.*

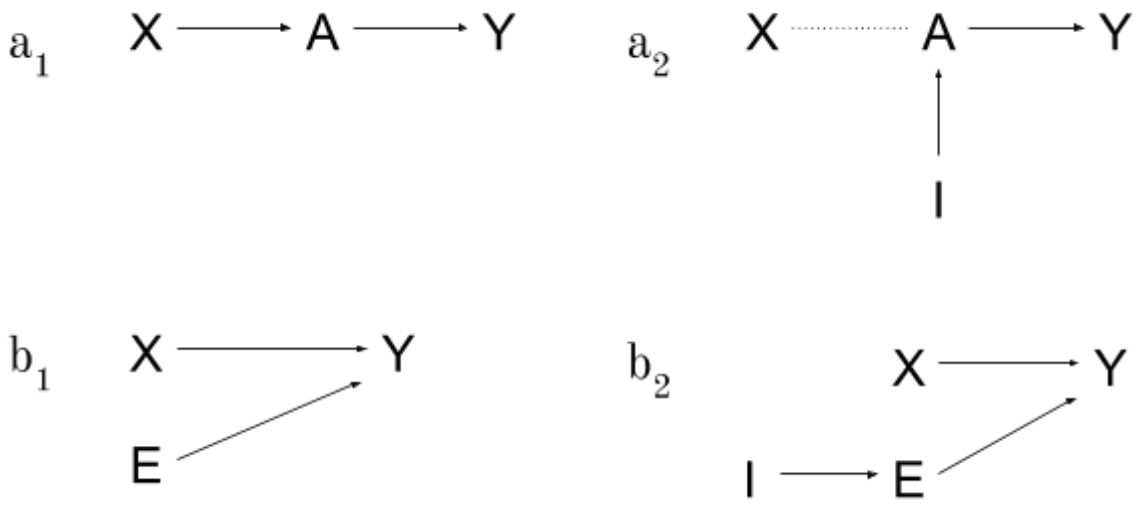

Figure 1. Adapter *A* couples two properties or processes by acting as an *intermediate* cause, while enzymes *E* (for example) act as *cooperating* causes. **(a_1)** *X*, *A*, and *Y* form a "chain", in Pearl's (2009) terminology. **(a_2)** Intervention *I* breaks the causal path from *X* to *Y;* X's effect on Y is thus fully accounted for by X's effect on A. **(b_1)** In contrast, *X*, *Y*, and *E* form a "collider"; *E* is a *cooperating* cause with respect to *X*→*Y*. **(b_2)** *X*'s effect on *Y* is *partly* independent of *E*, in the sense that intervention on *E* does not break the causal dependency between *X* and *Y*.

My claim is that what distinguishes adapters as a functional class is that they facilitate a causal relationship in the first way rather than the second, i.e. by acting as an intermediate, rather than a cooperating cause. In a given mechanism that underwrites a causal relationship $X \rightarrow Y$, a molecule plays the role of an adapter if it can be represented as an intermediate cause between X and Y. In turn, it is when at least one molecule acts as an adapter in this sense that we recognize the causal relationship $X \rightarrow Y$ as arbitrary.

This focus on the role of adapters applies to the examples seen so far. In translation for example, tRNAs act as the adapters in the mechanism by which messenger RNA specifies amino acid sequence (Figure 2). The mRNA recruits the corresponding tRNA to the ribosome, which in turn determines which amino acid is added at a given position. The arbitrariness of the translation process—the reason the genetic code can be altered *without altering the mechanism*—is because of the mechanism's use of tRNA molecules, and because these molecules constitute intermediates in the causal relation between mRNA and amino acid sequence.

