European Journal for Philosophy of Science (2022) 12:15
https://doi.org/10.1007/s13194-022-00445-1

PAPER IN THE PHILOSOPHY OF THE LIFE SCIENCES

The logic of explanation in molecular biology: historical-processual and logical-procedural aspects

Giovanni Boniolo [1] · Raffaella Campaner [2]

Received: 12 November 2020 / Accepted: 6 January 2022/Published online: 10 February 2022

Abstract

This work addresses biological explanations and aims to provide a philosophical account which brings together logical-procedural and historical-processual aspects when considering molecular pathways. It is argued that, having molecular features as *explananda*, a particular non-classical logical language – *Zsyntax* – can be used to formally represent, in terms of logical theorems, types of molecular processes (pathways), and to grasp how we get from one molecular interaction to another, hence explaining why a given outcome occurs. Expressing types of molecular biology processes in terms of the *Zsyntax* language allows us to represent causal interactions by taking into account their context-sensitivity, and amounts to partly reviving the spirit of the so-called received view of explanation – which aimed to capture scientific explanatory accounts in terms of their logical structure and their appealing to nomological relations. Such a partial revival is pursued by invoking here non-classical deductions and empirical generalisations, which are called to provide the epistemic norms to explain the behavior of molecular pathways.

Keywords Zsyntax · Non-classical deductions · Empirical generalisations · Historical explanation · Processual explanation · Processes · Pathways

Giovanni Boniolo and Raffaella Campaner contributed equally to this work.

✉ Giovanni Boniolo
giovanni.boniolo@unife.it

Raffaella Campaner
raffaella.campaner@unibo.it

[1] Dipartimento di Neuroscienze e Riabilitazione, Università di Ferrara, Via Fossato di Mortara, 64a, 44121 Ferrara, Italy

[2] Department of Philosophy and Communication Studies, University of Bologna, Via Zamboni, 38, 40126 Bologna, Italy

1 Introduction

As well-known, the philosophical debate on scientific explanation has gradually moved in a few decades from the initial search for a formal account which could apply to *all* explanations in *all* scientific fields, to an increasing stress on the peculiar aspects explanations could get in different scientific contexts.[1] Emphasis has been put on different sorts of *explananda*, in particular on differences between *explananda* in physics and those in the other special sciences, and on how they might require to be explained on the basis of different explanatory accounts. This general trend has been accompanied also by some shift from the idea that the scaffold of explanatory procedures is entirely logical, to the idea that a variety of different relations – e.g., statistical, causal, functional, etc. – can play a role. Furthermore, discussions have been largely pursued in the last decade or so on whether explanations are to be qualified as epistemic or ontic, and how exactly their epistemic and/or ontic features must be understood.

Such debates have significantly involved the biological sciences, and questioned what relations genuinely explain the occurrence of biological phenomena, given their peculiar features of being highly variable and dependent on the context in which they occur. The very idea that genuine biological *laws* hold has been questioned. Accordingly, also the idea has been challenged that we might adopt a nomological approach to explanation, meant to subsume the *explananda* under alleged "biological laws" included in the premises of a logical argument..However, even if "all generalizations in biology seem to admit of exceptions" (Waters, 1998, p. 5), and "the traditional notion of a universal law of nature has few, if any, applications in [e.g.] neurobiology or molecular biology" (Machamer et al., 2000, p. 7), biological phenomena exhibit patterns which are not merely accidental, and regularities hold which are "intermediate between mere observations of empirical regularities and universal propositions on nature" (Schaffner, 2006, p. 389). Given the high context-sensitivity of biological phenomena, generalizations will hold ceteris paribus, that is, in the presence of many enabling conditions and in the absence of interferences or counteracting conditions – which are not easy to fully specify in a significant, not vacuous or trivial, way.

Reflections on biological explanation have been intertwined with some wider rethinking of both the explanatory adequacy of a nomological approach to explanation in general and of the notion of "biological law", to make sense of the identification of forms of stability without conflating them with universal regularities. Much of the debate on explanation in biology has then been focusing on causal explanatory relations, and, hence, on how causal links are to be most properly conceived in biological contexts. In particular, much room has been devoted to mechanistic explanations of biological phenomena: to explain – it has been argued – is to identify how a given set of organized entities act and interact in bringing about a given outcome. Positions which, although in different forms, basically follow this line have been advanced in a number of works by, for instance, Machamer et al. (2000); Craver and Darden (2005); Darden (2008); Bechtel (2011), Bechtel and Abrahamsen (2005);

[1] A "classical" essay providing an extensive reconstruction of the philosophical debate on scientific explanation is Salmon (1989), which covers its first four decades. A good sense of the developments of the debate thereafter is given by, e.g., in Persson and Ylikoski (2007). Braillard and Malaterre (2015) effectively shows how a variety of explanatory accounts can be provided in, more specifically, the life sciences.

Brigandt (2013); Baetu (2019); and many others. Recent accounts of scientific explanation that have appealed to causation have mostly done so either by appealing to mechanisms, more often than not stressing the priority of mechanisms over nomological relations, or to counterfactual interventions, which allow us to identify generalizations that are invariant under intervention. Neither emphasizes the explanatory role of formal tools. What we want to stress is that causal *nomological* relations can be expressed in formal terms, and recognized to have explanatory power. Although it can be agreed that biologists are not primarily concerned with the identification of laws, the use of lawlike generalizations is not incompatible with the claim that biological phenomena work as causal mechanisms. As recently remarked by Cartwright, Pemberton and Wieten, very often

> "effects are just what is to be expected given the features of the parts in that arrangement and the covering laws in which these features figure. *The effects are just what is to be expected because that is what must happen if all those features act as they should under the general laws that govern them.* [...] Many satisfactory explanations – especially many mechanistic ones—are covering laws explanations and moreover part of their force as explanations derives from their use of covering laws to derive the explanandum phenomenon" (Cartwright et al., 2020, p. 7; 10; italics in the text).

In this scenario, we suggest a new explanatory perspective to philosophically address molecular biology. It takes as *explananda* the behaviour of molecular pathways, and re-evaluates the role of generalizations. We believe that this latter should be stressed, and that should be done with the joint recognition of the explanatory role of *causal* relations. Our proposal, on the one hand, revives the idea that a philosophical analysis of the structure of scientific explanation can satisfactorily capture molecular processes in logical terms, and, on the other hand, brings together some lessons we have learnt from the debate following the received view – more specifically, some of those regarding historical explanations, the explanatory role of causal relations, and the need to account for the context-sensitivity of the phenomena under investigation.

Our proposal has a logical core. Usually, whenever logic entered the debate on explanation, in particular when the nomological-deductive model was presented, it was meant in its classical formulation, that is, as a classical deduction. The difficulties this approach has had are well-known. In the following sections we will argue for the adoption of a different logical language, *Zsyntax*, which allows to effectively represent types of chains of molecular interactions of different length, i.e. processes or pathways, in terms of non-classical deductive theorems. We will show how *Zsyntax*, could be used to grasp the structure of the explanation of molecular *explananda* by taking into account causal relationships and their context-sensitivity. The deductions are representations of chains of biochemical reactions.

Note that in the case of molecular biology *Zsyntax* (analogously to classical logic in the case, for example, of nomological explanations in Newtonian mechanics via covering laws) is the formal language by means of which we can represent processes and, in such way, see their logical structure. Thus, as we will show in the following

sections, it does not describe or explain on its own, but allows for the description of the processual steps and, through this, for the explanation of the outcomes of these steps.

We are not arguing that the deductive-nomological approach and the received view are to be rescued as they stand, but, rather, that some of their features can be so-to-speak up-dated and set in a dialogue with more recent trends in the debate on explanation. In our proposal, the explanatory job is done by *Zsyntax*, which is one of many examples of logic of empirical theories. It has not been constructed asking what the logical structure of the molecular biology domain is per se, but, vice versa, asking lab scientists what the empirically valid biochemical reactions are. These have been considered the empirically valid formulas (EVF), which, then, have been logically connected – as we will see – via the more standard logically valid formulas (LVF). We interpret *Zsyntax* in a pragmatic way, that is, as a logico-*epistemological* tool to grasp the best current scientific understanding of the behavior of molecular processes.

