

# Biochemical Kinds

**Jordan Bartol**  
*University of Leeds*

*Chemical kinds (e.g. gold) are generally treated as having timelessly fixed identities. Biological kinds (e.g. goldfinches) are generally treated as evolved and/or evolving entities. So what kind of kind is a biochemical kind? This paper defends the thesis that biochemical molecules are clustered chemical kinds, some of which—namely, evolutionarily conserved units—are also biological kinds. On this thesis, a number of difficulties that have recently occupied philosophers concerned with proteins and kinds are shown to be resolved or dissolved.*

1. Introduction
2. Conflicting Intuitions About Kinds of Proteins
3. Against Permissive Pluralism
4. Against Nominalism, Toward a Duality of Kinds
5. Biological Kinds, Chemical Kinds, and Their Relations
6. Implications for Biological Individuals
7. Implications for Monism and Scientific Practice

## 1. Introduction

Protein molecules are an interesting case for philosophy of science because they are at once the objects of biology and chemistry, an in-between status that leads to conflicting intuitions. The first is that, qua chemical molecules, their physical structure is fundamental. The second is that, qua biological objects, their physiological roles are important to recognise and understand. The conflict manifests in a number of ways, but the principal problem is the classification of proteins into kinds. Consider the lens crystallin protein, which forms the lens of your eye but also ‘moonlights’ as a number of functionally distinct enzymes. A structural classification, like that used for simpler chemical molecules, will gloss over this biological diversity. A functional classification will separate these proteins and so fail to highlight structural similarities. Privileging one of these classifications of another appears at best subjective and at worst arbitrary.

Two issues arise. First, we might ask after the actual scientific practices of classification, focussing on how scientists can, should, or do address this problem. Caught up with this inquiry we find a mix of questions about the aims, norms, contexts, and limitations of investigation. Call this ‘the epistemic

question' about protein classification. Second, we might ask after the status of proteins as natural kinds. Is there a natural or correct arrangement of proteins into kinds, no such arrangement, or many? In case there are natural kinds, we want to know what sorts of kinds they are and how they relate to one another. This second line of inquiry asks after the ontological structure of the biomolecular world and the metaphysical relations therein. Call this 'the metaphysical question' about protein classification. Recent work on the epistemic question leads me to turn my attention to the metaphysical one. The present paper asks, in light of what we know about proteins and biochemistry, what can we say about nature's joints?

William Goodwin (2011) recently argued that the practice of protein taxonomy is 'fundamentally' grounded in structural considerations, which are then adjusted ad hoc as dictated by specific phenomena and scientific interests. Call this position 'pragmatic pluralism' about classificatory practice. Goodwin resists this label, but my use of it will become clear as we proceed. Pragmatic pluralism about taxonomic practice offers four metaphysical interpretations. First is monism; second is nominalism; third is what I will call permissive pluralism; fourth is a more conservative or restricted pluralism. Neither biologists nor philosophers hold out hope for a tidy ontological reduction to either biological or chemical kinds. This rules-out metaphysical monism. Nominalism and the two pluralisms remain. On its own, Goodwin's account privileges no particular interpretation. Two other recent papers, by Matt Slater (2009) and Emma Tobin (2010), also call for pragmatic pluralisms but invite metaphysical conclusions along side epistemic ones. Slater and Tobin's descriptions point toward a highly permissive brand of pluralism. In addition to the conflict between biological and chemical classification, these philosophers focus on the physical underdetermination between a protein's initial amino acid sequence (called 'primary structure') and its final folded three-dimensional state (called 'conformation'). Their problematisation is consistent with many scientific accounts (e.g. Copley 2012; Dyson & Wright 2005). If accurate, these accounts point to the conclusion that there are innumerable many ways to classify proteins; possibly as many ways as there are structural properties. This is what I mean by 'permissive' pluralism. In section 3, I argue that physical underdetermination is in fact not a problem. Microstructuralist accounts of chemical kinds are well equipped to treat higher-level chemical structures as constrained by the lower-level microstructure. The pressing issue remains the multiple-realisation of physiological roles by chemical microstructures.

With the conflict thusly framed, I dispense with the nominalist option as unmotivated and introduce my position, which is a far more restricted pluralism. I motivate a theory of two kinds: the biological and the chemical. Though the disconnect between structure and function is instructive, it would be a mistake to identify biological protein kinds with their functions. There are a number of well-known problems with functional kinds, and function appears to be an accidental property of chemical structures rather than a necessary property of any kinds. This becomes clear when we conceive of biological protein kinds as Darwinian. On the proposed view, chemical kinds are best viewed as pieces or tools that are picked up, shuffled, and recombined and sometimes acquire physiological functions. These pieces are occasionally conserved through evolution. It is through this process that biological kinds emerge. Biological kinds are what they are in virtue not of their physical structure, but their evolutionary history. Within that history chemical structures have some influence

on outcomes, but biological kinds are ultimately created and shaped by their histories of contingency, chance, and selection.

Separating the proposed duality from the permissive pluralist picture is important because the former is more fruitful. While the permissive pluralist position is forced to view different classifications as alternate ways of describing the same world, the theory I offer describes two fundamentally different kinds and their interactions. This account paves the way for more general discussions about the differences between the kinds of the biological world and those of the physico-chemical world.

Implications for metaphysical monism, on the one hand, and for biochemical practice, on the other, will come into full view in the concluding section. To begin, I will introduce the two intuitions and their conflict, which form the backdrop of my analysis.

## 2. Conflicting Intuitions About Kinds of Proteins

At base, the conflict between the biological and chemical is a clash of intuitions. Respecting both aspects of proteins comes at the expense of a single consistent classificatory scheme. While practical workarounds can and have been found, these might trouble the monist about classificatory metaphysics. I begin with the pre-theoretic conflict between these biological and chemical intuitions before expounding the precise nature of the relationship between these two sides of the protein world.