Here we should consider whether this result is a matter of how we choose to represent the process: Is it possible to truthfully represent this mechanism as a collider rather than a chain, with tRNA as a cooperating cause rather than an intermediate? The answer is no, because to do so would be to make a different empirical claim: To represent tRNA as merely a cooperating factor would be to claim that even once the tRNA is accounted for, the mRNA sequence still has some bearing on the resulting amino acid. Yet according to the accepted model, it does not: mRNA's effect on protein sequence is entirely due to its intermediate effect on tRNA recruitment. Whether this model is accurate is of course a defeasible matter of empirical fact, as it should be.

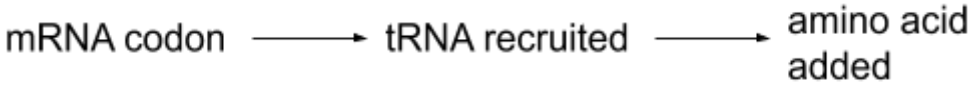

Figure 2: tRNAs act as an *intermediate* cause: it is through tRNAs that mRNA sequence causally specifies amino acid sequence.

The story applies to the role of the repressor in Monod's *lac* operon model (Figure 3): With respect to the causal coupling of *lac* gene expression to the presence of lactose, this is explained by lactose's effect on the conformation of the repressor. An implication of this model is that if we intervened on repressor's conformation, thus breaking its upstream causal connection, the presence or absence of lactose would no longer make a difference. According to this model, then, the Lac repressor acts as an adapter, which is why lactose's causing *lac* gene expression is arbitrary. Again, this model and one in which the repressor is a cooperating cause are distinct causal structures with definable empirical consequences, and thus the difference is an empirical matter rather than one of choice of representation.

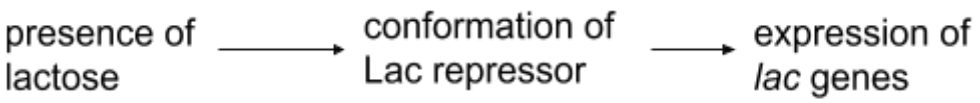

Figure 3: the Lac repressor is an intermediate cause that couples the presence of lactose with the expression of the *lac* operon genes

We can also see how paradigm examples of *non*-arbitrary mechanisms fail to meet the above conditions. The role of molecules such as enzymes cannot be accurately represented as intermediates, for example; instead, they act as *cooperating* causes. Consider the role of RNA polymerase in transcription, shown in Figure 4: While it may undergo a range of conformational changes and interactions with DNA[6], the DNA sequence retains an independent causal relevance. Note that "independent" here does not mean that the polymerase enzyme is unnecessary; it simply means that the role of the DNA in specifying mRNA sequence is not fully accounted for by its effect on the enzyme: no external manipulation of the RNA polymerase would render the DNA sequence irrelevant to the sequence of the mRNA transcript.

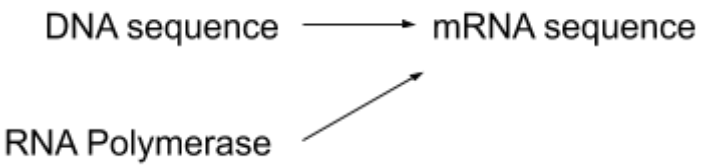

Figure 4: In the causal process by which DNA sequences produce mRNA sequences, RNA polymerase acts as a cooperating cause rather than an intermediate.

A final point requires clarifying: I have argued that 1) arbitrary relations are mediated by adapters, and that 2) an entity's role in a mechanism is that of an adapter when it acts as an intermediate in the cause-effect relation in question. It may be argued that this is simply a restatement of the claim that arbitrary relations are indirect which, as we've seen, threatens to make arbitrariness appear trivial.

Consider, for example, a metabolic pathway such as the one in Figure 5. S_1 to S_4 represent the substrates in this pathway. Suppose we wanted to know if the causal relation $S_1 \rightarrow S_4$ is arbitrary. To do so, we need to cite the molecular entities that play a role in generating this causal relationship—in other words, the list of things you need if you want to turn S_1 into S_4. In this case, the entities in question are the enzymes E_1 to E_3; each acts as a cooperating cause, and is hence not an adapter. But since the two intermediate variables S_2 and S_3 are intermediate variables, are these not adapters?