Coming to the organisation of this paper, in Section 2 we will first mention some recent trends in addressing biological explanation, and embrace a conception of molecular pathways (processes) as chains of causal interactions. We then move on to consider a few hints from the debate on historical explanations, and the role regularities and sequences of events can play therein, to shed light on, more in general, processual explanations and their historical dimension. In Section 3, we will present *Zsyntax* as a logical language, showing how it can be used to provide an adequate philosophical account of molecular explanations in terms of non-classical deductions. In Section 4, we will stress how *Zsyntax* proves capable to properly grasp a number of explanatory aspects in molecular biology at once. More specifically, it allows us to: i) capture the historical-processual dimension of the molecular behaviour; ii) recognize an explanatory role to causal interactions; iii) grasp the e*xplanans-explanandum* relation in logical terms; and iv) acknowledge the context sensitivity of molecular behaviour. On such grounds, we will argue that the adoption of this non-classical logical language fosters some original rethinking of the philosophical discourse on biological explanations, allowing for some reconciliation and joint addressing of trends that in the theoretical debate on scientific explanation have been largely proceeding separately so far.

2 The historical-processual aspects of causal explanation in molecular biology

How can we effectively explain the occurrence of particular molecular outcomes? This is, indeed, the main question we wish to face: the "philosophical rationale" of the explanation of singular molecular events. We will try to answer it by appealing to reflections from different and apparently distant philosophical debates, bringing together, on the one hand, what we regard as the epistemologically most adequate way to approach molecular pathways, and, on the other hand, some portions of the debate on the structure of historical explanation. We will show how the latter can provide interesting hints to address biological (molecular) phenomena and the explanatory stance in molecular biology. More specifically, here we point out possible analogies

with reflections related to a "classical" work on explanation in history, namely William Dray's, and some later views.[2] Such reflections will lead us to suggest that a formal approach to explanation might have a number of advantages in jointly capturing different features of an explanatory account. Although parallels and comparisons between explanations in biology and in history have already been elaborated in the literature (e.g. Kaiser et al., 2014), what we suggest hereinafter presents some original aspects insofar as it specifically focuses on molecular biology (rather than, for instance, evolutionary biology, see, e.g. Rosenberg, 2001), and brings together historical-processual and logical-procedural aspects of biological explanations in a novel way, by means of *Zsyntax*.

Among biological phenomena, let us take here the end points of molecular pathways as our *explananda*. As pointed out in a few loci in the recent philosophical literature (e.g. Duprè, 2013; Ioannidis & Psillos, 2017; Nicholson & Dupré, 2018; Boniolo & Campaner, 2018; Ross, 2018, 2020), molecular pathways should be most properly conceived as processes that bring about certain effects by virtue of a number of intermediate steps, causally related to one another. The current debate on processes and pathways has partly expressed metaphysical stances – as in the case of works by Dupré and colleagues – and partly strictly epistemological ones, stressing the role such notions play in causal investigations, and, in particular, how they figure in explanatory accounts (e.g. Boniolo & Campaner, 2018; Ross, 2020). Our take on processes is the latter. In other terms, we believe that from an epistemological standpoint, molecular outcomes should be regarded as the end points of series of causally connected steps. In particular, we wish to recall that: 1) any pathway is the representation of *a series of causal actions among molecules*; 2) any pathway is the representation of a series of actions among molecules *that end in the performance of a particular function*; 3) any pathway is the representation of a number of actions among molecules, *where such a number can vary from* $n = 1$ (there is just one action leading to one molecular bond) *to* $n = m$ (there are m bonding steps, depending on the length of the pathway considered); 4) any pathway is the representation of a *temporally continuous process* of bonding steps. Moreover, as Boniolo and Campaner (2018) have stressed, a focus on pathways brings with it a shift in conceiving explanation as well: it requires a focus on the performance of a specific function, which is enacted by virtue of specific molecular set-up conditions, i.e., of a *specific molecular context*. For each pathway, "we can specify the set-up conditions as the molecules that initiate the first reaction in the pathway, and the terminating conditions as the molecules that are produced by the last reaction in the pathway" (Thagard, 2003, p. 238). Changes in the molecular context constrain the performance of the function which that specific pathway is meant to act out and, hence, our understanding of such a function. Thus, different contexts allow for different functional explanations of the behavior of the same molecule. In other terms, the specific context of instantiation of the pathway's acting conditions the epistemic validity of its explanation in functional terms. This perspective suggests that the notion of function and of the context in which it is performed are to be analyzed in depth – where the context could be thought of as defining the conditions of adequacy of the explanation itself. In this sense, it is then to be remarked that context-dependency

[2] For a survey and discussion of Hempel's view of historical explanation and the debate which followed, see Dray (2000).

constitutes a crucial aspect in the behavior of molecular processes, and that explanations in molecular biology will thus be required to take into proper account, at once, causal, functional and context-related features of molecular processes.

As just recalled, molecular pathways unfold in time, as a sequence of causally related and context-dependent steps. Duprè and Nicholson state: "It is always possible, it seems to us, and ultimately even necessary, to treat biological structures as *explananda* as well as *explanantia*" (Dupré & Nicholson, 2018, p. 32). Biological structure will hence figure as both what explains – in our case, a given sequence of molecular pathways – and what is to be explained – a given molecular outcome. What about the *explanatory relation* linking a given *explanans* to its *explanandum*? Widening the scope of the discussion, is there any approach which can help us shed light on the most adequate ways to describe and explain how molecular pathways behave? We already remarked above that philosophers of biology have often insisted that generalizations in biology differ from what are genuinely taken to be laws of nature, claiming, e.g., that they are restricted to a particular space-time region, or to a specific kind of entities, or strongly dependent on background and initial conditions.[3] As known, the explanatory role of generalizations and laws has been greatly discussed with respect to historical investigations as well. We believe interesting analogies can be drawn between, on the one hand, explanatory issues in molecular biology and, on the other hand, considerations on strategies to explain envisaged in the debate on historical explanation which have appealed to generalizations covering unfolding events and their relations. More specifically, it seems that some fundamental points highlighted by William Dray in his *Laws and Explanation in History* (1957), and by some following works in the debate related to his and Hempel's work,[4] can prove interesting to address temporal features of the structures of molecular pathways.

Although we cannot get in all the details here, let us stress some crucial matters on which explanatory stances in history addressed by Dray and reflections on molecular biology can be made to converge. Dray denied the relevance, but not the logical necessity, of nomological generalizations in historical explanations, and he did not deny that historical events could be explained causally.[5] Generalizations are logically necessary, but they are not as historiographically relevant as the causal conditions that have led to the event to be explained. "Historical explanations are inferences about the causes of specific outcomes in particular cases. They are intended to explain outcomes that have already happened, whether in the distant or in the recent past" (Mahoney et al., 2009, p. 116). Historical accounts, in other terms, appeal to sequences of linked causal factors, explaining a given outcome by a chain of connected events that have unfolded over a given amount of time. Investigating the "logic of historical explanations" amounts to showing how they "explain [outcomes] by tracing the sequence of events that brought them about" (Roberts, 1996, p. 16). Therefore, the historian's job does not concern the search for nomological generalizations in the first place, but for the right succession of the conditions that have caused the event to be explained. From

[3] Main arguments are recalled – and partly refuted – in Reutlinger (2014). For a survey, see also Hamilton (2007).

[4] On the related debate see also Gardiner (1974), Ankersmit (1986), Murphey (1994), Ankersmit and Kellner (1995).

[5] See Dray (1957, pp. 22 and ff.). With respect to the debate on nomological generalizations in the historical explanations, see Hempel (1965, 1942) and Mandelbaum (1961).

this point of view, the objection brought up, for example, by Joynt and Rescher (1961), according to which history is made up of unique events, loses its strength: in a historical explanation we emphasize the particular conditions occurring in that particular moment precisely because we want to explain an event which is unique. Dray pointed out that generalizations are not to be excluded, but, if the way in which a certain event or state of affairs occurred can be made explicit by generalizations, "it is very difficult indeed to account completely for the historical narrative in terms of the covering law model *alone*" (Dray, 1972, p. 25, italics added). If we focus our attention on the succession of causal conditions that have led to the event to be explained, we need to make clear how such a causal imputation is made.[6] Quoting from Dray again, understanding a given outcome "is very directly related to the fact that I can now trace the course of events by which it came about", and can hence "envisage a continuous series of happenings" (1957, p. 68). A story is to be told which represents a given set of events as an intelligible sequence, causally unfolding in time, and contingent upon given contextual conditions.

Once this point is taken into account, we can grasp what Dray's continuous series model amounts to: the reconstruction of the continuous succession of causal conditions that from an initial event, occurring at time t_0, have led to the final event, occurring at time t_f. The *explanandum* will not be accounted for just by a general covering law by itself. Rather, also the reference to a series of facts constituting the story of what happened between a given initiating event and the outcome will play an explanatory role with respect to the latter. And "even if it were true that these smaller scale events were each covered by law in the sense that in every case I would be prepared to assent to a law corresponding to a sub-sequence, the laws involved would be, at most, *part of the explanation* of the gross event" (ibid., p. 70, italics added).