The biological intuition has a long history in the sciences that study proteins. This tradition emphasizes the importance of proteins in physiology. Though many of proteins' biological roles are newly discovered, their importance has long been recognized. The term 'protein' was coined in the mid-19th century from the Ancient Greek 'proteios' meaning 'primary' or 'in the lead', in order to emphasize their presumed essential role in micro-biological processes. Proteins are the most profuse macromolecule, occurring in all parts of all cells. Though they carry out a wide variety of functions and take on a staggering number of forms, all proteins are created from amino acids linked in linear sequences and then folded into complex shapes, called 'conformations'. There are at two varieties of protein. The first are fibrous proteins. These make up every tissue in organic bodies; common textbook examples include keratin and collagen. Second are globular proteins, which carry out important physiological roles as enzymes, antibodies, regulators, and cellular messengers. The importance of proteins in this regard should not be understated. Enzymes are necessary for the catalysation of nearly all organic chemical reactions and, as such, are involved in a wide variety of molecular biological processes—and this is not to downplay the biological importance of messengers, regulators, and antibodies.

Understanding protein function is a key part of understanding molecular biology. Not only do they comprise all organic bodies and play key roles in organic reactions, they have recently become the subject of re-focussed interest for their role in molecular evolution. Adaptations from the development of anatomy to alteration in metabolic processes involve changes at the protein level. As a result, a key tool in uncovering the progeny of extant physiology is the study of the proteins involved. Specifically, biochemists study the semi-autonomous 'domains' from which proteins are compiled. Conserved domains are shuffled, recombined, duplicated, and changed to carry out new functions. Tracking domains allows us to map the evolution of new traits. Understanding their physiology is key

to understanding ancestral functions of current protein molecules, which is key to understanding the process of molecular evolution.

It is fair to say that not only are proteins fascinating for their chemical complexity, but also for their biological significance, on multiple fronts. Any description of protein kinds should respect this—so the intuition goes.

The chemical intuition, on the other hand, is an extension of standard thinking about natural kinds in science, which has long been structured by treatments of chemical kinds. Indeed, the gold standard case of a natural kind is a chemical one: gold. Nearly all introductions begin with this elemental example. Even Marc Ereshefsky's (2009) reference article 'Natural Kinds in Biology' introduces kinds not with a biological example, but with the familiar chemical:

The traditional account of natural kinds asserts that the members of a kind share a common essence. The essence of gold, for example, is its unique atomic structure. That structure occurs in all and only pieces of gold, and it is a property that all pieces of gold must have.

Paul Griffiths (1999) similarly explains:

My gold watch resembles your gold navel ring [...] because the atoms of which both are composed share an essence: their atomic number. (p. 209)

The received view of chemical kinds is microstructural essentialism. Microstructuralism is a brand of reductive physicalism, which holds that objects can be individuated and explained solely in terms of their physical microstructural properties (Hendry 2006). Essentialism<sup>1</sup> is the claim that natural kinds are knit together in virtue of a shared property. Combined, these commit the natural kind theorist to taking microstructure as the essential property. Both Ereshefsky and Griffiths, referencing the atomic structure of gold, align themselves with this tradition. Simple chemical objects make ideal examples of kinds because they are neatly divisible and admit a clear physical reduction of their macro-level properties to some physical microstructure. More importantly, microstructure is a unique determinate of identity since any instance of a chemical kind cannot lose its microstructure without changing kinds.

Though elements provide the simplest cases, we can extend theories of chemical kinds to more complex molecules. Robin Hendry (2006) has recently shown how microstructuralist accounts can be scaled up from elements to molecules, arguing that just as the essence of Gold is represented by atomic number 79, so too might we identify the essence of carbon dioxide with its constituent atomic elements, represented by the formula CO<sub>2</sub>. This microstructure is causally efficacious, explains the relevant properties of carbon dioxide, and is necessary and sufficient for being carbon dioxide.

Microstructure becomes less clear as we move up the complexity scale. The microstructure of CO<sub>2</sub> might be its constituent atoms, the atoms and their bonds, or the atoms, their bonds, and their spatial relations. Chemists and biochemists describe a range of 'levels' of structural arrangements. The problem becomes considerably more complex when we get to proteins. Protein molecules are described at the level of primary structure, which includes just the linear sequence of amino acids;

---

<sup>1</sup> I do not wish this version of essentialism – the view that individuals have some shared property – to be confused with another view called essentialism – that individuals belong to their kinds *essentially*. The latter, owing much to Kripke, is concerned with the modality of kind membership and the reference of kind terms. The two are not the same and running them together causes great problems, (McOuat, 2009; Wilkins, 2013).

secondary structure, which describes stable recurring geometrical patterns in localised sections of the molecule; and tertiary and quaternary structures, which describe the geometric and bond structures of the whole molecule or the molecule plus bound partners, respectively. Forwarding a metaphysical thesis, the microstructuralist would do well not to attach to any one of these representations. For these are just that, representations, fallible attempts to capture some physical description of the protein. The physical facts that get included in a given representation are a function of goals and interests, but also of the context, since certain physical features are stable only in specific environments. Given present concerns, what is interesting about microstructuralism is the grounding of kinds in microphysical facts. How and whether we know or represent those facts is a separate matter. Rather than take a stance on which representation of microstructure is best, I will use the general phrase ‘chemical structure’.

There are several reasons why philosophers might expect and want a theory of chemical kinds to extend upward to proteins. First, from a purely physico-chemical point of view, proteins are simply very large chemical molecules; they are macromolecules. They can be annotated and described in much the same way as smaller molecules, but on a much larger scale.<sup>2</sup> If microstructuralism can handle one step up the complexity scale, from elements to molecules, then what’s a few more? Second, some may find it suggestive that microstructuralist individuation is a dominant method in protein classification today. When biochemists investigate proteins, they work largely at the levels of conformation and primary structure. This is the main way in which proteins are annotated and referenced. But the final and most philosophically forceful motivation behind extending the standard account of chemical kinds comes from the prospect of monism. As Slater (2009) explains,

Following the lead of natural kinds essentialists of old, one might suggest individuating proteins (and other biological macromolecules) on the basis of their chemical structure. At first glance, this stance affords a tempting monism about biochemical taxonomy. (p. 852)

Though he goes on to reject this possibility, the *prima facie* appeal of monism is worth understanding. The issue evinces a longstanding concern with the heterogeneity of the world investigated by the sciences. Monism has implications scientific realism (Chakravartty 2011), practical implications for the formulation of scientific laws (Millikan 1999), and more general ramifications for ontology. There is great appeal in the thought that molecules, atoms, and macromolecules are all fundamentally the same types of thing. A monist might envision a single (enormous) hypothetical taxonomy representing the varieties of chemical types, from hydrogen to ununocium, and on to molecules and proteins. Microstructuralism proffers one set of things with one type of kind membership conditions.

