Brit. J. Phil. Sci. **0** (2011), 1–31

Aspects of Reductive Explanation in Biological Science: Intrinsicality, Fundamentality, and Temporality

Andreas Hüttemann and Alan C. Love

ABSTRACT

The inapplicability of variations on theory reduction in the context of genetics and their irrelevance to ongoing research has led to an anti-reductionist consensus in philosophy of biology. One response to this situation is to focus on forms of reductive explanation that better correspond to actual scientific reasoning (e.g. part–whole relations). Working from this perspective, we explore three different aspects (intrinsicality, fundamentality, and temporality) that arise from distinct facets of reductive explanation: composition and causation. Concentrating on these aspects generates new forms of reductive explanation and conditions for their success or failure in biology and other sciences. This analysis is illustrated using the case of protein folding in molecular biology, which demonstrates its applicability and relevance, as well as illuminating the complexity of reductive reasoning in a specific biological context.

doi:10.1093/bjps/axr006

1 Introduction

Reductionism in the life sciences is a central issue in philosophy of biology (Weber [2005]; Rosenberg [2006]; Schaffner [2006]; Brigandt and Love [2008]). The most robust account of reductionism in the second half of the 20th century was Kenneth Schaffner's ([1967], [1976], [1993]) general reduction–replacement (GRR) model based on Ernest Nagel's discussion in *The Structure of Science* ([1961]). Schaffner argued that reduction was occurring in biology, especially in the molecularization of genetics ([1969]), but that this philosophical interpretation was not relevant to ongoing research methodology ([1974]). The ensuing discussion about the relationship between classical and molecular genetics is expansive but characterized chiefly by a polarity between those in favor of reductionism and those against it (e.g. Hull [1974]; Schaffner [1976]; Kitcher [1984]; Rosenberg [1985]). The so-called anti-reductionist consensus (Waters [1990]; Sterelny and Griffiths [1999], Chapter 7) coalesced around both the *inapplicability* of the GRR model (or some modification thereof) and its *lack of relevance* to actual research,[1] spurring the development of different ideas about theory structure (e.g. Kitcher [1984]). An oft-cited reason for this 'failure of reductionism', both in terms of application and relevance, is the misappropriation of models of reductive reasoning derived from the physical sciences (e.g. Mayr [1988]).

Some philosophers responded to this situation by focusing on different forms of reductionism that were more applicable to the reasoning present in genetics and molecular biology and not subject to the criticisms of the GRR model (Wimsatt [1976]; Waters [1990]; Bechtel and Richardson [1993]; Sarkar [1998]). Another response was to adopt different terminology, such as 'mechanisms' (Machamer *et al.* [2000]), even though themes associated with reductionism remained (such as mechanism descriptions 'bottoming-out'). Instead of answering the question of reductionism with an unqualified yes or no, the aim is to provide perspectives on reductive explanation that are

[1] 'Application': a philosophical account corresponds to actual scientific reasoning rather than a highly idealized version that scientists do not recognize (i.e. descriptive adequacy). 'Relevance': a philosophical account can be of use to scientific researchers in ongoing inquiry and explanation (i.e. prescriptive import). The importance of relevance is a function of whether philosophers take their task to include some form of normative guidance to scientists.

applicable if not also relevant to current research. The philosophical task is to explicate the reasoning in particular areas of biology and understand the diverse standards used by scientists to assess whether reductive explanations are successful, not to be for or against reductionism *per se*. Reductive explanation exhibits many forms in different areas of science.

Here, we take up this broad philosophical task of comprehending the heterogeneous nature of reductive explanation in biology. Our goal is to understand and explicate the diversity of reductionist reasoning practices in biology ('application'), with special attention to their utilization in ongoing research ('relevance'), not to produce an overarching theory of reductive explanation as a competitor to GRR models or other accounts.[2] Our strategy is to distinguish two core facets, composition and causation, and trace out the consequences for different kinds of reductive reasoning construed as part–whole relations. Three aspects of biological explanation emerge as salient: intrinsicality, fundamentality, and temporality. Intrinsicality and temporality have not received sufficient attention in analyses of reductive explanation although the reasons why might seem straightforward. First, part–whole spatial relations presume that parts are contained within or intrinsic to the whole. Thus, intrinsicality is a precondition for part–whole relations rather than an auxiliary element in their evaluation. Second, relations between parts and wholes pertain to composition and arrangement. Thus, temporal relations appear beside the point. It is our contention that considering both of these aspects along with fundamentality augments our understanding of reductionism in biology.

We begin by distinguishing between composition and causation. Reductive explanations involve claims about both the *composition* of higher level entities by lower level entities and the *causal* production of higher level entities by lower level entities. Next, we review themes in traditional theory reduction to show that many ideas forged in these debates are atemporal in nature (e.g. isolating correspondence relations) and disregard part–whole relations. As a consequence, they overlook aspects of reductive explanation found in scientific reasoning. To redress this situation, we analyze three aspects—intrinsicality, fundamentality, and temporality—in the context of part–whole relations. In keeping with our criteria of adequacy, we treat the applicability and relevance of our account in the context of an example: the protein-folding

[2] Another way to conceptualize the project is as 'new wave metascience': 'an exercise in "bottom-up" philosophy of science [. . .] "Bottom up" refers to my letting a sense of reduction emerge from the detailed investigations drawn from recent scientific practice, instead of "imposing" a general account of scientific reduction onto them from science in general, or the "top down." [. . .] The job of new wave metascience is simply to illuminate concepts like reduction as these imbue actual scientific practice. [. . .] because a reasonable explanatory goal is to understand practices "internal" to important current scientific endeavors and the scope of their potential application and development. [. . .] from a detailed investigation of real examples of current research and an attempt to extend these results and practices to issues traditionally reserved for (and by) philosophers' (Bickle [2003], pp. 31–2, 37).

problem. Our analysis of temporality and intrinsicality makes explicit why some physics-derived models mischaracterize part–whole reductive explanations found in biological reasoning; i.e. in what respect some explanations in biology and physics are different. Additionally, our account clarifies the relations between intrinsicality, fundamentality, and temporality, and demonstrates that problems for one aspect of reductive explanation do not imply a failure of reductionism *per se* or a failure of explanation.

2 Composition, Causation, and Varieties of Reduction

2.1 Composition versus causation

There are at least two major facets of reductive explanation. The first pertains to *composition*: higher level entities are composed of, realized by, or nothing but lower level entities. An anatomical unit, such as the heart, is composed of myocardial cells and other cellular entities. A myosin filament found in myocardial cells is composed of myosin proteins, which are in turn composed of amino acid residues. The second facet involves *causation*: some higher level entities are caused, brought about, or determined by lower level entities. The heart's rhythmic beating is caused by the contraction of its myocardial cells. Myocardial cells contract because myosin proteins ratchet along another set of proteins. The appropriate folding of a particular myosin protein is determined causally by its amino acid sequence.

Many reductive explanations in the biological sciences are a mixture of compositional and causal claims (Craver and Bechtel [2007]). Although these two facets are not always distinguished in biological discourse, they can be distinguished in retrospect. From a philosophical standpoint, keeping them distinct is important for understanding reductive reasoning in biology (and other sciences). Consider the protein-folding problem in molecular biology as interpreted by Alex Rosenberg:

> The watchword of reductionism in biology has been the slogan that 'function is a consequence of conformation, and conformation is specified by sequence.' That is, the 'sequence' or linear one-dimensional order of the atomic components of a biologically significant molecule causally determines its three-dimensional structure, its shape or 'conformation'; its three-dimensional structure causally determines its effects, and in particular all its biological functions. ([1985], p. 73)

These claims about reductive reasoning pertain to *causation*, as indicated by the choice of phrasing ('causally determines'). Rosenberg continues:

> The shape of the whole molecule [. . .] is determined exclusively by the order of the amino acids that compose it and by the chemical properties of these amino acids. Given a specification of the [. . .] amino acids [. . .],

> we can deduce the shape of the whole molecule from the fact that some amino acids are hydrophobic – not water soluble, some are hydrophilic – water soluble, some are charged negatively and some positively, and that some contain larger and some smaller atoms in their side-chains. (pp. 75–6)

Here we see a shift to determination in virtue of *composition*. Given a particular set of amino acid residues (i.e. the composition of the protein), one can deduce the three-dimensional conformation of the folded protein. The *causal determination* of folding by the amino acid sequence can be distinguished from the determination of folding by the amino acid *composition* of the polypeptide. Although these two facets travel together in scientific discourse, the question of whether native protein conformation is determined by the linear polypeptide *composition* is separate from the question of whether there is a *causal* explanation of how the folding occurs. An affirmative (or negative) answer to one does not imply an affirmative (or negative) answer to the other. The significance of this difference is discernable in further comments offered by Rosenberg: 'All we need [. . .] is information about the chemical milieu and about the chemical bond already available in physical chemistry' ([1985], p. 76). Information about the chemical milieu does not concern the composition of the protein but is relevant to a causal account of its folding, which involves interactions among its amino acid residues. These intertwined facets need to be distinguished explicitly in analyses of reductive explanation.

