This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

*Author:*
**Bellazzi, Francesca**

*Title:*
**Biochemical kinds and the unity of science**

**General rights**
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

**Take down policy**
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

•Your contact details
•Bibliographic details for the item, including a URL
•An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

# Biochemical Kinds and the Unity of Science

By

Francesca Bellazzi

Department of Philosophy
University of Bristol

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Arts.

December 2022

Word count: 74.885

*'Nature' is what we see -*
*The Hill - the Afternoon -*
*Squirrel - Eclipse - the Bumble bee -*
*Nay - Nature is Heaven -*
*Nature is what we hear -*
*The Bobolink - the Sea -*
*Thunder - the Cricket -*
*Nay - Nature is Harmony -*
*Nature is what we know -*
*Yet have no art to say -*
*So impotent Our Wisdom is*
*To her Simplicity.*
(Emily Dickinson, 668)

# Abstract

The present thesis explores some metaphysical issues concerning biochemical kinds and the relations between chemical and biological properties and phenomena. The main result of this thesis is that there is something *sui generis* about biochemical kinds. This result is motivated by two theoretical steps. The first is characterising biochemical functions as weakly emergent from the chemical structure [Chapter 3, Chapter 6]. The second is via an account for which biochemical kinds are natural categories [Chapter 4, Chapter 7].

The thesis comprises four parts. Part I [Chapter 1, Chapter 2] aims to offer the methodological and conceptual tools that underpin this research. Chapter 1 presents the debate on the unity of science and the motivation of the research. Chapter 2 presents the account of natural kinds that will be used throughout the thesis. Part II [Chapter 3, Chapter 4] offers a detailed metaphysical analysis of the molecular gene. Chapter 3 argues that molecular genes are weakly emergent from nucleic acids. Chapter 4 argues that the category "molecular gene" can be deemed a natural kind, following the account presented in Chapter 2. Part III [Chapter 5, Chapter 6, Chapter 7] explores some of the themes related to biochemical kinds and generalises some of the results from Part II. Chapter 5 considers biochemical functions and how functional attribution should be interpreted for biochemical molecules. Chapter 6 considers the relation between biochemical functions and structure, spelt in terms of weak emergence and explores unity in biochemistry. Chapter 7 considers the topic of biochemical kinds *per se* and argues that biochemical kinds are natural categories. Part IV [Chapter 8] elucidates the main outcome and the implications that this research can have for future discussion in the philosophy and metaphysics of biochemistry and the unity of science.

# Acknowledgements

There are no results that are individual results. As with biochemical kinds, context and relations are those that allowed this research to happen, grow and be so enriching and fruitful for me. So, while not a fan of listing writing, this is the space to present my list of thanks, which can only gesture towards the immense gratitude I feel towards everyone mentioned here.

This research has been funded by the European Research Council *MetaScience Project* (No 771509), and my deepest gratitude goes toward the whole MetaScience team. I thank my first supervisor, Professor Tuomas Tahko, for his guidance, support, leadership and mentorship during these years. I thank my second supervisor Professor Samir Okasha for discussing, reading and reviewing multiple versions of my work. I thank the whole MetaScience team: Vanessa, Alex, Toby, Sam, Will and Elle. You have been a constant source of inspiration and support that allowed me to grow philosophically in a way I could not have foreseen. Being in a team with you has been extremely precious to me. I need to thank also my two visiting supervisors and mentors, Prof. Maurizio Zuccotti and Prof. Jessica Wilson, for offering me the chance to visit their universities and for their support in writing some of the major parts of this thesis.

I would like to thank the Department of Philosophy and the Postgraduate community at the University of Bristol for being a wonderful and welcoming environment of which I really feel part. I thank the audiences of all the conferences I have attended in the past years for their insights and comments on this work. I thank the many colleagues and friends across the world that have supported me in this journey with their discussions and comments, and for reading and revising multiple versions of my writings and drafts. In particular, I thank Giacomo, Luca, Margarida, Lorenzo, Milenko, Noelia, Katie. I thank Oyku for being a true friend and colleague in this still gender-imbalanced field and discipline. I thank Annalisa for sharing the wonderful experience of visiting the University of Toronto with me and for our long-lasting philosophical (and not only) friendship. I thank Konrad for his invaluable support at the beginning of this journey and for all the discussions and papers we wrote together. I thank Dan for sharing the office space with me, for listening to all my thoughts every morning (and afternoon) and for all the collaborations and discussions. I thank my two PhD mates and best friends, Matteo and Simone, for their precious friendship, their extremely brilliant insights and their love and support: this PhD would not have been possible without you.

My deep gratitude also goes to everyone supporting me outside of the philosophical discipline. I thank Maria for being my guidance, mentor, friend and constant source of inspiration. I thank Giulia for the endless walks in the countryside, important to thinking and to us, and for always being there. I thank Stanislaw for his true and reliable friendship in every moment, for being home away from home and for teaching me not to take life too seriously (or personally). I thank Lydia for being the most wonderful housemate. I thank Dora for reminding me of the importance of being a true feminist, always. I thank Anna, Pauline and Victoria for being a bridge between all my lives. I thank Fr. Rupert, Andrea, Sam and the whole Bristol Catholic Chaplaincy for having been a safe harbour in the difficult years of the pandemic.

I thank my lifelong friends that have always been with me since I read my first Aristotle: Emma, Francesca, Carlo, Maria, Riccardo and Laura. Lastly, my love and gratitude go towards my entire family, whose passion for science and knowledge and whose love, care and support have been with me all these years without ever failing.

I love you,
*Francesca*

# Publications by Author

Chapter 3 is version of Bellazzi, Francesca 2022a: “The emergence of the postgenomic gene”, *The European Journal for Philosophy of Science* 12, 17.

Chapter 4 is a version of Bellazzi, Francesca *forthcoming*: “The gene as a natural kind”. Edited volume *Life & Mind* (Springer Nature) (22 pages). Parts of Chapter 2 are also reproduced from this paper.

Chapter 5 is an expanded version of Bellazzi, Francesca *forthcoming*: “Biochemical functions”, *The British Journal for Philosophy of Science.*

Parts of the Prologue and Chapter 1 have been published in an edited and shorter version in the blog post: “Marvellous mysteries and the unity of science”, *Arts Matter Blog* (2020).

Parts of Chapter 3, Chapter 4 have been published in an edited and shorter version in the blog post: “Genes: from instruments of thought to the base of life”, *Jargonium* (2022c).

Parts of Chapter 5, Chapter 6 have been published in an edited and shorter version in the blog post: “The superpowers of proteins”, *Jargonium* (2022b).

# Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: DATE: 1st of December 2022

# Contents

***Preface: Marvelous mysteries*** ***10***

**I The quest for unity and natural kinds 13**

**1 Introduction 14**
- 1.1 Structure of the thesis . . . . . 14
- 1.2 Context: unity of science and biochemistry . . . . . 18
  - 1.2.1 Unity of science: a brief historical overview . . . . . 18
  - 1.2.2 Biochemistry and biochemical kinds . . . . . 23
- 1.3 Methodology . . . . . 25
  - 1.3.1 Moderately naturalistic metaphysics . . . . . 26
  - 1.3.2 Reflective equilibrium and illuminating accommodation . . . . . 27
- 1.4 Outcome and implications . . . . . 28

**2 Natural kinds 33**
- 2.1 The naturalness of kinds . . . . . 36
- 2.2 Khalidi's view of natural kinds . . . . . 38
  - 2.2.1 Projectibility and causal networks . . . . . 40
  - 2.2.2 Viruses: an instance of natural kinds . . . . . 43
- 2.3 The unification and the ontological question . . . . . 44
  - 2.3.1 Unification principles and the four-category ontology . . . . . 45
- 2.4 A variety of kinds and natural kind monism . . . . . 49
- 2.5 Conclusion . . . . . 52

**II A metaphysical analysis of the molecular gene 54**

**An introduction to the molecular gene 55**
- A brief history of the gene . . . . . 57

**3 The emergence of the molecular gene 61**
- 3.1 Introduction . . . . . 61
- 3.2 The postgenomic gene . . . . . 64
  - 3.2.1 What is a postgenomic gene? . . . . . 66
- 3.3 Ontological emergence as novelty and robustness . . . . . 69

3.4 From complexity to emergence . . . 73
    3.4.1 The emergence of the gene . . . 75
3.5 Conclusion: flexibility and context dependency . . . 84

**4 The molecular gene as a natural kind 86**
4.1 Introduction . . . 86
4.2 The gene in the "postgenomic" world . . . 88
4.3 Natural kinds again . . . 90
    4.3.1 Projectibility and causal networks . . . 92
4.4 The molecular gene as a natural kind . . . 95
    4.4.1 Genes classifications and taxonomies . . . 95
    4.4.2 The gene *is* a natural kind . . . 97
4.5 Conclusion . . . 101

**III The Metaphysics of Biochemical Kinds 104**

**An introduction to biochemical kinds 105**

**5 Biochemical functions 108**
5.1 Introduction . . . 108
5.2 Biochemical kinds and biochemical functions . . . 110
    5.2.1 Vitamin B12 and its function . . . 112
5.3 Biological functions and chemical functions . . . 114
    5.3.1 Biological evolutionary functions . . . 115
    5.3.2 Two problems for biological evolutionary functions . . . 118
    5.3.3 Chemical functions . . . 119
5.4 Biochemical functions: a unified approach . . . 122
    5.4.1 Genetic functions . . . 127
5.5 Assessing the view . . . 129
    5.5.1 Requirements . . . 130
    5.5.2 Biochemical functions and biochemical kinds . . . 133
5.6 Conclusion . . . 135

**6 Biochemical functions and unity in biochemistry 137**
6.1 Introduction . . . 137
6.2 (Dis-)Unity of science . . . 139
    6.2.1 Multiple determinability and multiple realisability . . . 141
6.3 The double problem of biochemical function . . . 144
    6.3.1 Vitamin B12 . . . 145
    6.3.2 Biochemical functions . . . 147
6.4 Which unity? . . . 150
    6.4.1 Weak emergence and the "proper subset strategy" . . . 151
    6.4.2 Biochemical functions are weakly emergent . . . 152
    6.4.3 Unity within autonomy . . . 155
6.5 Challenges and clarificatory remarks . . . 156
    6.5.1 Multiple realisability and multiple determination *again* . . . 156

6.5.2 Biological context to the rescue . . . . . . . . . . . . . . . . . . . . . . 159
6.5.3 Two accounts of weak emergence . . . . . . . . . . . . . . . . . . . . 161
6.6 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 163

**7 Biochemical kinds 166**
7.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 166
7.2 The problem . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 168
7.2.1 The problem of definition . . . . . . . . . . . . . . . . . . . . . . . . 171
7.2.2 The problem of taxonomy . . . . . . . . . . . . . . . . . . . . . . . . 172
7.3 Kinds of kinds . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 173
7.3.1 Natural kinds . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 174
7.3.2 Chemical kinds and biological kinds . . . . . . . . . . . . . . . . . . 176
7.4 Some kinds options for biochemical molecules . . . . . . . . . . . . . . . 178
7.4.1 Biochemical kinds as chemical kinds . . . . . . . . . . . . . . . . . . 178
7.4.2 Biochemical kinds as biological kinds . . . . . . . . . . . . . . . . . . 180
7.4.3 Biochemical kinds as dual kinds . . . . . . . . . . . . . . . . . . . . 181
7.4.4 Pluralism . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 182
7.5 Biochemical kinds . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 183
7.5.1 Some biochemical kinds: vitamin B12 and the molecular gene . . . . . . 186
7.5.2 Chemical kinds, polythetic kinds and evolutionary properties . . . . . . 190
7.6 The problem of taxonomy . . . . . . . . . . . . . . . . . . . . . . . . . . 192
7.7 Conclusion: unity and biochemical kinds . . . . . . . . . . . . . . . . . . 195

**IV Biochemical kinds and the unity of science 198**

**8 Conclusion 199**
8.1 An inquiry into biochemical kinds . . . . . . . . . . . . . . . . . . . . . . 199
8.2 Outcome and implications . . . . . . . . . . . . . . . . . . . . . . . . . . 201
8.2.1 Emergence of biochemical kinds . . . . . . . . . . . . . . . . . . . . . 203
8.2.2 A unified view of the world . . . . . . . . . . . . . . . . . . . . . . . 206

***Parting thoughts 209***

**Bibliography 211**

# List of Figures

Figure 1: T structure of the thesis [Chapter 1]

Figure 2: Unity of science [Chapter 1]

Figure 3: Lowe's four category ontology [Chapter 2]

Figure 4: Alternative splicing [Chapter 3]

Figure 5: Glycine [Chapter 3]

Figure 6: Selection [Chapter 5]

Figure 7: Weak emergence1 [Chapter 6]

Figure 8: Multiple determinability [Chapter 6]

Figure 9: Cobalamin [Chapter 7]

# List of Abbreviations

RE= reflecting equilibrium

IA= illuminating accommodation

HPC= homeostatic property cluster theory

HT= hierarchy thesis

NKM= natural kind monism

Ft= property of being transcribable or actively involved in transcription

BC-function= biochemical function

MR= multiple realisability

MD= multiple determinability

Coenz-blood = function to contribute to erythropoiesis processes

WE= weak emergence

# Preface: Marvelous mysteries

*"But what vast gaps there were, what blank spaces, she thought leaning back in her chair, in her knowledge! How little she knew about anything. Take this cup for instance; she held it out in front of her. What was it made of? Atoms? And what were atoms, and how did they stick together? The smooth hard surface of the China with its red flowers seemed to her for a second a marvelous mystery."* (Virginia Woolf, The Years)

How can this China with red flowers be made of atoms somehow sticking together? - asks Eleanor. The cup, like any other ordinary material object, seems to be composed of smaller entities and these of something smaller all the way down to atoms and more fundamental physical particles. How can these different entities be unified in one single object? This goes hand in hand with the question of how different disciplines can study the different entities and properties in this cup: physics the fundamental components, chemistry the fragility, biology the bacterias on its surface, psychology and sociology its usage. How can these disciplines be linked in studying this same object?

These questions are related to those concerning the "unity of science".[1] The search for the unity of the world or for a principle that underpins everything that exists is as old as Western Philosophy. As far as Aristotle reports, the first philosophers tried to find the principles of all things in one or a few natural principles (*Metaphysics*, 1).[2] This can be seen as the very first

---

[1] Part of this introduction has been published in Bellazzi, Francesca 2020: "Marvellous mysteries and the unity of science". *Arts Matter Blog.*

[2] "Most of the earliest philosophers conceived only of material principles as underlying all things. That of which all things consist, from which they first come and into which on their destruction they are ultimately resolved, of which the essence persists although modified by its affections - this, they say, is an

search for unity: if we find one principle (or some principles) that are shared among all things, then there is a sense in which the world is unified. At the same time, this quest has been accompanied by the research for a form of unity of our knowledge. The tension between the two is given by the fact that on the one hand, the world seems to be one and unified; on the other, our knowledge seems to take different forms. As Plato reports in the *Sophist*: "Knowledge, like other, is one, but each separate part of it applies to some particular subject and has a name of its own; hence there are many arts, as they are called, and kinds of knowledge, or sciences." (Plato, *Sophist* 257c/d).[3]

While time has passed since the Greek philosophers, the quest and the search for a unificatory principle of both our knowledge and the world has not. The two topics have been linked: if knowledge is knowledge of the world, then the unity of the world and the unity of knowledge can be related. Given the importance of scientific knowledge, one could do a step forward and ask whether the unity of the world is connected to the unity of the sciences or how the sciences can be unified. This can be motivated by a possible correspondence between the layering of the world and the layered structure of scientific theories. Whether the sciences can be ordered and linked in this way can be explored from both an epistemic perspective, looking at how theories can be related or reduced to one another, or from an ontological perspective, looking at how different entities displaying different properties can be related or reduced (Tahko 2021). In this thesis, I focus on the latter sense of unity of science that concerns how the world can be maintained to be one while being so varied.[4]

In the search for a form of unity, it seems natural to start by looking at the relations that hold either from the realm of the smaller entities studied by parts of physics upwards or the

---

element and principle of existing things. Hence they believe that nothing is either generated or destroyed, since this kind of primary entity always persists." (*Metaphysics*, 1-983b).

[3] The context of the dialogue is the following: "Stranger: Let us consider another point and see if you agree with me. - Theaetetus: What is it? - Stranger: It seems to me that the nature of the other is all cut up into little bits, like knowledge. - Theaetetus: What do you mean?- Stranger: Knowledge, like other, is one, but each separate part of it applies to some particular subject and has a name of its own; hence there are many arts, as they are called, and kinds of knowledge, or sciences." (Plato, *Sophist* 257c/d).

[4] The ontological and epistemic dimensions are related: the latter provides us access to the former. Nevertheless, the unification (or its absence) of different scientific theories can be approached in different ways, and its possibility or impossibility does not exclude the correspondent ontological ones.

opposite. This can allow us to see how the entities associated with different disciplines interact with each other. While the traditional image of the physical world as a layered cake or a pyramid composed of different levels of entities does not hold, as it is difficult to identify to what levels exactly amount and precise reductive relations between different entities, it seems still possible to identify various points of contact for the entities that relate different domains or realms.[5] Instances of this are the relations between molecular structure and quantum particles, the ones between chemical molecules and biological processes or those between psychological states and economic behaviours. In this thesis, I have decided to focus on the relation between biological and chemical phenomena, exploring different aspects of the interactions between the two. This analysis can provide insights and implications for the discussion on unity and can hopefully guide future research concerning the unity of science. Biochemical kinds, such as DNA, vitamins, and proteins, play a crucial role in many fields, including medicine and molecular biology, and present both properties normally ascribed to chemical kinds and biological ones. Accordingly, these kinds can be considered philosophically interesting for our purposes because they are at the interface between chemistry and biology. Moreover, despite being foundational for understanding the living world, their philosophical study is still nascent. This knowledge gap offers the chance to explore a range of issues. Specifically, I will consider the following research questions:

- What are the properties at play at the biochemical level?
- How are these properties related?
- Which kinds of kinds do we find in the biochemical realm? Are they biological or chemical kinds?

The answers to these questions have implications for the view of the unity of science that can be offered given the relations at play at the biochemical level. While a complete answer to the problem of the unity of science requires a more extensive case-by-case analysis that considers different inter-level relations, looking at biochemistry can still offer interesting insights and has implications for the consideration of this problem. This is what this thesis does.

---

[5] More on this will be discussed in Chapter 1.

# Part I

# The quest for unity and natural kinds

# Chapter 1

# Introduction

## 1.1 Structure of the thesis

The research presented here develops in the general context of the quest for a form of ontological unity in science, as will be further presented in detail, but focuses specifically on one case of inter-level relations: the ones between the chemical and the biological. The main result concerns the nature of the entities and the relations at play in the biochemical domain and will be that there is something *sui generis* about them. The exploration of the biochemical world from a philosophical perspective provides novel results, such as the emergence of the molecular gene and of biochemical functions, and the naturalness of biochemical kinds. Moreover, this research has a series of more general implications for a unified picture of the world.

The present thesis comprises of four parts. Part I, composed of Chapter 1 and Chapter 2, is an introduction that aims to offer the methodological and conceptual tools that underpin this research. Chapter 1 presents and motivates the research structure and gives an overview of the context, the methodology, and the main outcome and implications. Chapter 2 is an overview of the position of natural kinds that is used in the detailed analysis of biochemical kinds presented in the rest of the thesis. After this introductory part, there are two core parts that follow a "T" shape structure [Figure 1].

Part II, composed of Chapter 3 and Chapter 4, offers a detailed metaphysical analysis of the molecular gene, which is the main case study of the thesis. This is motivated by the fact that molecular biology, particularly genetics, is an interesting interplay of chemical and biological phenomena. The molecular gene has a crucial role in protein synthesis and other biological processes, while being composed of nucleic acids. Accordingly, it represents a good starting point for discussing unity and inter-science relations between the chemical and the biological domain. The study of the molecular gene is developed in two chapters. Chapter 3 argues for a particular understanding of the relationship between instances of molecular genes and nucleic acids: molecular genes are weakly emergent from DNA or RNA. Chapter 4 considers the status of the category "molecular gene" and argue that it can be deemed a natural kind following the account presented in Chapter 2. Part II deepens the understanding of which properties are at play at the interface between chemistry and biology and has two important roles for the general argument of this thesis. The first is to point out that there are at least two main properties crucial for the characterisation of the molecular gene, molecular structure and biochemical functions. These two properties are shared across different instances of biochemical kinds, as will be discussed. For instance, proteins, vitamins, and other biochemical molecules seem to be characterised by a special relation between structure and function. Acknowledging the importance of functional and structural properties for biochemical kinds is the starting point for the third part of this research. The second is that Part II offers a case study oriented argument in favour of the main result of this research for which there is something specific about biochemical kinds and their properties: the molecular gene can be considered weakly emergent from nucleic acids and its category can be considered a natural kind.

Part III, composed of Chapter 5, Chapter 6 and Chapter 7, explores "horizontally" some of the themes related to biochemical kinds and generalises some of the results gained from Part II. Specifically, it focuses on biochemical functions, the relation between function and molecular structure and biochemical kinds by using a further case study, vitamin B12. Chapter 5 represents the main bridge between Part II and Part III. This chapter explores in detail the nature of biochemical functions. This follows from the importance of genetic functions for the molecular gene and the difficulties in identifying to what these functions amount. The role of biochemical functions calls for a more in-depth analysis, especially considering that other biochemical kinds are also characterised by their function. Having provided an account of biochemical functions, Chapter 6 considers the relation between biochemical functions and structure in terms of weak emergence and explores unity in biochemistry. Finally, Chapter 7 considers the topic of biochemical kinds *per se* and argues that biochemical kinds are natural categories.