There are two senses of ‘monism’, here. First, microstructuralism offers monism in the fashion after which molecules are naturally individuated. Call this ‘category monism’ (CM), since it is monism about the ontological category, ‘kind’. Microstructuralism holds that all chemical kinds are what they are in virtue of microphysical facts. Contrast this with functional kinds, historical kinds, or relational kinds. If these other brands of natural kind were admitted to our ontology then we would have a plurality of kind categories. Second, microstructuralism offers the promise of a single taxonomy. Call this ‘taxonomic monism’. Every kind in the microstructuralist taxonomy is unique. There is no worry

---

<sup>2</sup> My point is not that proteins *are* described in the exact same way as smaller molecules, but that they can be. The chemical formula of a protein is far too long and cumbersome to be of any use in talking about proteins and so is not used. A higher level of description, focusing on component motifs and domains, is much more practicable.

about one particular belonging to multiple incompatible kinds.<sup>3</sup> Both brands of monism are on the table. Slater is right to be tempted. I will return to the prospect of monism in the final section.

Hopes for grounding protein identity in microstructure are not just idle metaphysics. The tradition has a corresponding scientific history. Scientists long presumed that chemical structure determined biological properties. The study of proteins was once dominated by reductionist ideology, which claimed that the function of a protein was determined by its three-dimensional structure, itself determined by the protein's molecular composition. This came to be known as the 'Sequence-Structure-Function' paradigm (SSF). When SSF stood strong there was no problem of protein kinds, since, according to SSF, there was a straightforward link between molecular composition and physiological function. Whether individuated structurally or functionally, the result should have been the same. Unfortunately these canonical beliefs have proved false. Though various phenomena have caused doubts about SSF, multifunctional (or 'moonlighting') proteins are taken by many to be the nail in the coffin—so much so that one researcher recently declared, 'Moonlighting is mainstream: Paradigm adjustment required!' (Copley 2012).

Multifunctional proteins are a heterogeneous class. As the name suggests, these are cases where 'one' protein performs multiple functions. The relations that make these count as 'one' protein vary, but in general proteins are considered the same when they share an amino acid sequence. In some cases, proteins with identical sequence adopt different folds in different contexts in order to carry out different functions. Differential folding serves to utilize different functional domains, associated with different tasks. In other cases, conformationally-identical proteins carry out different but related roles in the same physiological process (Copley 2012). Regardless of the particulars, proteins that share some chemical structure but differ in physiology have left scientists without a theory of protein kinds. It is unclear whether a protein is what it is in virtue of chemical structure, biological function, or something else entirely. The result, among other things, is a serious disjoint in competing classificatory techniques (see report from Carr et al. 2004) leading to poor understanding of when two proteins are or are not the same.

Though protein taxonomists still tend to classify microstructurally, microstructuralism alone does not undergird taxonomic practice. Using only microstructural classification results in unhelpful and counterintuitive classifications, such as cases where proteins that appear wildly different at the biological level are grouped together at the chemical level, and proteins that fulfil the same biological role yet are grouped apart due to chemical dissimilarity. Microstructural classification is thus supplemented with biological considerations, when appropriate, to correct these irregularities. Thus Goodwin's (2011) concession: though he wishes biochemical classification to be based on physico-chemical structure, biological facts must be accommodated, often ad hoc, by augmenting or supplementing structural classifications. He explains,

While there is a fundamental, structural way of individuating proteins, there are also supplemental classifications introduced to address various biological interests. (p. 537)

---

<sup>3</sup> Since it is often unclear why taxonomic monism is desirable, it is likewise unclear what it entails about the particulars in the taxonomy. The point here is that no single object will be two incompatible kinds. An atom cannot be both hydrogen and oxygen – nor can it be both hydrogen and water, even though a water molecule might be composed partly of hydrogen.

Structural information about proteins may come close to a biologically meaningful classification, but it must be adjusted to highlight pertinent biological similarities or differences.

The point can be made salient with the example mentioned at the beginning: the multifunctional protein family known as ‘crystallins’. Crystallins are the transparent structural proteins found in the lens and cornea of the eye. There are many varieties of crystallin but nearly all demonstrate some multifunctionality. In chickens and ducks  $\alpha\beta$ -crystallin forms the refractive surface on the lens of the eye, yet also occurs as a heat-shock protein and an enzyme, called a ‘lyase’. This is mirrored in many other animals. In birds and crocodilians the crystallin that forms the lens also doubles as the digestive enzyme lactate dehydrogenase. The  $\alpha$ -crystallin present in all vertebrate lenses also functions as a molecular chaperone and may have an enzymatic role in digestive processes (Copley 2012). Standard physico-chemical classification leads us to say that we have one protein, but intuitions about biological function lead us to conclude otherwise. These intuitions come out when Slater (2009) insists that protein kind classifications preserve ‘important biochemical facts’ about the molecules, something that structural definitions fail to do. Though the chemical facts are presumably explained, many functional (often physiological) ones are not. The desire for category monism is at odds with the desire to respect the biology.

Scientists can describe structural proteins with multiple biological roles, or biological proteins with multiple structures. But no single scheme will perfectly categorize both. Hybrid schemes are needed. Different taxonomies, and different types of kind category, are necessary under specific disciplinary circumstances. Though some communicative problems may result (Carr et al., 2004), these are presumably resolvable with more specific language or better databases for classification. Biochemists face no in-principle problem, having developed a rich epistemic system of interwoven classificatory practices, which change with contexts and aims. The situation is only problematic if we hold the belief that there is a ‘correct’ or ‘natural’ way to classify proteins—to carve nature at its joints—and that biochemistry should aim at this ideal, but misses for all its pragmatism.