Although numerous issues accompany composition and causation in reductive explanation, we want to extract three key aspects. The first is 'intrinsicality' and relates to composition. Claims about reducing a higher level entity to its component parts involves an individuation of the higher level entity such that its components can be distinguished from other entities (a surrounding context). Intrinsicality has an epistemological (or pragmatic) aspect because what counts as intrinsic depends on the explanatory goals of researchers. If the goal is to explain (reduce) a cell in terms of its component parts, then the boundary between intrinsic and extrinsic is the cell membrane. If the goal is to reduce the heart to its component parts then the boundaries of the organ demarcate intrinsic from extrinsic. What counts as intrinsic is relative to local explanatory aims.[3]

The second aspect is 'fundamentality' and often relates to composition. Fundamentality corresponds to the assumption that higher levels of organization are reduced to lower levels of organization: parts are taken as more

[3] Biological researchers are explicit about this relativity to local explanatory aims; e.g. 'it is often beneficial to separate contributions arising from fluctuations that are inherent to the system of interest (intrinsic noise) from those arising from variability in factors that are considered to be external (extrinsic noise). [. . .] The definition of intrinsic noise is problem-dependent, and varies from one context to another' (Kærn *et al.* [2005], p. 456).

fundamental than the compound (Sarkar [1998]). This can be qualified so that only restricted sets of properties (e.g. biochemical moieties) of the parts count as fundamental. These kinds of qualifications are spelled out locally in the context of explanation. Qualified senses of fundamentality are important to distinguish because they can introduce qualifications about intrinsicality. For example, a qualified fundamentality might pick out biochemical properties as explanatory of cell properties, which implies that biochemical properties outside of the cell (i.e. extrinsic) are fundamental and that other intrinsic properties of cellular components (e.g. location) are not considered explanatory. Therefore, the character of a reductive explanation depends on whether and how it is assessed with respect to fundamentality, intrinsicality, or both simultaneously.

The third aspect, 'temporality', is associated with causation. Any causal explanation will involve some element of temporal duration, which may be operationalized in different ways depending on the explanatory goals in view (similar to intrinsicality and fundamentality). If the aim is to explain the increased rhythmic contraction of the heart by the cellular mechanisms that process adrenergic hormones, then the explanation requires (at least implicitly) an earlier time at which hormones are interacting with cellular receptors and a later time when the altered rhythmic contraction obtains.

These three aspects do not yield a new theory of reductive explanation but instead draw attention to its inherent diversity. A reductive explanation of the cell using only biochemical properties might fail in terms of intrinsicality but succeed with respect to fundamentality—success or failure of a reductive explanation is not an all or nothing phenomenon. For any reductive explanation, we must not only inquire whether each of these three aspects is applicable, but also characterize the details involved in order to evaluate whether there is success or failure of one kind or another. In other words, explanations are *reductive* explanations provided they conform to certain additional constraints, including intrinsicality, fundamentality, and temporality. Compositionally, this means ensuring that constituents mentioned in the *explanans* bear appropriate spatial relations to the *explanandum* phenomena. Causally, this means ensuring that causes, laws, and/or properties mentioned in the *explanans* are appropriately fundamental.[4]

[4] Explanatory reduction is only one dimension of epistemological reductionism. For example, the claim that it is a fruitful (or the most fruitful) heuristic strategy to investigate natural phenomena reductively is termed methodological reductionism. Explanatory and methodological reduction can be decoupled because they do not entail one another: methodological reductionism does not guarantee explanatory success and a successful explanatory reduction does not imply that methodological reduction is the most favorable strategy of inquiry.

2.2 The Nagelian framework and its aftermath

Before we continue this line of reasoning we must turn to the Nagelian framework of reduction in order to show why intrinsicality, fundamentality, and temporality have been neglected in prior analyses. The shared background for discussions of reduction in philosophy of science is Nagel's account of the *formal* criteria of reduction (Nagel [1961], Chapter 11). One of Nagel's primary concerns was whether an older theory is reducible to its successor. Reduction was conceived of as a special case of deductive–nomological explanation. If the old theory reduces to its successor then the laws of the old theory as well as its observational consequences can be deduced from the successor theory. Successful 'Nagel reduction' integrates the old theory into the successor theory and provides a clear sense in which the successor theory is better than its predecessor. Nagel reduction is an explanation because it claims that one theory can explain another theory via deductive derivation. It is a *reductive* explanation because the successor theory is more fundamental, where fundamentality is understood in terms of the theory's greater generality rather than in terms of the properties that the theory mentions. This approach provides a rationalization of the history of science but did not capture the bulk of attention in subsequent literature (and we ignore this historical dimension as well).[5]

What did exercise many philosophers was the problem of bridge laws. If in a putative Nagelian reduction the two theories invoke different terminology, as in the case of thermodynamics and statistical mechanics, connections need to be established that link terms such as 'temperature' or 'entropy' with those from statistical mechanics. According to one influential strand of thought (e.g. Schaffner [1967], [1993]), these bridge laws are biconditionals that express synthetic *identities*, such as between the temperature of an ideal gas and its microstate. The question of reduction translates into a question of whether there are such identities. Are the properties described by the reduced theory identical with those picked out by the reducing theory? Seeking these identities can be a method of exploring relations between two theories, both of which are accepted to some degree (cf. Wimsatt [1976]). If these identities are secured then the Nagelian account provides a rationale for any resulting theory replacement.

A number of key concepts were introduced in this context and became central to debates about reductionism. For example, the fact that properties picked out by the theory to be reduced can be multiply realized by properties described in the reducing theory was taken to preclude the identification of higher and lower level properties or types (see Brigandt and Love [2008]).

[5] The succession of theories in the history of science has been labeled 'diachronic' reduction by some philosophers (see discussion in Brigandt and Love [2008]).

The notion of supervenience offered some consolation by capturing the idea that a particular instantiation of one set of properties of a system (e.g. physical ones) 'in some sense' determines another instantiation from a different set of properties (e.g. mental ones), even if a demonstration of identities among types remains beyond reach. For example, Jaegwon Kim defines strong supervenience as follows: 'Mental properties supervene on physical properties, in that necessarily, for any mental property M, if any thing has M at time t, there exists a physical base (or subvenient) property P at t, and necessarily anything that has P at a time has M at that time' (Kim [1998], p. 9, underline added).

Three features of these concepts merit attention. First, the concepts of synthetic identity, (multiple) realization, and supervenience gained importance within the *formal* reductionist framework outlined by Nagel. Outside this framework, their significance is an open question (cf. Waters [1990]). Second, identity, realization, and supervenience focus on relations that obtain between two properties (different in kind; e.g. mental and physical) of one and the same object or system. The question is whether one state or property of a system is identical to, or at least metaphysically dependent on, physical states or properties of the *same* system (and not, for instance, on the states or properties of the parts).[6] Third, these concepts are *atemporal* because identity, realization, and supervenience relations obtain at a time slice.[7] This atemporality is consistent with the identity relation persisting over time, although this is usually assumed rather than articulated. In Kim's definition of supervenience, it is the mental properties *at t* and the physical properties *at t* that are taken into consideration. Either one property realizes another property only at a particular time or identities between properties are assumed to persist. This focus on atemporal relations that pertain to properties of one and the same system explains why causation and composition have not been distinguished in most analyses of reductive explanations.

Much of the discussion about reductionism in biology has revolved around the relations between classical genetics and molecular genetics. The primary context in which these debates occurred is the problem of bridge laws and the question of synthetic identities between properties referred to in different theories. This includes the presumption that synthetic identities hold between different properties of one system and the presumption of atemporality, i.e. relations such as identity, realization, and supervenience obtain strictly at a

[6] Gillett ([2003]), Hoffmann and Newen ([2007]), and others have identified this as a shortcoming and argued that the concept of realization should accommodate the relation between properties of compounds and properties of parts (i.e. a relation between properties of *different* systems).

[7] A *temporal* relation is one in which a property or state at t is related to another property or state at t^*, with $t^* \neq t$, as is usually the case in causal relations. Wilson and Craver ([2007]) broach the issue of whether realization is necessarily intrinsic and synchronic.

time slice (see, e.g. Rosenberg [1978], [1985], [1994]).[8] Much of the antireductionist consensus was forged in this same context, largely accepting a Nagelian framing of the issue. Although the question of whether predicates of different theories pick out identical properties could be relevant and applicable to reductive modes of explanation in scientific reasoning, it is of limited significance in biology and has not clarified the relations between classical and molecular genetics. Our understanding of reductive explanation in biology can be amplified when intrinsicality and atemporality no longer operate as presuppositions but are brought to the foreground for explicit consideration. This involves starting with the explanatory practices of biologists rather than notions such as identity, realization, and supervenience.

3 Part-whole Reduction: Intrinsicality, Fundamentality, and Temporality

In response to discussions about theory reduction, a variety of authors have conceptualized reductionism in terms of the relationship between parts and wholes (e.g. Wimsatt [1976]; Waters [1990]; Bechtel and Richardson [1993]; Sarkar [1998]).[9] These relations play methodological and explanatory roles in biology and bypass many issues that arise in the Nagelian context discussed above. Sahotra Sarkar ([1998]) argues that reductions in biology are explanations of phenomena from one non-fundamental realm (e.g. a whole) in terms of another more fundamental realm (e.g. the parts). These reductive explanations vary in strength based on how they fulfill criteria related to the representation of objects in a reductive explanation.[10] For example, 'physical' reductionism is the explanation of biological phenomena using the physical properties of constituent molecules and macromolecules: 'the behavior of wholes is supposed to be explained by those of their constituent parts' (Sarkar [1998], p. 136). Adopting this focus on part–whole relations, we turn to how intrinsicality, fundamentality, and temporality help to illuminate the diversity of reductive explanation and thereby afford us greater resources for characterizing biological reasoning.