Part IV concludes and Chapter 8 elucidates the main outcome and the implications that this research can have for future discussion in the philosophy and metaphysics of biochemistry and the unity of science.

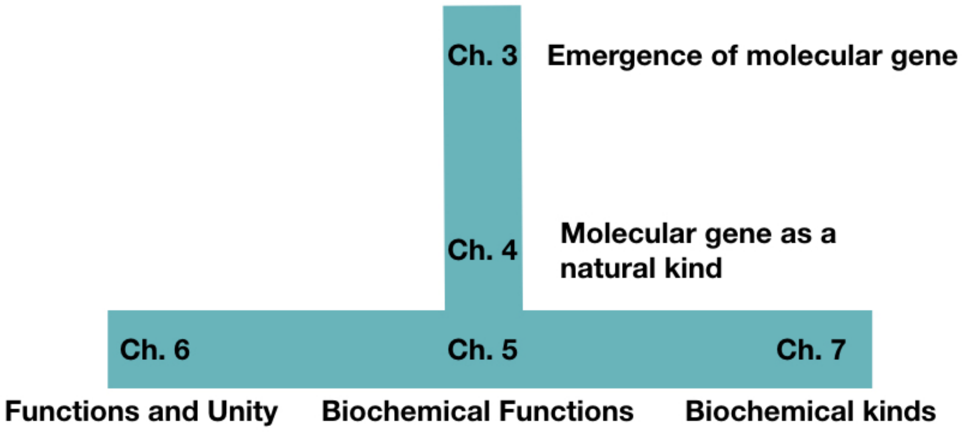

Figure 1. “T” structure of Part II and Part III of thesis. Chapter 3 and Chapter 4 represent the “vertical” case study on the molecular gene that lead to consider “horizontally” the metaphysics of biochemistry, as presented in Chapter 5, Chapter 6 and Chapter 7.

The central thesis defended here is that there is something *sui generis* about biochemical kinds and their properties. This result is motivated by two theoretical steps. The first is characterising biochemical functions as weakly emergent from a chemical structure [Chapter 3, Chapter 6]. The second is an account of natural kindhood, for which biochemical kinds are natural categories [Chapter 4, Chapter 7]. This analysis has implications for the quest for the unity of science as a form of ontological unity can be maintained by spelling out the main categories at play. For instance, unity can be achieved with a single account of natural kinds and via clarifying the relations that hold between the various entities of the world, such as weak emergence and causal networks.[1] While biochemical molecules are something *sui generis*, their very nature can be appreciated once one considers the level of interactions and connections present in the biochemical domain. This does not leave us with a scattered or "dappled world", as a patchwork of entities behaving in a diverse way (Cartwright 1999). Instead, it gives us a rather highly connected one. Moreover, the view presented here is in line with Mitchell's integrative pluralism, which has been presented as informed by biological complexity (Mitchell 2003). As in this account, also this thesis develops the idea that the variety of the biochemical stems out from the integration of the systems it is part of.

Moreover, the results presented in this thesis, as will be further detailed, are important for some debates in the philosophy and metaphysics of science. First, it offers a metaphysically informed study of the molecular gene in the postgenomic context, presenting a novel and original contribution to the literature in the philosophy of genetics and enriching the variety of case studies discussed in the metaphysics of science. Second, it offers a novel and detailed analysis of the metaphysics of biochemical kinds, clarifying the notion of biochemical functions, the relations between functions and structure and the naturalness of the biochemical kind category. Given that the discussion of the philosophical aspects of biochemistry is still in its nascence, this thesis also hopes to offer context and insights for future discussions on these topics. Lastly, this research has been developed as part of the ERC-funded "MetaScience Project" (n. 771509), which aims to study the unity of science from the perspective of metaphysics and philosophy by

---

[1] These notions will be further clarified throughout the thesis [Chapter 3, Chapter 4, Chapter 6, Chapter 7].

way of employing real-world case studies. Thus, the present work aims at contributing to the exploration of unity in biochemistry within this project and being aligned to its aims.

## 1.2 Context: unity of science and biochemistry

Biochemical kinds, such as proteins, vitamins, and nucleic acids, are fascinating. On the one hand, they can be seen as chemical compounds, complex molecules and, on the other hand, as entities with a specific role in biological systems. They are kinds at the borders between the chemical and biological realms and an interesting starting point to explore the unity of science. While the thesis focuses mostly on understanding the nature of the biochemical phenomena and the entities, properties and relations at play in this realm, the original motivation stems from an interest in the possibility of a form of unity of science. Accordingly, this section presents first a short historical overview of the debate on the unity of science. Then, it provides a definition of biochemistry and the related biochemical kinds, underlying why these kinds are interesting for the unity or the disunity of science debate.

### 1.2.1 Unity of science: a brief historical overview

The link between the unity of knowledge or science and the unity of the world by identifying a single principle has deep roots in the history of philosophy. Nevertheless, the discussion on the unity of science as it is meant in the contemporary debates has more recent origins in the discussions of the Vienna Circle in the 1920/30s. In detail, the idea of unity has been associated with the idea of reduction: if we show that we can reduce x to y, that x is nothing more than y or nothing over and above y, then we can say that x and y are unified (v. Riel, v. Gulick 2019). Given the successes of physics and physical sciences, some philosophers in the Vienna Circle put forward the idea that the sciences could have been unified when reduced to physics. Specifically, the aim was to “translate” all scientific language into physicalist language and physical statements. A summary of this position can be seen in Carnap: “all empirical statements can be expressed in a single language, all states of affairs are of one kind and are known by the same method." (Carnap 1934, 32). This project would have unified all the other sciences (or special sciences) into the most rigorous one: they would all be expressed in the language of physics.[2] The unity of language was also followed by a unity of theories or theoretical unity.

This project did not stop with its early developments in the Vienna Circle and various models of how to reduce entire theories to others have been proposed. Among them, Naglel's account of reduction stands out of his impact on the literature (1961). In the Nagelian account of reduction, a theory can be reduced to another when its laws can be seen as logical consequences (can be derived) from the theoretical assumptions of the primary science (Nagel 1961, 352). This account focuses on a specific form of reduction: a form of "theoretical reduction" that concerns our knowledge and the structure of our theories and language. This, which can be broadly seen as a form of "epistemic reduction", is distinct from the sense of unity this thesis is concerned with. On the opposite, the focus will be on the "ontological" sense of reduction that regards the structure of reality and the structure of the world: there is some basic class of entities to which all the others reduce. Whether the Vienna Circle project or the Nagelian project have ontological interests too is open to debate. However, as such, they did not put forward any specific ontological thesis about the world being unified or dis-unified.

A proposal of unity that combines theoretical and ontological considerations is the one of Oppenheim and Putnam in their article "Unity of Science as a working hypothesis", published in 1958. In this work, the authors draw conclusions about the structure of the world from the success of reductive explanations in scientific theories. Specifically, if we can explain some phenomena at a higher-level H2 in terms of others at a lower-level H1, then the entities at H2 are composed of H1 entities and accordingly reducible and identical to them. The relation of reduction, which they call *micro-reduction*[3], is seen to be applicable to each level and transitive.

---

[2] Specifically, Carnap writes: "We wish however to interpret the term 'the physical language' so widely as to include not the special linguistic forms of the present merely but also such linguistic forms as physics may use in any future stage of development." (1934, 54)

[3] "Micro-reduction": given two theories T1 and T2, T2 is reduced to T1 if and only if: i) the vocabulary of T2 contains terms not in the vocabulary of T1; ii) any observational data explainable by T2 are explainable by T1; iii) T1 is at least well-systematised as T2. This relation is defined as transitive, irreflexive and asymmetric (Oppenheim, Putnam 1958, 5).

Social entities are reducible to biological entities that are reducible to chemical entities that are reducible to physical ones, being themselves reducible to those of fundamental physics. This provided a rather intuitive, even if simplistic, view of the world: each block is composed of the lower-level one all the way down, and it is possible to identify identity relations between the components [Figure 2]. The world remains one because each level is seen as identical to its components, and everything can be identified, given the transitivity of micro-reduction, with the fundamental physical level. This idea can be seen as strict identity reductionism: phenomena at the higher level are strictly identical to phenomena at the physical level.

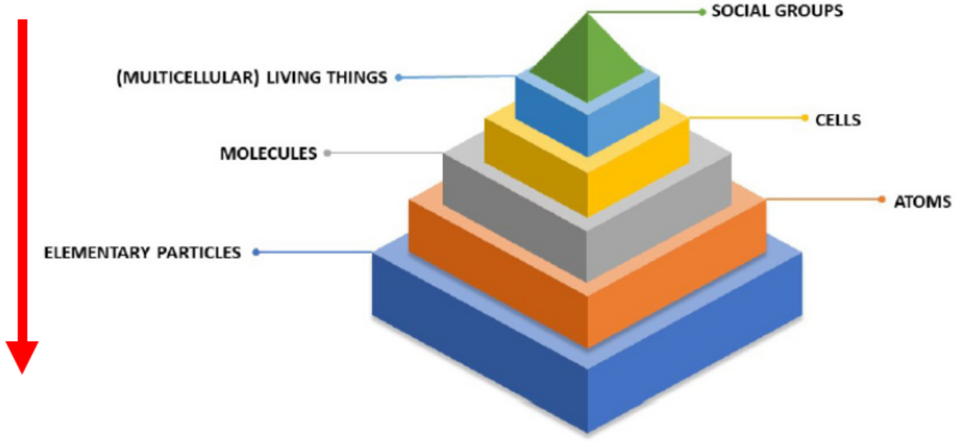

Figure 2. The unity of science in the model proposed by Oppenheim and Putnam can be achieved by many micro-reductions and is based on the fact that each level is composed of more basic entities to which it can be reduced. The arrow indicates the direction of the reduction relation. The figure is re-adapted from Tahko (2021).

To understand this view, let us Gresham's law from economics: "bad money drives out good money", discussed by Oppenheim and Putnam (1958, 16-17). In simple terms, bad money can be defined as money with a nominal value higher than the commodity value, simplifying, higher than the cost of the material it is made of. One can do more things with bad money at the same cost. Good money, on the opposite, has a nominal value similar or equal to the material it is made of. The Gresham law says that when both good and bad money is circulating, bad money will prevail in terms of usage. How can we explain this? The reductionist answer is that this law can be reduced or analysed in terms of psychological laws. Let us assume that it is a psychological law that choices are ordered by an individual preference function to maximise profit. Accordingly, the fact that bad money drives out good money can be explained by the fact that bad money maximises the profit of the individual agents compared to good money. Micro-reduction can be achieved if we can reduce the economics law to the psychological law, and, as such, we can see that economics and psychology can be unified.

However, a major problem for reductionism in these terms soon comes up. This form of reduction seems to presuppose a 1:1 relation between the higher-level and the lower-level and specifically between the *kind* of things that we find at the higher-level - in the example, money - and the *kind* of things that we find at the lower-level - in the example, objects and the psychological value associated with it. But this form of identity is fallacious due to widespread cases of multiple realisability, where different lower-level ones realise the one higher- level entity. As pointed out by Fodor in his paper “The Disunity of Science as a working hypothesis” (1974), monetary exchanges are realised by different kinds of things and behaviours. A variety of lower-level things can take the form of money and monetary exchange: we cannot simply point out a 1:1 relation. This can be defined as the challenge to reductionism from multiple realisation: the same higher-level feature can be realised by multiple lower-level ones, compromising the suggested micro-reduction. Money, in this case, is multiply realisable. Nevertheless, according to Fodor, money remains real enough, despite its irreducibility to lower-level features. We have all the reasons to retain the Gresham law as valid without reducing it to psychology. Accordingly, the failure of micro-reduction does not challenge the existence of money but rather corroborates the idea that the two sciences are dis-unified and so seem to be the world they aim to capture. As future discussions on these issues pointed out, the relations between entities in different levels or realms are complex in various instances and not only in this paradigmatic case (Polger, Shapiro 2016; Tahko 2021). The world present many cases of multiple realisability and also many cases of multiple determinability[4], where the lower-level can be determined in many different higher-level

---

[4] Multiple realisability refers to a phenomenon in which the same entity or property can be realised by different ones, where realisation can be defined as a “synchronic ontological dependence relation, distinct from identity, and that transmits physical legitimacy from physical realisers to what is realised” (Polger, Shapiro 2016, II, 4). Multiple determinability refers to the phenomenon opposite to multiple realisation: the same entity can determine different properties or other entities.

entities (as defined in Tahko 2020). The complexity of the world's relations has challenged a straightforward identity-reductionist approach to the unity of science for which the higher-level entities can be considered identical to the lower-level ones.

Given the difficulties in providing a unified framework, some philosophers have advocated for forms of disunity of the sciences and a pluralist view of reality: the world is varied, and so are the disciplines that study it (Duprè 1983, 1993; Cartwright 1999; Mitchel 2003).[5] Disunity is supported by methodological and ontological arguments, as suggested by Duprè (1983, 1993) and Cartwright (1999).[6] Methodologically, the different sciences present different methods of inquiry. Metaphysically or ontologically, there are two grounds for disunity. The first considers the plurality of classifications into various kinds of kinds: a pluralism of kinds leads to a dis-unified world. The second considers the irreducibility of the various entities, as already underlined by instances of multiple realisation and complex relations. These arguments applied also to the biochemical domain that can be seen as characterised by a duality of kinds: biological and chemical ones, and the two should be considered separate (as suggested by Bartol 2016; further discussion in Chapter 6). This would leave us with a disunified view of the world.

The question of whether a form of unity might still be maintained despite the complexity is open in the current metaphysics and philosophy of science. An extensive and detailed answer to how this can be achieved would require an overview of all the possible relations between disciplines. This would go beyond the scope and the space of the thesis. Accordingly, I have decided to explore one of the interplays between chemistry and biology: biochemistry. While this thesis will focus precisely on biochemical kinds, the research's interest and context is the debate on the unity of science and relations between different domains. In particular, I will consider the two arguments for disunity, the one concerning kinds and the one concerning the relation between the different entities, and argue that a form of unity can be maintained given how biochemical kinds and properties are characterised.

---

[5] This view is defended by what can be called the Stanford school, e.g., Duprè, Ian Hacking, Peter Galison, Patrick Suppes and Nancy Cartwright (Cat 2017).

[6] Cartwright's argument develops upon the consideration of laws of nature (1999). Given that the existence of laws in biochemistry, and generally in the special sciences, is a topic subject of controversy, this thesis will leave this debate aside.

### 1.2.2 Biochemistry and biochemical kinds

Before discussing issues concerning biochemical kinds in more detail, we should consider briefly how to identify the putative instances of biochemical kinds and what we mean by biochemistry. Following a moderately naturalistic approach (which will be presented shortly), I will take biochemical kinds to be those studied by biochemistry and that present properties partially associated with the chemical domain and partially associated with the biological one. Following the definition presented by the *Biochemical Society*, we can define biochemistry as:

> "The branch of science that explores the chemical processes within and related to living organisms. It is a laboratory-based science that brings together biology and chemistry. By using chemical knowledge and techniques, biochemists can understand and solve biological problems. Biochemistry focuses on processes happening at a molecular level. [...] Biochemists need to understand how the structure of a molecule relates to its function, allowing them to predict how molecules will interact."[7]

From this definition, we can understand two main aspects of biochemistry. The first is that biochemistry is an operational science, as it is based on chemical knowledge and techniques to deal with what is defined as a biological problem. This leaves the set of kinds relevant to biochemistry quite open. At the same time, the definition points out the crucial aspect of biochemical kinds that will be discussed in more detail in the thesis: the relation between the structure of a molecule and its functions. Given the nature of the problems posited by biochemistry and their orientation towards the relation between different sciences, the philosophical analysis of biochemical kinds and functions can inform the discipline of biochemistry itself, offering additional conceptual tools to deal with its core problems. Moreover, the philosophical analysis of biochemistry can also point the way to broader conclusions about the unity of science in general, as will be discussed in §1.5 and Chapter 8.

From the definition presented, the biochemical domain and the related biochemical kinds can be seen as the ones at the interface between chemistry and biology, and this can make them interesting to explore the relations between the two domains. Moreover, multiple determinability

---

[7] From website of the *Biochemical Society*: `https://biochemistry.org/`.

and multiple realisability seem common at the biochemical level. There are chemical molecules that can play many functions and be determined in many different biochemical kinds, and there are biochemical functions that can be realised by different chemical structures and molecules [Chapter 3, Chapter 5, Chapter 6, Chapter 8]. Biochemical kinds, exactly for these features, have attracted the attention of philosophers in recent debates (Slater 2009; Tobin 2010; Goodwin 2011; Bartol 2016; Havstad 2016, 2018; Kistler 2018; Tahko 2020, 2021). There are three main reasons why this is so and why these kinds are worth analysing when exploring questions of unity. The first is that there are controversies concerning taxonomies of biochemical kinds. Whether we favour a chemical or a biological approach, we can classify the same compound differently (see for more Slater 2009; Havstad 2016, 2018) [Chapter 6, Chapter 7]. The second is that it is difficult to understand whether there is something such as a biochemical natural kind or whether there are chemical or biological kinds only (see Bartol 2016). Third, discussing these kinds can inform debates concerning the unity of science and the relations between the chemical structure of biochemical kinds and their function in biological systems (Bartol 2016; Kistler 2018; Tahko 2020, 2021).[8] The purpose of this thesis is to contribute to the discussion that relates to the second and the third question, providing a metaphysical analysis of biochemical natural kinds and the relation between a chemical structure and a given function. This has been done by employing two main case studies, the molecular gene and vitamin B12. While the problem of taxonomy will not be considered in detail, the answers to the previous questions also offer more tools to answer that [Chapter 4, Chapter 7].

Before concluding, some readers might be concerned that the molecular gene does not represent a standard instance of a biochemical kind, as they may relate the entity to molecular biology instead. To answer this concern, I would like to stress that the main interest of the present research is to explore the relations and kinds in the biochemical realm with a focus on the chemical and biological properties that some molecules display. The molecular gene represents a good case

---

[8] These themes are related to broader discussions in the metaphysics and philosophy of science, such as whether we should be monist or pluralist about natural kinds and whether we should favour natural or pragmatic considerations when considering entities characterised by different properties (Goodwin 2011). Monists about natural kinds sustain that there is one way to carve nature at the joints, while pluralists maintain that it is possible to do it in a plurality of ways.

study for this endeavour. Moreover, while the inquiry takes contemporary science as its starting point, the division between domain and subdomains varies across scientific contexts and institutions and should not be taken as a rigid indicator for ontological explorations. Granted this, the relation between molecular biology and biochemistry is very tight. If biochemistry is what explores chemical processes within and related to living organisms, this includes the processes involved in DNA and RNA expressions. This is also acknowledged by the definition of the *Biochemical Society*, which reports: "Biochemistry covers a range of scientific disciplines, including *genetics*, microbiology, forensics, plant science and medicine" (emph. added).[9] Accordingly, I believe the molecular gene to be an interesting and informative case study for issues concerning kindhood and relations in the biochemical domain.

## 1.3 Methodology

In this section, I will briefly present the methodology of this research and its assumptions. This research developed as a growing dialogue between the study of detailed case studies and theoretical metaphysical reflections on them. In particular, I have been focusing on the molecular gene as a core informative case study concerning relations at the biochemical level. As such, the molecular gene represents an ideal bridge between chemistry and biology given its strict relation with nucleic acids and its crucial role in protein synthesis and other important biochemical phenomena. This case study has been particularly illuminating concerning some of the biochemical world's core features, which will be expressed in more detail in the thesis. Specifically, the gene allows us to understand the importance of the functional component and context dependence of biochemical phenomena, together with the relation with the chemical features realising them.[10] The study of these features has been further supported by a second case study, vitamin B12. Vitamin B12 has been chosen because its complexity makes it a good case to explore issues concerning biochemical functions and unity at the biochemical level. Specifically, vitamin B12

---

[9] From website of the *Biochemical Society*: `https://biochemistry.org/`.

[10] This has been done with a visiting period at the Developmental Biology Lab at the University of Pavia (IT). The exchange with scientists and practitioners has been essential for this project.

presents multiple realisation, as the different vitamers compounds are accumulated by having the same functions, and a case of multiple determination, as the vitamin plays different roles in different biological processes.[11]

Let us now consider some assumptions and methodological principles that underpin this work; these are a moderately naturalistic metaphysics framework and the application of principles that try to balance the input from science and case studies with philosophical and metaphysical analysis.

### 1.3.1 Moderately naturalistic metaphysics

In recent debates, methodological discussions on how to do metaphysics of science and the direction of informativeness between scientific research and philosophical one have been common (Ladyman 2012; Tahko, Morganti 2017; Wilson 2019). This is particularly relevant considering the interest in the unity of science and the naturalist approach to metaphysics (Ladyman, Ross 2007; Ladyman 2012). According to this approach, “metaphysics should be motivated exclusively by attempts to unify hypotheses and theories taken seriously by contemporary sciences” (Ladyman, Ross 2007, 1). While this thesis operates in the framework of discussing the implications of contemporary science concerning unity, here I have specifically endorsed a moderate version of naturalism, as summarised by Morganti and Tahko (2017).