At the same time, it can hardly be denied that generalizations exhibiting patterns of dependence do play a role.[7] As pointed out, much more recently, by Stuart Glennan (2010), an explanatory account can be deemed an historical "explanation" if it explains the occurrence of some particular event or state of affairs by describing how it came to be. If so conceived, it does not have to do only with the province of history as a discipline, but equally concerns both human and natural phenomena, and presents interesting common features across natural and human sciences by virtue, also, of the appeal to generalizations. Starting from the idea that the mechanistic approach can be stretched to prove adequate for singular causal sequences, Glennan stresses how historical explanations are peculiar insofar as

> "the circumstances that bring together the various entities whose interactions constitute the narrative [leading to the explanandum] are ephemeral. The

[6] This would open, in turn, a wider reflection that it is not possible to address here. Contemporary discussion of the topic may be dated back to Weber (1906); cf. Boniolo (2001).

[7] See also Nagel (1952), who, while stressing some peculiar features of explanations in history, also claims: "there appears to be no good reason for claiming that the general pattern of explanations in historical inquiry, or the logical structure of the conceptual tools employed in it, differs from those encountered in the generalizing and the natural sciences. The explanatory premises in history, as in the natural sciences, include a number of implicitly assumed laws, as well as many explicitly (though usually incompletely) formulated singular statements of initial conditions. The tacitly assumed laws may be of various kinds" (p. 163).

> difficulty of predicting the future course of history stems from the dependence of historical outcomes on chance conspiracies of circumstances, but once we have identified what those circumstances are, a good narrative explains by showing how, given those circumstances, there was a likelihood or necessity to the outcome" (2010, pp. 259-260).

While the configuration of the parts/events may be "the product of chance or exogeneous factors", still the way in which they interact can be characterized by generalizations which exhibit some recurrent patterns.

In our terms, even if the manner in which parts come together can be strongly accidental and context-dependent, the way in which they then interact can be captured by robust and reliable generalizations – which are taken by Glennan as somehow "reminiscent of Hempel's requirement that the links in narrative chains can be explained by appeal to covering laws" (ibid., p.261). Tackling the logic of historical explanation, and discussing the debate on Hempel's view, Roberts (1996) has argued that the covering-law model can be rescued when referring to the series of discrete events bringing about the macro ones. Covering laws are thus believed to play a role in an explanatory narrative, but one that needs to be elucidated in terms of the notion of "colligation", that is, historians explain the "occurrence of [historical events] by tracing the sequence of events that brought them about" (Roberts, 1996, p. 16). Roberts, hence, claims that his notion of colligation is analogous to Dray's model of the "continuous series". To overcome ambiguities and limits of Dray's – and others' – views, he suggests that Fig. 1 is capable to portray "the tracing of the causal connections between events" (ibid., p. 20) as a "colligatory explanation", with many subsequent steps.

In Fig. 1, L stands for law, C for conditions, and E for the event expressed by the *explanandum* sentence. Inferences are drawn from the law and conditions to the explanandum: the vertical line indicates the inference that the conditions to the left of the line *caused* the event to the right. For the purpose of our reflections here, what is to be stressed is that the arrow indicates that a given event becomes then a condition in the next explanans – while, as it will be stressed below, empirical generalizations take the place of Ls.

What tentative lessons can be drawn? Let us now address a few respects in which our focus on molecular biology can be related to the discussions of historical features of explanations recalled above. If we want to explain a molecular event

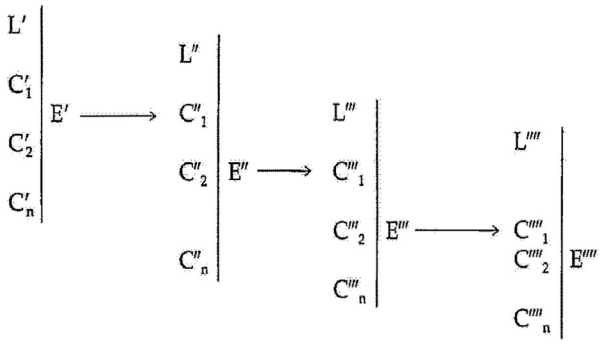

Fig. 1 From Roberts (1996, p. 22)

occurring at a time t_f (i.e. the *explanandum*) we must reconstruct the whole story that has led to that event by identifying the continuous chain of causal reactions, that is, the process, starting from a given initial time t_0, up to it. An example can better illustrate how such a continuous chain is unravelled. Suppose that we want to explain why in a cell at certain time t_f we have the degradation of the protein TP53, which is one of the most important cellular oncosuppressors. In other terms, the degradation of the protein TP53 at t_f is our *explanandum* (call it d(TP53)). We have to *reconstruct* the entire *process* that, starting from a certain set of molecules at an initial time t_0, has led to that final result. From the empirical world, that is from the lab, we know that: i) d(TP53) is the result of the causal interaction, at a certain time $t_{f-1} < t_f$, between the compound realized by TP53 and the protein Ubiquitin (U) (for simplicity we write this compound TP53*U) and the proteasome (P); ii) the compound TP53*U is the result of the causal interaction, occurring at a time $t_{f-2} < t_{f-1}$, between the compound realized by the protein MDM2 and TP53 (let us write this as MDM2*TP53) and the protein U; iii) the compound MDM2*TP53 is the result of the causal interaction, occurring at a time $t_{f-3} < t_{f-2}$, between the protein MDM2 and the protein TP53; iv) the compound MDM2 is the result of the causal interaction, occurring at a time $t_{f-4} < t_{f-3}$, between the gene *MDM2* and the protein TP53.[8] At this point, that is, after having reconstructed the entire process starting from an initial time t_{f-4}, we have the *explanans* of the degradation of TP53 occurring at a certain t_f. For simplicity, we could rewrite the causal chain getting from the causes to the effects of the causal interactions, and represent the process as follows:

$$t_0 - \mathrm{TP53}\,\&\,\mathit{MDM2} \rightarrow \mathit{TP53}\,{*}\,\mathit{MDM2}$$

$$t_1 - \mathrm{TP53}\,{*}\,\mathit{MDM2} \rightarrow \mathrm{MDM2}$$

$$t_2 - \mathrm{MDM2}\,\&\,\mathrm{TP53} \rightarrow \mathrm{MDM2}\,{*}\,\mathrm{TP53}$$

$$t_3 - (\mathrm{MDM2}\,{*}\,\mathrm{TP53})\,\&\,\mathrm{U} \rightarrow (\mathrm{MDM2}\,{*}\,\mathrm{TP53})\,{*}\,\mathrm{U}$$

$$t_4 - (\mathrm{MDM2}\,{*}\,\mathrm{TP53})\,{*}\,\mathrm{U} \rightarrow \mathrm{MDM2}\,\&\,(\mathrm{TP53}\,{*}\,\mathrm{U})$$

$$t_5 - (\mathrm{TP53}\,{*}\,\mathrm{U})\,\&\,\mathrm{P} \rightarrow (\mathrm{TP53}\,{*}\,\mathrm{U})\,{*}\,\mathrm{P}$$

$$t_f - (\mathrm{TP53}\,{*}\,\mathrm{U})\,{*}\,\mathrm{P} \rightarrow \mathrm{d}(\mathrm{TP53})\,\&\,\mathrm{U}\,\&\,\mathrm{P}$$

[8] Please note that MDM2 indicates a protein, while *MDM2* (italics) a gene.

where $t_n < t_{n+1}$; & means that there is the presence of two molecules (both as molecules participating in the reaction, and as molecules resulting from a reaction); * means that the two molecules form a compound; → means that there is a causal interaction which allows to get from what is indicated at its left to what is indicated at its right.

At this point it is easily recognizable that it is the *whole* historical process represented above (the *explanans*) that allows us to elaborate the explanation of the degradation of the TP53 (the *explanandum*). But what kind of logical structure does this historical-processual explanation have?

3 Zsyntax and the logical-procedural aspects of causal explanation in molecular biology

We have seen that a processual explanation in the case of molecular biology is a historical explanation, and have dwelled on some of its features through what has been suggested by Dray and others. Now we should address its procedural structure, and inquire if there is a logical language capable of formally representing the process leading to the *explanandum*. If we succeed, we have also a logical-procedural explanation which, on the one hand, grasps the historical-processual aspect of the explanation – described, step by step, in the *explanans* - and, on the other hand, allows us to understand the relation holding between the *explanans* and the *explanandum* in logical terms.