However, my claim is not that a causal relation is arbitrary whenever it has intermediate variables. Instead, my focus is on adapters as a functional class, and the fact that only certain mechanisms make use of them. While it is informative to understand the adapter role, and how this differs from other functional roles, in terms of the difference between chains and forks in causal graphs, there is more to adapters than is expressible in this way. But we are not forced to conclude that S_2 and S_3 in Figure 5 are adapters after all, or that

[6] It may be asked why there is no causal arrow connecting DNA and RNA polymerase in Figure 4, given that the two interact. The reason is that the particular *sequence* of DNA does not make a difference to how the enzyme behaves, and vice versa. Hence, there is no causal arrow to be drawn between the particular variables shown.

enzyme pathways are arbitrary; we all know that there is a difference between the machinery of a production line and its intermediate products. It's just that this difference is not easily spelled out in terms of causal structure[7]. In short, my claim is importantly different from the claim that arbitrariness is indirectness: I include an additional and important characterisation of adapters as a functional class.

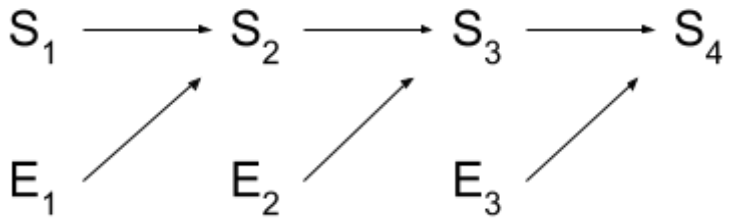

Figure 5: An enzyme pathway featuring substrates S_1 to S_4 and enzymes E_1 to E_3, which act as cooperating causes on the pathway. While S_2 and S_3 are intermediates in this causal chain, they are not adapters but merely intermediate products in the pathway.

In summary, the difference between causal relations that are typically recognized as "arbitrary", and those that aren't, is that the mechanisms for arbitrary causal relations rely on one or more entities that play the role of an adapter. In turn, to act as an adapter in a mechanism is to create a causal coupling between two other variables (or properties represented by variables) by acting as an intermediate cause between them. This feature of adapters explains why mechanisms that use them are amenable to the kind of alternatives suggested by the notion of chemical arbitrariness: altering certain features of the adapter can thus alter the causal relation.

In contrast, while non-arbitrary cause-effect relations may require a host of supporting molecules, the cause variable retains a causal relevance that is independent of all of these. In other words, there is a path from cause to effect that is not broken by intervention on any of the supporting molecular machinery. This places greater constraints on the ways in which the causal relation can be altered by altering the supporting machinery. One may wish to know more about the particular nature of this constraint in individual cases: In transcription, for example, DNA's independent role in the specification of RNA sequence comes from its specific binding of the ribonucleotides that are added to the chain. In the case of enzyme catalysis in metabolic chains, the substrate contributes some or all of the *material* that is present in the product. However, for the purpose of understanding the difference between arbitrary and non-arbitrary mechanisms, the reasons for non-arbitrariness can be subsumed under the more general principle of independent causal relevance.

This account adds the required qualifications to the modal aspect of arbitrariness as discussed in Section 2. That is, it provides an account of the characteristic features of mechanisms that instantiate arbitrary relations—to put it another way, an account of *what to look for*, empirically, when determining whether a given causal relation is arbitrary or not, or what it is about the mechanism of an arbitrary causal relation that permits alterations to that

[7] Another way to express this difference may be that intermediate products exhibit what Lauren Ross (2018) calls *material continuity*.

relation. These conditions may still imply that a great many causal relationships in biology are arbitrary. Importantly, however, it maintains that many of them are not.