“Moderately naturalistic metaphysics” is based on a continuity between metaphysics and science and a continuity between a priori and a posteriori reasoning, grounded in the idea that they cannot exist on their own independently from each other. This is based on the assumption that metaphysics and science have the same domain of inquiry (the world) but at a different level of generality. More precisely, metaphysics is the discipline that studies the nature and structure of reality, the ontological categories involved and the relations between them at the most general level. The sciences study the nature and structure of reality and focus on specific types of entities and domains such as the physical, chemical, biological, and so on. Moreover, the two disciplines differ in terms of the methodology used, one is mostly a priori, and the second is mostly a pos-

---

[11] More on vitamin B12 and its properties will be presented in Chapter 5, Chapter 6, Chapter 7.

teriori. However, mutual informativeness is exchanged between the two, and none of the two methods can be seen as applied on its own. For instance, science can represent a way of indirect testing for metaphysical hypotheses, and metaphysics can implement or provide an interpretation of our best scientific theories in dialogue with science. Accordingly, metaphysics done in a moderately naturalistic way considers these levels of mutual informativeness and proceeds in its analysis always being informed by both the scientific and the metaphysical framework. This approach has been used throughout the thesis and in the inquiry into the nature of biochemical kinds. I have used tools from metaphysics for the concepts of natural kindhood, emergence, and realisation (and others). These tools have been applied considering how biochemical kinds and their behaviour are described in the sciences via the usage and consideration of detailed case studies, particularly molecular genes and vitamin B12.

### 1.3.2 Reflective equilibrium and illuminating accommodation

Together with a moderately naturalistic approach, two further methodological points have been applied. The first is the one formulated by Goodman (1954/1979) and reported by Khalidi as "reflective equilibrium" (RE) (2013):

> **RE**: In order to answer a given metaphysical question, we need to consider different factors: i) philosophical usage of a given category; ii) pre-theoretic convictions regarding types of paradigmatic categories; iii) considerations coming from related philosophical studies; iv) scientific evidence and theorising on the given category.

This principle has been important to underpin the relevant concept of natural kinds, as in Khalidi (2013). This definition of a natural kind needs to consider i) philosophical considerations normally involved in natural kinds, such as being projectible categories; ii) pre-theoretic convictions about these kinds, such as the fact that electrons, chemical elements and species are normally paradigmatic instances of kinds; iii) considerations coming from philosophical reflections on laws of nature, causation and universals that come from other philosophical studies; iv) scientific evidence coming from the study of natural categories (as will be discussed in Chapter 2). This is particularly relevant when considering both which kinds we can deem natural and the properties relevant for the particular kinds to be so. In the first case, the main criterion for the naturalness of a kind can be given by the fact that they figure in scientific theories. In the second case, if some properties are scientifically relevant for a given kind to be what it is, such as chemical structure and function in biological processes in the case of biochemical kinds, then we need to consider these [Chapter 2, Chapter 4, Chapter 7].

The second principle that is methodologically relevant to this research is the "criterion of illuminating accommodation" (IA). Re-elaborating Wilson's formulation (2019, 2021), this criterion can be spelt out as:

**IA**: An adequate account of a metaphysical phenomenon (such as natural kindhood or emergence) should provide an illuminating basis (explanatory relevant) for accommodating the appearances of this phenomenon in a natural and straightforward fashion.

This principle has been applied to both natural kindhood and weak emergence. An account of natural kindhood should be informative and explanatory relevant regarding its appearance: it should provide explanatory reasons for a kind being natural. The same can be said for emergence, as an account of emergence should provide informative and explanatory reasons for which a feature or entity is emergent compared to being reducible to other lower-level features [Chapter 4, Chapter 6].

The combination of these two principles and a moderately naturalistic approach underpins the discussion and the arguments present in this thesis.

## 1.4 Outcome and implications

Before concluding this introduction, I would like to briefly summarise what I think is the main outcome of this project and its implications. This work's main conclusion is that biochemical kinds are *sui generis* natural kinds. Often in the literature, they are treated as chemical compounds (assuming a form of microstructuralism), or they can be considered biological kinds or instances of both (see Bartol 2016). However, I will argue that this characterisation does not do justice to their nature as chemical compounds that also play a role in a given biological system. Moreover, given the moderate naturalism assumed in the thesis, the growing relevance of biochemistry as a foundational science for the life sciences offers support for the ontological autonomy of biochemical kinds. This ontological autonomy is argued by two main argumentative strategies. The first considers the weak emergence of the instances of biochemical kinds, particularly biochemical functions [Chapter 3, Chapter 6]. As will be spelt out, weak emergence can be seen as a combination of dependence and autonomy of the emergent entities from the lower-level ones. In particular, I will consider weak emergence in terms of novelty and robustness (Franklin, Knox 2019), together with the proper-subset view presented by Wilson (2011, 2015, 2021). The second considers the natural kindhood of biochemical kinds [Chapter 4, Chapter 7]. This will be done following an account for which natural kinds are those categories postulated by science that are projectible and nodes in causal networks (as in Khalidi 2013, 2018) [Chapter 2, Chapter 4, Chapter 7].

This analysis can be seen to have at least three broader implications. While these will not be discussed in detail throughout the thesis, whose focus is on biochemical kinds, the reader might still find them interesting for the unity of science debate and other ones in the metaphysics of science.

One of the implications of this work concerns the context from which this research started: the unity of science. Specifically, the naturalness of biochemical kinds and the weak emergence of biochemical functions could be seen as a support the disunity of science advocated by Duprè (1983,1993). However, I will argue for the opposite and that this does not leave us necessarily with a dis-unified or disconnected world. This thesis suggests a form of unity with three grounds. The first way to maintain unity is via the form of material dependence that the biochemical has on the chemical features, which will be discussed in the account of weak emergence considered (as Wilson 2021) [Chapter 3, Chapter 6]. This offers us a view of the relation between the different entities that are not disunified. The second way to maintain unity is via Natural Kind Monism, for which if we can identify one notion of natural kinds that is applicable across the sciences, then a form of unity can be maintained (as Tahko 2021) [Chapter 2, Chapter 4, Chapter 7]. This contrasts with natural kind pluralism. The third is a sense of unity that is preserved when it is possible to identify how the different features of the world are connected. The identification of these relations (such as emergence, multiple realisation or multiple determination, casual networks) offers us a view of the world that is unified by their very existence, for how complicated these might be. In a recent book, Nancy Cartwright claims that denying unity is linked to denying the primacy of physics (2022). Here, I put forward a view of the unity of science that is not based on the primacy of physics or the reducibility of everything to physics but rather a different sense of metaphysical unity. This unity can be seen in acknowledging the connections and the relations between the different things in the world, together with identifying a single notion of natural kinds that can be applied across instances. This view develops in the same grounds as Mitchell's integrative pluralism, view proposed for the life sciences (2003). As a result, there is a form of autonomy for biochemistry - or each special science - and biochemical kinds and properties exist and are natural, but the final picture of the world is one unified by the relations that exist between objects. This can be seen as an encouragement for the exploration and analysis of how the different entities relate and are connected, leading towards an inter- and cross-disciplinary approach to understanding and explaining natural phenomena.[12]

The second implication of the thesis is that the discussions on emergence and natural kindhood at the biochemical levels can be done without committing to the existence of hierarchically ordered ontological levels corresponding to the domains of the different sciences. As one can see from the picture representing the Oppenheim-Putnam model of the world [Figure 2], it is intuitive and common to think of each scientific discipline as being associated with a level of reality (also discussed in Heil (2003) and Potochnik (2021)).[13] Physics corresponds to the physical level, which is also the most fundamental, chemistry to the chemical level and so on. These levels are often thought of as discipline relative, hierarchically ordered and neatly correspondent to types of kinds. They have been identified in terms of predicate and reference (as in Heil 2003), scale-relativeness both spatial (from the smallest entities at the fundamental level to the bigger ones in the higher levels) and temporal, composition, mechanisms together with reference to the various

---

[12] I thank Samuel Kimpton-Nye for noticing this aspect of the present work.

[13] As underlined by Heil, the idea that there are levels of reality is discussed in metaphysics, philosophy of science, and philosophy of mind, but it is also common among "biologists, psychologists, anthropologists, historians, journalists" (Heil 2003, 1).

disciplines or levels of explanations[14] (Potochnik 2021, Gillett 2021). However, while such level talk is ubiquitous, a precise and specific level concept and the identification of each ontological level remain difficult. It seems fair to ask if and how the world can be really layered into such levels (Potochnik 2021). For example, physics is concerned with both very big and very small objects, making the identification of the "physical level" difficult. Levels are also often thought of as being hierarchically ordered and ontologically neat, without kinds cross-cutting them or instances of the levels belonging to more than one. However, the exclusion of cross-cutting kinds (such as the various species kinds and kinds that can be associated with developmental stages in biology) is problematic because cross-cutting kinds are ubiquitous in the sciences [more in Chapter 2].[15] Following Khalidi, why should the world be hierarchically ordered and divided into neat and not-crosscutting categories (1998, 2013)? The combination of the difficulty in providing a good identification of the various levels and the difficulty in identifying the structure of the hierarchy makes the commitment to such levels philosophically complex. Accordingly, in the thesis, I will remain neutral towards the existence of levels. Level-talk will be present (such as relations held between higher/lower level). However, levels should be taken to be nothing but a tool and a heuristic[16] to explore different relations of dependence that exist between clusters of properties that are generally associated with given kinds of kinds, such as the features that are typical of biological or chemical kinds.[17] Avoiding commitment to neat, hierarchically ordered discipline relative levels does not leave us with a flat ontology but rather with an ontology that comprises different dependence relations and connections between chemical, biological and

---

[14] Levels of explanation are commonly discussed in the mechanistic literature and compositional explanations, as in Gillett (2021). However, he clarifies that the commitment to levels in compositional explanations might concern "internal ontologies" that is "the study of the ontological posits of certain successful scientific practices or products in models and explanations". He contrasts this with a different ontological discussion that he calls "ultimate ontology", which considers the nature of reality itself.

[15] Khalidi presents the example of insects that can be classified according to both their species and evolutionary history and according to the different developmental stages. In developmental biology, the reference to the kind "larva" is important for explanatory purposes, while in evolutionary biology, the reference to the different species might be the explanatory relevant one (Khalidi 2013, 70).

[16] Brooks and Eronen (2018) suggest that the concept and usage of levels can be as a useful heuristic for scientific biological practice; here, I have been using some level-talk to explore different relations of dependence.

[17] While the identification of a general and widely applicable definition of levels remains difficult, this does not imply that one should not appreciate the usefulness of other related or opposed concepts, as underlined by Potochnick 2021.

biochemical kinds. I deem biochemical kinds as natural and emergent (together with biochemical functions) and dependent and related to other kinds, such as biological or chemical ones. However, this does not lead us to the ontological thesis that there is a genuine biochemical level nor to commit to the existence of hierarchically ordered levels. Instead, these kinds exist in a network of dependence relations. The neutrality towards the existence of ontological levels might be considered an advantage and an interesting implication of this research project, whose results aim to remain valid independently of the commitment to hierarchically ordered levels. This can also invite, as suggested by Potochnick (2021), future analysis of the relations between different kinds of entities without employing ontological commitment to hierarchical levels.[18]

Lastly, this project can invite a further philosophical inquiry into the evolutionary properties of biochemical molecules [as will be discussed in Chapter 6, Chapter 7, Chapter 8]. As will be analysed, biochemical kinds seem to have also some forms of evolutionary properties that can be seen as more or less direct. These properties have an important role in their understanding and in their natural status. However, an analysis of the nature of the evolutionary properties that can be ascribed to biochemical molecules would require further studies on the philosophical implications of molecular evolution. Together with being an interesting domain of inquiry on its own, an analysis of the philosophical implications of molecular evolution can impact the understanding of the possible weak and strong emergence of biochemical kinds [as Chapter 8]. This could further help us to understand the relations between chemistry, biochemistry and biology.

Concluding, I hope that the reader will enjoy this journey into the metaphysical analysis and understanding of the biochemical realm and will appreciate how this can have different implications for the metaphysics of science and the relations between different entities of the world.

---

[18] I thank Jessica Wilson for underlying this implication of my work.

# Chapter 2

# Natural kinds

Division and classification of things into kind categories are common in science and daily life.[1] I recognise that bread is my favourite breakfast, and biologists study proteins and amino acids and find microscopes in their laboratories. We notice that some properties or features are co-occurrent in some individuals, and we cluster such groups of properties into kinds. These groupings are then associated with certain labels or predicates that allow the classification of individuals or relevant phenomena (as in Khalidi 2010; Tobin 2013). They allow us to make useful generalisations, being in our explanations and inductive inferences and proliferate from every corner of our lives. Moreover, kinds play a critical role in science. At least a part of scientific practice is based on clustering individuals into different categories and making explanations and predictions about them (Bird, Tobin 2022). For instance, a biologist is generally not interested in the individual instance of a protein she is studying or in the individual amino-acids string in her lab, but rather she aims at knowing something about the general category "protein". The knowledge of the kind allows her to make valid generalisations across different instances, contexts and laboratories. Also, identifying these categories should provide explanatory power and is often taken to be informative about the world.

Nevertheless, the ubiquity and variety of kinds brought the status of these categories into

---

[1] Parts of this chapter are taken from the article "The gene as a natural kind", forthcoming the Edited Springer Volume *Life and Mind.*

question. Philosophers asked first what these categories are and second which are natural and which are instead instrumental or artificial (for instance in Griffiths 2004; Khalidi 1998, 2013; Magnus 2015, 2018; Bird 2018; Crane 2021; Tahko 2022; Bird, Tobin 2022). In this respect, it is important to distinguish three different questions that one can consider concerning the topic of natural kind (Magnus 2015, 2018; Bird 2018; Tahko 2022, 144):[2]

- The"ontological question" asks which kind of entity (if it is an entity at all) is a natural kind. An answer might come from a theory of *sui generis* universals, clusters of universal properties or groups of similar individual objects.
- The "naturalness question" asks the difference between an arbitrary category from one that captures some genuine divisions in nature. What must a category do to "carve nature at its joints"?
- The "unification question" asks how the properties of given kinds are clustered together (as discussed in Tahko 2022).

In this introductory chapter and throughout the thesis, I will be concerned mostly with the second question, considering what makes a kind *natural*. Such analysis is often at the core of debates concerning reductionism, pluralism and unity of science - as seen in the traditional exchange between Fodor (1974) and Kim (1992) on these issues.[3] Specifically, questions about the ontological reduction of an entity are often questions concerning the type or kind that can be reduced (as presented by Tahko 2021). One can individuate a given natural kind, such as a "chemical molecule", and then see if it can be reduced to more fundamental kinds, such as elements or atoms. The topic of natural kinds is also relevant concerning forms of disunity advocated based on kinds pluralism. If there is a plurality of different kinds that present different features, then the sciences are not unified but disunified (see Duprè 1983, 1993). Accordingly,

---

[2] Tahko distinguishes three questions about kinds that he labels the what, the why and the how questions. The first asks *what* are the natural kinds as the type of real entities; the second asks *why* the properties of various members of a kind are clustered together; the third asks how such properties are unified in the various instances (2022,144).

[3] This debate considers whether kinds can be reducible or not upon the consideration of their multiple realisability. As mentioned in the previous chapter, the kind "money" for instance can be seen as multiply realisable. Other instances of multiple realisability will be discussed throughout the thesis.

the presentation of which notion of natural kind is at play is relevant to any discussion on unity in biochemistry and biochemical kinds. Moreover, it offers the context to consider whether biochemical molecules can be considered instances of biological or chemical kinds or whether they should be considered instances of the natural category "biochemical kind".

While here I do not aim to present a complete and extensive overview of the topic of natural kinds, this introductory chapter hopes to offer a context for the discussion that will be useful to the reader, especially concerning Chapter 4 and Chapter 7 of the thesis. The structure of this chapter is the following. Section §2 will consider in more detail the naturalness question. Section §3 will summarise the position of natural kinds I will use throughout the thesis, the one presented in various works by Khalidi (1998, 2013, 2016, 2018, 2020, 2021). This view has the benefit of answering the naturalness question capturing some of the features commonly ascribed to natural kinds while avoiding too heavy metaphysical commitments.[4] As will be further clarified, Khalidi's position is weak or moderate realist, for which kinds correspond to genuine joints in nature but not to distinct metaphysical categories. While metaphysical commitments are not a problem *per se*, their absence makes the view compatible with more accounts and different enquires in the philosophy of the natural world. This is because the account focuses mostly on stressing the existence of natural joints in nature and how we can identify which categories capture the joints, answering thus the naturalness question rather than answering the ontological one. Focusing on the naturalness question can make the view compatible with different answers to the ontological question. For instance, one can answer it with a form of nominalism, for which the genuine joints would correspond to clusters of similarity.[5] Or one can take a more "metaphysically loaded" view for which kinds can be considered clusters of properties or *sui generis* kind universals. This can be done while maintaining the criteria for naturalness set by the discussed account. Moreover, this view is compatible with the kinds discussed in different sciences. It can be applied to different kinds, such as historical and etiological ones, and to different instances of them (Khalidi 2018).

---

[4] The importance of the projectibility and the causal role of kinds is also underlined in Tobin (2013) and the overview on Natural kinds done by Bird, Tobin (2022).

[5] For instance, according to resemblance nominalism, kinds and properties correspond to classes of particulars that resemble each others, where resemblance is a primitive relation (see Rodriguez-Pereyra 1999, 2015).

In this chapter, I will illustrate this theory of natural kinds with the example of viruses. This will be used to present the benefits of his position. In section §4, I will also offer two possible answers to the unification and ontological question in a way compatible with the account. Specifically, while taking a stronger metaphysical view on kinds is not necessary, having an answer to the two questions can offer a more detailed account of the projectibility of natural kind categories and their causal relations. Lastly, in section §5, I will present some concluding remarks considering how natural kinds can relate to the unity of science and the debate on biochemical kinds.

## 2.1 The naturalness of kinds

Broadly, one can identify two main ways to answer to the naturalness question: conventionalism and various forms of realism.

Conventionalism provides a negative answer: there are no categories that can correspond to the joints of the world because there are only conventions that suit different purposes. In this framework, kinds are seen as categories that depend upon the choice of the "naturalist" and, accordingly, do not seem to track or map natural groupings of similarity but rather various criteria of utility (Bird 2018; Tahko 2022). For example, the division of fruit and vegetables in cooking practice seems to be different from what we would consider a natural cluster based on empirical data (following Bird 2018). Conventionalism (also called constructivism or constructionism) can be framed in terms of instrumentalism, for which all the categories used in science or other practices are groups that serve the various purposes of science (Bird, Tobin 2022). These groups might not be completely arbitrary but can be seen as reflective of particular human criteria and interests rather than reality or metaphysical structures in the world (as Hacking 1999). For instance, Brigandt has recently stressed the relevance of humans interest for kinds classification and the usage of kinds in practice, underlying how many kinds considered natural are actually based on contingent social processes (2022). This lead him to argue that philosophers should stop using the term "natural" kind in general.

In contrast with forms of conventionalism, there are various forms of realism. According to realism, the ascription of naturalness to a kind is related to the idea that such a category captures some real structure in the world rather than the interests or conventions of different epistemic practices (Bird, Tobin 2022). This view can be formulated in terms of strong and weak realism (Bird 2018). Strong realism claims that natural kinds are genuine entities of the world and is metaphysically committed to the existence of such entities. For instance, kinds can be considered *sui generis* universals or clusters of universal properties. It is defined as strong because of the metaphysical commitment it takes. Weak realism agrees that the structure of reality presents divisions among things and that our actual categorisations can succeed in matching them but does not take further metaphysical commitments. Some realist approaches to natural kinds can also be formulated as *naturalist*. According to a naturalist approach to natural kinds, the categories that capture the divisions in the structure of the world are those postulated by successful science.[6]

Generally, it is common to take a hybrid answer to the naturalness question and accept that some kinds that we find in scientific practice or daily life are conventions. In contrast, others might correspond to real features of the world. For instance, one can consider the kind "iron" as a natural kind, tracking some division in the world, and the kind "vegetable" in cooking as a category only useful for some practices. For the sake of the present analysis, I will assume a form of minimal naturalism for which at least some of the categories postulated by science correspond to genuine features of reality and might be candidates for natural kinds (Khalidi 2013; Bird, Tobin 2018). However, not all categories we identify as kinds seem to correspond to such divisions. In this context, how do we distinguish natural kinds from mere human categories? Various answers have been provided to this question. Some of them share some common features: natural kinds should be predictively and explanatory successful in a projectible way, they have a link with causation and causal structures, and they should be robust enough across different

---

[6] Weak realism as a form of naturalism can be associated to the following criteria for natural kinds (Bird, Tobin 2022): 1) Members of a natural kind should share some natural properties ; 2) Natural kinds should permit inductive inferences; 3) Natural kinds should participate in laws of nature; 4) Members of a natural kind should form a kind. Some authors also endorse forms of hierarchy and categorical distinctness between kinds. This thesis will be mostly concerned with the role of natural kinds in inductive inferences and in causal networks, as will be specified.

generalisations (Tobin 2013; Khalidi 2013, 2018, Bird, Tobin 2022). Here, I have decided to focus on a recently proposed account of natural kinds, the one presented by Khalidi. His proposal summarises these features of natural kinds while applying them to different instances in different contexts, in both physical and special sciences, and thus is a good starting point to inquire into the naturalness of biochemical categories. Moreover, as mentioned above, his theory can also be compatible with different metaphysical commitments. While he presents his view as a form of weak realism, this account of the naturalness of kinds does not exclude different answers to the ontological question and might be seen as compatible with more nominalist or more realist approaches.

## 2.2 Khalidi's view of natural kinds

Khalidi's theory of natural kinds aims at answering the natural question (1998, 2013, 2016, 2018, 2020, 2021). In his account, he subscribes to the methodological principle of "reflective equilibrium" [Chapter 1]. Accordingly, his proposal balances i) properties normally ascribed to natural kinds in philosophical debate, such as being projectible categories; ii) pre-theoretic convictions about paradigmatic instances of kinds, such as chemical elements or biological species; iii) considerations coming from philosophical reflections related to the topic of natural kinds, as the debate on laws of nature, causation and universals; iv) scientific evidence coming from the study of natural categories.