Notwithstanding, as emphasized, some reluctance in philosophy of biology to accept formal tools, in 2010, a logical language for molecular biology was proposed. It has been considered and adopted as a framework for formal reasoning by some biologists (see Ahmad et al., 2014a, b, 2015; Marchi et al., 2021), and had some positive feedbacks also from the community of logicians and philosophers of science (see D'Agostino et al., 2014). This language, called *Zsyntax*, belongs to a family of non-classical resource-aware systems, called "substructural logics".[9] *Zsyntax* has been elaborated to represent types of biological processes – which can be as long as needed by the analysis at stake – as logical theorems, by taking into account context-sensitivity as well. Initially it was thought especially to deal with text mining (see Boniolo et al., 2010), then improved in certain aspects concerning non-monotonicity (see Boniolo et al., 2013, 2015, 2021), and finally expanded and provided with an automated theorem prover (see Sestini & Crafa, 2018). Now, given such developments, it is meant as the core of a research program arguing for the place logic should have as the formal counterpart (and representation) of what occurs at the empirical level in molecular biology.

It works thanks to three logical operators: Z-interaction (denoted by the symbol *), Z-conjunction (denoted by the symbol &) and Z-conditional (denoted by the symbol →).

[9] This family has received a good deal of attention in the field of computational logic; see, for example, Girard (1987), Dosen and Schroeder-Heister (1993), and Restall (2000, 2008). On the resource logic, see Hoare (1985) and Lafont (1993)

Z-interaction is a binary operation that is defined only for pairs of molecules A and B that form a compound. In general, the operation * *is not associative*, since it may happen that the compound (A*B)*C is different from the compound A*(B*C). For, although A interacts with B, and the resulting product A*B interacts with C, it may not be true that B interacts with C, and so neither B*C nor A*(B*C) exist. For example, in the case of the *Trp Operator* of *E. coli*, the Trp-repressor does not bind to the Operator if it is not bound to Tryptophan. In other words, (Tryptophan*Trp-repressor)*Operator ≠ Tryptophan*(Trp-repressor*Operator), the latter being a condition that does not exist. Z-interaction thus allows us to grasp the causal interactions thanks to which compounds are formed.

Z-conjunction is a binary operation. Thus A & B denotes that the two molecules A and B are considered together, that is, they are an aggregate. The molecules in an aggregate do not necessarily react. However, if A and B are types of interacting molecules, any aggregate of type A & B, under suitable bio-physical conditions, and given an adequate amount of time, will yield a compound molecule of type A*B. It should be noted that, differently from the classical conjunction, the Z-conjunction, while commutative, *is not idempotent* - that is, A & A is not the same as A. This allows us to represent the fact that, for example, an aggregate with two Adenosine Trisphosphate (ATP & ATP) is not the same as one molecule of ATP.

Z-conditional It is indicated by → and represents the transition, usually due to a causal interaction, between an initial aggregate of molecules A&C, and a final aggregate B. For example, if the initial aggregate is given by the gene *MDM2* and the protein TP53, the causal interaction between them gives the molecular compound *MDM2*TP53. This is represented by *MDM2*&TP53 → *MDM2*TP53.

Given the syntax of formulae, *Zsyntax*'s logical system provides a set of logical rules that can be used to distinguish valid formulae from invalid ones. We do not recall here the deductive system of *Zsyntax*, we just remark the *Empirically Valid Formulae* (EVF). These represent causal interactions, and their validity depends only on empirical information acquired in the laboratory. Examples of basic EVFs include: (i) two molecules that interact giving a new molecular compound (e.g. *MDM2*&TP53 → *MDM2*TP53); (ii) two molecules that interact to deliver a molecular product (e.g. D-Glucose-6-phosphate*Glucose-6-phosphate isomerase → D-Fructose-6-phosphate), or to deliver a molecule encoded by a gene (e.g. *MDM2*TP53 → MDM2; where MDM2 is the protein encoded by the gene *MDM2* after its interaction with the protein TP53). The validity of the EVF is, obviously, *context-dependent*, that is, they are valid ceteris paribus. Note again that EVFs describe the unfolding of the behaviour of causal processes discovered through lab work and really occurring in the cells. For example, the formula (ATP & ATP) → (ATP * ATP * ATP) is not empirically valid, since there is no real causal process of this kind in any cell.

Beside the EVFs, in Zsyntax there are the *Logically Valid Formulae* (LVF) which govern the correct inferences. We do not dwell here on them (they are discussed at length in the original papers on *Zsyntax*), but, for our aims, it is worth recalling the general pattern of an inference rule that allows for the transition from an aggregate of type A to another aggregate of type B:

$$\text{If } S, \text{then } \frac{\Delta, A}{\Delta, B}$$

Here Δ stands for the molecular context in which the reaction is supposed to take place. S, instead, stands for the aggregates of molecules, present also in Δ,that could block the reaction. Thus it allows us to highlight the sensitiveness to the context. Differently said, if S is empty, then the reaction occurs (that is, in the context Δ there is no interfering molecules); if S is not empty, then in the context Δ there are molecules that impede that reaction. To illustrate this rule, let us consider the synthesis of ATP consisting in the transitions from the adenosine diphosphate (ADP), the inorganic phosphate (Ph), and the enzyme catalyzing the reaction (ATPsynthase) to the adenosine triphosphate (ATP). In formal terms:

$$\text{If } \{\text{Oligomycin}\} \notin S, \text{then } \frac{\Delta, \text{ATPsynthase\& ADP\&Ph}}{\Delta, \text{ATP}}$$

Or, in the more explicit form,

$$\text{If } \{\text{Oligomycin}\} \notin S, \text{then } \frac{\Delta, \text{ATPsynthase\& } ADP\&Ph, (\text{ATPsynthase\& } ADP\&Ph) \rightarrow \text{ATP}}{\Delta, \text{ATP}}$$

In this formula, which is clearly an instance of the logical rule of *modus ponens* (we can call it the *Z-modus ponens*), ATP is the outcome of a series of causal interactions starting from the aggregate (ATPsynthase & ADP & Ph). Moreover, the condition S allows us to take into account the very fact that there are inhibitors of the ATP synthesis, like Oligomycin (typically used as an antibiotic), which bind the ATPsynthase, thus preventing the synthesis. Therefore, in the rule above, S encodes the information telling us that the reaction occurs if there is no Oligomycin in the context represented by Δ. Thus, if it is empirically true that the causal interaction among ATPsynthase, ADP, Ph causes ATP (that is, if the following EVF is valid: ATPsynthase & *ADP* & *Ph* → ATP); if it is empirically true that in that particular molecular context there are the ATPsynthase, ADP, and Ph and there is no Oligomycin, then it is empirically true that we obtain ATP.

One of the main arguments against the use of logic in (molecular) biology was its incapacity to grasp the variety and the contextualization of the (molecular) biological events and processes. *Zsyntax* has been constructed exactly to successfully cope with this issue. The *Zsyntax* formalization allows for the contextualization of any EVF. In (molecular) biology what is relevant, from an epistemological point of view, is not the "universality" of the empirical generalizations (that is, the EVFs), but the "universality" of the context in which the empirical generalizations are valid. With a motto, we can say that: same contexts, same EVFs; different contexts, different EVFs. We will get back shortly to context-sensitivity; let us focus here on the fact that biological processes

can be represented as formal theorems of *Zsyntax*. For example, the process, already mentioned, leading to the degradation of TP53 is represented by the following deductive theorem: TP53 & TP53 & *MDM2* & U & P ├ d(TP53); where ├ indicates the deduction from an initial aggregate (IA), composed of all the molecules that have to be present in the cellular environment to instantiate the process (that is, the gene *MDM2*, the protein TP53, the protein U, and the protease P), to the final product, that in this case is the degradation of TP53 (note that, as before, *MDM2* - in italics - is the gene and MDM2 – in roman - is the protein encoded by that gene). That is,

Theorem

TP53& TP53& *MDM2*& U& P ├ d(TP53)

Demonstration

1. TP53 & *MDM2* → TP53**MDM2*
2. TP53 * *MDM2* → MDM2
3. MDM2 & TP53 → MDM2*TP53
4. (MDM2 *TP53) & U → (MDM2*TP53)*U
5. (MDM2*TP53)*U → MDM2 & (TP53*U)
6. (TP53*U) & P → (TP53*U)*P
7. (TP53*U)*P → d(TP53) & U & P
8. d(TP53)

Actually, for the sake of simplicity, this is the abridged version of the theorem. Only the EVFs figure, and not the LVFs, have allowed the passage from one EVF to another (the complete form of the theorem can be found in the Appendix, where the role of the LVF is clearly shown). However, this is exactly the formal representation of the biological process under analysis.