One reason why this should be seen as a refinement of the modal account, rather than an alternative to it, is that it does not aim to avoid modal reasoning entirely; counterfactual thinking remains in reference to the effects of hypothetical interventions on parts of the system. However, it is worth noting that this refinement has the benefit of reducing its demands on our counterfactual imagination: We need only think about what would happen to a given mechanism if we altered or intervened on it, not about radical alternatives to that mechanism as a whole. Of course, even this relatively modest counterfactual thinking will require support from theoretical principles provided by chemistry. Hence this is not a replacement of existing modal treatments but an elaboration on them—one that blocks problems raised by the possibility of wholesale changes to the mechanism.

4. The Biological Importance of Adapters

Suppose we accept the above proposal—that a molecular mechanism, or the causal relation it sustains, is "arbitrary" when and because it relies on a molecular adapter. What, then, is biologically important about arbitrariness understood in this way? How does an understanding the role of adapters in arbitrary causal relations lend itself to a fuller understanding of living processes and their achievements? Recall that Monod's view of the importance of arbitrariness was that it freed biological processes from the dictates of chemistry—a grand claim on its face, and one that invites a modal interpretation. My claim has been that, even granting those modal features, one cannot properly understand the arbitrariness of arbitrary mechanisms without clarifying what concrete mechanistic features give rise to that freedom. Now we have an idea of what these features are, what bearing does this have on Monod's claim? It is this question I will explore in this section: I consider why molecular adapters are key to understanding the biological importance of arbitrariness, and in particular what this has to do with their being intermediate rather than cooperating causes in the arbitrary relations that they underwrite.

The answer, I propose, is that elaborating on the mechanistic underpinnings of arbitrary relations prompts a subtle but important shift in emphasis on the nature of the problem at hand. As we've seen, the usual framing of the arbitrariness problem is to take an observed relation in biology, such as between DNA and mRNA or mRNA and protein, and ask whether, and in what sense, it could have been different than it is—whether the same cause could have produced a different effect, or the same effect could have been produced by a different cause. If so, the relation is arbitrary in the sense that it is chemically contingent, and by implication a result of the molecular machinery evolution has produced.

But while this reference to chemical contingency is by no means false, and is indeed an important aspect of arbitrary phenomena, there is as yet no reference to the *functional* importance of those relations: In every example considered in this paper—conditioning *lac* gene expression on the presence of lactose, meeting pathogens with an immune response, and so on—the cause-effect relation in question is valuable, even indispensable, to meeting the

needs of the organism. When we remember this, the fact that there is no chemical "law" or "principle" guaranteeing those responses is not a metaphysical mystery. In fact, it may even be thought of as the fundamental problem of life: that physics and chemistry are indifferent, or even hostile, to the needs of living things. With that in mind, forging causal connections that are not "predestined" is an engineering problem that living things have to solve. So when we observe that organisms succeed in coupling causal processes according to functional need, and responding to stimuli in advantageous ways, the question this raises is what kinds of mechanism make this possible.

Put in this way, we can understand the functional class of adapters as a vital solution to problems of that kind. If functional classes are characterised by the kind of problem they solve, we can view adapters as a distinct kind of solution to the problem of how to couple properties or processes according to functional needs. We could call these *coordination* problems—problems about how to coordinate form and/or behaviour with often-unpredictable changes in external conditions (see e.g. Bechtel 2009). Given a coordination problem, the question to ask is *how* that coupling is or could be achieved. In the synthetic case study in Section 3, for example, the researchers began with a causal coupling that they *wanted* but did not yet exist—that is, between the presence of various analytes on one hand and measurable changes in conductance on the other. Similarly, the researchers in Case 2 observed an existing causal coupling in nature, noted that it could not be explained by the signalling pathways that were known so far, and thus wondered how else it was achieved; this suggested that there must be an as-yet-unknown adapter forging this connection. A similar story applies to the discovery of tRNAs: It was known that mRNA sequence determined protein sequence before it was known how this was done. In all these cases, the answer to the question included a critical role for a molecular adapter.