3.1 Intrinsicality and fundamentality

Compositional part–whole reductions are reductive in two respects. First, they appeal to *intrinsic* features of the compound system in question; parts are

[8] A few philosophers have noted the difference between atemporal and temporal types of reduction (e.g. Sober [1999a], [1999b]).

[9] Issues related to parts and wholes were present in Nagel's original discussion ([1961], pp. 380–97), but overshadowed in subsequent developments of reductionism by philosophers.

[10] 'Representation' refers to how natural phenomena are symbolized, embodied, pictured, or designated such as in equations, scale miniatures, or abstract diagrams.

presumed to be *contained within* or *internal to* the whole by some boundary or line of demarcation. No extrinsic features are invoked to explain the property of the whole. Secondly, they appeal to a more *fundamental* realm or lower level features (the parts), or a restricted set of properties within this realm, as compared with the whole (the non-fundamental realm). Thus, compositional part–whole reductions can fail as *reductive* explanations either because intrinsicality is violated (for a particular individuation), fundamentality is violated (in the broad sense or for some restricted set of properties), or both are violated (for a particular individuation and for some restricted set of properties).

The intuition behind intrinsicality is that 'a property is intrinsic if anything that has it has it regardless of what is going on outside itself' (McKitrick [2003], p. 158). In the present context, we only need to add that the application of this concept is relative to explanatory (or pragmatic) goals. Researchers must demarcate a system of interest from its environment (thereby deciding what is 'inside' and 'outside') because natural phenomena do not come with labels attached to indicate their boundaries. This specification of boundaries predisposes researchers to focus on particular causal factors in a reductive explanation. Once the system–environment distinction has been drawn and the system is decomposed into parts, explanations will concentrate on the causal powers of the parts ('intrasystemic'—intrinsic) rather than the role of factors external to the system ('extrasystemic'—extrinsic) (cf. Wimsatt [1980]). This observation has been deployed to emphasize how explanations of biological phenomena may be biased toward attributing causal responsibility to intrasystemic features (e.g. genes) rather than extrasystemic ones (e.g. predation) (Robert [2004]). What has gone unnoticed is that this bias is neutral with respect to whether extrinsic features also meet a qualified sense of fundamentality. The bias applies not only to attributions of causal responsibility for compounds at the same 'level' within the environment, but also for extrasystemic, fundamental realm entities exhibiting the restricted set of properties. Therefore, fundamentality is decoupled from intrinsicality because the failure of the latter does not imply a failure of the former in a part–whole reductive explanation.

The significance of this decoupling is muted if part–whole relations are always atemporal. A violation of intrinsicality seems impossible because the very idea of part–whole relations presumes that parts are contained within the whole (i.e. a nesting of parts). But if part–whole relations are considered temporally, then it is possible for intrinsicality to be violated because a part of a whole at time t may no longer be a part of the whole at t^*. Consider a system S that is individuated from an environment E and partitioned into components $(s_1, \ldots, s_n)$ at time t. Let S be the heart, E be the rest of the body cavity, and $s_1, \ldots, s_n$ be cells within the heart. If the components are blood

Table 1. Forms and Aspects of Reductive Explanations

Forms of explanation	Fundamentality	Intrinsicality	Temporality
I	Y	Y	N
II	Y	Y	Y
III	Y	N	Y
IV	N	N	Y

Different forms of explanation that result from combinations of the three aspects (fundamentality, intrinsicality, and temporality). 'Y' or 'N' refers to whether the aspect is present (Y) or absent (N). When temporality is present, it is possible for fundamentality and intrinsicality to be absent in specific ways (see discussion in text). Other combinations are omitted intentionally (see footnote 12)

cells coursing through nascent atrial chambers during ontogeny, then their presence in the heart at time t meets the intrinsicality condition even though at t^* they have passed out of the heart.[11] Because blood flow is a key factor in the proper shaping of the heart during embryogenesis (Hove *et al.* [2003]), blood cells at t can have a causal effect on the shape of the heart at t^* when these cells are no longer intrinsic to S. This is still a kind of reductive explanation because properties of entities at the qualified fundamental level of cells account for changes in the properties of entities at the non-fundamental level (organs).

Alternatively, adrenergic hormones secreted in E can modulate heart rhythm. These hormones are extrinsic to the heart (S), even though they meet the fundamentality condition, and exert an effect through a temporally extended process. Another possibility is that different organs interact directly with the heart to bring about changes in its morphology during embryogenesis (e.g. through physical contact). These other organs are not only extrinsic to S but also non-fundamental because they are at the 'same' non-fundamental level as S.

We can summarize these combinations of the three aspects in a table of reductive explanatory forms (Table 1). The availability of different combinations indicates the heterogeneous ways that explanations can succeed or fail *reductively*, especially given that intrinsicality, fundamentality, and temporality each need to be characterized individually for any reductive explanation. The first case (I) corresponds to what can be recovered from prior discussions of reduction. The second case (II) includes extending compositional relations through time, but also encompasses other possibilities, such as explanations that appeal to fundamental, intrinsic causes at t bringing about an effect in the non-fundamental level whole at t^* while ignoring compositional relations that

[11] Intrinsicality is construed spatially because of the explanatory context of part–whole relations. Other interpretations of intrinsicality might be more appropriate in different explanatory contexts where part–whole relations are not in view, such as when explaining the heart's function rather than its structure.

obtain at other times between t and t^*. The third case (III) includes an extrinsic entity from a fundamental level explaining a system property (e.g. the adrenergic hormone case). Although it is reductive in the sense of appealing to properties at the fundamental level, it fails in the sense of a whole being causally explained by its parts. The fourth case (IV) corresponds to the possibility of one organ interacting with another organ—the entity is extrinsic to S and also resides at the same level of non-fundamentality. Here, there is a failure of the parts to explain the whole and a failure of the fundamental-level properties to explain non-fundamental-level properties. Once part–whole relations are treated temporally, intrinsicality and fundamentality take on independent significance in reductive explanations.[12]

If we recall that what counts as intrinsic is relative to explanatory aims, then a novel interpretation of the 'context objection' to reduction emerges (cf. Delehanty [2005]). The context objection claims that a reduction can be blocked because of an ineliminable appeal to contextual factors. The standard reductionist rejoinder is to pursue a reduction of this context. By distinguishing two different ways a reductive explanation can fail as a *reductive* explanation (III—intrinsicality fails; IV—fundamentality and intrinsicality fail), we are able to recover a more nuanced reading of the context objection: the reductionist rejoinder maintains the fundamentality condition even when intrinsicality is violated. But even if the standard rejoinder preserves fundamentality, and thus a reductive explanation in one aspect, the context objection retains some validity because another aspect of part–whole reductive explanation, intrinsicality, is violated. If an attempt is made to recover intrinsicality by redrawing system boundaries (e.g. treat the body cavity as the system so that the adrenergic hormone is a part of this new whole), then we have changed what counts as whole and parts, thereby altering the aspect of intrinsicality. Whether a part–whole reductive explanation succeeds or fails with respect to intrinsicality depends on how wholes are individuated, which means any redrawing of the boundaries constitutes a change of the original question about whether a whole can be explained reductively by its intrinsic parts. The failure of a reductive explanation for the aspect of intrinsicality may be an empirical indicator that redrawing boundaries is epistemologically warranted. Even if a new reductive explanation succeeds as a consequence of different individuation criteria, this is compatible with the claim that the behavior of interest cannot be explained in terms of its intrinsic features under the original individuation criteria.

[12] The table omits two possibilities ruled out in discussion: (a) atemporal part–whole reductions that violate intrinsicality or fundamentality, and, (b) fundamentality failing when intrinsicality holds, because if a feature is intrinsic to S then in order to be contained within S it must be instantiated at a more fundamental level than S itself.

3.2 Temporality

Temporality refers to the aspect of reductive explanation whereby a property of a whole at t^* is explained in terms of properties of its parts at an earlier time t. If a *temporal* relation is one in which a property or state at t is related to another property or state at t^*, then a *causal* relation is one in which a property or state at t determines or influences another property or state at t^*; the state or properties of the parts and their interactions at t bring about a change in the state or properties of the compound at time t^*. In addition to its role in illuminating the significance of intrinsicality and fundamentality, the import of temporality expands when reductive explanations in biology are contrasted with those found in physical science. Part–whole explanations in physics are primarily atemporal (see Hüttemann [2005] for further discussion), whereas in biology part–whole explanations are often temporal.

3.2.1 Atemporal part-whole reduction

For the behavior of a physical system, part–whole reductions can pertain to either its states or its temporal evolution (i.e. dynamics). A part–whole reduction of a *state* explains the state of a compound system at a time on the basis of the states of its parts at the same time. For example, we might explain why a compound system (e.g. an ideal gas) has the determinate energy value E (the macro-state) by appeal to the determinate energy values of its constituents (e_1 to e_n; the states of the parts). One can reconstruct this as a deductive–nomological explanation that relies on the states of the parts (particular facts) and a law of composition that tells us how the states of the parts contribute to the state of the compound. If we assume that interactions can be neglected, the kinetic energy values simply add up. The explanation is reductive because it only uses states of the parts and a law of composition in the *explanans*, but it is also atemporal.