Two points are worth noting here. The first concerns the fact that each EVF in the demonstration above works *in the proper molecular context*, that is, each one is *context-dependent* and for each one our Z-*modus ponens* works. For example, in order to show the context dependency, the third step can be written as

$$\text{If } S, \text{ then } \frac{\Delta, \text{MDM2\&TP53}, (\text{MDM2\&TP53}) \rightarrow \text{MDM2*TP53}}{\Delta, \text{MDM2*TP53}}$$

Said differently, any EVF is an empirical generalization representing a real molecular reaction, and this empirical generalization is valid only in the proper molecular context Δ where there is no S. While Δ can hardly be specified, since it has to do with the whole inter- or intra cellular environment, S can be specified considering what we empirically know about the possible molecules whose presence in the environment can hinder the reaction - as we have above illustrated with the synthesis of the ATP, which can be hindered by the presence of Oligomycin.

The second point regards the fact that Roberts' graphical representation of the *role of colligation* in historical explanation (see Fig. 1) can be easily reformulated, mutatis mutandis, in terms of *Zysntax*, as may be seen in Fig. 2, which adjusts it to the process/theorem just considered.

In Fig. 2, the vertical lines indicate the *causal interactions* (allowed by the EVFs working in the proper set of conditions Δ) getting from what is at the right to what is at the left. The arrows indicate the transition from one step to the other. If, instead, one prefers the formal interpretation, the vertical lines can be easily taken to designate the *logical inferences* (always allowed by the EVFs) from what is at the left to what is at the right. Of course, the entire sequence of steps is in time and, therefore, the entire process is historical. *Zsyntax* allows us to explain one molecular outcome at a time, on the basis of both an understanding of general behaviours of molecular processes, and of the initial conditions and the contingent features of the molecular environment on which the outcome at stake depends. Here above, we have shown that any Z-conditional is the formal representation of a causal interaction. Moreover, we have argued that the demonstration of the thesis of the theorem (which represents the final product of the process) is the formal representation of the steps of the process, up to the *explanandum* (e.g., the degradation of TP53 at a given moment in time). This means that, while the transitions from one molecular step to another give a *historical-processual explanation*, their formal representation via *Zsyntax* provides us also with a *logical-procedural explanation*. That is, thanks to *Zsyntax*, we can hold that any historical-processual explanation in molecular biology has a non-classical deductive procedural explanation as its formal counterpart. In other terms, this can be taken to mean that *Zsyntax* permits to show the non-classical logical structure of a molecular biology explanation.

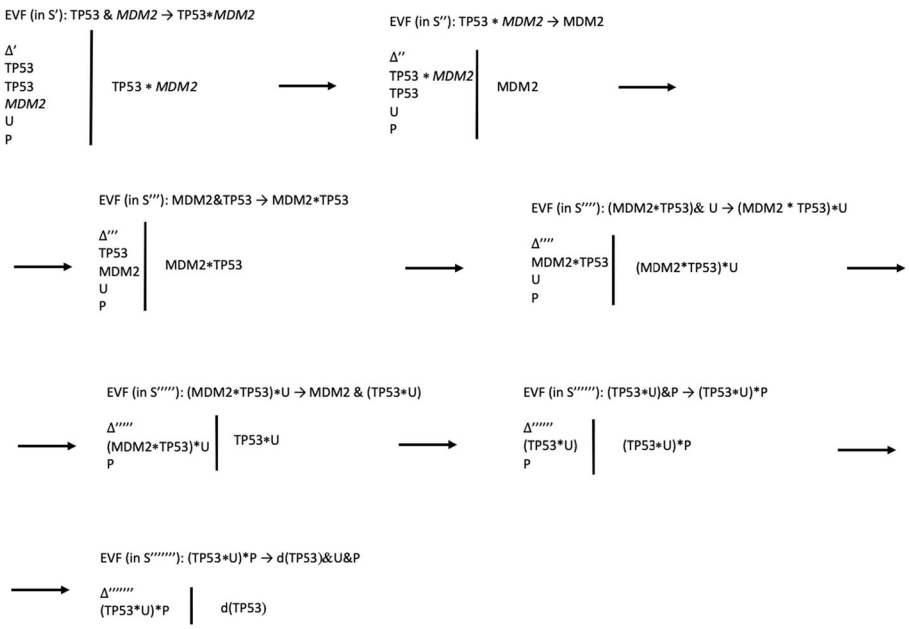

Fig. 2 Reformulation, in terms of Zsyntax and in the context of molecular biology, of Roberts' graphical representation of the role of colligation in historical explanation

The biological processes constituting the *explanans* can be straightforwardly represented as formal theorems of *Zsyntax*, as can the explanatory relation holding between the *explanans* and the *explanandum*.

This encourages us to put some emphasis on the possibility to appeal to formal tools to philosophically grasp what explaining amounts to in molecular biology. On the one hand, recurring to *Zsyntax* allows us to retrieve the idea, starting from Neopositivism and conveyed by the received view, that logic is entitled to play a crucial role in the unraveling of the structure of scientific investigations – more specifically here, *explanatory* scientific investigations, including their causal and temporal dimensions. However, such retrieval needs to be qualified. On the other hand, in fact, the way in which we suggest *Zsyntax* is to be employed for explanatory purposes stands *against* two features characterizing the received view on explanation – that is: i) the fundamental explanatory role played by *genuine* scientific *laws*, and ii) the exclusion of contextual elements.

With respect to i), as sketchily remarked in introducing our work, there is a strong position in the philosophy of biology arguing that there are no *biological laws strictly speaking*, since – it is argued –statements concerning what occurs in *biological* phenomena do not share the very same features with those statements describing *physical* phenomena that are labeled as genuine "laws". Without getting deeper in the debate, let us stress that, as already pointed out above, that, however, it is undeniable that empirical generalizations (as the EVFs witness) hold which govern the behavior of the molecular processes, and that present some degree of stability. As also stressed in the literature,

> "many philosophers and biologists would agree that scientific understanding requires general statements of an empirical nature regardless of whether these statements fit all the criteria traditionally attributed to 'laws'. Yet, once philosophers decided that biology lacked genuine laws, they seem to have lost interest in analyzing the empirical generalizations of the science. Meanwhile, biologists continue to generalize. […] Surely, empirical generalizations, even if not formulated as true universal statements, play important roles in the scientific investigation and understanding of the biological world" (Waters, 1998, p. 6),[10]

and "biological explanations invoke empirical generalizations that refer to causal regularities exhibited by various kinds of biological entities" (ibid, p. 8). An emphasis is put on the fact that what counts from an explanatory point of view is not the universal form of the generalizations, but their describing the distribution of a given set of biological entities, and some causal relation holding between them. In this perspective, it can be easily argued that *Zsyntax* provides us with a tool to address how they can be

[10] Waters, for instance distinguishes between *two types of empirical generalizations*. Those of the first type are "historically-based contingencies which represent current or former *distributions* of biological entities of various kinds". The second type of generalizations "presupposes the existence of *causal regularities*", and are clearly identified by "explanatory and investigative practices" (ibidem, italics added). Both kinds of generalizations can happen to be expressed by a single sentence, which will tend to be especially employed for explanatory purposes.

effectively represented, and their role properly grasped, in formal terms, since the EVFs aren't but empirical generalizations.

With respect to ii), it is worth stressing that *Zsyntax* allows us not only to adequately represent molecular transitions, which can be expressed by empirical generalizations, but also to capture a further crucial aspect of real molecular processes, that is, their being dependent on the context: the EVFs are context-sensitive empirical generalisations, that is, they are valid given a range of specified circumstances. Far from being a marginal aspect, context-sensitivity is central, since all biochemical reactions and biochemical processes always occur in a specific molecular environment (see Boniolo et al., 2021). *Zsyntax* is able to grasp this feature, for example through the *Z-modus ponens*, showing that the validity of the empirical generalizations (the EVFs) depends on the particular molecular contextual conditions (given by Δand S) that realize the environment in which the latter take place.

4 *Zsyntax*: A way to an integrated approach to explanation in molecular biology

What has been illustrated above envisions *Zsyntax* as an effective tool to formally account for a range of explanatorily relevant features, and to build bridges between different stages the philosophical debate on scientific explanation has gone through. As anticipated in the Introduction, the adoption of *Zsyntax* for explanatory purposes presents at least four simultaneous advantages.