At this point we can also see the importance of the *intermediate* causal role that adapters play, and which set them apart from other functional classes of molecule. Recall that in the case of enzyme catalysis, there is a pre-existing relevance of cause to effect that is in a sense independent of that enzyme: a chemical complementarity in the case of replication, for instance. In those cases the enzyme's role, however necessary, is simply to complete the causal picture rather than forging it *ex nihilo*. In other cases, however, *there is no such pre-existing relationship to exploit*. In those cases, the only possible solution is a mechanism that involves an intermediate—an entity that is both affected by the cause and causes the effect, thus connecting the two. In short, adapters are necessary in cases where the intended outcome requires more than just causal cooperation.

What's more, even if it is *possible* in principle to achieve a functional goal through cooperating causes alone, this is not necessarily the most effective means of doing so. As Crick (1958) acknowledged, for example, it might have turned out that translation relied on direct chemical contact rather than the intervention of adapters. Yet this would not necessarily have been preferable. In fact, it would largely defeat the purpose of using DNA as a carrier of information for proteins in the first place: As Bergstrom and Rosvall (2009; 2011) argue, transmitting copies of the proteins themselves would be cumbersome—while very good at being biological agents, proteins are less good at being copied and transported. DNA, on the

other hand, “is exquisitely fashioned so as to (1) encode lots of sequence information in a small space, (2) be incredibly easy to replicate, (3) be arbitrarily and infinitely extensible in what it can say, and (4) be structurally very stable and inert” (*ibid.* p.167). However, using DNA as a transmission solution poses the further problem of how to retrieve the protein sequences from nucleic acid sequences once transmitted. tRNA adapters constitute a key part of the answer. Even if the particular assignments of the genetic code had turned out to be a frozen accident, the functional problem of encoding protein sequence in DNA would remain: even if different assignments were functionally equivalent, there are very good functional reasons for storing protein sequence specificity in DNA in the first place. In turn, doing so involves a key functional role for tRNA adapters regardless of which particular code they use.

A further advantage to adapter-based mechanisms over non-arbitrary ones, even if the latter are possible, is that adapters once in place are amenable to further change. The evolutionary search for an optimal genetic code, made possible by translation’s use of adapters, is just one example. As another, the synthetic molecular sensor in Case 1 can be “programmed” to detect a variety of different analytes, making this same system adjustable for a variety of different purposes. In general, adapters can be added, removed or tweaked through mix-and-matching of modular receptor and effector domains, according to functional needs. And because they act as causal intermediates—wholly rather than just partly responsible for the causal relation they forge—this can be done in a way that is less constrained than if they were merely cooperating causes.

In summary, considering the distinct role of adapters in phenomena we recognize as arbitrary, especially their role as intermediate causes, allows us to connect the chemical contingency implied by arbitrariness to the biological importance that has been granted to this property. While chemical contingency is a critical piece of the puzzle, understanding this importance also requires an account of what underwrites this contingency. When Monod says that biological systems are able to “transcend the laws of chemistry”, he evidently does not mean that arbitrary relations confound chemical understanding. Instead, he simply means that through the use of adapters (or “transducers”), life can “choose” what is coupled to what according to functional requirements. Because of this, understanding the freedom of choice implied by arbitrariness requires clarifying the adapter molecules underwriting this freedom.

5. Conclusion

In this paper, I’ve proposed an elaboration of the definition of arbitrariness as the chemical possibility of alternatives. This elaboration serves to bring out the biological importance that Monod and others have granted this property. My proposal has been that what makes chemical arbitrariness functionally important is that arbitrary causal relations are mediated by molecular adapters, which couple a cause to an effect by acting as intermediates in causal chains between them. This establishes chemically arbitrary causal relations as an important class of molecular phenomena from a functional perspective: they constitute a set of functional problems for living things to which adapters are the best, or even the only, solution. By understanding the engineering problems adapters solve—such as coordination, or

reading the "information" for protein synthesis—we can understand the significance granted to their results. It is in this sense that arbitrariness frees living things from the constraints of chemistry: whether or not "anything is possible" with them, as Monod suggested, far more is possible with them than without.