His view can be considered a form of weak realism for which kinds are objective features of reality but not necessarily corresponding to distinct metaphysical categories. Nature has some joints, objective and discoverable features of the world, and our best theorising should aim at "carving up" these joints. Natural kinds are not invented but rather discovered. However, how do we distinguish between natural and non-natural or conventional kinds, given that not all of our categories seem to correspond to natural kinds? One obvious place to look at to answer the naturalness question should be science and scientific practice (Khalidi 2013, 2021). In detail, taking a realist stance towards science, one of its aims can be seen as identifying the kinds existent in the world and not mere and useful theorisations. And the success of science seems to support this point. This is acknowledged by various accounts of natural kinds in the philosophy of science, which seem to agree that kinds can be discovered by our best scientific theories (Khalidi 2013; Bird 2018; Bird, Tobin 2022). Accordingly, scientific theories or established knowledge represent our access and guide to the existence of such kinds, while they do not determine their existence (Khalidi 2013).

Nevertheless, even within the best scientific theories, the history and philosophy of science have shown that not all categories present in the discipline can be considered as capturing something in the world. For instance, a category like "hysteria", which has been used as a scientific category in the past, has proved not to be a natural kind of disease and was abandoned as a kind (Khalidi 2013, 59). Moreover, some categories can have an instrumental role or cannot be considered stable or robust enough to be informative about the world.[7] An example can be the case of "hardness". This property is used in some scientific explanations in given sciences. However, it seems to refer to various conditions in each instance and use and might be considered a conventional (even if useful) category (Wilson 1985).[8] In order to distinguish natural from non-natural kinds, even within those postulated by science, we need something more. So, in this regard, Khalidi proposes a solution that elaborates on the "homeostatic property cluster" (HPC) theory of kinds proposed by Richard Boyd (1989, 1991) and the "simple causal theory" of natural kinds proposed by Craver (2009).

The HPC theory of natural kinds can be considered one of the main theories of natural kinds in the contemporary debate. According to this view, natural kinds are associated with a cluster of properties that are kept in homeostasis (equilibrium) with each other by *causal mechanisms* whose instantiation is caused by various mechanisms. This view does not postulate any property as necessary for the membership of the kind but rather *mechanisms* of equilibrium between the

---

[7] Griffiths (2004, 907) distinguishes between kinds and "investigative" kinds that are useful for the investigation of the world and can be seen as tentative categories. According to the presented account, all kinds are, in a way, investigative in that we discover them in our attempt to know about the world.

[8] A discussion of hardness is present in Wilson (1985, 30). He writes: "Although hardness certainly seems like a perfectly clear and distinct notion (otherwise Descartes wouldn't have liked it), in reality, it happens that it represents only a loose family of macroscopic, instrument linked traits." (Wilson 1985, 30).

main properties associated with a given category. For instance, biological species are seen as paradigmatic cases of HPC kinds: the mechanism of interbreeding together with genetic descent or environmental pressure can be seen as what clusters the properties together. Thanks to the postulation of simply mechanisms of equilibrium of properties of the considered category, this account applies to different instances of natural kinds. Nevertheless, some philosophers have criticised the current view - as summarised by Khalidi (2013) - because the kind might risk being associated with the mechanism itself. This can compromise the generality of this account as there are instances of kinds that are not associated with a precise mechanism. Moreover, one could doubt whether homeostasis is necessary for kind membership. Some kinds, such as species, seem to be characterised by having changing and dynamic properties instead of stable ones. So, while the HPC stresses the importance of causality for the naturalness of kind, it might be doing too many steps in the direction of homeostasis and mechanisms to be widely applicable. The relevance of causality has been recognised and put forward in Craver's "simple causal theory of natural kinds", a re-elaboration of HPC (Craver 2009). He states that natural kinds are those "kinds appearing in generalisations that correctly describe the causal structure of the world" without committing to homeostatic mechanisms (Craver 2009, 579).

From these two theories, Khalidi suggests that natural kinds are categories postulated by the natural sciences that capture the world's causal structure and present two broad features: i) genuine projectibility and ii) being a "node" in a causal network. Genuine projectibility is the capacity to project some properties from an instance of a kind to another: the properties of the kind can point to other properties reliably, allowing us to make strong inductive inferences. Being a node in a causal network implies that the core properties of a natural kind should be causally linked with each other and the kind should have a causal impact on its surroundings. Let us look at these two features in more detail.

### 2.2.1 Projectibility and causal networks

Natural kinds are often considered particularly efficacious categories when framing inductive inferences, and they feature in many empirically verifiable generalisations. This characteristic of natural kinds can be presented in terms of projectibility: the observation of some natural kind K allows the inference that other instances of K have and will have the same properties or features (Griffiths 2004; Tobin 2013; Khalidi 2013; Crane 2021). More precisely, the projectibility of a natural kind can be defined as follows:

> "when it comes to a natural kind predicate K, there is no shortage of other predicates, P1, P2, ..., Pn, and so on, such that we can reliably assert that if x is K, then x is P1, x is P2, ..., x is Pn and we can do so with a high degree of generality" (Khalidi 2018, 1385).

This projectibility can be taken as a sign of the naturalness of the kinds. These categories have explanatory and predictive power across different contexts and circumstances by being projectible. They are so because they capture some natural carvings present in the world (Khalidi 2013, 2018, 2020, 2021). Nevertheless, while projectibility can be a good hint about the naturalness of the considered category, it still requires an explanation: why is it possible to draw these inferences? This can be answered by adding a second requirement that natural kinds should satisfy: they are "nodes" in causal networks. In fact, projectibility is possible because it results to be a "reflection" of the causal network in which instances of the kinds are involved. Some kinds of categories are particularly successful because their properties are causally clustered (Khalidi 2013, 2018, 2020). The joints that natural kinds carve so successfully can be found in the world's causal structure.

Together with providing an ontological reason for the projectibility of kinds, causal relations provide further grounds to answer the naturalness question. Specifically, considering kinds that correspond to causal nodes in causal networks offers us two main reasons for distinguishing natural from conventional kinds. First, natural kinds do not only present a set of projectible properties but a set of properties that are hierarchically ordered as "causes and effects in recurrent causal processes" (Khalidi 2018). Second, natural kinds are those categories that represent nodes within broader causal processes: they are causally efficacious on other kinds and are intersections within the webs of causal relations. This causal profile should be identified as being robust and long-lasting, underpinning genuine projectible generalisations (Khalidi 2016).

Let us look at these features in more detail. Natural kinds are not taken to present any cluster of properties or merely correlated ones but those properties that are linked to each other in a robust and identifiable way. Among these properties, some are “derivative,” and others are “core” properties. The instantiation of the derivative properties of a given kind is caused by a set of core ones upon which they depend. A natural kind is then associated with a series of properties, some of which are core and others that causally depend on them. Let us take a standard example, such as lithium (discussed in detail by Khalidi 2013, 166-171). Lithium is a chemical element with a series of properties, such as having atomic number 3, being an alkali metal and so on. A detailed analysis of these properties, all associated with the kind, underlines that the property of having atomic number 3 seems to cause the instantiation of the other ones, such as being solid at room temperature, being corrosive, being an alkali metal, and having a precise melting point. While the detailed and complex characterisation of the derivative properties can be provided, they seem to be at least partially caused by some microstructural properties that depend on the atomic number that might be considered the core property of lithium. Accordingly, we can track a causal link between the core properties of lithium and its other properties. These properties also make the instantiations of the kind “lithium” causally responsible for different causal processes, making the kind a node within different causal processes.

Before concluding, it is also important to present a distinction between kinds of kinds according to the properties they possess: the one between “monothetic” and “polythetic kinds”. A monothetic kind is “associated with a property or set of properties of which is singly necessary for membership in the kind and all of which are jointly sufficient” (Khalidi 2013, 16). While “polythetic kinds” (or cluster kinds) do not satisfy this condition because we cannot identify a clear set of necessary and sufficient conditions in every instance. Instances of monothetic kinds will then possess all and only the same properties in quality of being members of the kind. Instances of a polythetic kind can possess different properties. Monothetic kinds can be found in fundamental sciences, such as physics; for example, members of the kind “proton” share the properties “*positive charge 1.6 x* $10^{-19}$ *C, mass of 1.7x* $10^{-27}$*kg, and spin 1/2*”. Polythetic kinds can be found across different sciences, and a common example can be seen in biological species given the difficulty (or impossibility) of identifying a well-established set of necessary and sufficient properties shared by all the instances. For example, most of the instances of the chipmunk species (*Tamias striatus*) share some of the phenotypic properties associated with the kind, such as having a striped back, and these properties are derivative of the "core" evolutionary properties of the chipmunk. Nevertheless, not all instances of the species do share the same set of properties.

Summarising, natural kinds are scientific categories with a set of projectible properties whose co-instantiation causes the instantiation of other properties and have a role in the causal network of the world (Khalidi 2013, 2018). Being nodes in a causal network is the ontological principle for distinguishing natural from unnatural kinds. Moreover, it is the underpinning of the projectibility of such categories.

### 2.2.2 Viruses: an instance of natural kinds

To further elucidate the account, let us briefly consider an instance of a natural kind presented by Khalidi: the case of viruses (2013, 180). Viruses represent an established category of a proper sub-discipline, virology. Virions, individual particles of the virus, are characterised by some synchronic causal properties. First, they are protein particles (more or less complex) containing a genome capable of making an mRNA readable by the ribosome of a host cell. Second, the virions of a given virus display a specific infectious cycle comprised of well-understood causal relations. Following Khalidi's methodology, we need to consider whether viruses present the three main features of naturalness: i) being present in scientific theories, ii) being projectible, and iii) being nodes in causal networks. The first criterion is easily satisfied, as our analysis started with a scientific category present in an established enough discipline. Let us consider the other two.

The infection cycle and the life cycle of viruses are understood in terms of common and repeatable causal processes that display a given hierarchy of relations. These causal processes allow virologists to make empirical generalisations about the different instances of viruses and their infective cycle. Such generalisations can then be projected from viruses already observed to the ones that have not been. The projectibility is broad and stable across contexts: different kinds of viruses in different circumstances present a cycle that is re-conductible to the general causal one identified by virology. This makes “virus” a projectible kind: it allows explanatory and predictive power across different instances of the kind. Furthermore, projectibility is sustained because viruses can be considered appropriate nodes in causal networks. First, the core properties of the virus, such as being a protein containing a genome capable of making an mRNA readable by a host cell, cause the instantiation of other derivative properties, such as a given infection rate and behaviour within the host cell. Accordingly, we can order the properties of the virus category following a causal hierarchy. Second, viruses enter into causal interactions in a uniform or similar pattern and have a causal impact on the network of relations in which they are embedded. Viruses are causally efficacious categories within the causal network in which they interact. Accordingly, the category virus is a natural kind as it is the object of a successful part of contemporary science, it is a projectible category, and such projectibility is based on causal relations. Specifically, the properties that pertain to the kind are causally related to each other’s and lead the members to enter into causal interactions uniformly.

Concluding, Khalidi’s theory offers a starting point to answer the naturalness question. Thanks to his account, we can identify which categories can be deemed natural by analysing categories present in scientific theories and considering their projectibility and role in causal networks.

## 2.3 The unification and the ontological question

Khalidi’s theory presents many benefits, and his naturalist approach creates a virtuous dialogue with the sciences. First, this account is a good example of the consensus reached by philosophers on how to answer the naturalness question, providing us with a way to disentangle natural from unnatural kinds (see also Tobin 2013; Bird 2018; Bird, Tobin 2022). This is done without embracing an essentialist theory that tries to provide a specific set of properties for which an individual is a member of a kind[9] that makes the theory more metaphysically committed and less

---

[9] Essentialism, as the theory for which kinds are identified by essential properties that all their members or instances share, is often considered incompatible with most kinds coming from the special sciences and often incompatible with kinds categories coming from scientific practice (Khalidi 2013; Bird, Tobin 2022).

applicable to kinds in special sciences. Moreover, it applies the principle of reflective equilibrium between scientific input and philosophy. It is a combination of convictions on categories generally regarded as paradigmatic kinds, often stable categories in scientific theories, philosophical discussions on natural kinds and a set of considerations drawn from scientific practice. Lastly, this account accepts the “naturalness” of many different kinds. This is possible because it applies well to concrete and varied case studies from different scientific disciplines, such as etiological and historical kinds as natural kinds. For instance, in these cases, one can consider a particular origin or genealogical history as the core properties that cause the instantiation of other ones (Khalidi 2021). Moreover, the combination of naturalism with projectibility and causality allows the theory to be applied to concrete case studies within fundamental physics and the special sciences, such as lithium, cancer cells and viruses (examples from Khalidi 2013).

Nevertheless, his account focuses on the naturalness question and can be complemented by answering two other questions concerning natural kinds, the ontological question and the unification question. This could provide a more complete answer to the problem of natural kinds, and this is what I will consider in this section.

### 2.3.1 Unification principles and the four-category ontology

As summarised at the beginning of the chapter, there are three different enterprises that one can explore concerning the topic of natural kind: the ontological question, the naturalness question and the unification question. Khalidi’s account can be seen as mostly oriented towards answering the naturalness question, offering a way to distinguish natural from non-natural categories. Here, I will suggest two possible ways to answer the two other questions that can be seen as compatible with Khalidi’s account. These two proposals to the ontological and unification question are not the only possible ones and aim only to provide a possible more complete account of natural kinds.

Let me start with the unification question (discussed in Tahko 2022). This question asks how the properties of given kinds are clustered together. An answer to this question is informative in assessing the mind-independence and naturalness of the various kinds. In Khalidi’s account, the properties are clustered together thanks to the causal connections between them. However, as suggested by Tahko (2022), if we do not add a more detailed answer to the unification question, this theory can incur into the "generality problem": it picks up every cluster of causally efficacious properties and there might be categories that could respect the criteria for natural kindhood but not be intuitively natural kinds. This objection is particularly relevant for social kinds or kinds whose mind independence might be difficult to assess. For instance, let us imagine a specific scientific community that holds beliefs about ghosts in order to explain some psychological phenomena related to the loss of beloved ones. In this case, it seems possible that such a scientific community can project properties from one instance to another and identify a series of causal patterns concerning them. However, we would not accept ghosts as genuine natural kinds.[10] This problem has been recently considered by Tahko (2022), who suggests that we need to answer the unification question by identifying mind-independent unification principles. The identification of this principle can be different according to the instances of kinds, and it can include HPC mechanisms, evolutionary history or any other trackable unification principle. This can provide further grounds to differentiate any bundles of causally efficacious properties from natural kinds (as discussed in Tahko 2022).[11] For instance, in the case above, it would be difficult to identify how the properties associated with the kind "ghost" are clustered together. Instead, this problem does not hold for other cases, such as the chipmunk species *Tamias striatus*, for which we can identify a particular evolutionary history and mechanisms that cluster the properties together.

The ontological question differs from both the unification and the naturalness question and asks which kind of entity (if it is an entity at all) is a natural kind. As already mentioned, one of the advantages of endorsing a weak realist view of natural kinds is it, in principle compatible with different metaphysical commitments and views on natural kinds.[12] A weak realist approach to natural kinds takes kinds as objective features of reality but leaves space to ask what these kinds correspond to (as in Tobin 2013). An answer might come from a theory of *sui generis* universals,

---

[10] The similar example of tooth-fairy is discussed in Tahko (2022).

[11] The problem of mind independence would require a more extensive consideration (see Khalidi 2016; Ereshefsky 2018; Tahko 2022). For the thesis, it suffices to identify the mind-independent unification principle in the causal connections between properties of the kind.

[12] The difference between the ontological and the naturalness question allows the possibility of genuine kinds that are not natural, such as artefacts or abstract kinds.

clusters of universal properties or groups of similar individual objects. The clarification of the ontological commitments of the position and the identification of the real features of the world to which the kinds correspond further support the answer to the naturalness question. Specifically, natural kinds can be seen as efficacious in inferences and correspond to causal networks of the world because there are natural boundaries between things, and this categorising is happening because of "the real nature" of the world and the various different kinds. Among various options, this can be accounted for because natural kinds belong to a priori distinct ontological categories, such as kinds universals (Lowe 2001, 2005, 2011; Tobin 2013; Bird 2018). More in detail, natural kinds have their role because they not only play a classificatory and inferential role, but they also have one in individuating the *sort* of thing that the instances considered are and accordingly can be considered substantial universals on their own. Considering natural kinds as sortals universal or kind universal offers further support to the roles associated with natural kinds. Among various ways to frame the answer to the ontological question by taking a stronger metaphysical commitment, one can be seen in Lowe's four-category ontology (2001, 2005, 2011).

In this ontology [Figure 3], there are four basic categories of entities: objects, kinds, properties and modes. Objects and modes are particular entities, while kinds and properties are universal. In this case, natural kinds would correspond to kinds that display a series of causally linked properties, and these are instantiated in objects and modes.

$$\begin{tikzcd}
Kinds \arrow[r, "\text{characterised by}"] \arrow[d, "\text{instantiated by}"'] & Properties \arrow[d, "\text{instantiated by}"] \arrow[dl, "exemplified\ by"] \\
Objects \arrow[r, "\text{characterised by}"'] & Tropes(Modes)
\end{tikzcd}$$

Figure 3. *Lowe's four-category ontology*: at the top level the universals and at the bottom level the particulars (Lowe 2005, 26)

The relations of dependence are a crucial feature of this ontological square (Lowe 2001). In particular, universals depend for their existence on the categories at the bottom level. The Principle of Instantiation expresses the latter relation: kinds or universal properties cannot exist un-instantiated, i.e. all universals exist in space and time, or wherever and whenever their particular instances are (Armstrong 1983; Lowe 2005). When discussing the naturalness question, it is still possible to avoid ontological commitments in the context of the metaphysics of science. However, an answer to the ontological question in terms of a four-category ontology has some advantages. First, natural kinds play a distinct individuating role via the instantiation of the kind universal, providing a criterion for the continued existence of the kind.[13] This can support the role of kinds in induction and counting different instances (as summarised by Tobin 2013). Second, a four-category ontology can offer a stronger metaphysical framework for Khalidi's account. More precisely, it could provide a clearer picture of the relations between instances of the kinds, the kinds and their properties.

Let me consider an example, an instance of chipmunk, a member of the kind species *Tamias striatus*, characterised by a particular evolutionary history and a striped back, among other properties. In this case, the kind *Tamias striatus* is characterised by the historical evolutionary properties and the various derivative properties that can derive from it. The instance of chipmunk instantiates the kind and is characterised by the modes of the historical evolutionary properties and the other properties, such as having a striped back. The modes characterising the kind are causally related and ordered in core and derivative properties. Having an answer to the ontological question in these terms also offers further support to the properties that support the naturalness of the kind. Projectibility is supported by the fact that the same kind universal is instantiated in different instances and so are its properties. This allows the inference of the properties from one instance to another because the same kind is instantiated. The role in causal networks can be seen as properties and modes causal relations (both between the core and the

---

[13] In the context of the analysis of natural kinds this role is played by natural kinds. However, Lowe's account of kinds is compatible with the existence of non-natural kinds, such as social kinds or artefact kinds.

derivative properties and between the instances of the kind and the surroundings). This can provide a clearer view of how kinds and their properties are related.

Concluding, while this is not the only ontology compatible with the framework, a four-category ontology can be a way to illuminate the interplay between kinds, instances, properties and modes to complement the answer to the naturalness question. Nevertheless, the present thesis will focus mainly on the naturalness question for biochemical categories and leaves more considerations regarding the answers to the unification and ontological question for future research.

## 2.4 A variety of kinds and natural kind monism

As previously introduced, the topic of natural kinds is relevant to debates concerning unity. Specifically, whether kinds are reducible to each other or which view of kinds one subscribes to can influence our approaches to whether a form of unity can be preserved. In recent debates, the defence of a disunity of science is often linked to kind pluralism: the world is dis-unified because a plurality of kinds is discussed in the various disciplines, as in Duprè (1983, 1993). These kinds can be taken to be different entities characterised by different properties and are not related to each other. As we shall explore in Chapter 6 and Chapter 7, biochemical kinds can be seen as an instance of this kind of disunity upon these grounds: they can be seen as instances of chemical and biological kinds in a dual, disunified way (as in Bartol 2016). However, both chemical and biological kinds can be considered natural, according to Khalidi's account, and if biochemical kinds can be considered natural kinds themselves, then this could support a monist view of natural kinds rather than a pluralistic one [as in Chapter 7].

Let me briefly present chemical and biological kinds and how these categories can be deemed both natural according to Khalidi's account. Chemical kinds are generally characterised in terms of a variety of microstructural properties that can then lead to a series of other instantiated properties related to the kind (Hendry 2006; Slater 2009; Goodwin 2011; Tobin 2010; Havstad 2018;

Bird, Tobin 2022;).[14] For lithium, for instance, some properties as the number of protons are considered core in both the causal structure and projectibility of these kinds. In virtue of the presence of these properties, we can both project the kind from one instance to another and predict and explain the behaviour of chemical kinds. Moreover, the microstructure causes, at least partially and in the right context, the instantiation of other properties and the causal role that the instances of the kind have in various causal networks. It is important to notice that, according to this account of natural kinds, microstructural properties and other macrostructural properties are associated with the kind. However, it is possible to track a series of dependence relations for which the derivative properties associated with the kind depend on the core properties

Biological kinds instead are associated with various evolutionary or diachronic properties.[15] Specifically, the evolutionary historical properties of these kinds can be considered the core properties for which the kind's instances possess various properties and can play a role in the causal network (Khalidi 2013, 2021; Bartol 2016; Bird, Tobin 2022). These evolutionary properties also make the kinds projectible and allow the inference of one property to another. Common examples are biological species identified by a specific evolutionary history. The instances of the species *Panthera tigris* share a common core set of evolutionary properties that then cause, in a varied and gradable way, the presence of the other properties associated with the kind.[16]

Biochemical kinds present properties typical of both chemical and biological cases. They have structural chemical properties, biochemical functional ones and some forms of evolutionary properties. The discussion of these kinds will be detailed in Chapter 7. However, for the purposes of this chapter, it is important to notice that their naturalness might be questioned as they can be seen at the intersection between the chemical and the biological kinds (as discussed in Bartol

---

[14] An alternative account of chemical kinds discussed is a macroscopical approach for which also macroscopical properties can be allowed for the identification of chemical kinds, such as thermodynamic properties (Serri 2020; Bird, Tobin 2022).