The second situation is a part–whole reduction of the *dynamics* of the physical system; the dynamics of a compound system can be explained in terms of its parts. For instance, the Hamiltonian for a compound system can be analyzed in terms of those for the parts along with interaction terms and a law of composition. The relation between the Hamiltonian for the compound and those for the parts (i.e. a part–whole relation) is reductive but *a*temporal because the Hamiltonian of the compound at time t^* is not calculated in terms of the Hamiltonians of the parts and their interactions at *another* time t. (Typically, the Hamiltonians of the parts and the compound are time independent.) A part–whole reduction of the dynamics of a compound is often integrated into a causal explanation. In classical dynamics, if we want to explain why the state of the solar system Z develops over time into state Z', we

can appeal to the dynamics of its parts and their interactions since the solar system is a compound. The two dimensions of the explanation are separable: (i) the non-reductive, temporal (causal) dimension; and, (ii) the atemporal, part–whole reductive (compositional) dimension.

The export of atemporal models of reductive explanation from philosophical reflection on the physical sciences has hindered our analyses of reductive part–whole explanations in biology.[13] It has long been recognized that relations between parts and wholes in biological hierarchies involve temporality and causal interaction or interdependence (Nagel [1961], Chapter 12). In the part–whole reductive explanation for a physical system described above, the temporal (causal) and the reductive (part–whole) dimensions can be separated neatly in contrast to the hybrid compositional and causal claims found in biology or situations where the compound does not exist at the earlier time (e.g. during embryological development). Furthermore, physics is concerned typically with isolated systems and, as a consequence, it is often assumed that new parts are *not* added and that parts do not 'leave' the compound (as might happen with cellular parts that undergo apoptosis). In other words, it is assumed that the compositional relations remain constant. Applicability and relevance require that we focus on the reductive explanations biologists put forward, not the ultimate possibility that biological processes will be explained solely in physical terms.

3.2.2 Temporal (causal) part–whole reduction

Part–whole reductions (and explanations more generally) in biological science are often *temporal*. Properties of a whole at t^* are explained in terms of properties of parts at an earlier time t; the behavior of the parts at t *causes* the compound to have a certain behavior or property at a later time t^*. Temporal or causal part–whole reductions can be described differently depending on how the behavior of the parts is described. In those areas of biology where physical laws play a role, causal part–whole reduction involves the laws and initial conditions that pertain to parts of a whole at t explaining the behavior of the whole at t^*. Causal part–whole reductions also can be described in terms of causal powers: the behavior of a whole at t^* is explained in terms of the causal powers of its parts at t. Causal part–whole explanations are reductive because they constrain the *explanans* to only laws or causal

[13] But this is not because atemporal part–whole explanations in physics appeal to universal laws. It is possible to articulate atemporal part–whole explanations in terms of causal powers or invariance relations (e.g. Cartwright [1999]; Mitchell [2000]; Woodward [2003]). The forms of reductive explanation described do not rely on a specific commitment to the nature of explanation. Our terminology reflects an expository preference.

powers of the parts (and maybe also compositional rules); i.e. they require intrinsicality and fundamentality.

Examples of causal part–whole reductions abound in biology, such as explaining muscle tissue activity (behavior of a system) at a later time t^* by appealing to the properties of muscle cells (the parts) composing the tissue (the whole), which contain special motor proteins (e.g. myosin) that contract by molecular ratcheting at an earlier time t. Temporal part–whole reduction explains the state of a compound or whole (muscle contraction) in terms of states of the parts at earlier times (myosin ratcheting) and must, for several reasons, be distinguished from atemporal part–whole reductions, which explain the behavior of a compound in terms of the behavior of the parts at the same time.

First, atemporal part–whole reductions focus on *composition* whereas temporal part–whole reductions take into account *causation* as well. Atemporal part–whole reduction corresponds to most reductive explanations in physical science and temporal (causal) part–whole reduction accurately captures most reductive explanatory practices of biology.[14] Second, part–whole reduction must be distinguished from identity-reduction, which focuses on the relation of two kinds of properties that pertain to the same system rather than explaining properties of a compound system in terms of its parts.[15] These differences clarify why multiple realization and supervenience do not transfer easily from the context of identity-reduction to temporal part–whole reduction. Even attempts to define supervenience or realization for part–whole relations (e.g. Menzies [1988]; Gillett [2003]) are not applicable because they are about atemporal relations. Causes do not *realize* their effects, and the existence of multiple causal pathways from earlier states of the parts to a later state of the compound does not undermine causal part–whole reductions.[16]

[14] We leave open the possibility that biological temporal part–whole reductions might eventually be explicated in terms of atemporal part–whole reductions in physics (according to the standards of the scientific community).

[15] Explaining the property of a whole in terms of the properties of its parts and explaining a property of the whole in terms of another property of the whole by identity are distinct explanatory questions. Consider an atemporal case where the property of a whole (the temperature of an ideal gas) is explained in terms of the combined properties of the parts (kinetic energies of the molecules). First, we can ask whether it is possible to explain why an ideal gas with a certain mean kinetic energy has a specific temperature. This involves two kinds of properties of the same system—the specific mean kinetic energy and temperature of the ideal gas at a time. A bridge law may link the mean kinetic energy (a property of the gas as a whole) to the temperature (a property of the gas as a whole). A second question is whether it is possible to explain why an ideal gas has a specific mean kinetic energy on the basis of the behavior of its constituent molecules. This involves the behavior of the components of the ideal gas and its behavior as a whole: how do the individual kinetic energies add up to the kinetic energy of the whole? The relation is between parts and wholes, not two properties within the same system.

[16] Multiple causal pathways from parts to compounds would undermine the attempt to reduce special science causal *laws* to physical science causal *laws*. Jerry Fodor ([1974]) discussed whether a special science causal law, (1) $S_1(x) \rightarrow S_2(x)$ ('all S_1 situations bring about S_2 situations'), can be reduced to a physical science causal law, (2) $P_1(x) \rightarrow P_2(x)$. The reduction of (1) to

Finally, we should be wary of assuming that temporal part–whole reductions can be subsumed under an atemporal account. This would collapse cases II, III, and IV (Table 1), which all appear in scientific reasoning, into the situation of case I. The criterion of applicability and relevance encourages preserving these distinctions in order to better limn the contours of reductive reasoning in biology. The default stance is to recover the actual reasoning patterns found in science rather than beginning with a reconstruction that deviates from the spatial and temporal individuation choices of scientists. In fact, attending to intrinsicality, fundamentality, and temporality has generated a new perspective on the context objection and the differences between reductive explanations in biology and physics. Now it is time to see their applicability and relevance in a specific example from biology.

4 The Protein-Folding Problem

4.1 Background and significance

Researchers working on various aspects of protein folding have recognized explicitly the centrality of reduction in the context of their inquiry.

> The protein folding problem [. . .] represents an unusually concrete and limited case of the whole problem of reductionism. An unfolded protein is clearly a chemical object [. . .] its properties are relatively dull and quite predictable by summing up the properties of its components. A folded protein on the other hand, in addition to complexity and unpredictability, has acquired meaning: unity, controlled interaction with other systems, and biologically significant function. [. . .] Understanding the rules of [folding] would teach us worthwhile lessons about hierarchical organization, cooperative properties, and exactly how an organic whole becomes so much more than a sum of its parts. (Richardson [1982], p. 1)

Philosophical terminology appears regularly in textbooks: 'the function of a protein [. . .] is an emergent property resulting from exquisite molecular order. [. . .] We have taken a reductionist approach in dissecting proteins to their four levels of structural organization' (Campbell and Reece [2002], pp. 74, 78). The protein-folding problem remains an outstanding question in molecular biology pertinent to part–whole explanatory reduction (Sarkar [1998], p. 169).

(2) is an identity issue. If the cause S_1(x) is multiply realized (different constituents p_1 to p_n and p'_1 to p'_m), then it would undermine the reduction of (1) to (2). It would not undermine temporal part–whole reduction of S_2(x) at t^* in terms of the causal powers of either p_1 to p_n or p'_1 to p'_m at t (cf. Sober [1999b]). The significance of multiple realization for reductive explanation depends on the meaning of 'reduction' in use.

4.2 Reductive explanation in molecular biology

Proteins are composed of amino acid components ('residues') that are linked by covalent peptide bonds into a chain ('polypeptide'). This linear chain ('primary structure') is produced from a process termed 'translation' whereby specific cellular constituents (ribosomes, themselves proteins) translate a linear stretch of RNA with a triplet code of nucleotides (e.g. AAG) into amino acid residues for a linear polypeptide (e.g. AAG = lysine). Secondary structure refers to repeating patterns of coiling or folding (α-helices or β-pleated sheets) that obtain as a consequence of regular hydrogen bonding; i.e. not due to the specific chemical moieties of each amino acid residue. Nearly all proteins adopt a three-dimensional structure (tertiary structure) in order to be functional. This conformation of the polypeptide is understood in terms of interactions among its amino acid residues (e.g. hydrophobic residues avoid interaction with surrounding water by segregating to internal regions). The protein-folding problem (hereafter, PFP) concerns explaining how this active conformation is achieved for polypeptides subsequent to translation from RNA in the cellular context.[17] 'A functional protein is not *just* a polypeptide chain, but one or more polypeptides precisely twisted, folded, and coiled into a molecule of unique shape' (Campbell and Reece [2002], p. 74). The linear sequence hypothesis (hereafter, LSH) holds that the three-dimensional folding of a protein results from the properties of the amino acid residues in the polypeptide and their interactions alone. A folded protein is explained by the chemical interactions of its component residues as ordered in a linear polypeptide—the whole is a 'sum' of the interaction of its parts.