In the final section, I will argue that scientific practice need not utilize a taxonomy of natural kinds. Yet without being normative about scientific practice, we can still ask what pragmatic pluralism in practice means for the metaphysical question. It is still possible to ask what structure of kinds is compatible with known phenomena and would support the pragmatic pluralism that characterizes taxonomic practice.

Pragmatic pluralists see multiple kind classifications as alternate and equally legitimate descriptions of the same entities. If there is no sense in which one such scheme is fundamental, or privileged, then there exist no grounds on which to say, ‘this classification describes how the kinds really are’. From here, there are still a number of answers to the metaphysical question. One option is nominalism. Perhaps pragmatic pluralism reveals the poverty and scholasticism of natural kinds talk altogether (Hacking 2007). Not only is there no single way that the kinds are, perhaps there is no way at all! Another option is highly permissive pluralism. Perhaps each type of kind category comes from a different taxonomy of natural kinds, and we simply pick and choose from different natural kind structures as situations dictate. On this somewhat deflationary view, there are as many natural kinds as there are natural properties from which to classify (see Chakravartty’s 2011 ‘socialbility based

pluralism' for a recent example of this type of view). A third option is to accept a more conservative pluralism. Perhaps there are not innumerable many kinds, just a select few.

### 3. Against Permissive Pluralism

Setting aside for the moment the nominalist option, the two realist pluralisms have similar appeal. Both concede the force of the pre-theoretic observation that one classification will not do. The choice between the two is a matter of just how many types of classification might lay claim to naturalness or primacy. Many treatments of proteins emphasize the physical underdetermination of final conformation. These accounts draw attention to the number of distinct 'levels' of structural arrangement of molecules, suggesting that each might be a unique physico-chemical kind. Such descriptions strongly legislate for permissive pluralism. This is a red herring.

Imagine a protein family where a single sequence of amino acids results in a number of distinct final conformations under different conditions. Many such cases exist. Classification according to primary structure would yield one scheme of kinds; classification according to final conformation would yield another. These cases are frustrating to practicing taxonomists and appear to have implications for the philosophical discussion, as well. As Tobin (2010) claims,

If two chemical kinds can have the same [microstructural] essence, yet are considered distinct at the macrostructural level, then the ... microstructure would seem insufficient for macromolecular classification. (p. 53)

On the surface, physical underdetermination appears equally problematic as the disconnect between biological and chemical classification. The claim is that these cases recommend a collection of different physico-chemical classification schemes, appealing to many different levels of structural organization.

Yet microstructuralism is well equipped to handle these cases. We can maintain that conformation is selected extrinsically from within a space of possibilities imposed by the intrinsic microstructure. Some comparatively simpler cases from chemistry will help clarify. Many molecules share a chemical composition but exist in different states at higher levels of structure. One such group are conformational isomers, called 'conformers', where one set of component atoms, with just one arrangement of bonds, can exist in multiple conformations. This happens in relatively simple molecules and is also a common feature of proteins. Conformational isomerism is a property of bond structures that permit movement, usually around single bonds, enabling multiple geometries. The particular geometry that obtains is a function of external forces; temperature is perhaps most commonly discussed. Certain possible conformations are stable under common conditions and these are the conformations that are recognized in practice. Given free reign over extreme temperatures, electrostatic forces, and other conditions scientists can bring about additional marginally-stable forms. The familiar n-butane ( $C_4H_{10}$ ) is commonly recognized to have two conformers (trans and gauche), but these are just the most common and stable in our world. At least two more isomers are possible yet difficult to isolate and stabilize in the lab and even more intermediate forms might be possible. But however many possibilities exist, they are finite, constrained by the bond structures that must realize them together with the laws of physics. Possible conformations are constrained by the microstructure. For this reason, conformers pose no threat to microstructuralism. Extrinsic



determination of geometry should be viewed as the selection of one possible geometric state from an internally-constrained space of possibilities.

A second type of isomer might be thought more problematic, and indeed more similar to the troublesome protein cases. These are structural isomers: cases where the same component atoms, represented by the same chemical formula, can be arranged in unique bond structures, yielding unique geometric shapes. Chemists often regard structural isomers as being of different kinds. Consider again n-butane. In addition to its two conformers, n-butane also has a structural isomer, methylpropane. Both have four carbons and ten hydrogens, but n-butane is a linear structure and methylpropane is a branching structure. Colloquially, the conformers and structural isomers are all called 'butane', but the International Union of Pure and Applied Chemistry (IUPAC) separates the structural isomers into two types. This is in contrast with its treatment of the two conformers, which are viewed as two instantiations of the same type, n-butane.

It is important to consider IUPAC's rationale. The rules for dividing and grouping isomers are complicated, often tied up with concerns about nomenclature, but the relevant concern here is practical: n-butane and methylpropane behave differently, are independently stable in experimental contexts, and are used separately. Contrast this with the conformers of n-butane, which rapidly flip back and forth between conformations and exhibit relatively similar properties. In practical applications chemists simply do not work with pure solutions of a single n-butane conformer; nor would they need to, given the negligible difference in behaviour. Different IUPAC stances on conformational versus structural isomers reflect practical demands of science. These classificatory norms have an important place in the practice of chemistry and are accordingly relativised to the contexts of the human pursuit of chemistry. Radically different contexts, new uses, or more stringent acceptability standards for difference could lead to different decisions (notice for instance that the stance on n-butane is relativised to human timescales).

While IUPAC's practices do count against microstructuralism about classificatory practices (as a descriptive or normative claim), they should not count against microstructuralism as a metaphysical thesis. Both conformational and structural isomers exhibit the same type of relation between higher and lower levels of organisation. Though structural isomers admit a greater space of possible geometries, the relation is still one of internal constraint and contextual determination. The atoms in butane are subject to the electrostatic properties of the collective component atoms and within that space of possibilities the physical environment (understood as a number of various forces over time) determines which possible arrangement the molecule can actually take. The molecular essence provides a disposition to act this way or that, depending on relevant context.