[15] It is possible to debate whether *all* biological kinds are historical or etiological kinds; some traits, for instance, might be identified in terms of present causal powers rather than evolutionary history, such as eyes (see Khalidi 2013). Currently, most of the literature would agree that most biological kinds are individuated by evolutionary history. Other historical kinds can be found in other disciplines, such as geology.

[16] An extensive analysis of historical and etiological kinds is Chapter 4 of Khalidi 2013 and Khalidi 2021.

2016). In particular, the naturalness of biochemical kinds can be challenged if one wants to subscribe to what is called the hierarchy thesis (HT). HT, as formulated by Khalidi (1998, 35), states that "natural categories cannot crosscut one another"; that is that one instance cannot belong to two or more kinds unless these kinds can be subsumed one after the other.[17] The HT allows only those kinds that can be properly located in the hierarchy of kinds, which might exclude biochemical kinds. The commitment to this thesis can follow from two criteria that can be associated to natural kinds in a naturalistic framework. These are that natural kinds should form a hierarchy and that natural kinds should be categorically distinct, that is there should be no smooth transition between kinds (as in Ellis 2001, 2002; Bird, Tobin 2022). However, these two criteria and the HT have been criticised in the literature, as by Khalidi (1998, 2013) and Havstad (2021), because many instances of kinds in the sciences seem to be cross-cutting. For instance, as discussed by Havstad (2021), the kind "nuclear hormone receptor superfamily" be classified simultaneously and interactively by the same scientists in the same labs within the same in a variety of ways, challenging, therefore, the possibility of one system of classification.[18] The rejection of the hierarchy thesis will be one of the starting points to discuss the issues of natural kinds in the sciences. Specifically, I will argue that biochemical kinds can be genuine natural categories even if their instances can display properties ascribable to chemical and biological kinds [Chapter 7]. This will be done without committing to the hierarchy thesis, the absence of crosscutting kinds or the presence of levels, but with the application of one notion of natural kinds presented.

The application of the same account of natural kinds across different instances from different sciences is relevant for the inquiry of unity in the biochemical world because natural kind monism can support forms of unity in science in contrast with natural kinds pluralism and the correspondent disunified view of the world. This hypothesis has been recently presented by Tahko (2021) with the principle of *Natural Kind Monism*:

NKM: "The view that there is a single notion of "natural kind" and anything falling under that

---

[17] The term of "crosscutting kinds" is original of Khalidi, as discussed in his "Natural Kinds and Crosscutting Categories" (1998) and throughout his book on the topic (2013).

[18] For more on crosscutting kinds and the detailed case study of the nuclear hormone receptor superfamily see Havstad (2021).

> notion can be defined in terms of the same general set of identity criteria." (ibid. 2021)

In this framework, a view of natural kindhood that can be applied in more contexts offers us reasons to find a form of unity: the same notion of natural kinds unifies the different sciences. This contrasts the pluralist view of kinds for which biochemical kinds would have a dual nature because they instantiate biological and chemical kinds and these kinds are different and irreducible. Instead, according to NKM, while different natural kinds can display different properties, they share the same general structure. The application of one single account of natural kinds throughout different instances supports a unified view of the world. This will support the acceptance of biochemical kinds as existent and natural together with a unified picture of the world, as we will see in the conclusions of this thesis [Chapter 8].

## 2.5 Conclusion

In this introductory chapter, I have provided a short overview of the position of natural kinds assumed in the thesis to offer some context to the following chapters. There are three questions one can deal with when considering the topic of natural kinds: the ontological question, the naturalness question and the unificatory question. In this thesis, I am concerned mainly with the naturalness question that asks how we distinguish natural from non-natural categories. This is so because the naturalness status of biochemical kinds can be challenged by the ascription of chemical and biological properties to these molecules. Moreover, this possible duality of biochemical kinds can be seen as a source of disunity in science (Bartol 2016). Accordingly, this chapter presented the view that will allow me to explore the issues concerning biochemical kinds and how a form of unity can be preserved. The view of natural kinds assumed throughout the thesis is presented in various works by Khalidi (1998, 2013, 2016, 2018, 2020, 2021). In this account, natural kinds are those categories postulated by science that are projectible and nodes in causal networks. After presenting his account, with the aid of the example of viruses, I also offered two possible ways to answer the unification and ontological question in a way compatible with the account. These additions can offer a more detailed account of the projectibility of natural kind categories and their causal relations.

Lastly, while this chapter does not aim to be definitive about an account of natural kinds, the defended view has the benefit of capturing some of the features commonly ascribed to them while being compatible with different instances and kinds of kinds. It presents a clear and easily applicable view of natural kinds that can allow us to explore different issues concerning the molecular gene [Chapter 4] and biochemical kinds in general [Chapter 7]. In particular, it will allow us to consider whether biochemical molecules are instances of chemical or biological kinds or whether they should be considered a self-standing category [Chapter 7]. Moreover, as suggested, a view of kinds compatible with different instances can provide further grounds for unity, given NKM, as will be suggested in the last chapter [Chapter 8].

Concluding, Part I represents the general introduction to the present research and has presented the context, the methodology and the main conceptual tools that underpin it. Chapter 1 has introduced the unity of science debate as the general research framework for the development of the thesis. Then it presented the methodological principles together with the primary outcome and implications. Chapter 2 has offered an introductory overview of the account of natural kinds that will be used in the more detailed analysis of the molecular gene [Chapter 4] and biochemical kinds [Chapter 7]. The account of natural kind presented will also be used to argue in favour of natural kind monism as a support for the unity of science [Chapter 7, Chapter 8].

# Part II

# A metaphysical analysis of the molecular gene

# An introduction to the molecular gene

> “Genes are ways in which cells utilise available template resources to create the biomolecules that are needed in a specific place at a specific time: *genes are things an organism can do with its genome.*”
>
> (Griffiths, Stotz 2013, 75; emph. added).

This thesis aims to investigate the relations between the biological and chemical realms [Chapter 1]. In this respect, the complex interplay between DNA, RNA and the molecular gene stands out for its implications for living organisms. This interaction is crucial for protein synthesis and other biological processes. Accordingly, this part focuses on the molecular gene as a main case study, considering the ontological status of the molecular gene and the naturalness of this category.

The molecular gene is standardly defined as a stretch of DNA that encodes the primary structure of a polypeptide or a functional RNA molecule.[19] This implies that it can be seen as composed of a part of a complex macromolecule, DNA. Three chemical components form

---

[19] As will be further specified, the gene can also be seen as composed of RNA.

the DNA itself: deoxyribose, phosphate and four nitrogenous bases that are adenine, guanine, cytosine, and thymine. These chemical components are organised into nucleotides composed of a phosphate group, a deoxyribose molecule, and any of the four bases. The DNA is then structured as two side-by-side chains of nucleotides, where thymine is always bound to adenine and guanine is always bound to cytosine. However, this account of DNA is not enough to account for the gene. As a recent analysis of the topic has shown, the gene seems something more than just a complex chemical molecule: it also needs to fulfil a *particular function*, namely encoding the primary structure of a polypeptide or a functional RNA (for instance El-Hani 2007). The function cannot be ascribed simply to linear sequences of DNA because the genome is characterised by both multiple determination and multiple realisation, i.e. the same stretch of DNA can be determined into different genes, and the same gene can be realised by different stretches of DNA [more in Chapter 3]. Accordingly, the function is ascribed to a complex entity: the gene. If we follow contemporary genetics, the molecular gene is characterised by both functional properties for its role in biological processes and by chemical structural properties, composed of chemical compounds and nucleic acids.

Here, we are at the intersection between chemistry and biology. On the one hand, the gene can be regarded as a complex chemical molecule; on the other, a proper analysis of the gene seems not exhausted by a mere chemical account. How can we provide an exhaustive analysis of the gene? Are the instances of the genes existent as something *sui generis*, and can the molecular gene be considered a natural kind? I will approach these questions in the following two chapters arguing that molecular genes are emergent from the nucleic acids they are composed of, and that the category molecular gene can be considered a natural kind. This part presents novel results as this case study has not been considered yet in the recent philosophical literature from the perspective of the metaphysics of science and can provide interesting insights for the metaphysical analysis of the biochemical domain. First, I will present a short history of the gene. This gives the reader an overview of the discovery and conceptualisation of this entity and why it is an interesting case study for the metaphysics of science. Chapter 3 will argue that the molecular gene can be considered an entity that is weakly emergent from the nucleic acids, using an account of weak emergence that will be shortly defined. Then, Chapter 4 will consider specifically whether the gene can be considered a natural kind according to Khalidi's account of natural kinds [summarised in Chapter 2].

This case study will serve two main purposes for the general argument of the thesis. The first is to show that there are two main properties at play when considering kinds at the interface between chemistry and biology, those related to molecular structure and biochemical functions. This result will be the starting point for Part III. The second is to offer an argument in favour of the main result of this research that there is something specific about biochemical kinds and their properties, by employing a specific case study. This part presents the two main theoretical steps that underpin this result in relation to the molecular gene [as in Chapter 1]: the characterisation of the relation between biochemical functions and chemical structure in terms of weak emergence [Chapter 3] and an account of natural kindhood, for which the molecular gene is a natural category [Chapter 4].

# A brief history of the gene

The last century has been rightly called "The Century of Gene" by Fox-Keller (2000).[20] After the disruptive discovery and spread of the theory of evolution, the 20th century started with the aim of solving the puzzle of the stability of traits and their transmission. In 1900, three journal articles were published by d. Vires, Correns, v. Tschermak exposed the laws of genetics, further developing and re-interpreting Mendel's studies on hybridisation (El-Hani 2015). It was clear that something was transmitted discretely from parents to offspring in a way that respected precise probability distributions. What was unclear was the nature of the entity under discussion. There were many candidates at the time, such as gemmules proposed by Darwin, d. Vires' pangenes or Weismann's determinants, possible entities characterised in different ways that could have been

---

[20] A version of this section is part of the article "The gene as a natural kind", forthcoming the Edited Springer Volume *Life and Mind* and from the short article "Genes: from instruments of thought to the base of life" (Bellazzi 2022c).

what is transmitted respecting the laws of genetics (Fox-Keller 2000, Beurton 2010, Falk 2010).[21] In response to this conceptual unclarity, Johannesen decided to introduce the new term "gene" to disentangle the discussion and to facilitate scientific practice. In 1909, he presented the gene as what explains the transmission of traits from parents to offspring and should have been "free from any hypothesis", with no theoretical pre-assumptions. It was considered only a "convenient notational concept" (Falk 2010, 321). First introduced as an instrumental theoretical entity, one whose observation wasn't needed, but whose postulation could improve the explanations and predictions of a theory, the gene changed our way of doing biology.

Nevertheless, not all scientists were aligned with an instrumentalist view of genetics. With its development, genes started to have properties in their own right, even if they were detectable only indirectly. The tensions between concrete experimental demands and the specification of the phenomenon under consideration led towards a more specific material identification of the gene (Fox Keller 2000; El-Hani 2007; Falk 2010). In particular, Muller, a student of Morgan, supported a realist hypothesis of genes as material units, possibly chemical molecules, in disagreement with most scientists and his master (El-Hani 2007). This realist understanding supported the research for a material molecular basis of the gene. In the early 1940s, the one-gene-one-enzyme hypothesis was rising, supported by the discovery in 1944 of DNA as the substance of heredity by Avery and his team (Beurton 2010). Finally, in 1953 Watson and Crick built a model for the structure of the DNA molecule and its replication, thanks to x-ray diffraction images of DNA taken by Gosling, Franklin and Wilkins (Clark, Pazdernik 2012). The molecular basis of the gene was found. From this moment, the history of the gene should be considered two-fold. On the one hand, the gene remained the so-called Mendelian gene, the unit of trait transmission and the object of classical genetics; on the other hand, the gene became identifiable with a specific chemical molecule (Kitcher 1984). The relations between the gene as the unit of transmission and the molecular gene are complex, and I will not enter into the details as this would go beyond the scope of this chapter. Suffice it to say that Mendelian genes are not easily reducible to the

---

[21] Gemmules were thought to be shed by various organs and then transmitted via reproduction; pangens were thought to be particles bigger than chemical molecules but smaller than cells that could multiply and grow; determinants were entities that present in every cell, but only one of them - the relevant one - was active (Bellazzi 2022c).

molecular genes, and they can be considered as something distinct with different functions and properties (Kitcher 1984; Okasha 2019). Here, I will focus only on the molecular gene, the object of molecular genetics (Waters 2007).

In the 1960s, it seemed clear that genes were nothing more than segments of DNA located on a chromosome that give rise to a particular amino acids sequence. The search for a material basis for the gene led scientists to postulate a correspondence between genes, segments of DNA and amino acids. This correspondence was formulated as the Crick-sequence hypothesis: each codon, a sequence of three bases, specifies only one amino acid, and a gene is a sequence of codons that specify a polypeptide. The "material molecular gene" was born. Soon, a molecular understanding of the genes was accompanied by the idea that they were just open reading frameworks: DNA frameworks open to be read (ORFs). This facilitated research at the time, as the gene was identified as a well-defined and structured stretch of DNA with clear borders and a singular function. The success of molecular genetics was read in terms of eliminativist reductionism. In 1969, Schaffner proposed to apply the models for epistemic reduction used in physical sciences to the case of the gene. If an identity relation between genes and DNA molecules had been found, then an eliminativist reduction would have been accomplished. The conclusions of this reductionist interpretation were more than epistemic, as they aimed at eliminating the gene, and there was empirical and experimental support. A gene was claimed to be identical to a precise and defined stretch of DNA, gene-1=DNA-1, as a simple 1:1 co-linear eliminativist reductive relation. "In the light of the Watson-Crick model, Benzer considered the possibility of translating his biological genetics into chemical terms" (Schaffner 1969, 339), and Schaffner regarded the entire reduction of molecular biology to chemistry and physics as something not only possible but very close in time.

However, things turned out to be more complicated than Schaffner and many biologists thought in the 1970s. The production of new technologies to sequence genomes and the advances in molecular biology of the last forty years have disclosed peculiar genetic phenomena (Fox-Keller 2000; Hall 2001; El-Hani 2007; Griffiths, Stotz 2013; Meyer et al. 2013). The sequence of entire genomes and the study of eukaryotic genomes revealed that the Crick-sequence hypothesis was simplistic, and further developments in genetics have made it impossible for genes to be a merely contiguous DNA segment co-linear with the product derived (Fogle 2010; Perini 2011). These results compromised the material identity of the gene as a discrete stretch of DNA and showed the inefficiency of its identification in mere material terms (Falk 2010). They also clarified how the gene operates thanks to a series of complex mechanisms and processes surrounding it. Particularly relevant to the beginning of the new phase of molecular biology is 2001 and the publication of the first Human genome draft sequence, as a result of the Human Genome Project, and the ENCODE Project.[22] This can be considered a "threshold year", and from this point onwards, genetics has entered the "postgenomic era" (as in Griffiths, Stotz 2013). In 2007, El-Hani spoke of the crisis of the gene that finds itself between "the cross and the sword" because of the identification of a series of complex phenomena, such as split genes, alternative splicing, overlapping and nested genes. This will be the starting point for this part of the thesis. Specifically, I will consider two questions. First, I will explore whether the gene instances actually exist, given the complexities of the behaviours of genetic phenomena [Chapter 3]. Second, I will consider whether the category molecular gene can be deemed a natural kinds [Chapter 4].

---

[22] The Human Genome can be considered the largest and most expensive biological collaboration in history, with the purpose and aim to sequence the full Human Genome and was developed between 1990 and 2003. The ENCODE Project, Encyclopedia of DNA elements, followed with three projects from 2003 to 2020.

# Chapter 3

# The emergence of the molecular gene

## 3.1 Introduction

Do genes really exist? With the discoveries that followed the genome project, this question has proved difficult to answer due to the complexity of genetic phenomena.[1] The various processes involving nucleic acids have shown the limits of characterising genes only in terms of a determinate physical stretch of DNA and a specific molecular product (also Hall 2001; Fogle 2010). This has resulted in a sceptical or deflationary approach to genes, which can be considered mere useful tools for genomic analysis. Nevertheless, it is hard to accept such a view, as genes still retain a central role in many sub-disciplines of contemporary biology. In this chapter, I will argue in favour of a positive answer: genes exist, but they emerge during the transcription processes. The result elaborates on the postgenomic awareness that genes have a proper identity only within the broader context of general molecular interactions and cellular processes (Hall 2001; Burian 2004; El-Hani 2007; Scherrer, Jost 2007; Fogle 2010; Griffiths, Stotz 2013). In order to understand what it takes for something to be a gene and then assess its existence, we need to stop focusing on its molecular components and we should consider the interactions that concern gene expression (El-Hani 2007; Fogle 2010; Griffiths, Stotz 2013). The gene has its proper home in the cell and

---

[1] A version of this chapter is published in Bellazzi F., 2022a: "The emergence of the postgenomic gene". *European Journal for Philosophy of Science* 12, 17.

cannot be understood without it. This will be the starting point for the present analysis.

My argument in favour of the existence of genes as emergent entities will proceed in two steps. First, I will defend and clarify a definition of postgenomic genes presented in the literature by Gerstein and colleagues (2007) and Griffiths and Stotz (2013). According to this proposal, genes are not mere linear sequences of DNA but complex entities that depend on a context of inter-, intra-, and extra-cellular factors. Then, I will argue that genes so-characterised are weakly emergent from the DNA basis insofar as they display explanatory novelty and robustness. This account of emergence will be ontological and motivated on moderately naturalistic grounds.[2] In detail, this research has different implications. First, it applies the metaphysics of science and the concept of emergence to a complicated issue in the philosophy of biology and might help disentangle some of the tensions within the gene debate. This approach has been producing fruitful results in relation to other sciences, and it can be the same for the philosophy of biology. Furthermore, biologists who have been experiencing the problems of the gene concept might benefit from this account. It will provide them with a novel framework that allows for both gene identity and contextual dependence, and this can provide a better understanding and clarity of the phenomena under discussion. Specifically, considering the context in which genes exist can be relevant when theorising on genetic phenomena. Also, it can provide theoretical grounds to accept the existence of genes within the manifested complexity to scientists with a realist tendency towards their discipline. Second, this case study can be of interest to all the philosophers working in the field of emergence in the sciences. I apply here a novel account of weak emergence mostly proposed in the philosophy of physics, as in Franklin and Knox (2018), to cases in biology. This view is a significant proposal of weak emergence because it is sensitive to scientific practice, it is easily applied to concrete examples from science, and it is compatible with non-eliminativist ontological reductionism. This approach will allow us to retain the gene as materially constituted by nucleic acids while maintaining a type difference between the gene and its underlying basis.

---

[2] Reference for moderate naturalism is Tahko, Morganti (2017), as summarised in Chapter 1, for which ontological considerations should be done based on a dialogue between metaphysics and the scientific postulation of entities.

The structure of the chapter is the following. In section §2, I will present the gene in the context of the postgenomic era. In the 1960s, the gene was defined as a precise and ordered sequence of nucleotides that encodes the primary structure of a polypeptide or a functional RNA molecule. Nevertheless, discoveries of the last four decades made the identification of a 1:1 correspondence between genes and stretches of DNA impossible. The postgenomic gene is highly dependent on the molecular and cellular context. The complexities involved in the current status of the postgenomic genes can be taken into account because, I will argue, genes weakly emerge from the underlying nucleic acid basis. In section §3, I will introduce the concept of ontological weak emergence relevant to the case in question, presented by Franklin and Knox (2018). According to this account, a phenomenon is claimed to be weakly emergent when it has properties that are (i) explanatory novel, as they improve explanations; (ii) robust relative to some lower-level perturbations, here interpreted in relation to multiple realisability and multiple constitution.[3] This type of emergence is ontological and incompatible with eliminativism but nevertheless compatible with ontological reductionism: the token of the emergent phenomenon will be identical with some lower-level phenomena at the time considered. This is coherent with the strict relation between DNA and genes acknowledged by scientific practice. Furthermore, the focus on robustness, a well-acknowledged feature of genetic phenomena, makes it particularly suitable for its application in the life sciences (see Eronen 2015; Boone 2018).