Interestingly, there is an ambiguity in the LSH claim that polypeptide primary structure contains all of the information required for achieving three-dimensional conformation (cf. Freedman [1999]). One interpretation sees the LSH as a claim about *inferring* the three-dimensional structure of a protein from its linear sequence of amino acids, which focuses on the phrase 'containing all of the information required'. A second interpretation concentrates on the phrase 'achieving three-dimensional conformation' and takes the claim to be about the kinetic, thermodynamic, and structural aspects of the protein-folding process. The first interpretation construes the PFP in terms of prediction from composition—*that* you can make the correct inference ('predictive construal'); the second interpretation glosses the PFP as a causal question about *how* the folding occurs ('folding construal').[18] Whether a protein

[17] The PFP is not isomorphic to the question of protein function because the latter is often due to quaternary structure, which refers to the further aggregation of tertiary structured proteins.

[18] One way these two interpretations are conflated is with talk of the 'translation of information' from the linear amino acid sequence (parts) into the three-dimensional conformation of the protein (whole).

folds only as a consequence of its amino acid residues is a causal question involving reductive explanations of wholes in terms of parts.

The (ambiguous) LSH began to crystallize with the unraveling of the genetic code (Morange [1998]; Tanford and Reynolds [2001]). Biochemists held that there were yet-to-be discovered rules that governed three-dimensional protein conformation, such as regularity of amino acid sequence. This expectation was a form of reductionist reasoning (cf. Jaenicke [2005]) and its first expression came from Francis Crick: 'It is of course possible that there is a special mechanism for folding up the chain, but the more likely hypothesis is that the *folding is simply a function of the order of the amino acids*' (Crick [1958], p. 144, emphasis in original). Philosophical commentators saw protein folding as susceptible to the trend of reductionism in biology via molecularization (Schaffner [1969], p. 344; cf. Rosenberg [1985]).

Some of the strongest evidence in favor of the LSH was derived from experiments on the denaturation and refolding of ribonuclease proteins *in vitro* (Anfinsen [1973]). Ribonucleases subjected to denaturing conditions were able to refold rapidly into the proper configuration rather than the myriad of other biochemical possibilities.[19] In accord with the LSH, correct refolding occurred as a function of the linear sequence of amino acid residues composing the polypeptide.

> The three-dimensional structure of a native protein in its normal physiological milieu (solvent, pH, ionic strength, presence of other components such as metal ions or prosthetic groups, temperature, and other) is the one in which the Gibbs free energy of the whole system is lowest; that is, that the native conformation is determined by the totality of interatomic interactions and hence by the amino acid sequence, in a given environment. (Anfinsen [1973], p. 223)

But the precise role of extrinsic factors in the folding process ('normal physiological milieu' and 'in a given environment') was not explored, even though Anfinsen's group found that folding took an hour or longer rather than several minutes or less without an enzyme from the endoplasmic reticulum (a cellular organelle where much translation occurs).

In recent research, the role of extrinsic factors in protein folding has grown in importance. Although it has been claimed that 'today's view [is] that *whatever* its amino acid sequence, a polypeptide chain will *spontaneously* fold into a stable configuration' (Morange [1998], p. 123, emphasis in original), the situation is more complex. First, many denatured proteins do not refold as cleanly

[19] 'A chain of 149 amino acid residues [...] would be able to assume on the order of 4^{149} to 9^{149} different conformations in solution. The extreme rapidity of the refolding makes it essential that the process take place along a limited number of "pathways" ' (Anfinsen [1973], p. 228). This is known as the 'Levinthal paradox' (Levinthal [1969]): how is the native protein conformation rapidly 'determined by the totality of interatomic interactions' if a random search through the vast array of possible biochemical pathways is impossible? (Honig [1999]).

as those studied by Anfinsen's group. 'Many cell biologists, having been taught that polypeptide chains can spontaneously fold to the native state, have been frustrated to discover that, although spontaneous folding can occur for small simple proteins [. . .], spontaneous, high-yield folding to the native state might be the exception, rather than the rule' (Clark [2004], p. 527). Second, spontaneity of folding is very dependent on the 'normal physiological milieu', which includes more than just physico-chemical components of the environment (e.g. pH, salt concentration, or temperature). Chaperone proteins guide folding during and after polypeptide synthesis (Frydman [2001]), as well as in refolding subsequent to stressful conditions such as heat shock (Feder and Hofmann [1999]): 'Proteins need the assistance of molecular chaperones and folding enzymes to reach their native structure efficiently' (Liscalijet *et al*. [2005], p. 78). Third, the term 'spontaneous', which is often found in descriptions of protein folding, should not be confused with instantaneous.[20] There is a (rapid) temporal dimension in protein folding and it is not a static achievement. The conformation necessary for functionality is dynamic, fluctuating around one or more preferred states, and is affected by molecular interactions such as ligand binding or phosphorylation (Eisenmesser *et al*. [2005]).

One reason why molecular chaperones must provide oversight in the process of protein folding is that the cellular environment is crowded (Ellis [2001]; Liscalijet *et al*. [2005]; Homouz *et al*. [2008]). Research has uncovered a complex set of causal roles for biological macromolecules in the physico-chemical milieu. Distinct functional groups of chaperones monitor protein folding during *de novo* synthesis, quality control, and the response to stress (McClellan *et al*. [2005]; Albanese *et al*. [2006]). Chaperones work in different ways, as well as cooperatively, sometimes providing a sequestered environment for folding (e.g. as a nanocage), and at other times actively facilitating folding (Ellis [1998]).[21] An experimental change to the volume of the cavity inside a nanocage increases the folding speed for small proteins by modifying the 'energy landscape' of the protein (Tang *et al*. [2006]). Multiple amino acid residue interactions between an already functional, folded protein (the chaperone) and the not-yet-folded polypeptide underlie the process of correct folding (Tang *et al*. [2008]). Even when mutations are introduced that lead to

[20] For example, 'A polypeptide chain of a given amino acid sequence can spontaneously arrange itself into a three-dimensional shape' (Campbell and Reece [2002], p. 78). 'Spontaneous' usually denotes something happening apart from external stimulus when conditions are 'just right'. It might be said that a protein has a disposition to fold under appropriate environmental triggering conditions. The difficulty is that chaperones are often specific to particular proteins and operate over extended durations of time. This is different from the many non-specific triggers one could use to manifest the fragility of a glass window.

[21] 'Unlike [nanocages], GimC/prefoldin plays a more active role in protein folding by interacting with unfolded proteins and stabilizing them against aggregation for subsequent folding' (Siegers *et al*. [2005], p. 756).

altered amino acid components in a polypeptide, which should as a consequence prevent correct folding, correct folding can be induced by the overproduction of molecular chaperones (Maisnier-Patin *et al*. [2005]).

5 Philosophical Evaluation

According to Sarkar ([1998], pp. 169–70), the claim that correct folding is due to the linear order of amino acid components in a protein might fail in two ways: (1) given the laws of macromolecular physics, the linear order of the amino acid residues may be insufficient to explain the three-dimensional conformation of the folded protein, and/or (2) the laws of macromolecular physics may be insufficient to explain the three-dimensional conformation of the folded protein. Because the LSH focuses on the adequacy of linear amino acid structure for protein folding *assuming* macromolecular physics, we ignore (2) and concentrate on (1) to show the applicability and relevance of the three aspects described earlier.

Explanations become reductive when they conform to certain constraints, which depend on our explanatory goals. A reductive explanation may violate one constraint and not another since there may be more than one constraint in operation. Two constraints play a role in the context of explaining protein folding in terms of the LSH. The first is that only properties of the parts are required for a part–whole reductive explanation of protein folding (intrinsicality). Contextual or extrinsic causal factors besides the amino acids and their interactions are not supposed to play a role or contribute to correct folding. The second constraint is about what counts as the fundamental or reducing realm (Sarkar [1998], Chapter 3). In this case, it is the amino acids (parts), as well as macromolecular laws that describe their interactions. System properties due to a complex three-dimensional structure that are absent from the linear polypeptide are not fundamental. This is a qualified sense of fundamentality because properties of three-dimensional protein structure governed by macromolecular laws are excluded.

5.1 Application: intrinsicality and fundamentality

The folding construal of the LSH is a causal (temporal) part–whole reductive explanation. Tertiary structure of a three-dimensional protein whole is explained by the interaction of its component parts at earlier times; the amino acid residues interact causally to bring about the state of the whole ('correctly folded'). In the context of the PFP, a property is intrinsic to the linear polypeptide if it is a property of one of its amino acid components. (Relational properties between amino acid residues count as intrinsic.) Extrinsic properties are anything in the 'normal physiological milieu',

inclusive of physico-chemical components, such as H_2O, as well as other proteins (e.g. chaperones) or nucleic acids (e.g. RNA). This accords with standard accounts of the LSH: 'The specific function of a protein [or functional shape] is an emergent property that arises from the architecture of the molecule [. . .] the information for building specific shape is intrinsic in the protein's primary structure' (Campbell and Reece [2002], p. 78). The folding construal of the LSH is a claim about how properties of a linear polypeptide confer a disposition to fold into a three-dimensional protein; the disposition manifestation of folding is purely a function of the intrinsic properties (causal powers) of the linear polypeptide.