There is no reason why this strategy cannot scale up to proteins. We can treat different conformations adopted by multifunctional proteins as a function of extrinsic factors, constrained by the possibility space imposed by the physical microstructure of the macromolecule.

The reader might justifiably wonder why we cannot extend this strategy to explain the underdetermination of physiological function. If conformation is constrained intrinsically and determined by environment, why not say the same thing about physiology? We might think, for instance, that just as a structure contains the potential for many conformations so too any given structure contains the potential for innumerable physiological roles. The roles that get selected are a

function of extrinsic determining factors. Indeterminate intrinsic physical microstructure at the lower level, plus context, equals determinate outcome at the higher level. On the face of it this seems very much like the problem of isomerisation with a larger space of possibilities and a lower likelihood of actualization. But there is an important difference. In the case of isomerisation, the phenomena at the higher level are not multiply-realised at the lower level. Any molecule that is n-butane or methylpropane is necessarily  $C_4H_{10}$ . Nothing could be one of these structural isomers yet originate from a different underlying microstructure. The same cannot be said of physiological roles. Phenomena at the biological level are multiply realisable at the structural level. The lens crystallin role may be filled by the  $\alpha\beta$  variant, but so too could it have been filled by a number of other crystallins. This possibility is clear from the large numbers of species utilizing different crystallins in their lenses. Many molecules are suited to this biological role. Molecular structures are surprisingly functionally flexible. The molecule that gets the job is the one that happened to have been evolutionarily conserved, which is a matter of great chance and contingency.

A stronger (and more loaded) way to say this is that in all possible worlds the chemical structure we recognize as n-butane is realized by  $C_4H_{10}$ , but there are many possible worlds in which the biological role 'lens crystallin' is realized by structures other than the  $\alpha\beta$ -crystallin protein. While there are commonalities between the underdetermination of conformation and the underdetermination of physiology, the difference lies in the existence of multiple-realizability in the opposite direction.

The conflict is much deeper than an observed incongruence between microstructural and biological classification, since this weak observation would also lead us to claim that microstructure cannot account for final conformations, which also appear quite different from the microstructures that bear them. Much more strongly, the claim is that microstructure lacks the bidirectional relations with physiology that are in place between microstructure and conformation.

Without the problem of physical underdetermination, there is no push left toward the highly-permissive pluralist interpretation. Classification need only accommodate the microstructuralist, on the one hand, and the biologist, on the other. But further elimination is not possible. The multiple-realization problem shows that it would be folly to attempt to privilege one of these considerations over the other. It would be no more than a trading of intuitions over the relative significance of biology versus chemistry. Any such decision would be metaphysically arbitrary. The phenomena are best respected by a dual theory, comprised of biological kinds and chemical kinds. Though chemical kinds are well described by microstructuralism, biological kinds are more difficult. It is less-clear just what a biological kind is. Functional classification, though tempting, will not suffice. Not only are there well-rehearsed difficulties with functional individuation (Slater (2009) rehearses some of these with regard to proteins), the functional flexibility of molecular structures, the multiple realization of biological functions, and the evolutionary contingency of function suggests that function is an accidental property of molecules, not an essential property of any kinds. Instead, I suggest conceiving of biological kinds as historical kinds. I will now unpack this account while defending it against the nominalist option.

#### 4. Against Nominalism, Toward a Duality of Kinds

With the structural underdetermination problem dispensed with, two answers to the metaphysical question remain: nominalism and restricted pluralism. The most plausible case for nominalism about protein kinds derives from the observation that microstructuralist classification can neither explain nor capture certain characteristic properties of proteins. Granted, chemical structure can explain certain of the physical properties of proteins, but it cannot explain everything about the biological functions. It cannot tell us, for instance, why a certain biochemical performs the specific multiple functions that it does. For a nominalist, this limitation forces scepticism about proteins as microstructural kinds and probably about proteins as kinds altogether. Both of these conclusions are misguided.

To introduce my position, consider the following extended analogy:

Take a solid gold necklace, a solid gold ring, and a solid gold electronic connector pin. It is perfectly acceptable to tell a story about atomic structure according to which the gold of the jewellery and the gold of the electronic pins are all members of the same kind; we might take this to be a story about why all of this gold is indeed the same kind of thing. This story would be one about natural kinds.

Now suppose you were asked about the other kinds in this scenario: viz. gold jewellery and gold electrical components. Your account might include some facts about the gold from which they are created, including facts that account for its ductility, conductivity, malleability, and colour, which explain why gold makes useful electrical pins and attractive jewellery. Yet these facts would not tell us why humans chose to make jewellery and electrical pins, nor would they tell us why we chose to make these things from gold rather than palladium, silver, cadmium, or platinum. These facts would not tell us why these very different kinds of things happen to have been made from the same material, nor would these facts tell us much about the uses to which jewellery and electrical pins are put. In short, the physico-chemical facts about gold are helpful, but they do not tell the full story. A better account would include any number of historical, economic, and anthropological facts. These facts are about humans, not nature, and for this reason we do not call kinds like jewellery and electrical pins natural kinds. If they are kinds then humans, not nature, make them so. Yet the fact that the microstructure of gold cannot account for the existence or form of the kinds 'jewellery' and 'electrical pins' would not prevent us from saying that gold was a natural kind.

Consider now a more difficult case:  $\alpha\beta$ -crystallin. Recall that in addition to serving as the lens in ducks and chickens,  $\alpha\beta$ -crystallin also occurs as a lyase enzyme. If we wanted to tell a story about how duck-lens protein, chicken-lens protein, and lyase protein were all similar, we would appeal to their physico-chemical microstructure, much as we did with gold. A shared microstructure is why all instances of  $\alpha\beta$ -crystallin are generally taken to be instances of the same kind. We might take this to be a story about natural kinds.