In section §4, I will show that genes are weakly emergent in the considered sense. Even if there is substantial conceptual unclarity about the definition of genes, two crucial features are identifiable. Broadly, a gene is a part of DNA that has the property of being transcribed into specific molecular products (Griffiths, Stotz 2013). This entails that there are two crucial properties of the gene, one material and one functional. First, it has to be composed of nucleic acids (DNA or RNA)[4], second it has to be actively transcribed. This allows us to identify a

---

[3] Further analysis will be provided in §3, §4

[4] I thank an anonymous reviewer for the EJPS version of this chapter and Dr. Margarida Hermida for pointing out that genes can be composed of both DNA and RNA accordingly to the phenomenon or the account under consideration (as in Scherrer, Jost 2007). Specifically, Scherrer, Jost 2007 suggest that genes as images of the final molecule emerge in eukaryotes in the final mRNA. Or, if we accept that also viruses have genes, they might be composed of RNA. For simplicity's sake, throughout the chapter, I will consider genes as composed of DNA, being aware that they could also be composed of other nucleic acids such as RNAs.

material component of the gene: the sequence strictly involved in transcription, and according to most models, this also includes the promoter region (or TATA box)[5] (Fogle 2010, 6; Griffiths, Stotz 2013, 71). And it permits the identification of a functional component: the gene has the function of being transcribable. This is what discriminates genes from other regions of DNA. Furthermore, this second property is not an intrinsic property of a contiguous linear sequence of DNA, but it depends on a full set of molecular and cellular interactions. Here, I will propose how the gene can retain its identity despite this context dependence. First, I will address some of the complicated phenomena that made the identification of genes with linear stretches of DNA molecules more complicated. Then, I will argue why genes satisfy novelty and robustness. More clearly, a gene emerges from a precise stretch of DNA when it is expressed, thanks to the interactions that make it novel and robust. I will conclude in section §5 that the emergence of the postgenomic gene allows to retain the gene as an existent phenomenon, embracing its flexible and context-dependent identity.

For the purposes of the thesis, this conclusion will provide two main results. The first one is that the gene is weakly emergent from the DNA, and this can be seen as a first support in favour of the *sui generis* nature of biochemical kinds. The second is that it shows which properties are relevant for its characterisation: chemical structure and biochemical function. This latter result will be used both in Chapter 4 and in Part III to explore further aspects of the metaphysics of biochemical kinds.

## 3.2 The postgenomic gene

The 1960/70s were the golden years of molecular genetics, and it seemed clear that genes were nothing more than segments of DNA located on a chromosome that give rise to a particular amino acid sequence. This correspondence was formulated as the Crick-sequence hypothesis: each codon, a sequence of three bases, specifies only one amino acid, and a gene is a sequence of

---

[5] See Fogle 1990 for the different models of what constitutes the material component of the genes.

codons that specify for a polypeptide (Griffiths and Stotz 2013). With subsequent discoveries of the 1980s, genes were understood as "open reading frameworks" (ORFs): DNA frameworks open to be read (Gerstein et al. 2007). This facilitated research at the time, as the gene was identified as a well-defined and structured stretch of DNA, with clear borders and a singular function. However, things turned out to be more complicated. The production of new technologies to sequence genomes and the advances in molecular biology of the last forty years have disclosed peculiar genetic phenomena (Fox-Keller 2000; Hall 2001; El-Hani 2007; Griffiths and Stotz 2013; Meyer et al. 2013). Specifically, the sequence of entire genomes and the study of eukaryotic ones revealed that the Crick-sequence hypothesis was simplistic and further developments in genetics have made it impossible for genes to be a merely contiguous DNA segment co-linear with the product derived (Fogle 2010; Perini 2011). These developments compromised the material identity of the gene as a discrete stretch of DNA and showed the inefficiency of its identification in mere material terms (Falk 2010). With the new century, the gene is now a concept in tension. In 2007, El-Hani spoke of the crisis of the gene that finds itself between "the cross and the sword" because of a series of complex phenomena, such as split genes, alternative splicing, overlapping and nested genes (Griffiths and Stotz 2013).

The phenomena discovered in the postgenomic context illustrate how the gene is embedded in a series of complex and different interactions that surround it. And despite the temptation to consider the gene a dead road, there is a consensus on the need of relocating the gene within its cellular and organismal context (as Beurton 2010 argues). Given the number of publications, experiments and biological scientific practice that even today involves gene-talk, it seems mandatory to understand the gene keeping together its molecular aspect and its context dependency (as Burian 2004 argues).[6] It seems true that "the gene is not dead, but alive and well, even though orphaned, homeless, and seeking a haven from which to steer a course *to its natural home, the cell* as a fundamental morphogenetic unit" (Hall 2001, 225-228; emph. added). As a reaction, many have been re-thinking and re-defining the gene concept in a variety of ways,

---

[6] Context does not mean different disciplines or different scientific practices, but different environmental contexts such as the nucleus, the cell and the overall organism with its environment. As in Griffiths and Stotz (2013), the regulatory molecules and mechanisms are intra-cellular, inter-cellular and extra-cellular environmental systems.

trying to accommodate theoretical and practical requirements.[7]

Generalising, we can indicate two main ways to re-think the gene concept. The first can be referred to as "the nominal gene" (Burian 2004; Griffiths and Stotz 2006, 2013). This is a "relatively conservative conception of gene" as stretch of DNA with precise nucleotides sequences that encode a specific product (Griffiths and Stotz 2013, 66). As suggested by its nominal component, such a view of genes is highly operational and has a deflationary connotation. Related to this, we find the so-called "consensus gene" as any "general pattern of biochemical architecture and process" that shares the features of exemplary gene cases according to scientific practice and empirical evidence (Fogle 2010). The second is a more realist understanding of the "postgenomic genes" as "images of the target produced molecules" (Griffiths and Stotz 2013, 75). Such genes can only be fully understood in their environment and in their context of action (Griffiths and Stotz 2006; Falk 2010). Here, I will start from this second approach to the postgenomic genes. Such a conception allows us to retain a correspondence between genes and products, but it does not imply the fixity of identifying the gene with a given DNA sequence.

### 3.2.1 What is a postgenomic gene?

Let us now look more closely at postgenomic genes. In contemporary biological textbooks, genes are broadly defined by their function and by their composition. For instance, "The gene is a unit of information and corresponds to a discrete segment of DNA that encodes the amino acid sequence of a polypeptide" (Fletcher and Hickey 2012).[8]. However, such a definition provides just a starting point to understand the gene in detail, and it is close to the mentioned nominal gene concept. Let us now look at a more advanced definition proposed in the context of the ENCODE analysis: "A gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products" (Gerstein et al. 2007). Something is defined as *gene* if it

---

[7] It is common to advocate for a pluralist view of genes for which there are a variety of valid gene concepts in various disciplines, as sustained by Hall (2001), or Fogle (2010) and as reported by Griffiths and Stotz (2006, 2013). For instance, we can distinguish "instrumental genes" as "factors in a model of the transmission of a heritable phenotype" from the "nominal" and "postgenomic genes" (as in Griffiths and Stotz 2006).

[8] Another source of reference for the discussion on the definition of genes in textbooks is Gericke et al. (2014), as suggested by an anonymous reviewer for the EJPS version of this chapter.

is composed of a genomic sequence and it is transcribed into a coherent set of transcripts. Nevertheless, the gene is not a linear sequence but a union of different genomic sequences, and transcription cannot be understood without the complex molecular and cellular context that allows for it. Trying to account for such a complexity, Griffiths and Stotz (2013) elaborate on the proposal of Gerstein et al. (2007) and define postgenomic genes as "images of the target molecule" that can be produced only in a wider system of interactions and environment. This makes the genes not simple contiguous stretches of DNA but rather sets of DNA sequences with a specific function that are not necessarily contiguous; as far as their definition is concerned, "function and structure are inseparable" (Fogle 2010, 5). The structural or material component of the gene normally includes not only the finally transcribed sequence, but often also the promoter (or TATA box) is accepted as a consensus feature of a gene (Fogle 2010, 6). In addition to it, there might be other regions involved that are essential for the activation of the gene and the regulation of its transcription (Griffiths and Stotz 2006; 2013).

Here, two questions might arise. The first is which parts of the DNA should be considered genes. I claim that such an answer can only come from empirical scientific practice. The empirical evidence concerning which parts of the DNA regions are counted as genes is constantly changing and improving, and the answer to this question has to be provided by science.[9] For instance, it is actual scientific practice that can tell when some regulatory parts of the DNA can be included in the different genes. This might also be dependent on individual cases. The second question asks what it takes for something to be considered a gene at all and whether such entities exist. This is what I will consider here.

For the purposes of the chapter, we are interested in identifying a definition that points toward what makes a phenomenon different from something else. Combining the definitions above, we notice that the gene has a two-fold one. The gene has a structural or material component as a region of DNA. Conversely, not every region of DNA is a gene. Only those

---

[9] Many have tried to account for the precise material basis of genes. Fogle (1990) proposed four different models to accommodate the concept of gene: Model A includes the transcribed region of the DNA and all neighbouring sequences which influence the process; Model B considers only the transcribed region; Model C includes only the set of exons derived from a pre-mRNA; Model D is limited to the coding exons of a primary transcript, excluding non-coding leader and trailer sequences. Here, I leave open the question of the exact basis from which the genes emerge.

regions actively involved in transcription as "images" of the target molecules are genes. And a gene is what has the property of being transcribed into a RNA, which can then encode the primary structure of a polypeptide or can play a regulatory, structural or catalytic function. Thus, there are two crucial properties of the gene: i) it has to be composed of DNA; ii) it has to be transcribed or strictly involved in transcription (Fogle 2010; Griffiths, Stotz 2013). Transcription, however, is not a self-subsistent phenomenon, but it is rather a reactive one: it happens only within the right circumstances and thanks to a set of interactions that operate at different levels (Griffiths, Stotz 2006; 2013). In functional terms, a gene is a part of DNA that has the function of being transcribed or being involved in transcription (according to the accounts considered). This implies that the gene is a context-dependent phenomenon: "a function is always a role in something and a contribution to something" (Germain et al. 2014), and it acquires its proper identity only when considered within the cellular context in which it exists. In the case of genes, the transcriptional aspect makes them an "image" of the molecular product. Its involvement in transcription makes some parts of the DNA a gene. Further support for the relevance of the functional component comes from the fact that genes are normally divided into two functional subtypes. The first are genes that encode the RNA for the amino acid sequence of a polypeptide; the second are genes that encode an RNA that regulates various cellular processes or plays other functional roles (as in Fletcher and Hickey 2012; Meyer et al. 2013). Accepting that genes are functionally defined does not imply that they are *exclusively functional*. But here, I claim that a minimum requirement for something to be a gene is to have a functional property and *some* underlying material DNA basis.[10] The importance of the functional aspect brings considerations about the context and conditions in which such a function is realised. Without molecular interactions, DNA is an inert molecule that cannot be transcribed. The gene can only be properly understood by taking into account the system that surrounds it. This implies that the core property of genes is relational: it exists only when specific interactions between the DNA and the context happen.

Such complexity and context dependency might lead towards a more deflationary view of

---

[10] A gene is different from other merely functional kinds, such as chairs or bicycles, which are what they are *only* because they perform a certain function.

the genes. However, the present chapter wants to argue in the opposite direction. It aims at providing a novel framework that allows us to accept the existence of genes as entities with a material component (for instance, DNA) and a context-dependent functional one. I will do so by arguing that genes are more complex than their DNA bases and, in having a functional component, they are weakly emergent.[11] This approach will interest scientists with a more realist attitude towards their discipline as it will provide them with an account in favour of genes' existence. Moreover, such a view can provide more clarity when theorising genetic phenomena: genes exist as emergent entities in specific contexts.

## 3.3 Ontological emergence as novelty and robustness

In which sense, then are the genes emergent? The debate on emergence is wide in approaches and topics. For the purposes of this chapter, I will focus only on an account recently proposed by Franklin and Knox (2018).[12] This has been advanced as a form of ontological weak emergence, and it elaborates on previous works of Knox (2016) and Butterfield (2011). I find it among the best recent formulations of weak emergence, as it is sensitive to scientific practice and compatible with non-eliminativist ontological reductionism.[13]

Emergence can be generally defined as a combination of autonomy and dependency: the emergent phenomenon has to be autonomous in some sense but nevertheless dependent on its basis. The formulation of Franklin and Knox (2018) captures well these two aspects. The autonomous component of emergence is given by the features of novelty and robustness. Nevertheless, such an account is compatible with forms of non-eliminativist reductionism, which allows for the dependency component.[14] This is because higher-level entities exist (so they are

---

[11] I consider here instances of the general kind "molecular gene", even if, in nature, there are specified genes with specific functional properties.

[12] I am aware of the difficulties of defining emergence unambiguously, as reported in Wilson (2015, 3). This section does not aim to elucidate the general discussion on emergence and its metaphysical significance, but to present the account relevant to our discussed topic.

[13] In Chapter 6, I will discuss emergence more in detail, and I will combine the presented view with the proper-subset view of emergence, as presented by Wilson (2021).

[14] This account is also compatible with forms of epistemological or methodological reductionism, but this will not be considered here.

not *eliminable*), but they are reducible to the lower-level entities that realise them: the emergent entities will be identical at the time considered with some of the lower-level features.[15] Moreover, the combination of epistemic and ontological criteria, novelty and robustness, makes it coherent and sensitive to scientific practice and scientific discoveries. This is because it is easily applied to concrete case studies. Specifically, it can be applied to instances in which scientific practice takes the lower-level to be token identical with the emergent entities, even if there is a difference in type, as it can happen in genetics. Furthermore, it acknowledges the relevance of robustness, a feature of biological systems that plays a crucial role in the discussion in the biological sciences (see Eronen 2015; Boone 2018). This contemporary proposal of weak emergence has been presented as the only one able to capture a complicated phenomenon, the emergence of phonons, and the authors claim that this account may be extended and applied to "many other instances of emergence across science" (Franklin, Knox 2018, 68). Even if phonons and genes are different in many aspects, they both share a complexity that calls for further analysis, and I will show that this account applies to the case of the gene, supporting its emergence.

Some clarifications are needed before presenting the details. Here, I will assume a simple two category ontology of properties and entities that bear the properties, and claim that something is emergent when its definitional properties are emergent (as in Wilson 2015, 4).[16] So, taken a given phenomenon (in this case, the gene), this phenomenon will be considered emergent when its definitional properties satisfy the given criteria for emergence. Moreover, when considering the relations between levels, I will assume that higher-level properties are *realised* by lower-level features, while higher-level entities are *constituted* by lower-level entities (see Gillett 2013; Kistler 2018).[17] Let us now move to emergence. Weak emergence refers to the existence of higher-level features that are "realised by the lower-level ones" in a genuine way, even if every token of the property of the emergent feature is identical with *some* lower-level feature at the time considered

---

[15] Provided that the lower-level features are physical, this position can also be referred to as non-reductive physicalism. Moreover, the novelty of a phenomenon and its emergence does not lead to "any kind of failure of theoretical reduction" (Franklin, Knox 2018, 74). Ontological reductionism is considered compatible with realism about a phenomenon, as in Brock, Mares (2014).

[16] The term entity is left willingly unspecified, as it can be whatever bears a property according to the ontology considered. An entity can be a concrete particular object, an abstract one, a process or an event.

[17] This precision of terminology is justified by its coherence to scientific practice, as in Gillett (2013).

(Wilson 2015). Furthermore, Franklin and Knox (2018) specify that a weakly emergent phenomenon also has to display properties whose higher-level postulation improves our explanatory power, and these properties are robust. In detail, the defining properties of the phenomenon under consideration are considered emergent when they are characterised by two features:[18]

- Novelty. This feature implies that it is possible to identify the emergent property in a distinctive way from the properties held by the lower-level entities, and the consideration of such a property improves explanations, leading to new ones. Accordingly, a phenomenon is emergent when it has a property whose postulation and usage in scientific theories leads to novel explanations (see also Knox 2016).[19]
- Robustness. This feature implies that the emergent property displays stability within a certain range of perturbations and relatively to some lower-level properties (Butterfield 2011; Franklin and Knox 2018). Even if the system undergoes modifications, the property nevertheless realises. Accordingly, a higher-level phenomenon is said to be robust when it displays a property that is stable within a certain range of perturbations, here interpreted in terms of multiple realisation and multiple constitution.

Let me consider the two relevant features in more detail. Novelty has a crucial epistemic

---

[18] It is legitimate to ask about the relation between these features and the phenomenon's emergence. Here the two features, novelty and robustness, should be considered the *ratio cognoscendi*, i.e. how to know that something is emergent. But from the ontological point of view, it is the fact that the phenomenon is emergent that represents the *ratio essendi* of the features, novelty and robustness. We can know that something is emergent when it displays novelty and robustness. But something displays novelty and robustness because it is emergent (see also Eronen 2015, 3966). To summarise, emergence represents the feature thanks to which we can both postulate the existence of something and its knowledgeability. So, when studying an emergent phenomenon, we can know that it is such because it displays novelty and robustness: we have epistemic access to its emergence thanks to its being novel and robust. But, from an ontological point of view, the phenomenon can display novelty and robustness because it is emergent and so existent, in the first place. I thank an anonymous reviewer for the EJPS version of this chapter for pushing me on this point.

[19] I thank an anonymous reviewer for the EJPS version of this chapter for pointing out that improving explanations and providing new ones are two different epistemic moves. However, if we are considering explanations about the *same* phenomenon, then the difference between the two might be on a continuum. Here novelty is interpreted as what leads towards novel explanations and these can be seen as an improvement or "empowerment" of the explanations already present (as in Franklin, Knox 2018). These improvements might also lead to new explanations about the same phenomenon. Although this raises some interesting issues, this topic would require a separate paper, and there is no space to develop it in more detail.

component, as it is what permits *novel explanations* (Knox 2016): the postulation of the property allows for new and better explanations compared to the postulation of only the basis that realises the property. The same can be said for the emergent phenomenon: it is novel when its postulation improves explanations, compared to postulating only the components. Even if novelty is epistemic, assuming a form of realism in science, it has further value as the explanatory novelty of properties is a good hint about the real existence of such a property in the world. This is in line with a form of moderate naturalism, for which ontological considerations should be made in dialogue with the scientific postulation of entities. The best way to account for the explanatory power of a phenomenon in a scientific theory is to postulate its ontological existence. Nevertheless, novelty is not enough on its own and it needs robustness.[20]

Robustness is then the ontological feature of this account, in the sense that it concerns the existence of the property rather than its role in our scientific theories. And robustness implies that the property is present even under perturbations and changes in the environment. It is legitimate to ask which kind of perturbation is relevant for robustness. Differently from the original proposal[21], I will refer with "stability under perturbations" to cases in which the emergent features present *multiple realisability* and *multiple constitution*, as this reflects the meaning of robustness across the biological sciences (as in Boone 2018).[22] This is a relevant aspect of the proposal, as the robustness of the phenomenon can be read as a *genuine discriminator* between what is merely postulated by a theory and what the theory effectively captures of the world (see Weisberg 2013, 156-170; Eronen 2015). In the debate of emergentism, robustness might distinguish mere epistemic emergence from ontological emergence, that is the difference between the irreducibility of a theoretical phenomenon in a theory from its real existence in the world.

---

[20] As pointed out by an anonymous reviewer for the EJPS version of this chapter, novelty can also be provided by idealisations that play an important explanatory role in science. However, idealised phenomena are not what usually counts as possibly emergent. This is one of the reasons why the feature of novelty is not sufficient on its own to provide emergence, and it has to be combined with robustness. Much more could be said about the role of idealisations in science and their ontological impact, but this topic cannot be explored here due to the focus of this chapter.

[21] In Franklin, Knox: "To show that a phenomenon is robust is to show that its description and dynamics are stable with respect to perturbations in the underlying physics; in order to be emergent, a phenomenon must not be too fragile or too fleeting." (2018, 73).

[22] For "multiple constitution" see Gillett 2013; Kistler 2018. In Kistler, a phenomenon is multi-constituted when there are two or more microscopic structures that obtain the phenomenon (2018, 18).

Moreover, it acknowledges an important feature of genetic phenomena that have always been characterised as robust in the debate on the philosophy of genetics (Falk 2010). I will come back to elucidate the specific features of robustness in §4.1.

## 3.4 From complexity to emergence

Let us now consider whether genes are weakly emergent from their DNA basis according to the presented account. I will first present some of the phenomena that make genes more complex entities than what was thought in the 1970s. Then, I will present why genes satisfy the two requirements for weak emergence, novelty and robustness.

As mentioned, the main properties of a gene are i) the property of being constituted of some nucleic acids and ii) the property of being transcribable or actively involved in transcription (from now on **Ft**). Specifically, **Ft** has a special role as it allows for the gene's being an "image" of the target molecule (as Griffiths, Stotz 2013 points out) and is the defining property of the gene, relevant in identification and explanation in scientific practice.[23] At first, **Ft** property had been ascribed to well-individuated stretches of DNA, with barriers and, consequently, the gene had a precise structural and material identity. However, a series of complex phenomena challenged the identification of the gene with a precise continuous stretch on the DNA. In particular, there are three broad classes of phenomena that compromised this characterisation (El-Hani 2007; Meyer et al. 2013):

- There are **one-to-many** correspondences between DNA segments and RNAs/ polypeptides. This means that the same stretch of DNA can give rise to different molecular products via complex mechanisms. Given the functional definition of the gene, this can be interpreted as the possibility of the same stretch of DNA constituting different genes

---

[23] This property is identified *ahistorically*, that is genes are (mostly) individuated by what they do now, rather than their evolutionary history (the distinction from ahistorical and historical individuations is from Gillett 2021). As indicated by an anonymous reviewer for the EJPS version of this chapter, the material identity of the gene is also ontologically relevant and represents a condition of being. However, it is not sufficient on its own to identify the genes, given that genes are not only stretches of nucleic acids, but specific ones that are "images" of a given molecule. Both aspects of the gene definition (the material and the functional one) are ontological. However, given that the functional one allows us to distinguish what is a gene from what is not, I have focused the discussion on emergence on the functional component.

with different functions.[24] For instance, the discovered discontinuous structure of genes can allow one gene to be contained inside another one's intron (Gerstein et al. 2007). DNA seems to have *multiple determinations*, where multiple determinability is the possibility of one microstructure, one stretch of DNA, to realise multiple biological properties and, in this case, to compose multiple genes (Tahko 2020).