The significance of molecular chaperones now becomes apparent; if their activity contributes necessary properties required for folding and is not just appropriate environmental background, then the folding construal of the LSH is false. 'The manner in which a newly synthesized chain of amino acids transforms itself into a perfectly folded protein depends both on the intrinsic properties of the amino-acid sequence and on multiple contributing influences from the crowded cellular milieu' (Dobson [2003], p. 884). The intrinsic properties of the linear polypeptide arising from its amino acid residue parts are not sufficient to explain the protein-folding manifestation in the cell. Chaperones do not merely trigger the disposition manifestation but contribute specifically to its manifestation. The temporally extended process of folding not only requires appropriate environmental conditions but also the contribution of extrinsic chaperones; i.e. there is a failure with respect to the aspect of intrinsicality. Additionally, the causal contribution of chaperones in protein folding is a consequence of three-dimensional structure, a kind of property that the amino acid parts clearly lack. Thus, the best explanation of protein folding also involves a failure with respect to the aspect of fundamentality (case IV in Table 1). Systems with properties due to complex three-dimensional structure, rather than systems that lack it, are necessary to produce the native conformations of proteins *in vivo*.

Recall that there are two basic ways for a temporal part-whole reduction to fail as a *reductive* explanation (Table 1): either the behavior of an *extrinsic* element from the fundamental realm is necessary (intrinsicality fails – case III) or the behavior of an extrinsic system with a *non-fundamental* property is necessary (intrinsicality and fundamentality fail – case IV).[22] Case III would obtain if the relevant extrinsic properties were not folded proteins but rather free floating amino acid residues or unfolded polypeptide chains (elements of the fundamental realm). Chaperone function corresponds to case IV because another extrinsic system with a non-fundamental property (a folded,

[22] Fundamentality alone cannot fail because if intrinsicality holds then the parts must be instantiated at a more fundamental level than the whole (see footnote 12).

tertiary-structured protein) is required for the parts of a new linear polypeptide chain to become a folded protein. Systems with non-fundamental properties are necessary to bring about the native conformations of proteins *in vivo*. The parts alone in combination with the macromolecular laws of composition are not enough—temporal part–whole reduction fails with respect to both aspects as a *reductive* explanation.[23]

But couldn't a 'reductionist' adopt the rebuttal to the context objection here ('just reduce the context also')? For example, chaperones are composed of 'parts' and therefore we can 'reduce' the operation of an extrinsic chaperone protein 'whole' to its parts. This is similar to the strategy of preserving a reduction by 'extending the mechanism' backwards in time (Delehanty [2005]). Unfortunately, the maneuver faces a basic difficulty in the present situation: chaperone proteins require other chaperone proteins for their own proper folding, so the attempt to reduce the extrinsic chaperone (or extend the mechanism) leads to a type of explanatory regress. According to the individuation schemes adopted by scientists, extrinsic, non-fundamental wholes (folded proteins/chaperones) are required for the proper folding of another whole (folded protein). A related objection is to suggest a new individuation scheme: the cell as a 'larger' whole contains the protein *and* the crowded cellular milieu, thereby making the molecular chaperones intrinsic and fundamental. But even if a causal part–whole explanation of the behavior of the *cell* is feasible in terms of *its* intrinsic parts, this would change the question of what parts and wholes are being reduced. It would still remain a fact that the folding of a *protein* cannot be explained solely in terms of *its* amino acid parts; both intrinsicality and fundamentality are violated because the features appealed to over and above the intrinsic parts in the original explanation are extrinsic and not located in the fundamental realm. Shifting to a larger whole simply changes the *explanandum* and does not undermine our assessment.

Although we have a better understanding of the process of protein folding, the explanation does not meet the expectations of early researchers who put forward the LSH (see above, Section 4.2). Scientists now recognize that the causal powers requisite for folding are not all contained within the parts of the linear polypeptide. Molecular chaperones are extrinsic, non-fundamental causes that make a necessary contribution in the folding process and therefore the folding construal of the LSH has not been vindicated. This is because chaperones are tertiary-structured proteins that have undergone correct folding themselves.

[23] This claim is relative to the individuation and decomposition of the system offered by scientists, and pertains to the process of *bringing about* the three-dimensional protein structure. Whether the pertinent causal powers of the molecular chaperones are truly novel *vis-à-vis* the causal powers of its parts (as discussed in metaphysical contexts) concerns constitutional reductionism in an atemporal sense rather than causal part–whole reductive explanation.

It might seem odd to emphasize that the folding construal of the LSH fails with respect to intrinsicality and fundamentality when all the factors that go into the explanation are non-mysterious (i.e. macromolecules of different types). While understandable, this misses the point of the present analysis and misunderstands the original reductionist hypothesis for protein folding. The goal was not to show *that* reductive explanations fail but illuminate *how* they might succeed or fail in different ways. Our analysis provides resources for showing that the inability of intrinsic properties of the parts of a linear polypeptide to causally explain the folding process is a failure of causal reductive explanation *with respect to* intrinsicality and fundamentality. Precisely because only macromolecules are involved, this 'unintuitive' failure of a temporal part–whole reduction demonstrates why a failure of reduction is not necessarily a failure of explanation, or a capitulation to enigmatic non-physical factors.

5.2 Relevance: temporality

The conceptual distinctions we have developed are applicable regardless of whether the folding construal of the LSH is vindicated because intrinsicality, fundamentality, and temporality are aspects of reductive explanations in biology. A further advantage of our account is a perspicuous reconstruction of the scientific discourse about protein folding. Our analysis connects directly with the vocabulary used by researchers: 'there is a need for molecular chaperones because the intrinsic properties of proteins assure that incorrect interactions are possible' (van der Vies *et al*. [1993], p. 73). Including temporality facilitates an interpretation of paradoxical statements made by protein-folding researchers:

> At first glance, the concept [of chaperone assisted folding] is in conflict with the paradigm of autonomous, spontaneous protein folding established by Anfinsen and many others. The apparent contradiction is that if the acquisition of the unique three-dimensional structure of a protein is governed by its amino acid sequence and the resulting interactions between amino acid side chains, there should be no need for molecular chaperones. However, additional factors have to be taken into account. [. . .] Molecular chaperones do not provide steric information for the folding process and thus do not violate the concept of autonomous folding. (The information for folding is encoded solely in the amino acid sequence.) (Buchner and Walter [2005], pp. 163–4)[24]

[24] This tension has been present for more than a decade. 'The early observations made by Anfinsen and colleagues that denatured ribonuclease refolds spontaneously in the absence of any other proteins into an active enzyme, formed the basis for the [. . .] hypothesis [. . .] that all the information required to assemble a polypeptide chain into a biological three-dimensional structure is contained within the amino acid sequence [. . .]. Results of studies in a number of different experimental systems have recently led to the realization that protein assembly *in vivo* is more

The paradox of chaperones threatening to falsify the LSH alongside the claim that the amino acid sequence 'contains all the relevant information' is a result of the ambiguity between the predictive and folding construals of the LSH (Section 4.2). Focusing on the aspect of temporality prevents this conflation and removes the paradox.

Chaperone-assisted folding is in conflict with the LSH construed as a causal process *in vivo* because it is a claim about temporal (causal) part–whole reduction. The key role played by extrinsic, non-fundamental entities means that the amino acid parts are inadequate to explain appropriate folding in the cellular context—intrinsicality and fundamentality are violated in the context of the folding construal only. But chaperone-assisted folding is not in conflict with the predictive construal of the LSH, which is not a claim about *causal* part–whole reduction. Rather, it concerns the ability to infer or predict the native conformation of a protein from the linear polypeptide *composition*.[25] Thus, our analysis of causal part–whole reductive explanation is *relevant* to the discourse of scientists working on the PFP. It indicates why the failure of the folding construal due to chaperone-assisted folding is compatible with the successes associated with the predictive construal, which involves a separate community of researchers with distinct modeling practices (Ramsey [2007]). Robust inferences of three-dimensional conformation from linear sequence information are possible in the absence of details about the causal (temporal) process of protein folding in the cell, and researchers focused on the predictive construal need not deny that extrinsic molecular chaperones are necessary and specific causal factors for protein folding *in vivo*.

Our analysis of reductive explanation has the added advantage of providing increased comprehension of philosophical claims about reduction in biology. The ambiguity of the LSH is contained within Rosenberg's discussion of protein folding (see above, Section 2.1), and is indicated in the switch between *causal* (temporal) *determination* of folding by the amino acid sequence in the polypeptide and the determination of native conformation by the amino acid *composition* of the polypeptide. Causal (temporal) determination corresponds to the folding construal of the LSH and involves causal part-whole reduction. But the determination of protein conformation by amino acid composition

complex than was originally thought, and requires the involvement of other proteins that have collectively been termed molecular chaperones' (van der Vies *et al.* [1993], p. 72; cf. Jaenicke [2005]). Some researchers disagree with the claim that chaperones never provide steric information (e.g. Ellis [1998]).