Now suppose you were asked about the other kinds in this scenario: the kinds 'duck-lens protein', 'chicken-lens protein', and 'lyase enzyme'. You could appeal to some facts about the  $\alpha\beta$ -crystallin molecule, explaining why it happens to be well suited to refracting light and binding various substrates. These facts tell us why  $\alpha\beta$ -crystallin makes effective eyes and also why  $\alpha\beta$ -crystallin makes useful enzymes. Yet, just as we saw with jewellery and electrical pins, these physico-chemical facts about  $\alpha\beta$ -crystallin will not give us the full story about these kinds. They do not, for instance, tell us why  $\alpha\beta$ -crystallin is used to make duck-lenses, rather than  $\epsilon$ -crystallin,  $\tau$ -crystallin, or  $\alpha$ -crystallin. Though helpful, the physico-chemical facts about  $\alpha\beta$ -crystallin do not tell us everything about the various uses to which it is put.

I take it that no one will believe my tale to have proved that gold is not a natural kind. Rather, the point is that we would not take our inability to account for the existence of the kinds 'gold connector pin'

and ‘gold jewellery’ as evidence against gold’s status as a natural kind. By parity of reasoning, I suggest, we should view biological facts about chemical kinds to be non-problematically beyond the pale of microstructural kinds.<sup>4</sup> We should not take the incongruence of chemical and biological classification to count against the natural kind status of the chemical kind  $\alpha\beta$ -crystallin. The correct move is to retain the microstructural chemical kind and search for a second set of kinds. In the gold case the second set of kinds were artefacts, in the protein case the second set of kinds are biological kinds.

The kinds ‘gold jewellery’ and ‘gold electrical pin’ are not natural kinds. They are human kinds. How about lenses and lyases? We should view them as evolutionary or historical kinds. This description is in keeping with a recent tradition of philosophers who have described biological species as essentially historical entities (Griffiths, 1999; Millikan, 1999; see also Okasha, 2002). Are they natural kinds? Like others, my argument for their naturalness will be left implicit, an appeal to the naturalness of natural selection.

## 5. Biological Kinds, Chemical Kinds, and Their Relations

The microstructuralist versus historical kinds debate has traditionally been rehearsed in the context of biological species (see Griffiths 1999; Millikan 1999). Even among those not engaged in the historical kinds discussion, there is near-consensus that microstructuralist accounts of species will not work (though see Devitt, 2010). The claim is that there simply are no microstructural essences. Whether we think the relevant structure is genetic, morphological, or somewhere in between, we are bound to be disappointed. Species demonstrate considerable structural heterogeneity. There are no shared microstructures because the morphology, genetic structure, and all other physical structures of biological kinds change constantly. The microstructures of organisms are just parts of the organism and, as parts, they are subject to the same forces that shape the organism as a whole. They change in response to selection and they change by chance.

According to proponents of historical biological kinds, the only essential properties for species are their unique histories. It is these histories of selection and chance that have made them what they are. Compared to species, biochemical molecules are a more instructive case, owing to their comparative simplicity. Unlike species, the case of biomolecules helps illustrate how microstructural and historico-evolutionary kinds relate. This can be seen already from the gold analogy. Gold is a chemical kind that can be picked up and used by humans in the service of creating new kinds of things. These take on a life of their own, independent of the materials from which they were originally created. We view the chemical kind ‘gold’ as a tool or part. Human kinds like jewellery and electrical pins are created with or from these tools, by design or by happenstance, and are subsequently maintained or changed by innumerable forces, both intentional and accidental, using new and different chemical kind tools. If we wanted to ground the stability of these human kinds in spite of structural and functional changes, we would need to take recourse to their trajectory through human history. It is this unique history that has shaped the contemporary kinds.

---

<sup>4</sup> Morange (2012) provides an excellent discussion of the limitations of both chemical and biological explanation. He claims that biological—specifically evolutionary—explanations provide a sort of historical explanation that fills in the sorts of details left out of a chemical explanation, such as why a molecule performs this or that function.

We can think of proteins the same way. Swap gold for chemical macromolecules, rings for enzymes, and humans for evolution and you have an account of biochemical kinds. There are chemical kinds that get picked up, used, and changed, by selection and drift in the service of biological kinds. Different chemical kinds come in and out of this process as genes mutate. Different functions emerge and disappear as contexts change and selection pressures emerge. Through all of this change the closest thing to a constant is the biological protein's historical trajectory. Current chemical microstructures and current physiological functions are simply the latest stage in an ongoing history.

Conceiving of proteins qua biological kinds as essentially historical entities helps avoid several well-known problems with the microstructuralist account. Over time a set of genes coding for a protein will mutate, leading to change in protein structure. Many of these will have no impact on the protein's physiological function yet are stabilized over time. How are we to conceive of these? Should we say that the protein has changed kinds? Other mutations may be more severe, inducing physiological changes slight enough to register as 'change', but not enough to remove the protein from the proteome altogether. In case of functional alteration, should we consider it to be a new protein? Biological species pose these same two problems. They exhibit change in genetic and morphological structures and also gain, lose, and alter behaviours during and between generations. These problems strain microstructural accounts.

In the case of species, the reply to both of these worries is to reject the microstructuralist account. The same is true of proteins. The microstructure of a biological individual does not make it what it is. The microstructure is just one part of the biological individual. The parts, like the whole, change. We can think of the chemical kinds from which proteins are compiled as sets of often interchangeable parts, with varying effects on functionality. We can likewise think of the function of a biological protein kind as just another property, subject to sporadic change and change in response to force. Neither of these are 'essential' properties of the biological kind. Changes in biological protein kinds only constitute changes of kind when a 'speciation' event occurs, at which point two proteins share only part of an evolutionary history.

So when we ask: Are proteins that share a function but differ in structure the same natural kind? How about proteins that share a structure but differ in function? The answer will be: It depends on the historical details.