In the first place, it captures the historical-processual dimension of molecular behaviour. In line with other recent contributions in the literature in philosophy of biology and medicine, we have argued that, in order to explain the end points of molecular pathways, the latter are most adequately conceived as processes bringing about given effects in a number of steps unfolding in time. The molecular outcomes we aim to explain aren't but the end points of such steps. In the second place, the explanatory account we have presented recognizes the explanatory role of causal interactions: the series of intermediate steps in molecular pathways leading to a given outcome are causally related to one another. Any pathway is the representation of a series of causal interactions among molecules performing a given function, jointly responsible for a certain molecular end point. At the same time, the perspective suggested here brings to the fore the fact that from a philosophical standpoint an (appropriate) logical language can grasp explanatory relations between phenomena occurring in molecular biology. Moreover, *Zsyntax* constitutes a logical tool capable of capturing the relation holding between some initiating events, the sequences of following ones, up to a certain end point which is taken as the explanandum. Finally, as the examples above clearly show, the logical form of such explanations does not deny – but, rather, allows us to emphasize – other characterising features of molecular biology. It is in this respect that we have pointed out how *Zsyntax* allows us to acknowledge the context-sensitivity of molecular behaviour: while representing molecular pathways in historical-processual terms, and logically grasping the explanans-explanandum relation, the present account also stresses that the performance of a given function by a series of causally related steps is enacted by virtue of specific conditions – the molecular

behaviour being highly context-sensitive. The performance of a given function by a given pathway is constrained, in other terms, by a given molecular context.[11]

Considering all of this, *Zsyntax* can be employed for genuine *explanatory* purposes, and (re)establish logic (in this case, non-classical logic) as playing a core role in representing what explanation is about, while not excluding other, non-logical, elements from the explanatory picture, such as the genuinely causal interactions figuring in the *explanans* and their context-sensitivity inside a historical-processual account.

In sum, in line with a few aspects stressed in Dray's and others' positions on historical explanations widely conceived, *Zsyntax* suggests that particular molecular outcomes can be explained in terms of the unfolding in time of sequences of causal interactions, which depend on contextual conditions and whose overall behaviour can be described by empirically-based generalizations (the EVFs). Moreover, the occurrence of the *explanandum* is logically derived from the formal representation of such relations. *Zsyntax* is thus capable of combining the purely logical features – namely, tautological statements and truth-preserving inference laws - and the biological content – namely, extra-logical information coming from the laboratory. As the Hempelian tradition would have had it, the empirical content of the *explanans* is warranted, together with the logical derivation of the *explanandum*. At the same time, and setting some distance with such tradition, contextual aspects are included, insofar as *Zsyntax* acknowledges that all biochemical reactions and pathways cannot occur as they do but in a specific molecular environment.

Let us also consider three possible plausible objections. First of all, somebody could wonder whether there really is a strong substantial analogy between historical processes and molecular processes. Actually, we believe that this needs to be qualified. Single historical processes present unique features, are highly contingent and usually conjecturally re-constructable only a posteriori. Philosophical views which sustain that there are (more or less general) historical laws will also have to admit that such laws hold only with respect to a coarse-grained description of historical processes. Differently, the molecular processes inter or intra cells are not unique in exactly the same sense, since they re-happen any time the proper molecular and energetic conditions allow them to[12]; they are not contingent in the same sense, since any time we have the same proper molecular and energetic conditions, the same process occurs; and usually they can be traced a priori*,* since if we know a priori the proper molecular and energetic conditions, then we know how that process is going to develop in time. Any historical process is a type instantiated in a token a single time and in a unique manner, whereas any molecular process – generally speaking – is a type which can be instantiated in as many tokens as the proper molecular and energetic conditions allow.[13] Therefore, when we speak about an analogy between historical processes and molecular processes, what we mean is that both are characterised by a given set of temporal and causally-related steps. Differently said, the analogy lies in the temporal and causal structure, and - of course - not in the content, nor in their peculiar features.

[11] What proposed above is an account of explanation of singular molecular outcomes, and not of empirical generalizations, that is, of EVFs. Actually, the latter too could be explained (that is, they could be *explananda*), but their explanation should receive a totally different discussion.

[12] This means that they are strongly context dependent.

[13] This, of course does not mean that we share the idea that a cell, or a population of cells, could be thought of as a machine (see Nicholson, 2019).

Second, why should we use logic to represent molecular processes? There has been some reluctance for quite some time to recur to logic in the field of biology. But, setting some distance from the scepticism which many attempts to introduce logic into biology in the past decades received,[14] we should be ready to acknowledge that now, in the age of precision medicine (with its use of big data coming from new sequencing and imaging technologies, and from clinical and lifestyle records), a deep change of perspective on the matter is progressively taking place. There is, indeed, an almost hidden but extremely important role for logic in computational biology and bioinformatics. To become aware of the enormous and extremely relevant position that logic currently has, it suffices to think about the deep philosophical and technical relevance of the Curry-Howard isomorphism, which states a correspondence between a well-defined class of computer programs (essentially those developed by means of functional programming languages) and logical proofs or, more generally, between logical systems and models of computation like typed λ-calculus and related variants (see Curry, 1934; Howard, 1980). Thus, considering the actual modes of advancement in scientific practice goes hand in hand with recognizing that logic is highly relevant to biological and biomedical sciences, in particular whenever they must face issues stemming from molecular research. It is hence high time that the philosophers of the life sciences rethink the role of logic in this domain, investigating, among others, which formal tools could best capture a number of biological relations. Our aim here is not to tackle such a wide task in general, but, more limitedly, to focus on molecular biology (or molecular biomedicine), and present *Zsyntax* as one of the possible logical languages to be adopted in the field to address explanatory stances. There is a new place for logic related to the rising role of computational approaches in the empirical sciences, in particular in molecular biology and biomedicine. It is in this spirit that in what follows we draw attention to the core features causal explanations in molecular biology exhibit, to then suggest a unified way to address them.

Third, does appealing to logic in dealing with explanatory accounts amount to reviving an aspect of a philosophical tradition – like the Neopositivist one - which has given an enormous emphasis on logic, but which belongs to the past? Belonging to the past does not necessarily mean being obsolete. Moreover, the Neopositivist and Post-Neopositivist studies on the logic of explanation have shed a lot of conceptual light on how sciences explain, and this is clearly still a direction worth pursuing. As we have shown, we believe that some of the features exhibited by DN should be preserved into an account of explanation, and conveyed through *Zsyntax*. The debate has often overemphasized the limits of the deductive-nomological explanation debate, which still provides interesting hints as to what explaining amounts to. We are not suggesting that the deductive-nomological approach should be rescued in its original form, but that some of its aspects still make sense when addressing explanation of biological pathways. Among such aspects, we especially want to stress the explanatory role generalizations have, and the idea that explanatory procedures rely on the recognition of a logical relation holding between the *explanantia* and the *explananda* – such that the latter can be inferred from the former. In other words, we believe that explanations are to be understood by appealing to a logical structure, holding by virtue of nomological

[14] Recall, for example, the dismissive comments wrote by Ruse (1975) and Hull (2000). See also Woodger (1937) and Nicholson and Gawne (2014).

relations, which in this case happen to be causal. We do not think it is possible just to rescue DN as it stands, since, e.g., it does not assign any fundamental explanatory role to causal interactions, does not include any temporal requirement, does not account for context-sensitivity, and paves the way for a unificatory account – which we are not advocating. Our aim is not to revive the deductive-nomological model per se, but to reposition a range of elements in the philosophical discourse on explanation, treasuring, on the one hand, current lessons from the zooming in on specific cases, and, on the other hand, lessons from past approaches which can now be seen through different lenses.