Acknowledgements

I'm grateful to Marc Ereshefsky, Samir Okasha, Ulrich Stegmann, C. Kenneth Waters, colleagues at the University of Calgary, and two anonymous referees for invaluable comments on this paper. This work was supported by the European Research Council Seventh Framework Program (FP7/2007-2013), ERC Grant Agreement No. 295449, and by funding from the John Templeton Foundation ("From Biological Practice to Scientific Metaphysics," grant no. 50191).

References

Bechtel, W. (2009). Constructing a philosophy of science of cognitive science. Topics in Cognitive Science 1, 548–569.

Bergstrom, C. T., & Rosvall, M. (2009). The Transmission Sense of Information. Biology & Philosophy, 26(2), 159–176

Bergstrom, C. T., & Rosvall, M. (2011). Response to commentaries on "The Transmission Sense of Information." Biology & Philosophy, 26(2), 195–200.

Butler, T., Goldenfeld, N., Mathew, D., & Luthey-Schulten, Z. (2009). Extreme genetic code optimality from a molecular dynamics calculation of amino acid polar requirement. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 79(6), 1–4.

Craver, C. F. (2001). Role Functions, Mechanisms, and Hierarchy. Philosophy of Science, 68(1), 53–74. https://doi.org/10.1086/392866

Crick, F. H. (1958). On protein synthesis. Symposia of the Society for Experimental Biology, 12, 138–163. https://doi.org/10.1038/227561a0

Cummins, R. (1975). Functional Analysis. The Journal of Philosophy, 72(20), 741–765.

Dretske, F. I. (1988). Explaining Behavior: Reasons in a World of Causes. MIT Press.

Godfrey-Smith, P. (2000). Information, arbitrariness, and selection: Comments on Maynard Smith. Philosophy of Science, 67(2), 202–207.

Gu, L. Q., Braha, O., Conlan, S., Cheley, S., & Bayley, H. (1999). Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. Nature, 398(6729), 686–690. https://doi.org/10.1038/19491

Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I., & Zamecnik, P. C. (1958). A Soluble Ribonucleic Acid Intermediate in Protein Synthesis. Journal of Biological Chemistry, 231, 241–257.

Horng, T., Barton, G. M., & Medzhitov, R. (2001). TIRAP: an adapter molecule in the Toll signaling pathway. Nature Immunology, 2(9), 835–841.

Koonin, E.V. and Novozhilov, A.S. (2008) Origin and evolution of the genetic code: The universal enigma. *Life* 61(2): 99–111.

Kumar, B. and Saini, S. (2016) Analysis of the optimality of the standard genetic code. *Molecular BioSystems* 12, 2642–2651.

Maynard Smith, J. (2000). The concept of information in biology. Philosophy of Science, 67(2), 177–194.

Maynard Smith, J. (2000). Reply to commentaries. Philosophy of Science, 67(2), 214–218.

Maynard Smith, J., & Szathmáry, E. (1995). The Major Transitions in Evolution. Oxford University Press.

Monod, J. (1971). Chance and Necessity. New York: Alfred A. Knopf.

Ross, L. N. (2018) Causal Selection and the Pathway Concept. Philosophy of Science 85(4), 551-572

Sarkar, S. (2000). Information in genetics and developmental biology: Comments on Maynard Smith. Philosophy of Science, 67(2), 208–213.

Sarkar, S. (2005). How genes encode information for phenotypic traits. In Molecular Models of Life: Philosophical Papers on Molecular Biology (pp. 261–283). MIT Press.

Stegmann, U. E. (2004). The arbitrariness of the genetic code. Biology & Philosophy, 19(2), 205–222.

Woodward, J. (2002). What Is a Mechanism? A Counterfactual Account. Philosophy of Science, 69(S3), S366–S377.

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