[25] One reason why the predictive construal is successful despite the failure of causal part–whole reduction is because the former *assumes* the extrinsic features in the models used to predict three-dimensional conformation. This is seen in Anfinsen's articulation of the LSH (see above, Section 4.2), as well as in more recent research. Some models that predict the 'kinetics' of protein-folding incorporate extrinsic causal contributions (e.g. denaturant concentration) by adjusting pre-exponential factors (Fierz and Kiefhaber [2005], p. 840).

corresponds to the predictive construal of the LSH. The question of whether native protein conformation can be inferred from the linear polypeptide is distinct from the question of whether there is a *causal* part–whole reduction of how the folding occurs. An affirmative answer to one does not imply an affirmative answer to the other. The violation of intrinsicality and fundamentality in a causal part–whole reduction does not entail non-supervenience (or non-identity)—no controversial mereological claim must be made (e.g. the folded protein is something over and above its amino acid components).

6 Conclusion

The goal of this article was to explore aspects of part–whole reductive explanation that were both applicable to actual scientific reasoning and relevant to ongoing research. The applicability of our analysis is illustrated by how the different kinds of reductive explanations that emerge by considering the aspects of intrinsicality, fundamentality, and temporality correspond to the reasoning practices of working scientists. The neglect of these aspects was attributed to specific features of philosophical thought about reduction over the past few decades. The relevance of our analysis is observable in how it characterizes the folding construal of the LSH and keeps the predictive construal separate. Distinguishing intrinsicality, fundamentality, and temporality is useful for categorizing and prosecuting further research. For example, many diseases arise from the misfolding and aggregation of proteins, which is the province of the folding construal of the LSH where causal part–whole reduction is appropriate (Dobson [2003]; Mu *et al*. [2008]). Distinguishing the three aspects of reductive explanation also captures differences in reasoning between physics and biology, thereby explaining why the appropriation of 'atemporal' physical science reasoning for comprehending 'temporal' biological reasoning is often problematic.

Reductionism debates in the philosophy of biology have focused on theory reduction and its formal strictures for too long. Both the antireductionist consensus and attempts to argue for reductionism in terms of supervenience fall into this tradition. We have argued that a conception of reduction that involves the distinct aspects of intrinsicality, fundamentality, and temporality in the context of part–whole relations is a fruitful approach to analyzing reductionist reasoning in biology (and elsewhere). Our perspective apprehends how the folding construal of the LSH is a failure of reductive explanation in the context of the PFP, even though it is not a failure of reductionism in the more traditional sense or a failure of explanation *per se*.

The arc of reductionist research into the mechanisms of protein folding conforms to general claims about the development of scientific research

programs (Wimsatt [1980], [1997]). An initial choice is made to divide a system, the linear polypeptide, from its environment and then research concentrates on the entities and interactions that are *intra*systemic (the amino acid residues) to explain protein folding. Simultaneously, environmental conditions are simplified (e.g. *in vitro* denaturation experiments). These heuristic biases of reductionist methods could be cataloged more precisely, but the crucial point is how researchers responded to their lack of success—they learned from it. Reductionism is honored in the breach; the failure of causal (temporal) part–whole explanation occurred by way of prosecuting a form of methodological reductionism.[26] The initial success of the LSH arose in part from a bias against perceiving *extrasystemic* features (deemed extrinsic in the original demarcation of the system) as significant causal factors in the process of folding. The failure of reductive explanations with respect to intrinsicality and fundamentality did not arise from *anti*reductionism, but because a successful non-reductive explanation of protein folding was secured.

Acknowledgements

Order of authorship is alphabetical; both authors contributed equally to this paper. We are grateful to the comments and feedback we have received from many individuals (Ingo Brigandt, Tom Doyle, Susan Hawthorne, Marie Kaiser, Peter McLaughlin, Ken Schaffner, Ken Waters, Marcel Weber, and four anonymous referees) and at various venues (Department of Philosophy, St. Cloud State University; Seminar on 'Reductionism in Biology' at the University of Minnesota; Department of Philosophy, University of Münster; Seminar on 'Philosophy of Biology' at the University of Münster; 'Emergence and Reduction in the Sciences' [Second Pittsburgh-Paris Workshop], Center for Philosophy of Science, University of Pittsburgh; Center for Philosophy and Ethics of Science, University of Hannover; Minnesota Center for Philosophy of Science, Biology Interest Group at the University of Minnesota; Workshop on 'Explanation, Confirmation and Prediction in Biology and Medicine' at the University of Konstanz). We also want to express our appreciation for support from the Alexander von Humboldt Foundation and the American Philosophical Society to pursue this collaboration, which grew out of our mutual participation in the Second German-American Frontiers of Humanities Symposium, Hamburg,

[26] 'We tend to start with simple models of complex systems—models according to which the parts are more homogeneous, have simpler interactions, and in which many differentiated parts and relationships are ignored. [. . .] But then as our models grow in realism, we should both capture more properties and see more of them as organization dependent' (Wimsatt [1997], p. S382).

Germany, October 2005. Funded by an Alexander von Humboldt Foundation CONNECT Fellowship (to A.H. and to A.C.L.) and the University of Minnesota McKnight Land-Grant Professorship (to A.C.L.).

Philosophisches Seminar,
Universität zu Köln, Albertus-Magnus-Platz 1, 50923 Köln, Germany
ahuettem@uni-koeln.de

Department of Philosophy and Minnesota Center for Philosophy of Science,
University of Minnesota, 831 Heller Hall,
271 19th Ave. S, Minneapolis, MN 55455, USA
aclove@umn.edu

References

Albanese, V., Yam, A. Y.-W., Baughman, J., Parnot, C. and Frydman, J. [2006]: 'Systems Analyses Reveal Two Chaperone Networks with Distinct Functions in Eukaryotic Cells', *Cell*, **124**, pp. 75–88.

Anfinsen, C. B. [1973]: 'Principles That Govern the Folding of Protein Chains', *Science*, **181**, pp. 223–30.

Bechtel, W. and Richardson, R. [1993]: *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton: Princeton University Press.

Bickle, J. [2003]: *Philosophy and Neuroscience: A Ruthlessly Reductive Account*, Dordrecht: Kluwer Academic Publishers.

Brigandt, I. and Love, A. C. [2008]: 'Reductionism in Biology', in E. N. Zalta (*ed.*), *Stanford Encyclopedia of Philosophy*, <plato.stanford.edu/archives/fall2008/entries/reduction-biology/>.

Buchner, J. and Walter, S. [2005]: 'Analysis of Chaperone Function *in vitro*', in J. Buchner and T. Kiefhaber (*eds*), *Protein Folding Handbook. Part II (Volume 1)*, Weinheim: WILEY-VCH Verlag, pp. 162–96.

Campbell, N. A. and Reece, J. B. [2002]: *Biology*, 6th edition, San Francisco, CA: Benjamin Cummings.

Cartwright, N. [1999]: *The Dappled World: A Study of the Boundaries of Science*, New York: Cambridge University Press.

Clark, P. L. [2004]: 'Protein Folding in the Cell: Reshaping the Folding Funnel', *Trends in Biochemical Sciences*, **29**, pp. 527–34.

Craver, C. F. and Bechtel, W. [2007]: 'Top-down Causation without Top-down Causes', *Biology and Philosophy*, **22**, pp. 547–63.

Crick, F. H. C. [1958]: 'On Protein Synthesis', in *The Biological Replication of Macromolecules, Symposia of the Society for Experimental Biology XII*, Cambridge: Cambridge University Press, pp. 138–63.

Delehanty, M. [2005]: 'Emergent Properties and the Context Objection to Reduction', *Biology and Philosophy*, **20**, pp. 715–34.

Dobson, C. M. [2003]: 'Protein Folding and Misfolding', *Nature*, **426**, pp. 884–90.

Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E. and Kern, D. [2005]: 'Intrinsic Dynamics of an Enzyme Underlies Catalysis', *Nature*, **438**, pp. 117–21.

Ellis, R. J. [1998]: 'Steric Chaperones', *Trends in Biochemical Sciences*, **23**, pp. 43–5.

Ellis, R. J. [2001]: 'Macromolecular Crowding: Obvious but Underappreciated', *Trends in Biochemical Sciences*, **26**, pp. 597–604.

Feder, M. E. and Hofmann, G. E. [1999]: 'Heat-Shock Proteins, Molecular Chaperones, and the Stress Response: Evolutionary and Ecological Physiology', *Annual Review of Physiology*, **61**, pp. 243–82.

Fierz, B. and Kiefhaber, T. [2005]: 'Dynamics of Unfolded Polypeptide Chains', in J. Buchner and T. Kiefhaber (*eds*), *Protein Folding Handbook. Part I (Volume 2)*, Weinheim: WILEY-VCH Verlag, pp. 809–55.

Fodor, J. [1974]: 'Special Sciences (or: the Disunity of Sciences as a Working Hypothesis)', *Synthese*, **28**, pp. 77–115.

Freedman, R. B. [1999]: 'Protein Folding in the Cell', in T. E. Creighton (*ed.*), *Protein Folding*, New York: W.H. Freeman and Company, pp. 455–539.

Frydman, J. [2001]: 'Folding of Newly Translated Proteins *in vivo*: The Role of Molecular Chaperones', *Annual Review of Biochemistry*, **70**, pp. 603–47.

Gillett, C. [2003]: 'The Metaphysics of Realization, Multiple Realizability, and the Special Sciences', *Journal of Philosophy*, **100**, pp. 591–603.

Hoffmann, V. and Newen, A. [2007]: 'Supervenience of Extrinsic Properties', *Erkenntnis*, **67**, pp. 305–19.