- There are **many-to-one** correspondences between DNA segments and RNAs/ polypeptides. This means that several different DNA segments or sequences can realise the same functional product. There are different stretches of DNA that can be "images" of the same transcribed product. Furthermore, there are cases in which small modifications in the underlying gene's sequence do not change the transcribed product and the realisation of **Ft**. An instance of this phenomenon are synonymous mutations, changes in the DNA sequence that codes for a specific amino acid without affecting the final product, the encoded amino acid. **Ft** results then *multiply realisable*, where multiple realisability is the capacity of one property to be realised by a variety of microstructures or mechanisms. In parallel, it can be said that the gene is *multiply constituted*, as the same gene can be constituted by different stretches of DNA realising the appropriate **Ft**.

- And, at last, there can be a **lack of straight** correspondence between DNA segments and RNAs/ polypeptides. This means that there are functional products that do not arise from any specific and straightforward DNA sequence. An instance of the phenomenon is mRNA editing. In mRNA editing, the messenger RNA molecules are modified by enzymes after their synthesis on specific nucleosides. In this case, **Ft** is realised by whatever part of the DNA encodes the then modified mRNA. This might make the "image" within the postgenomic gene more distant from the final molecule, but it is nevertheless present. An instance of modifications of the mRNA is *trans-splicing*, in which the final mRNA is obtained from "two or more independently transcribed pre-mRNAs" (Griffiths, Stotz 2013). In some of these cases, we can even notice a **many-to-many** relation in which the same sequences of DNA can then be modified and transcribed in different ways, fusing or

---

[24] For the importance of functional similarity of products in genes classification, see Fogle 2010.

"scrambling" the exons (*ibid*). These phenomena happen thanks to the complex molecular interactions that involve the gene and make it context-dependent.

A reaction to these phenomena can be a deflationary or merely nominal account of the genes, as they do not anymore satisfy the 1:1 relation that is supposed by classic molecular genetics, and their clear identification is difficult.[25] However, an alternative answer can be provided if we embrace the context dependency of the genes and consider its definitional property **Ft** as relational. Genes are genes because of a system of molecular patterns and relations around them. Such context dependency should not be discouraging, as the dynamical aspect of biochemical phenomena should allow for no strict requirements on precise material barriers. If the material identification of the gene is problematic, we can still retain that transcription happens in particular contexts and not all parts of the DNA are transcribed. These complexities are the starting point from which I will argue that genes are novel and robust, and so emergent.

### 3.4.1 The emergence of the gene

Following the definition of the postgenomic gene, a full understanding of it should consider its material composition in DNA sequences together with its functional significance. The utility and relevance of this definition, together with the three macro-patterns identified before, will be used to argue that **Ft** is novel and robust. As a result, I will conclude that the gene is emergent.

Here, I will first briefly present the precise context I will use to support the thesis. Then, I will consider why the functional property defining the genes can be considered novel and robust. A consequence will be that genes emerge during gene expression, that is in the moment in which transcription is happening, and they are constituted by *some* nucleic acids. However, this proposal remains weak and coherent with non-eliminativist ontological reductionism. When I write of genes as emergent entities, I do not mean they are concrete separated individual objects, but simply that they exist as something qualitatively distinct from the mere DNA bases. The

---

[25] On the lack of straight linear correspondence between bio-molecular entities and functions see also Kistler: "There is a second and complementary reason for which there is no 1:1 relation between biomolecules and functions. The first is that a given type of molecule can have several functions. The second reason is that several molecules can share a function." (2018, 16).

components of the gene are identified with specific tokens of chemical molecules involved at the time considered, but they are not any molecules but the ones with the property **Ft**.[26] This view implies a temporal and contextual connotation of the gene, as its weak emergence is supported by the one of **Ft** realised specifically during transcription. Moreover, this is coherent with genetic practice as often the DNA is manipulated to manipulate the genes, assuming a token identity between the lower-level and the emergent one, despite a difference in type.

### Genes and protein synthesis

Even if genes can also encode RNA molecules with regulatory and functional roles within the cell, their scientific importance is particularly significant in consideration of protein synthesis. Thus, I exemplify my argument by focusing specifically on genes that encode the primary structure of a polypeptide.

Protein synthesis is the process that produces proteins inside the cell. It can be divided broadly into two main phases: transcription and translation. Transcription is the first phase in which a section of DNA "becomes" the gene: a sort of template molecule, or "image", for the messenger RNA (mRNA). The transcription of a gene into a mRNA sequence (which can then be further elaborated into a mature mRNA) is carried out by RNA polymerases. The action of this enzyme is possible firstly thanks to the detection of the promoter region and the action of enhancer or silencer regions (Griffiths, Stotz 2013). Then, it also depends on molecular conditions of the relevant section of DNA, such as chromatin remodelling and the action of other proteins, and general cellular interactions. Translation is the second phase in which the mRNA is translated by the ribosomes that use the sequence to order the sequence of amino acids for a polypeptide chain.

For the gene's analysis, let us focus on transcription. As already underlined, the gene has a specific existence within the right cellular environment, and it is intertwined with its transcription. The complex context in which transcription happens makes the core property of the gene **Ft** relational, as it needs a set of interactions for it to be present. This is pivotal to understanding

---

[26] I consider the identification of the exact bases from which genes emerge an empirical enterprise that should be determined by scientific practice.

genes' emergence. The gene is existent thanks to the action of RNA polymerases, together with promoter regions and other regulatory regions of DNA, the right cellular environment and the unfolding of DNA. Let us now consider more in detail the reasons why the genes satisfy novelty and robustness and can be considered properly emergent during protein synthesis.[27]

### Novelty

As previously introduced, novelty is i) what makes it possible to identify the emergent property distinctively from the properties held by the lower-level entities; and ii) the consideration of such a property improves explanations, leading to new ones. The novelty of the definitional property of the gene **Ft** is found in both these aspects.

First, that **Ft** is not a property of precise stretches of DNA with definite barriers should be clear from the multiple realisability of **Ft** and the multiple constitution of the gene mentioned in §4.[28] Genes result in a discontinuous structure that can make one gene completely contained inside another one's intron, or one gene overlapping with another (Gerstein et al. 2007). Considering any specific transcribed product, this can be realised by different stretches of DNA and, conversely, different stretches of DNA can realise similar products. **Ft** is present when there are complex interactions between different parts of the genome, as gene regulatory networks, and depends on the action of RNA molecules and enzymes. This property is specifically of the gene and it is what distinguishes other regions or sequences of DNA from genes regions.

Second, the novelty of **Ft** is justified by the improvements the gene brings to our explanations in science. If the structural-material conception of genes, for which genes are defined only by their material DNA components, cannot account for the complexity of the transcription and

---

[27] As suggested by an anonymous reviewer for the EJPS version of this chapter, the process of transcription can be identified as the ground of becoming or *ratio fiendi* of the gene that emerges during this specific process and in a dynamic context.

[28] Claiming that **Ft** is not a property of a precise stretch of DNA *does not imply* that it can never be considered, for scientific practice or in simple cases such as in prokaryotes, a property of a precise stretch of DNA. According to weak emergence, the property and the emergent phenomenon need a synchronic basis, in this case the DNA molecule, from which they are realised and in simple cases the identification of the gene with a well-delimited stretch of DNA can work. Nevertheless, a strict identity relation rules out many interesting genetic phenomena and is incompatible with the assumed functional definition of the gene (see Fogle 2010; Griffiths. Stotz 2013).

the first stage of protein synthesis, the consideration of its functional property helps scientific practice. Interesting cases are represented by monogenetic diseases or alternative splicing. It is thanks to the conceptualisation of the gene as a functional unit that "mirrors" a molecular product (even if more or less), rather than a "mere" straightforward DNA stretch, that one can explain and discover these complex phenomena or the interactions among different components necessary to encode a polypeptide. Even more, it is the conceptualisation of the gene as a functional unit that allows us to see its importance at the cellular and organismal level.

As mentioned earlier, one of the main explanatory roles genes play concerns the one they have during protein synthesis. Genes are the basic building blocks to explain why some proteins are encoded rather than others and allow for the conceptualisation of the transcription of some parts of DNA rather than others (even if they are not the only main agents[29]). This role is important in different domains, from molecular biology to medicine. For instance, genes have an important explanatory role in the case of the monogenetic disease cystic fibrosis. In this disease, the mutation of a single gene *CFTR* on chromosome 7 in humans can cause a disruption in the production of the Cftr protein, which regulates the flow of salt and fluids in and out of the cells and regulates the levels of mucus within the body. Considering this single gene as a combination of the sequence and specific **Ft** allows for a better explanation. The mutation in the gene *CFTR* alters the result of the final molecular product, as it is a different "image", and explains the possible disease. In "normal" cases, on the contrary, it is the absence of such mutation in the gene *CFTR* that allows to explain the normal levels of Cftr proteins. Scientists and doctors speak of mutations within the gene (and not sequence-mutation) and of *mono-genetic disease* because it is the consideration of such a single gene (with the mutation) that explains the presence of the disease.

The novelty of the gene can also be noticed considering a lower-scale genetic phenomenon, such as alternative splicing. In this phenomenon, the patterns of introns and exons are rearranged so that the same DNA stretch can be transcribed into different mRNAs encoding different proteins. Nevertheless, this case is often considered one of a single gene if the proteins encoded are

---

[29] See Griffiths, Stotz (2013) for a more detailed analysis.

functionally similar enough, and so they can both count as "mirroring" the same gene (Gerstain et al. 2007; Fogle 2010; Griffiths, Stotz 2013). Again, the functional aspect plays an important role. If one considered only the material stretch of DNA, alternative splicing would be a bizarre phenomenon in which the gene is cut and modified to produce different products.[30] With a partially functional view of the gene, however, one can theoretically elaborate alternative splicing as a case in which the same gene can be transcribed in different RNAs, realising different (but similar) functional outcomes (see Fogle 2010).[31] Moreover, **Ft** explains why certain polypeptides are obtained from alternative splicing, instead of others and can even help predict these results.

Alternative splicing is a highly frequent phenomenon, and one elucidating example is given by the *PTC7* gene in the genome of *S. cerevisiae*.[32] The *PTC7* is a gene that, after the transcription in RNA, can be alternatively spliced to generate two different mRNAs that can be translated into distinct but functionally similar, proteins. In particular, the relevant gene is transcribed into two mRNAs, codifying different proteins, Ptc7s and Ptc7u [Figure 4]. In the case of *PTC7*, "one isoform of *PTC7* is created by the removal of the intron, while the other isoform results from intron retention" (Jeneau et al. 2009, 186). This case of alternative splicing supports the novelty of **Ft**. It is thanks to the functional properties of making two different, but similar proteins (**FPtc7u**, **FPtc7s**) that it is possible to conceptualise alternative splicing as the presence of two different products of the same gene. Moreover, alternative splicing illustrates the multiple determinability of the DNA that makes different proteins. On the grounds of these phenomena, I can conclude that the gene with its property **Ft** is novel as it displays qualitative

---

[30] El-Hani writes "alternative RNA splicing requires that the conceptualisation of a gene [i.e. material gene] moves far beyond the simple scheme captured in formulas such as one gene [i.e. material gene]-one protein or polypeptide" (2007, 4).

[31] The interpretations of alternative splicing are complicated due to the issue of "similarity" between the functional outcomes. If the functional outcomes are similar enough, then there is one gene mirroring two different products, while if the similarity is too weak, then it can be a case of two genes. However, given that similarity is a vague relationship, drawing the line should be done case by case and should be based on scientific practice. What remains relevant for the current discussion is the importance of **Ft** for the identity of the genes.

[32] Another interesting case is the *DSCAM* in *Drosophila* (Kashyap and Tripathi 2008). *Drosophila* contains 116 exons of which 17 are maintained in the final mRNA. Kashyap and Tripathi (2008, 3) write that "theoretically this system is able to produce 38,016 different proteins. And, in fact, over 18,000 different ones have been found in *Drosophila*", supporting the multiple determinability of the stretch and the realisation of different functional products.

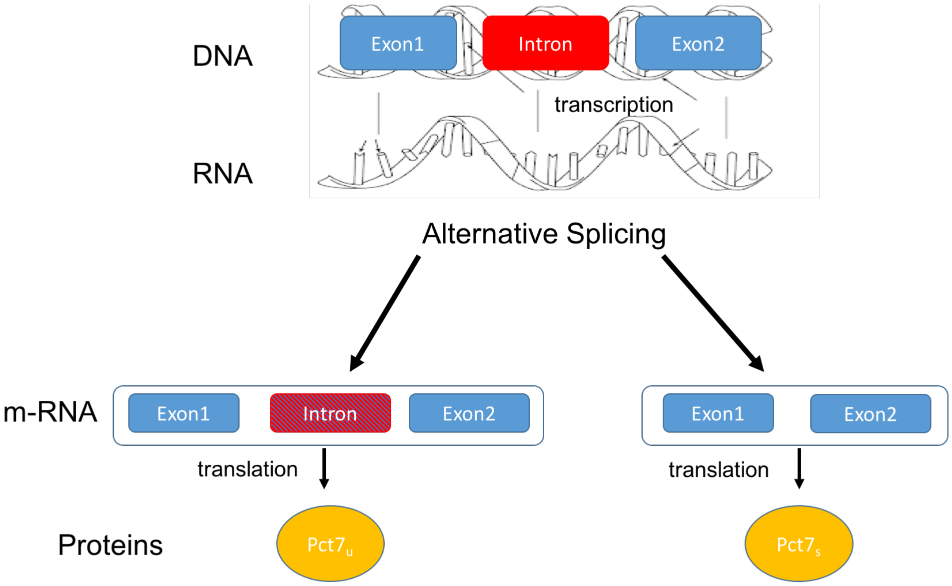

Figure 4. Alternative splicing in *PTC7* in *S. cerevisiae*.This is a simplified representation of alternative splicing at the *PTC7* gene and the encoding of the proteins Ptc7u, with the intron, and Ptc7s, without the intron. It helps to visualise the difficult identification of the gene with a precise stretch of the DNA and its multiple determinability. For the full information about alternative splicing of *PTC7*: Jeneau et al. (2009). This figure is taken from Bellazzi (2022a).

difference and the consideration of it improves scientific explanations.

### Robustness

Further, genes can be considered a relatively robust phenomenon. Robustness is the core ontological feature of emergence, guiding ontological commitment and supporting novelty further. Specifically, a phenomenon is robust when it displays stability within a certain range of perturbations, and its properties are realised even if the system undergoes some modifications (Butterfield 2011; Franklin, Knox 2018).

In the genetic context, perturbation can be read in terms of mutations or changes that can affect the realisation of **Ft** and the transcribability of a given region of DNA. This type of robustness can be indicated as a form of *functional robustness*[33] that is “the robustness of some

---

[33] This is a type of *structural robustness*, as a change in the attributes relevant to the structure of the system under consideration, and in the biological cases, it includes the environment as well (Weisberg 2013, 161). It differs from *parameter robustness*, i.e. changes to the parameters of the model’s description, and *representational robustness*, i.e. changes to the representation of the attributes in a model (Weisberg 2013, 160-163).

function or effect produced by a system over variation in or perturbations to the components and properties of that system" (Boone 2018, 81).[34] Briefly, this means that the functional property is realised despite a range of modifications at the underlying lower level. Robustness is so related to multiple realisation, through which the property is realised by multiple lower-level features, and to multiple constitution, for which the phenomenon defined by the functional property can be constituted by a variety of bases within different circumstances (Gillett 2013; Kistler 2018; Boone 2018). In order to understand the robustness of the gene, let us recall the definition offered by Gerstein et al.: "A gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products" (Gerstein et al. 2007). As evident from the relations mentioned in §4.1, **Ft** of a given union of genomic sequences (the gene) is robust within limits of perturbations of the DNA stretch. This is because there are many complex mechanisms, such as reparatory or alternative ones, that maintain the transcription of a given set of potentially overlapping functional products stable. Consequently, the gene is said to be robust because it can be constituted by different stretches of the DNA and obtains despite some ranges of modifications in its components.[35] Robustness is evident in a variety of genetic phenomena and I will consider some of them in what follows.

A first evidence comes from "synonymous mutations", changes in the DNA sequence that codes for a specific amino acid without affecting the final product, the encoded amino acid. This is possible thanks to the redundancy of the genetic code, an adaptive feature for which multiple codons can code the same amino acid. Broader synonymous mutations can also code for the same polypeptide, supporting the robustness of the functional property **Ft**. An interesting example is glycine, a proteinogenic amino acid (an amino acid integrated into a protein during translation). Glycine is codified by GGT and any change in the third position of the codon, either with A, C or G (resulting in GGA, GGC, and GGG) will result in the same amino acid in the right position, even when coding a more complex protein sequence (Waters et al. 2016, 3) [Figure

---

[34] This elaborates on Mitchell (2008, 698), where robustness is defined as "the ability of a biological system to maintain normal functioning in the face of internal or external perturbations".

[35] Note that this section is not arguing that the gene is more robust than the DNA, but that the gene displays some robustness.

5].[36]This phenomenon is called single-nucleotide polymorphism, because the mutation involves only a single change in the codon. This synonymous mutation is simple, highly frequent and shows a first level of robustness across perturbations.

Moreover, robustness is found in broader phenomena where the context and mechanisms around the gene allow for the property **Ft** to be present despite changes in the DNA. Cells have evolved an "orchestrated interplay of various DNA repair mechanisms to prevent the life-threatening disruption of replication and transcription by DNA damage" (Walmacq et al. 2012, 1). This type of robustness represents a necessary feature for the cell to perform protein synthesis and other activities. If the cell were to be disrupted by any underlying perturbation, it would probably stop living very soon. As Griffiths and Stotz (2013) illustrated, the genome is highly reactive to a system of different mechanisms that allows for the transcription of the gene and maintains such transcription stability despite underlying perturbations. In so far as the

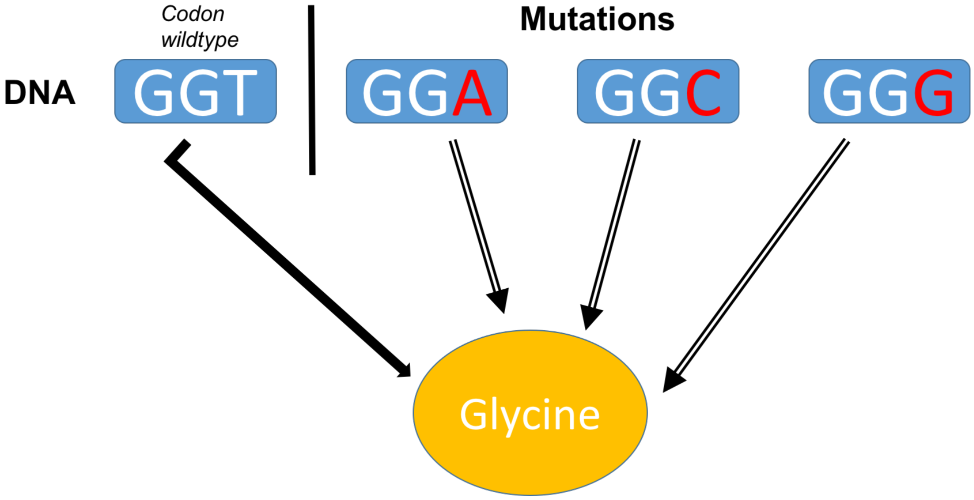

Figure 5. Silent mutations in codons for the amino acid glycine.This is a simplified representation of the modification of a base in the wild type, GGT, into three silent mutations. The black arrow stands for codification in the wildtype, while the empty black arrows stand for codification in the cases of mutation. The amino acid is made in all four cases. The gene encoding a protein containing glycine and the function remain robust because the function is realised even in the presence of mutations. The figure helps visualise the *multiple realisablity* of the functional property of the gene, in this case the property "encoding a protein with glycine". This figure is taken from Bellazzi (2022a).

---

[36] GGT is the most common wildtype for glycine, even if there are variations across different species. For instance, some *Bacteria* will use GGA more often.

transcribed products are potentially overlapping and functionally similar, the considered union of genome sequences can be deemed as the same gene (as Gerstain et al. 2007; Fogle 2010). An instance where robustness is evident is in mechanisms of translesion transcription by RNA polymerase II in *S. Cerevisiae*. In these cases, the translesion RNA polymerases II is able to transcribe the gene with or without repair of the damages created by exposure to UV lights (Walmacq et al. 2012).[37] This is relevant because it allows for the property **Ft** characterising the gene to be realised, despite the underlying modifications. Furthermore, given the relational nature of such a property, it shouldn't be surprising that this realisation happens thanks to the interaction with an external actor, the RNA polymerases.

To sum up, all these cases of robustness are accounted for by the multiple realisability of **Ft**, the multiple constitution of the gene and the multiple determinability of the DNA. Such multiple realisation and multiple constitution concerning genetic phenomena are possible thanks to all the interactions that are pivotal to the defining property of the gene **Ft**. The possibility of transcription becomes real within the right environment and system of relations.

The robustness of the gene has interesting consequences for the tensions concerning the gene concept. First, robustness is a feature present thanks to a full set of interactions and mechanisms that surrounds the gene and allows its stability. This is in line with the idea that the identity of the gene could be understood only within the right context, the cell and the overall environment. The defining property of the gene **Ft** is a relational property and this makes it dependent on different contexts. Such context dependence of the gene should not represent an obstacle to its existence but should be accepted as a crucial feature of its identity. Second, a given gene's existence has a temporal connotation. Using Griffiths and Stotz's expression: "genes are ways in which cells utilise available template resources to create the biomolecules that are needed in a specific place at a specific time: genes are things an organism can do with its genome" (2013, 75; also in Griffiths, Stotz 2006). The definitional property of the gene, being relational, can only be realised when there is the possibility for such relation to hold and when there is the need

---

[37] UV exposure can cause cyclobutane pyrimidine dimers (CPDs) in the DNA strand that stall transcription elongation by RNA polymerase II (Pol II). For full reference on translesion RNA polymerases II see Walmacq et al. 2012.

for it to hold. This implies that the gene itself exists within protein synthesis or, more general, within the process of transcription. Thus, it is dependent on a variety of different mechanisms that make the genome react in different ways, allowing for the emergence of the genes in specific moments.