A wider consideration of the philosophical debate on explanation requires to clarify where this view stands in the epistemic/ontic conceptions debate. Hempel's account is strictly related to epistemic procedures – so much so that the received view is included in the so-called "epistemic conception" of explanation. This is not the spirit of our proposal, which captures features of biological phenomena that are independent of our getting to know them or not. As for the relation between explanation and prediction, in our view an adequate understanding of the contextual conditions allow us to appeal to the proper generalization to predict what will happen next. The occurrence of the effect will be the result of a series of causal steps unfolding in time, which the formal tool is able to express. Hence, some of the features of deductive-nomological approach are still worth recalling, without the need to embrace the entire epistemic conception of explanation it belongs to. What we want to emphasize are the roles it assigns to inferential logical procedures and the appeal to laws, that also happen to be among the features that have been criticized the most by the debate on explanation which has followed the so-called received view. We do not believe that our proposal is to be inserted in the purely, "old", epistemic conception. It is rather aimed at grasping actual molecular features. In this sense, it can amount to a contribution in the ontic field. However, the debate is not as clear-cut as it was back in the Seventies-Eighties, when an ontic position *à la* Salmon would claim that explanations are existent *entities in re* (Salmon, 1984; see also Wright, 2015). A number of positions that hybridize the two poles of the debate, in different ways, have been put forward (see, e.g., Illari, 2013; Wright, 2015; van Eck, 2015; Sheredos, 2016). We agree with the view according to which both stances – the epistemic and the ontic – are present in explanatory accounts, where the latter, though, constrains the former. To be adequate, an explanation needs to somehow get things right, which cannot depend just on epistemic norms (see also Illari & Williamson, 2012). At the same time, the world does not explain itself: it is cognitive subjects who engage in explanatory activities and elaborate explanatory accounts. In other words, questions about the adequacy of an explanation are largely settled by features of the world, but to elaborate and grasp an explanation are epistemic procedures. *Zsyntax* is meant to grasp objective molecular processes, but, as a logical tool, it has been elaborated by a group of researchers; as such it is the product of a cognitive – epistemic – process. In sum, it can be conceived as a pragmatic tool to grasp our best current understanding of the actual features of some sets of molecular processes.

Our proposal addresses a specific field, and, in this sense, it works in the spirit of large portions of the recent and current debates on scientific explanation, where the stress has been put on scientific practice and on the dependence of an explanatory account on the specific object, field and purpose of the explanatory investigation. We are not arguing here that *Zsyntax* can provide a valuable answer to *any* explanatory

issue in any disciplinary context, nor are we excluding that it can be applied elsewhere too. We are exhibiting a context in which it can be fruitfully employed.

As for any explanatory procedure whatsoever, the adequacy of the explanatory account which is proposed has to be evaluated with respect to the exact explanatory question it is meant to address. *Zsyntax* is here proposed to tackle a specific explanatory question, which has to do with linear molecular processes, and what is stressed is that this formal tool – whose relevance for the sciences has been already discussed elsewhere (see, e.g., Boniolo et al., 2010; Marchi et al., 2021) – can be interpreted as connectable in different ways to various lines of research – older or more recent – elaborated in the philosophical debate on explanation. This is obviously not to say that biological systems cannot be addressed with different computational tools. Features of complex systems and biological organization are largely investigated by means of computational tools, that are simply not the object of this paper. Nevertheless, coming to feedback processes (more generally to control processes) and networks biology, it is to note that very recently some papers have been published exactly on these aspects showing how *Zsyntax* is able to capture them too (see Boniolo et al., 2021; Boniolo & Lanfrancone, 2016). In sum, the adequacy of the proposal put forward in this paper is to be measured with respect to the *exact job* that *Zsyntax* is called to play here.

In short, *Zsyntax* is an example of how a logical framework can be constructed and used for explanatory purposes, thus partly rescuing some arguments in favor of the place of logic in the philosophy of scientific explanation. Further on, this retrieval of the place of logic is inextricably accompanied by a proper emphasis on empirical constraints on explanatory modeling and scientific practice, perfectly in line with the most recent developments of philosophical reflections on explanation.

Appendix

In the text we have seen the abridged version of the theorem TP53 & TP53 & *MDM2* & U & P ⊢ d(TP53), in which only the EVFs are present. Actually, as cursorily said, Zsyntax has also a set of LVFs. They are *formally* valid, and their validity depends only on the definitions of the Z-operators, regardless of the molecular context. Some of these are similar to the correspondent classical ones and give the rules governing the transition from one EVF to the next one. In particular, they are:

- *Elimination of the Z-conditional* (that we indicate by $\rightarrow\boldsymbol{E}$). If A → B can be derived from C and A can be derived from D, then B can be derived from C&D.
- *Introduction of the Z-conditional* (that we indicate by →I). If B can be derived from C&A, then A → B can be derived from C alone.
- *Elimination of Z-conjunction* (that we indicate by $\&\boldsymbol{E}$). If the Z-conjunction of A and B (A&B) can be derived from C, then both A and B individually can be derived from C.
- *Introduction of Z-conjunction* (that we indicate by $\&\boldsymbol{I}$). If A can be derived from C, and B can be derived from D, then the conjunction of A and B can be derived from C&D.

It is to note that there are no logical rules for the Z-interaction *, since its behavior depends on empirical information acquired in the laboratory. Therefore, * cannot be governed by *formal* rules, but rather, it can be introduced and eliminated only via EVFs of, respectively, the form A&B → A*B (representing that we obtain a compound from initially separate molecules) and A*B → C&D (representing that we obtain two products from the division of an initial compound).

Once stated this, the complete demonstration of the mentioned theorem is the following:

Theorem

TP53& TP53& *MDM2*& U& P ├ d(TP53)

Demonstration

1. TP53 & TP53 & *MDM2* & U & P	Initial Aggregate
2. TP53	From 1 by &*E*
3. TP53	From 1 by &*E*
4. *MDM2*	From 1 by &*E*
5. U	From 1 by &*E*
6. P	From 1 by &*E*
7. TP53 & *MDM2*	From 2,4 by &*I*
8. TP53 & *MDM2*→TP53**MDM2*	EVF
9. TP53**MDM2*	From 7,8 by →*E*
10. TP53**MDM2*→MDM2	EVF
11. MDM2	From 9,10 by →*E*
12. MDM2 & TP53	From 3,11 by &*I*
13. MDM2 & TP53→MDM2*TP53	EVF
14. MDM2*TP53	From 12,13 by →*E*
15. (MDM2*TP53) & U	From 5,14 by &*I*
16. (MDM2*TP53) & U→(MDM2*TP53)*U	EVF
17. (MDM2*TP53)*U	From 15,16 by →*E*
18. (MDM2*TP53)*U→MDM2 & (TP53*U)	EVF
19. MDM2 & (TP53*U)	From 17,18 by →*E*
20. TP53*U	From 19 by &*E*
21. (TP53*U) & P	From 6,20 by &*I*
22. (TP53*U) & P→(TP53*U)*P	EVF
23. (TP53*U)*P	From 21,22 by →*E*
24. (TP53*U)*P→d(TP53) & U & P	EVF
25. d(TP53) & U & P	From 23,24 by →*E*
26. d(TP53)	From 25 by &*E*
27. (TP53 & TP53 & *MDM2* & U & P)→d(TP53)	From 1 to 26 by →*I*
28. d(TP53)	QED

From this complete version, the formal role of the LVFs and the empirical role of the EVFs (that is, how *Zsyntax* can put together formal and empirical aspects) can be easily grasped.

Acknowledgments We would like to thank both the reviewers, whose suggestions and criticisms have been precious, and the friends who have read and commented on versions of this paper: Konstantina Antiochou; Massimiliano Carrara; Marcello D'Agostino; Stavros Ioannidis; Dan Nicholson; Stathis Psillos; Pavlos Silvestros; Petri K. Ylikoski.

Declarations

Ethical approval Not necessary.

Informed consent Not necessary.

Conflict of interest No conflict of interest.

References

Ahmad, S., Hasan, O., & Siddique, U. (2014a). Towards formal reasoning about molecular pathways in hol. In *2014 IEEE 23rd international WETICE conference*, pp. 378-383. IEEE.

Ahmad, S., Hasan, O., Siddique, U., & Tahar, S. (2014b). Formalization of Zsyntax to reason about molecular pathways in hol4d. In *Brazilian symposium on formal methods* (pp. 32–47). Springer.

Ahmad, S., Hasan, O., & Siddique, U. (2015). On the formalization of Zsyntax with applications in molecular biology. *Scalable Computing: Practice and Experience, 16*(1), 37–52.

Ankersmit, F. R. (1986). The dilemma of contemporary Anglo-Saxon philosophy of history. *History and Theory, 25*, 1–27.

Ankersmit, F. R., & Kellner, H. H. (Eds.). (1995). *A new philosophy of history*. Reaktin Book.

Baetu, T. (2019). *Mechanisms in molecular biology*. Cambridge University Press.

Bechtel, W. (2011). Mechanism and biological explanation. *Philosophy of Science, 78*, 533–557.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421–441.

Boniolo, G. (2001). Fictions, judgments of possibility and counterfactuals: Is there room for hermeneutics? In *Normativity and legitimacy*, ed. R. Dottori, , Yearbook for Philosophical Hermeneutics, Münster: LIT, pp. 55–17.