Homouz, D., Perham, M., Samiotakis, A., Cheung, M. S. and Wittung-Stafshede, P. [2008]: 'Crowded, Cell-like Environment Induces Shape Changes in Aspherical Protein', *Proceedings of the National Academy of Sciences of the United States of America*, **105**, pp. 11754–9.

Honig, B. [1999]: 'Protein Folding: From the Levinthal Paradox to Structure Prediction', *Journal of Molecular Biology*, **293**, pp. 283–93.

Hove, J. R., Köster, R. W., Forouhar, A. S., Acevedo-Bolton, G., Fraser, S. E. and Gharib, M. [2003]: 'Intracardiac Fluid Forces are an Essential Epigenetic Factor for Embryonic Cardiogenesis', *Nature*, **421**, pp. 172–7.

Hull, D. [1974]: *Philosophy of Biological Science*, New Jersey: Prentice-Hall Inc.

Hüttemann, A. [2005]: 'Explanation, Emergence, and Quantum Entanglement', *Philosophy of Science*, **72**, pp. 114–27.

Jaenicke, R. [2005]: 'Paradigm Changes from "Unboiling an Egg" To "Synthesizing a Rabbit"', in J. Buchner and T. Kiefhaber (*eds*), *Protein Folding Handbook. Part II (Volume 1)*, Weinheim: WILEY-VCH Verlag, pp. 3–31.

Kærn, M., Elston, T. C., Blake, W. J. and Collins, J. J. [2005]: 'Stochasticity in Gene Expression: From Theories to Phenotypes', *Nature Reviews Genetics*, **6**, pp. 451–64.

Kim, J. [1998]: *Mind in a Physical World*, Cambridge, MA: The MIT Press.

Kitcher, P. [1984]: '1953 and All That: A Tale of Two Sciences ', *Philosophical Review*, **93**, pp. 335–73.

Levinthal, C. [1969]: 'How to Fold Graciously', in P. Debrunner, J. C. M. Tsibris and E. Münck (*eds*), *Mossbauer Spectroscopy in Biological Systems*, Urbana: University of Illinois Press, pp. 22–4.

Liscalijet, I. M., Kleizen, B. and Braakmen, I. [2005]: 'Studying Protein Folding *in vivo*', in J. Buchner and T. Kiefhaber (*eds*), *Protein Folding Handbook. Part II (Volume 1)*, Weinheim: WILEY-VCH Verlag, pp. 73–104.

Machamer, P., Darden, L. and Craver, C. F. [2000]: 'Thinking About Mechanisms', *Philosophy of Science*, **67**, pp. 1–25.

Maisnier-Patin, S., Roth, J. R., Fredriksson, A., Nystrom, T., Berg, O. G. and Andersson, D. I. [2005]: 'Genomic Buffering Mitigates the Effects of Deleterious Mutations in Bacteria', *Nature Genetics*, **37**, pp. 1376–9.

Mayr, E. [1988]: *Toward a New Philosophy of Biology: Observations of an Evolutionist*, Cambridge, MA: Harvard University Press.

McClellan, A. J., Scott, M. D. and Frydman, J. [2005]: 'Folding and Quality Control of the Vhl Tumor Suppressor Proceed through Distinct Chaperone Pathways', *Cell*, **121**, pp. 739–48.

McKitrick, J. [2003]: 'A Case for Extrinsic Dispositions', *Australasian Journal of Philosophy*, **81**, pp. 155–74.

Menzies, P. [1988]: 'Against Causal Reductionism', *Mind*, **97**, pp. 551–74.

Mitchell, S. D. [2000]: 'Dimensions of Scientific Law', *Philosophy of Science*, **67**, pp. 242–65.

Morange, M. [1998]: *A History of Molecular Biology*, Cambridge, MA: Harvard University Press.

Mu, T. W., Ong, D. S., Wang, Y. J., Balch, W. E., Yates, J. R., Segatori, L. and Kelly, J. W. [2008]: 'Chemical and Biological Approaches Synergize to Ameliorate Protein-folding Diseases', *Cell*, **134**, pp. 769–81.

Nagel, E. [1961]: *The Structure of Science: Problems in the Logic of Scientific Explanation*, New York: Harcourt, Brace & World, Inc.

Ramsey, J. [2007]: 'Calibrating and Constructing Models of Protein Folding', *Synthese*, **155**, pp. 307–20.

Richardson, J. S. [1982]: 'What do the Folds in Proteins Look Like?', in D. B. Wetlaufer (*ed.*), *The Protein Folding Problem*, Boulder, CO: Westview Press, pp. 1–28.

Robert, J. S. [2004]: *Embryology, Epigenesis, and Evolution: Taking Development Seriously*, New York: Cambridge University Press.

Rosenberg, A. [1978]: 'The Supervenience of Biological Concepts', *Philosophy of Science*, **45**, pp. 368–86.

Rosenberg, A. [1985]: *The Structure of Biological Science*, Cambridge: Cambridge University Press.

Rosenberg, A. [1994]: *Instrumental Biology or the Disunity of Science*, Chicago: University of Chicago Press.

Rosenberg, A. [2006]: *Darwinian Reductionism: Or, How to Stop Worrying and Love Molecular Biology*, Chicago: University of Chicago Press.

Sarkar, S. [1998]: *Genetics and Reductionism*, New York: Cambridge University Press.

Schaffner, K. [1967]: 'Approaches to Reduction', *Philosophy of Science*, **34**, pp. 137–47.

Schaffner, K. [1969]: 'The Watson-Crick Model and Reductionism', *British Journal for the Philosophy of Science*, **20**, pp. 325–48.

Schaffner, K. [1974]: 'The Peripherality of Reductionism in the Development of Molecular Biology', *Journal of the History of Biology*, **7**, pp. 111–39.

Schaffner, K. [1976]: 'Reductionism in Biology: Prospects and Problems', in R. S. Cohen, C. A. Hooker, A. C. Michalos and J. W. van Evra (*eds*), *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1974, Dordrecht: D. Reidel, pp. 613–32.

Schaffner, K. [1993]: *Discovery and Explanation in Biology and Medicine*, Chicago: University of Chicago Press.

Schaffner, K. [2006]: 'Reduction: The Cheshire Cat Problem and a Return to Roots', *Synthese*, **151**, pp. 377–402.

Siegers, K., Bracher, A. and Hartl, F. U. [2005]: 'Structure and Function of Gimc/Prefoldin', in J. Buchner and T. Kiefhaber (*eds*), *Protein Folding Handbook. Part II (Volume 2)*, Weinheim: WILEY-VCH Verlag, pp. 756–67.

Sober, E. [1999a]: 'Physicalism from a Probabilistic Point of View', *Philosophical Studies*, **95**, pp. 135–74.

Sober, E. [1999b]: 'The Multiple Realizability Argument against Reductionism', *Philosophy of Science*, **66**, pp. 542–64.

Sterelny, K. and Griffiths, P. E. [1999]: *Sex and Death: An Introduction to the Philosophy of Biology*, Chicago: University of Chicago Press.

Tanford, C. and Reynolds, J. [2001]: *Nature's Robots: A History of Proteins*, New York: Oxford University Press.

Tang, Y. -C., Chang, H. -C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M. J., Hartl, F. U. and Hayer-Hartl, M. [2006]: 'Structural Features of the GroEL-GroES Nano-Cage Required for Rapid Folding of Encapsulated Protein', *Cell*, **125**, pp. 903–14.

Tang, Y. -C., Chang, H. -C., Chakraborty, K., Hartl, F. U. and Hayer-Hartl, M. [2008]: 'Essential Role of the Chaperonin Folding Compartment *in vivo*', *EMBO Journal*, **27**, pp. 1458–68.

van der Vies, S. M., Gatenby, A., Viitanen, P. V. and Lorimer, G. H. [1993]: 'Molecular Chaperones and Their Role in Protein Assembly', in J. L. Cleland (*ed.*), *Protein Folding: in vivo and in vitro*, Washington, DC: American Chemical Society, pp. 72–83.

Waters, C. K. [1990]: 'Why the Antireductionist Consensus Won't Survive the Case of Classical Mendelian Genetics', in A. Fine, M. Forbes and L. Wessels (*eds*), *PSA: Proceedings of the Philosophy of Science Association*, Volume 1, East Lansing, MI: Philosophy of Science Association, pp. 125–39.

Weber, M. [2005]: *Philosophy of Experimental Biology*, New York: Cambridge University Press.

Wilson, R. A. and Craver, C. F. [2007]: 'Realization', in P. Thagard (*ed.*), *Philosophy of Psychology and Cognitive Science (Handbook of the Philosophy of Science, Volume 12)*, Amsterdam: Elsevier/North Holland, pp. 31–54.

Wimsatt, W. C. [1976]: 'Reductive Explanation: A Functional Account', in R. S. Cohen, C. A. Hooker, A. C. Michalos and J. W. van Evra (*eds*), *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*; 1974, Dordrecht: D. Reidel, pp. 671–710.

Wimsatt, W. C. [1980]: 'Reductionistic Research Strategies and Their Biases in the Units of Selection Controversy', in T. Nickles (*ed.*), *Scientific Discovery: Case Studies*, Dordrecht: D. Reidel, pp. 213–59.

Wimsatt, W. C. [1997]: 'Aggregativity: Reductive Heuristics for Finding Emergence', *Philosophy of Science*, **64**, pp. S372–84.

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