## 6. Implications for Biological Individuals

Attention to evolution recommends one additional change: a refocusing of particulars away from whole molecules and toward evolutionarily-conserved units. The commonsense focus for a theory of biological protein kinds is the protein molecule itself. Intuitively this seems rather simple: why wouldn't you focus on the spatially delimited molecule? This sort of physical delimitation is often a good strategy when it comes to chemical kinds. But a *prima facie* problem should give pause: Biomolecules are often ever-changing composites, made up of smaller proteins and amino-acid residues. Though these parts converge onto one chemical molecule, they will almost certainly be of different evolutionary origins. It is unclear where and when one molecule ends and another begins, and it seems that solving this problem by appealing to the entire composite as a single molecule runs afoul of the historical kind theorist's appeal to evolutionary history.

To solve the mereological quandary we might borrow a trick from certain discussions of biological organisms and draw physical limits according to whatever composition is required to achieve physiological integration. Yet this appears to inherit the general problem of finding a mind-independent sense of 'function'. If physical composition is judged against functional integration, then there must be a privileged sense of physiological function. But philosophers have long-strained to find any such notion. Slater (2009) has already shown how difficult it is to find the function for a given protein molecule. Which of many actions and interactions we take to be the function is a matter of explanatory context. An appeal to physiological integration will not work.

Following Hull (1980), many philosophers began to take seriously the claim that processes of evolution and selection should impinge on our ontology of biological individuals. In its simplest form, this tradition disregards common-sense physical boundaries and identifies individuals as whatever it is that has evolved or has been conserved over time. Call these 'Darwinian Individuals' (DI). Since biologists disagree over just what gets conserved over time, or just what counts as a target of evolution, DI theories are varied. They are necessarily tied to specific views on evolution and levels of selection (Godfrey-smith 2012). The variety of DI theories cannot be surveyed, here. But applied to proteins, a general DI approach recommends focus on conserved units, rather than spatially-delimited ones. It is common to call these sequences 'domains', but since that term is also used to refer to units of function it is perhaps better to call them 'conserved domains', reserving 'functional domains' for the other use.

Taking conserved domains as the individuals is not without precedent in practice. Biochemists do not explicitly conceive of conserved domains as the individuals, but it is nevertheless conserved domains, not whole molecules, that serve as the focus of work on molecular evolution. Biochemists recognize that one spatially delimited molecule may contain multiple amino acid strings of different evolutionary origins. Thinking of conserved domains as potentially distinct from the whole molecule is necessary in order to explore the evolutionary history of protein molecules and establish cladistic relationships.

This approach need not face the mereological problems imposed by the need for static physical constraints, or the function problem imposed by the need for physiological integration. This approach is not challenged by the existence of molecules that contain conserved domains of different historical origins, since it regards these as separate individuals and so they are able to be members of different kinds. A DI approach allows us to say that, in many cases, different biological kinds converge into a single chemical molecule with a single set of physical limits, to participate in the same or different physiological performances. We can treat the molecule and all its parts as chemical kinds without being committed to taking those same objects as the particulars that feature in biological kinds.

## 7. Implications for Monism and Scientific Practice

The treatment I offer requires conceiving of proteins in a very different way. A single chemical molecule may contain multiple biological individuals. Moreover, the same biological kinds will often exist on different chemical kinds. This shift in focus leaves the door open to a form of monism about taxonomies. Proteins consist of two different types of objects, with similar extensions. There is no

cross-classifying the same object. In the chemical case, we have whole molecules characterized by physical microstructure. In the biological case, we have conserved domains characterized by evolutionary relations. Though we must admit a duality of kind categories, there is a strong sense in which they do not categorize the same objects. But taxonomic monism is just a door prize. The more important implication is that this avoids a potentially unsavoury consequence of permissive pluralism. On that view, one will have to concede that there are many different but equally natural ways to categorize the same protein molecules. Instead, we have an account of two different types of kinds and their relations.

What is lacking in this account is a single taxonomy of biochemical kinds. It is not that we have multiple taxonomies of biochemicals, of course, it is rather that biochemical kinds appear not be kinds at all. Biochemicals are at the nexus of two kinds (of kinds); but this is no skin off the nose of the monist.

Though this position is overtly category dualist, notice that the chemical side of the protein case appears consistent with other chemical kinds. One type of chemical kind—that described by microstructural essentialism—seems perfectly equipped to describe atoms and molecules of all shapes and sizes. The protein case offers no reason to suspect that there are limitations to the scope of microstructuralism within the world of chemical molecules. Insofar as this is the monism behind the chemical intuition, the desire appears stated.

Some philosophers of science might be troubled that my theory of protein kinds diverges radically from actual scientific practice and that my theory cannot take the place of current taxonomies. While we should allow scientific knowledge to guide investigation into kinds, it is certainly not the case that scientific practices should straightforwardly dictate metaphysical conclusions. Nor is it the case that the conclusions I offer should be taken to recommend the revision of scientific practice.

Epistemic barriers constrain classificatory practices. These are a function of the means of acquisition of human knowledge and so should not constrain classificatory metaphysics. To take a simple example, we do not know the evolutionary histories of most proteins; this would preclude my biological classification. But a more subtle point is also worth considering. In order to begin an investigation of evolutionary origins, proteins must first be carved up into operational types. Those types should be carved according to their evolutionary relationships, but that would be putting the cart before the horse. In order to investigate the evolutionary history of a protein type, we need to have marked off that type to facilitate investigation. The best option is to use structure. One might try to classify in a way that approximates physiological similarity or phylogenetic relationships, but even this would be grounded in the relevant structural similarities. It is for precisely these reasons that biochemists use structure as a primary investigative tool in the understanding of physiological function. Structure provides the only currently accessible epistemic handle for thinking about proteins. The tools and techniques of biochemistry are accordingly built around structure. This is how I interpret Goodwin's (2011) finding that biochemical classification is 'fundamentally' grounded in structure. As Goodwin writes, "one of the enduring goals of biochemistry has been to explain the function of proteins in terms of their structure" (p. 534). It would be wrong to read this as commitment to reductionist metaphysics. This is simply a response to epistemic barriers.