Boniolo G., & Campaner, R. (2018). Molecular pathways and the contextual explanation of molecular functions. *Biology & Philosophy, 33*(3–4). https://doi.org/10.1007/s10539-018-9634-2

Boniolo, G., & Lanfrancone, L. (2016). Decomposing biological complexity into a conjunction of theorems. The case of the melanoma network, *Humana. Mente Journal of Philosophical Studies, 30*, 19–35.

Boniolo, G., D'Agostino, M., & Di Fiore, P. P. (2010). Zsyntax: A formal language for molecular biology with projected applications in text mining and biological prediction. *PLoS ONE, 5*(3), e9511.

Boniolo, G., D'Agostino, M., Piazza, M., & Pulcini, G. (2013). A logic of non-monotonic interactions. *Journal of Applied Logic, 11*(1), 52–62.

Boniolo, G., D'Agostino, M., Piazza, M., & Pulcini, G. (2015). Adding logic to the toolbox of molecular biology. *European Journal for the Philosophy of Science, 5*, 399–417.

Boniolo, G., D'Agostino, M., Piazza, M., & Pulcini, G. (2021). Molecular biology meets logic. *Context-sensitiveness in focus. Foundations of Science*. https://doi.org/10.1007/s10699-021-09789-y

Braillard, P.-A., & Malaterre, C. (2015). *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences*. Springer.

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*, 477–492.
Cartwright, N., Pemberton, J., & Wieten, S. (2020). Mechanism, laws and explanations. *European Journal for Philosophy of Science, 10*, 25 (pp. 1-19).
Craver, C. F., & Darden, L. (Eds.). (2005). Mechanisms in biology. *Special issue of Studies in History and Philosophy of Biological and Biomedical Sciences, 36*(2), 233–244.
Curry, H. (1934). Functionality in combinatory logic. *Proceedings of the National Academy of Science, 20*, 584–590.
D'Agostino, M., Piazza, M., & Pulcini, G. (2014). A logical calculus for controlled monotonicity. *Journal of Applied Logic, 12*(4), 558–569.
Darden, L. (2008). Thinking again about biological mechanisms. *Philosophy of Science, 75*, 958–969.
Dosen, K., & Schroeder-Heister, P. (Eds.). (1993). *Substructural logics*. Clarendon Press.
Dray, W. (1957). *Laws and explanation in history*. Oxford University Press.
Dray, W. (1972). Explanatory narrative in history. *The Philosophical Quarterly, 4*(14), 15–27.
Dray, W. (2000). Explanation in history. In J. Fetzer (Ed.), *Science, explanation and rationality: Aspects of the philosophy of Carl G. Hempel* (pp. 217–242). Oxford University press.
Dupré, J. (2013). I—John Dupré: Living causes. *Aristotelian Society Supplementary Volume, 87*(1), 19–37. https://doi.org/10.1111/j.1467-8349.2013.00218.x
Dupré, J., & Nicholson, D. (2018). A manifesto for a processual philosophy of biology. In *Everything flows: Towards a processual philosophy of biology* (pp. 3–45). Oxford University Press.
Gardiner, P. (Ed.). (1974). *The philosophy of history*. Oxford University Press.
Girard, J. Y. (1987). Linear logic. *Theoretical Computer Science, 50*, 1–102.
Glennan, S. (2010). Ephemeral mechanisms and historical explanation. *Erkenntnis, 72*, 251–266.
Hamilton, A. (2007). Laws of biology, laws of nature: Problems and (dis)solutions. *Philosophy Compass, 2*(3), 592–610.
Hempel, C. G. (1942). The function of general laws in history. *Journal of Philosophy, 39*, 35–48.
Hempel, C. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. Free Press.
Hoare, C. A. R. (1985). *Communicating sequential processes*. Prentice-Hall, Inc.
Howard, W. A. (1980). The formulae-as-types notion of construction. In J. P. Seldin & J. R. Hindley (Eds.), *To H. B. Curry: Essays on combinatory logic, lambda calculus and formalism* (pp. 479–490). Academic press.
Hull, D. L. (2000). The professionalization of science studies: Cutting some slack. *Biology and Philosophy, 15*(1), 61–91.
Illari, P. (2013). Mechanistic explanation: Integrating the ontic and epistemic. *Erkenntnis, 78*, 237–255.
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.
Ioannidis, S., & Psillos, S. (2017). In defense of methodological mechanism: The case of apoptosis. *Axiomates, 27*(6), 601–619.
Joynt, C. B., & Rescher, N. (1961). The problem of uniqueness in history. *History and Theory, 1*, 150–162.
Kaiser, M., Scholz, O. R., Plenge, D., & Hüttemann, A. (Eds.). (2014). *Explanation in the special sciences. The case of biology and history*. Springer.
Lafont, Y. (1993). *Introduction to linear logic. Lecture notes from TEMPUS Summer School on algebraic and categorical methods in computer science*. Czech Republic Yves.
Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.
Mahoney, J., Kimball, E., & Koivu, K. L. (2009). The logic of historical explanation in the social sciences. *Comparative Political Studies, 42*, 114–146.
Mandelbaum, M. (1961). Historical explanation: The problem of covering laws. *History and Theory, 1*, 229–242.
Marchi, S., Zanella, M., Pinton, P., Crafa, S., & Boniolo, G. (2021). MitopatHs: a new logically-framed tool for visualizing multiple mitochondrial pathways. *iScience, 24*, 102324. https://doi.org/10.1016/j.isci.2021.102324
Murphey, M. G. (1994). *Philosophical foundations of historical knowledge*. State University of New York.
Nagel, E. (1952). Some issues in the logic of historical analysis. *The Scientific Monthly, 74*(3), 162–169.
Nicholson, D. (2019). Is the cell really a machine? *Journal of Theoretical Biology, 477*, 108–126.
Nicholson, D., & Duprè, J. (Eds.). (2018). *Everything flows: Towards a processual philosophy of biology*. Oxford University Press.
Nicholson, D. J., & Gawne, R. (2014). Rethinking Woodger's legacy in the philosophy of biology. *Journal of the History of Biology, 47*(2), 243–292.

Persson, J., & Ylikoski, P. (Eds.). (2007). *Rethinking explanation*. Springer.
Restall, G. (2000). *An introduction to substructural logics*. Routledge.
Restall, G. (2008). Substructural logics, in *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/logic-substructural/
Reutlinger, A. (2014). The generalizations of biology: Historical and contingent?", in eds. Kaiser et al., 131-153.
Roberts, C. (1996). *The logic of historical explanation*. The Pennsylvania State University.
Rosenberg, A. (2001). How is biological explanation possible? *British Journal for the Philosophy of Science, 52*, 735–760.
Ross, L. (2018). Causal selection and the pathway concept. *Philosophy of Science, 85*, 551–572.
Ross, L. N. (2020). Causal concepts in biology: How pathways differ from mechanisms and why it matters. *British Journal for the Philosophy of Science, 72*(1), 131–158.
Ruse, M. (1975). Woodger on genetics a critical evaluation. *Acta Biotheoretica, 24*(1), 1–13.
Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton University Press.
Salmon, W. 1989. Four decades of scientific explanation, in *Scientific explanation*, *Minne-sota studies in the philosophy of science*, XII, eds. Kitcher P. and Salmon W.C., 3-219, University of Minnesota Press.
Schaffner, K. (2006). Reduction: The Cheshire cat problem and a return to roots. *Synthese, 151*(3), 377–402.
Sestini, F., & Crafa, S. (2018). Proof search in a context-sensitive logic for molecular biology. *Journal of Logic and Computation, 28*(7), 1565–1600.
Sheredos, B. (2016). Re-reconciling the epistemic and ontic views of explanation (or, why the ontic view cannot support norms of generality). *Erkenntnis, 81*(5), 919–949.
Thagard, P. (2003). Pathways to biomedical discovery. *Philosophy of Science, 70*(2), 235–254. https://doi.org/10.1086/375465
van Eck, D. (2015). Reconciling ontic and epistemic constraints on mechanistic explanation, epistemically. *Axiomathes, 25*, 5–22.
Waters, K. (1998). Causal regularities in the biological world of contingent distributions. *Biology and Philosophy, 15*, 5–36.
Weber, M. 1906. 'Kritische Studien auf dem Gebiet der kulturwissenschaftlichen Logik'; now in M. Weber, *Gesammelte Aufsätze zur Wissenschaftslehre*. Mohr, 1922.
Woodger, J.H. (1937). The axiomatic method in biology. CUP Archive.
Wright, C. (2015). The ontic conception of scientific explanation. *Studies in History and Philosophy of Science, 54*(2015), 20–30.