Supposing we could perfectly refine a biological taxonomy, perhaps based on god's eye view of evolutionary history, it is still not clear that this would provide the sort of taxonomy that scientists need. Natural kind taxonomies are insensitive to the contexts of investigation, whereas actual taxonomies need to be pragmatically tailored. While metaphysicians want their results to hold over all possible worlds, across all possible conditions, real-life scientists tend to work in just one actual world, and even then in a fairly circumscribed range of actual conditions. It is perfectly acceptable if they fine-tune their taxonomy to this world and those conditions. Yet when we set practice to one side we can see that, when it comes to biochemicals, nature has two sets of joints.

## Acknowledgments

*I have benefited greatly from discussions with and feedback from Juha Saatsi, Greg Radick, Stefan Linquist, Thomas Brouwer, Ageliki Lefkadiou, Jonathan Banks, and Joyce Havstad. I am also very grateful to two anonymous and assiduous referees from this journal.*

## Jordan Bartol

University of Leeds

[phjnb@leeds.ac.uk](mailto:phjnb@leeds.ac.uk)

[www.jordanbartol.com](http://www.jordanbartol.com)

## References

- Boyd, R. (1999) 'Homeostasis, Species, and Higher Taxa', in R. A. Wilson (ed), *Species: New Interdisciplinary Essays*, Boston: MIT Press, 141-85.
- Carr, S., Aebersold, R., Baldwin, M., Burlingame, A., Clauser, K., and Nesvizhskii, A. [2004]: 'The need for guidelines in publication of peptide and protein identification data: Working group on publication guidelines for peptide and protein identification data', *Molecular & Cellular Proteomics*, 3(6), pp. 531-3.
- Chakravartty, A. (2011) 'Scientific Realism and Ontological Relativity', *Monist* 94, pp. 157-80.
- Copley, S. D. (2012) 'Moonlighting is Mainstream: Paradigm Adjustment Required', *BioEssays* 34(7), pp. 578-88. doi: 10.1002/bies.201100191
- Devitt, M. (2010) 'Species Have (Partly) Intrinsic Essences', *Philosophy of Science*, 77(5), pp. 648-61. doi:10.1086/656820
- Dyson, H. J., & Wright, P. E. (2005) 'Intrinsically Unstructured Proteins and their Functions', *Nature Reviews Molecular Cell Biology*, 6(3), 197-208. doi:10.1038/nrm1589
- Ereshefsky, M. (2009) 'Natural Kinds in Biology', in *Routledge Encyclopedia of Philosophy*.



- Ghiselin, M. T. (1974) 'A Radical Solution to the Species Problem', *Systematic Zoology*, 23(4), 536-44.  
doi:10.1093/sysbio/syp096
- Godfrey-smith,. (2012) 'Darwinian Individuals' in F. Bouchard and P. Huneman (eds), *From Groups to Individuals: Perspectives on Biological Associations and Emerging Individuality*. Boston: MIT Press, 17-36.
- Goodwin, W. (2011) 'Structure, Function, and Protein Taxonomy', *Biology & Philosophy*, 26(4), pp. 533-45.  
doi: 10.1007/s10539-011-9252-8
- Griffiths, P. (1999) 'Squaring the Circle: Natural Kinds with Historical Essences', in R. A. Wilson (ed), *Species: New Interdisciplinary Essays*, Boston: MIT Press, pp. 209-28.
- Hacking, I. (2007) 'Natural Kinds: Rosy Dawn, Scholastic Twilight', *Royal Institute of Philosophy Supplement*, 61, pp. 203-39.
- Hendry, R. F. (2006) 'Elements, Compounds, and Other Chemical Kinds', *Philosophy of Science*, 73(5), pp. 864-75.
- Hull, D. (1980) 'Individuality and Selection', *Annual Review of Ecology and Systematics*, 11, pp. 311-32.
- Janzen, D. H. (1977) 'What are Dandelions and Aphids?', *The American Naturalist*, 111(979), pp. 586-9.
- LaPorte, J. (1997) 'Essential Membership', *Philosophy of Science*, 64, pp. 96-112.
- McOuat, G. (2009) 'The Origins of "Natural Kinds": Keeping "Essentialism" at Bay in the Age of Reform', *Intellectual History Review*, 19(2), pp. 211-30.
- Millikan, R. (1999) 'Historical Kinds and the "Special Sciences"', *Philosophical Studies*, 95, pp. 45-65.
- Morange, M. (2013) 'Explanatory Relationships Between Chemical and Biological Sciences', in R. F. Hendry, P. Needham, and A. I. Woody (eds), *Handbook of the Philosophy of Science Vol 6: Philosophy of Chemistry*, Amsterdam: Elsevier, pp. 509-18.
- Okasha, S. (2002) 'Darwinian Metaphysics: Species and the Question of Essentialism', *Synthese*, 131(2), pp. 191-213.
- Slater, M. (2009) 'Macromolecular Pluralism', *Philosophy of Science*, 76(5), pp. 851-63.
- Sober, E. (1980) 'Evolution, Population Thinking, and Essentialism', *Philosophy of Science*, 47(3), pp. 350-83.
- Tobin, E. (2010) 'Microstructuralism and Macromolecules: The Case of Moonlighting Proteins', *Foundations of Chemistry*, 12(1), pp. 41-54. doi: 10.1007/s10698-009-9078-5
- Weadick, C.J., and Chang, B.S.W. (2009) 'Molecular evolution of the betagamma lens crystalline superfamily: evidence for a retained ancestral function in gamma N crystallins?', *Molecular Biology and Evolution*, 26(5), pp. 1127-42. doi: 10.1093/molbev/msp028
- Wilkins, J. (2013) 'Essentialism in Biology', in K. Kampourakis (ed), *The Philosophy of Biology: A Companion for Educators*, Dordrecht: Springer, pp. 396-418.
- Wilson, R., Barker, M., and Brigandt, I. (2007) 'When Traditional Essentialism Fails: Biological Natural Kinds', *Philosophical Topics*, 35, pp. 189-215.
- Wistow, G. (1995) *Molecular Biology and Evolution of Crystallins: Gene Recruitment and Multifunctional Proteins in the Eye Lens*. Austin: R.G. Landes.