Taking stock, a gene emerges from a precise stretch of DNA when it is transcribed, and this happens when complex mechanisms are acting in the surrounding context. It is this possibility that permits its novelty in scientific explanation and it is such a system of mechanisms that allows for its robustness.

## 3.5 Conclusion: flexibility and context dependency

The account presented in this chapter wants to accommodate the tensions concerning genes' identity and existence. Specifically, these tensions are generated by the complexity of the genetic phenomena, as highly context-dependent, and the importance of genes across different biological disciplines, from molecular to medical biology. Here, I have argued that they can be solved by elucidating the existential characteristics of genetic phenomena: genes are existent weakly emergent entities. This result has been reached in two steps. First, I have clarified and defended a definition of postgenomic genes as "images of the molecular products", re-elaborating the ones of Gerstein and colleagues (2007) and Griffiths, Stotz (2006, 2013). Then, I have argued that genes are emergent as they display novelty and robustness. A consequence of this is that a gene emerges during transcription. Furthermore, this emergence remains weak since it is compatible with a form of ontological reductionism. Thanks to the token identity between the gene and the DNA, we can maintain it as materially constituted of nothing more than chemical molecules. However, the straightforward linear relation between DNA segments and genes is untenable. The gene has a proper type identity that is given within the cellular environment.

The emergence of genes is an important conclusion as it offers a new conceptualisation of the complex genetic phenomena, allowing us to accept their existence despite the unclear material constitution and the multiply-realised transcription factor. Postgenomic genes exist within the different circumstances in which transcription happens, and the identification of the precise basis should be highly flexible (Fogle 2010). Such flexibility, crucial for the topic, cannot be accounted for if genes are considered merely materially, as mere ORFs. Nor we have to be satisfied with the simple nominal and operational definition of the genes. On the opposite, we might consider emergence as what allows the existence of the genes with more flexibility in defining their borders and permits the consideration of their core relational and context-dependent property. Genes can thus result to be both flexible and highly context-dependent, and nevertheless existent, as they depend on the right environment that allows their definitional property to be robust. Furthermore, this account provides an ontological and not merely theoretical reason for which gene-talk in life sciences is justified as well as the important role that genes play in contemporary biological disciplines. If one of the aims of science is to unmask which kinds of things exist in the world, then the present account offers more support for the fact that science has reached this goal. Moreover, the importance of considering the context-dependency of genes might be relevant for those scientists that are trying to conceptualise and model genetic phenomena. This case can also be of interest to philosophers working on emergence and inter-level relations in the sciences. It presents the gene case in support of a recently proposed account of ontological emergence that was first applied to physical phenomena but aimed to be more general. Particularly, it fits biological cases thanks to the relevance of robustness and its attention to the importance that specific scientific concepts have in explanations.

Concluding, this chapter presents some results that are relevant for the development of this thesis. The first one is that the characterisation of gene as weakly emergent from the DNA is one of the two steps of the general strategy that is used to argue in support of the *sui generis* nature of biochemical kinds. The second is that it shows which properties are relevant for its characterisation: chemical structure and biochemical function. This latter result will represent the starting point for Chapter 4, which will continue the argumentative strategy by discussing whether the molecular gene can be considered a natural kind despite its complexity. Moreover, the identification of the core properties of the gene will be used in Part III to explore further aspects of the metaphysics of biochemical kinds.

# Chapter 4

# The molecular gene as a natural kind

## 4.1 Introduction

In the previous Chapter 3, I addressed the question concerning genes' existence despite the complexity of genetic phenomena.[1] The main result has been that instances of genes are weakly emergent entities during the process of transcription. This has been argued starting from the fact that the gene has two properties, a chemical structural one and a functional one, and the functional property respects the criteria set in the previous chapter for weak emergence. In this chapter, I will consider whether the category molecular gene can be considered a natural kind or whether it is just a useful tool or a conventional category for genetics and genomics. The naturalness of the gene category can be challenged upon the consideration of the developments of genetics of (at least) the last 20 years and the complexity of the genetic phenomena preciously discussed (Griffiths, Stotz 2006, 2013). The identity relation between individual genes and precise stretches of DNA has proved impossible and this might support a deflationary or nominalist view of the gene (Griffiths, Stotz 2006; El-Hani 2007; Fogle 2010). However, as expressed in the previous chapter, a realist approach to the gene can still be maintained while better understanding the cellular context in which it operates and embracing its complexity (El-Hani 2007 et al.;

---

[1] A version of this chapter is forthcoming the Edited Springer Volume *Life and Mind* with the title "The gene as a natural kind."

Griffiths, Stotz 2013; Bellazzi 2022a). But is this enough for the gene count as a *natural category* despite such complexity and context dependency?

Here, I will consider this question, and I will argue in favour of a positive answer as the gene satisfies some requirements for natural kindhood, following Khalidi's account of natural kinds (Khalidi 2013, 2018, 2020, 2021) [Chapter 2]. I will argue here that the category "molecular gene" used in scientific practice corresponds to a natural kind despite the complexity of the properties characterising it, and it captures some objective features of reality.

This inquiry has some conclusions of interest. First, this case study is interesting for the general debate on natural kinds because it illustrates how natural kinds can be found even within complex and highly interactional systems. Second, understanding whether something is a natural kind or not is important because the naturalness of a given category can provide us with a further justification for why we can make more robust inferences from it. In doing so, the identification of something as a natural kind can support the justification of a theory that presents such a kind. Third, a natural kind is more than a theoretical entity whose properties are postulated for practical purposes and this can direct research into discovering (rather than merely postulating) features about it. This supports the role that they also have in the process of discovering new information about the category under consideration.[2] Lastly, and more generally, this project represents an instance in which biology and biological practice inform the philosophical debate on what counts as a natural kind.

In respect to the results of the thesis, this chapter represents an extended case study of the account of natural kind assumed in the thesis and can support a form of unity of science via Natural Kind Monism [as defined in Chapter 2]. This is particularly relevant given the complexity of genetic systems and interactions. Such complexity together with the multiple realisability and multiple determinability advocated in the previous Chapter 3 might be taken to support a disunity in genetics or maybe a plurality of kinds at play in this domain. Here, instead, I will argue that the molecular gene can be considered a natural kind according to one single account

---

[2] On this, a reference is the short article Bellazzi 2022c, "Genes: from instruments of thought to the base of life" published on *Jargonium*.

and in a way that acknowledges the relations that exist between function and structure.[3]

The structure of this chapter is the following. In section §2, I will provide a brief summary of the definition of gene presented in the previous Chapter 3. This will serve as a reminder to the reader of the status of the controversy and the starting point to explore whether the molecular gene can be considered a natural kind. In section §3, I will summarise again the relevant account of natural kinds considered, as presented in Chapter 2. The account of natural kinds used here is the one presented by Khalidi (as in 2013), for which a natural kind is a category postulated by science that is projectible and a node in causal networks, as will be further clarified. Having provided the metaphysical ingredients, in section §4 I will move on to consider whether the gene is a natural kind. I will argue that this is the case as it satisfies the criteria for natural kindhood mentioned above.

## 4.2 The gene in the "postgenomic" world

As presented in the previous Chapter 3, various attempts have been made to (re)-define the gene concept in the "postgenomic" world, trying to accommodate both practical and theoretical requirements (as Beurton 2010). Summarising, we can highlight two main ways to re-think the gene concept within the molecular context [Chapter 3] (Bellazzi 2022a). The first is a deflationary instrumentalist approach that allows to retain gene-speech and use without further ontological commitment. The second is a realist approach that tries to define the gene by embracing its complexity and context dependency. While both these approaches are informative about gene individuation, i.e. how to identify individual genes, here we are concerned with the gene definition or gene characterisation, i.e. what it takes for a given entity to be a gene.[4]

The deflationary approach to the gene is often referred to as "the nominal gene" and iden-

---

[3] More on natural kinds in biochemistry will be discussed in Chapter 7.

[4] Havstad defines three classificatory practices concerning kinds: classificatory characterisation or definition, individuation and organisation (2021, 7670). The first - the one we are concerned with -focuses on the definition of the kind. The second focuses on identifying which tokens belong to a given kind. The last focuses on organising taxonomies.

tifies it in an operational way based on actual scientific practice and conditioned to research needs. According to this account, a gene can be any stretch of precise nucleotide sequences that encode a specific product (Burian 2004; Griffiths, Stotz 2006, 2013, 66). This approach leads to the "consensus gene", as whatever pattern of "biochemical architecture and process" presents the features of the exemplary gene, according to empirical evidence and scientific pattern (Fogle 2010; Bellazzi 2022a). This approach to the gene remains nominalist or deflationary because it does not commit to the existence of the gene as something *sui generis* nor ontologically distinct: genes are any stretches of DNA that we find useful to identify as such given a particular model of gene. Moreover, the identification of the gene remains conditional on research interests, maintaining a minimal commitment to the entity.

In contrast to this approach, we also find a realist one, presented in the previous Chapter 3 and the context of the ENCODE analysis. Within this latter project, Gerstein and colleagues take the gene to be an existent entity (and so independent from our research interests) and define it as "a union of genomic sequences encoding a coherent set of potentially overlapping functional product" (Gerstein et al. 2007). This has been further re-elaborated by Griffiths and Stotz, who understand the "postgenomic genes" as existent "images of the target produced molecules" (Griffiths, Stotz 2013, 75). While this latter is not a definition, it represents a helpful metaphor to understand the gene: it should present a non-necessarily contiguous sequence that is similar enough to the one of the transcribed molecule. According to this realist view, the gene is not simply identical to a linear and contiguous sequence, but a union of different ones. This union normally includes the finally transcribed sequence and the promoter region (or TATA box) (Fogle 2010, 6). In addition, the gene often comprises those regions essential for its activation and the regulation of its transcription; these might be contiguous or not (Griffiths, Stotz 2013). Following this realist approach in the formulation presented in the previous Chapter 3, a gene is an entity composed of those parts of the relevant nucleic acids that are transcribable (or involved in transcription) and encode a given mRNA. This second aspect, being transcribable, represents a functional component in the definition, as the gene is a union of sequences with the function of encoding a target molecule. While the gene has a material component, according to this approach, it is not only characterisable materially or as a material entity. The functional characterisation of the gene also allows us to embrace the context-dependency pointed out by the postgenomic analysis: "a function is always a role in something and a contribution to something" (Germain et al. 2014). Transcription is not a self-subsistent phenomenon but rather a reactive one: it happens only within the right circumstances and thanks to a set of interactions that operate at different levels. Accordingly, the gene can be fully understood only within such a context of action and interactions, and it is an entity defined materially and functionally (Falk 2010; Bellazzi 2022a).[5]

In this chapter, I will start from the second approach, in which the gene category refers to an existent union of sequences that are transcribable (or involved in transcription) and that transcribe precise genomic products. This definition is coherent with contemporary genomics, which can identify some (often not-contiguous) unions of sequences that take part in the transcription of given molecules (either amino acids chains or RNA molecules). But is the gene a natural kind? Or is it just a convenient category that groups some existent phenomena for practical needs? In order to answer these questions, we need to recall how we can distinguish natural from non-natural kinds.

## 4.3 Natural kinds again

As summarised in Chapter 2, there are three questions one can ask when exploring the topic of kinds: i) the *ontological question* that asks which kind of entity is a natural kind; ii) the *naturalness question* that asks what makes a kind natural[6]; iii) the *unification questio*n that asks how the properties of the kind are unified together (Tahko 2022; Magnus 2015, 2018). In

---

[5] As will be further clarified, the functional component of the gene definition allows for both multiple realisation and multiple composition (in contrast with a materialist only view). A gene would be any entity that is composed of the relevant material aspect, nucleic acids - either DNA or RNA - and that has the relevant function. However, for individual genes, the correspondence does not need to be 1 stretch: 1 function, as the function could be multiply realised by any stretch that realises it. For further references, see Chapter 3, Bellazzi 2022a.

[6] This question can also be referred to as the "taxonomy question". However, I prefer to avoid this terminology as it seems to constrain natural kinds to biological taxonomies or to identify kinds with taxonomical classifications.

this chapter, we are inquiring into whether genes are more than a useful tool to do genomics and we will focus on the naturalness question to do so.

We can identify two broad strategies to answer the natural question. A first is offered by conventionalism, for which there are only conventional categories that suit different purposes, as recently proposed by Brigandt (2022). A second is offered by various forms of realism: at least some of our categories correspond to natural kinds and the objective features of reality. Often, it is common to take a hybrid position. Some kinds that we find in scientific practice or daily life are conventions, while others might correspond to real features of the world. For the sake of the present analysis, I will assume a form of minimal realism for which at least some categories correspond to genuine features of reality and might be candidates for natural kinds (Khalidi 2013; Bird, Tobin 2022). However, not all categories we identify as kinds seem to correspond to such divisions. How do we distinguish "natural kinds" from mere "human categories" (Khalidi 2013)?

In the biological sciences, this question is particularly relevant as the study of life appears to be the reign of taxonomies and classifications. Historically, species have been considered a paradigmatic case of kinds, and the concerns on whether all of the *Linnean taxa* correspond to actual divisions in reality have been widely debated. In order to identify natural categories, some accounts have been proposed, among which are essentialism in intrinsic and historical forms, HPC cluster theories and others.[7] However, most of these accounts are concerned with the question of whether any of the *taxa* can be considered a natural kind (Slater 2013). This makes most of the accounts discussed in the literature on biological kinds interesting for the species or higher taxonomies debate but makes these accounts less applicable to other candidates of natural kinds in biology, such as kinds in biochemistry or at other levels (Slater 2013; Khalidi 2013; Kistler 2018). Accordingly, in this chapter, I will consider the naturalness question concerning the molecular gene by following the account of natural kinds presented in Chapter 2. Specifically, the inquiry is into whether unions of genomic sequences identified as a gene can be considered instantiations of a natural kind or not.

---

[7] See Khalidi 2013; Bird 2018; Bird, Tobin 2022 for an extensive overview of these accounts. A summary of the status of the species controversy can be found in Ereshefsky (2017).

### 4.3.1 Projectibility and causal networks

The debate on natural kinds is wide in approaches and topics. Nevertheless, there is a consensus that kinds should allow making reliable explanations and predictions across instances (Tobin 2013; Khalidi 2013, 2018, Bird, Tobin 2022). This feature is represented and discussed, together with other accounts, in Khalidi's approach presented in Chapter 2 (1993, 1998, 2013, 2018, 2020, 2021). His view presents kinds as those categories present in scientific theories that are projectible and capture nodes in causal networks of the world. This approach stands out because it is a realist account, for which kinds track objective features of reality, while at the same time avoiding excessive metaphysical commitments; as already discussed, it does not commit to the existence of universals or similar categories [Chapter 2].[8] Moreover, it is applied and applicable to various kinds from the physical and special sciences, taking into account both structural and historical or etiological properties. Accordingly, I take his view as a starting point to explore whether the gene could be deemed a natural kind. While my argument is conditional upon such account, I do not think that this compromises its validity as this account has commitments on projectibility and causal efficacy, which are often considered valid criteria for natural kinds (as in Bird, Tobin 2022). Let me briefly recall the account presented in Chapter 2 (Khalidi 2013, 2018, 2020, 2021).

This account answers the naturalness question, inquiring how we can distinguish conventional, non-natural and natural groupings from objectively existent ones. Natural kinds are objective features of reality that can be revealed by our systematic attempts to gain knowledge of nature. One of the best ways to attain this knowledge is through science and scientific practice. Accordingly, looking at the kinds present in science should be the starting point for the inquiry into natural categories (Khalidi 2013, 2021). Nevertheless, even within the best scientific theories, the history and philosophy of science have shown that not all categories present in the discipline can be considered as capturing something in the world [Chapter 2]. Moreover, some

---

[8] This view can be seen as a form weak or moderate realism (r) for which kinds are objective features of reality, but do not correspond necessarily to distinct metaphysical categories such as universals. This answers the ontological question rather than the natural one, and will not be considered further in this chapter.

categories can have an instrumental role or cannot be considered stable or robust enough to be informative about the world. This led Khalidi to add two further requirements for naturalness. With being present in scientific theories as the first criterion (i), these categories also need to satisfy ii) genuine projectibility and iii) being a node in a causal network.

Projectibility can be defined as the feature for which we can make stable and robust generalisations about kinds across instances, and it is often assumed as a feature that natural kinds should display(Tobin 2013; Khalidi 2013, 2018; Bird, Tobin 2022). Natural kinds are particularly efficacious categories when framing inductive inferences and feature in many empirically verifiable generalisations. This means that kinds are projectible, so they can be projected from one instance to another successfully. More precisely, the projectibility of a natural kind can be defined as follows:

> “when it comes to a natural kind predicate K, there is no shortage of other predicates, P1, P2, ..., Pn, and so on, such that we can reliably assert that if x is K, then x is P1, x is P2, ... , x is Pn and we can do so with a high degree of generality” (Khalidi 2018, 1385).

Kinds provide explanatory and predictive power across different contexts and circumstances because they allow us to project a set of properties from one instance to the other and to predict that such properties will be present (Khalidi 2013, 2018, 2020, 2021).

Moreover, this kind of projectibility requires an explanation: why is it possible to draw these inferences? What is the ontological ground for which the kind can be applied in an explanatory way to many instances? Khalidi answers these questions by adding a second requirement that natural kinds should satisfy: they are “nodes” in causal networks.[9] Projectibility results to be a “reflection” of the causal network in which instances of the kinds are involved and some kind categories are particularly successful because the properties of the natural kinds are causally clustered (Khalidi 2013, 2018, 2020). The joints that natural kinds carve so successfully can be found in the world’s causal structure. Together with providing an ontological reason for the

---

[9] It is important to notice that natural kinds display role in the causal network can be seen when considering instances of the kinds due to the nature of the causal relation. I thank Jessica Wilson for suggesting this important clarification and an anonymous reviewer for the volume *Life&Mind* for insisting on this aspect.

projectibility of kinds, causal relations also answer the naturalness question. Specifically, they play two main roles in distinguishing natural from conventional kinds. First, natural kinds do not present only a set of projectible properties but a set of properties that are hierarchically ordered as "causes and effects in recurrent causal processes" (Khalidi 2018). They present a set of "core" properties that cause the instantiation of other derivative properties of the instance of a given kind. Natural kind results then to be those categories with a set of properties discoverable by science and whose co-instantiation causes the instantiation of other properties (Khalidi 2013). Second, natural kinds are those categories that represent nodes within broader causal processes: they are causally efficacious on other kinds and are inter-sections within the webs of causal relations. The causal relations among property instances and the causal cores of natural kinds represent the ontological principle in virtue of which we can distinguish natural from unnatural kinds. Moreover, it is the underpinning of the projectibility of such categories.

As summarised in Chapter 2, this theory presents some advantages. First, this account is a good example of the consensus reached by philosophers on how to answer the naturalness question, providing us with a way to disentangle natural from unnatural kinds. This is done without embracing a theory that tries to provide a specific set of necessary and sufficient properties for which an individual is a member of a kind. Moreover, it is an application of a form of reflective equilibrium between scientific input and philosophy. It is a combination of convictions on categories generally regarded as paradigmatic kinds, often taken to be stable categories in scientific theories, philosophical discussions on natural kinds and a set of considerations that are drawn from scientific practice. Lastly, this account can be applied to a variety of different types of kinds. It is able to accept etiological and historical kinds as natural kinds, considering a particular origin or genealogical history as the core properties that cause the instantiation of other ones. Moreover, the combination of naturalism with projectibility and causality allows the theory to be applied to concrete case studies within fundamental physics and the special sciences, such as lithium, cancer cells, viruses, ADHD (examples from Khalidi 2013). Concluding, this account answers the naturalness question and will be used in the next section to explore whether genes can be considered a natural kind.

## 4.4 The molecular gene as a natural kind

The naturalness of the gene category has been questioned because of the history of the category mentioned at the beginning of Part II, the complexity of the genetic phenomena and the context-dependency that is implied by genes' functional aspect. Genes might not seem "neat" enough to count as a natural category or might be too ambiguous.

In this section, I will argue for the opposite, defending a view for which the gene can be deemed a natural kind. Before pursuing my thesis, some clarifications are in order. This is because the genomic and biochemical domain is rich in systems of practice and taxonomies, and clarity is particularly needed to avoid ambiguities. According to Havstad we need to consider three classificatory practices concerning kinds: classificatory characterisation or definition, individuation and organisation (2016, 2021). The first focuses on the definition of the kind, already presented in §2. The second focuses on identifying which tokens belong to a given kind. The last focuses on organising taxonomies. Accordingly, before assessing the naturalness status of the gene category, I will briefly consider gene classifications and taxonomies. Then, I will explore whether the gene can be considered a natural kind.

### 4.4.1 Genes classifications and taxonomies

As presented in section §2, the gene is defined the postgenomic analysis as "a union of genomic sequences encoding a coherent set of potentially overlapping functional product" (Gerstein et al. 2007). According to this approach, genes are those (not necessarily contiguous) sequences that have a specific function in encoding a given molecular product. The classification of tokens or instances of genes into types or families is mostly done according to the product: unions of sequences encoding the same product (or products that are similar enough) are clustered into the same type of genes . The specific function of a gene within transcription is what allows to classify them.[10] Moreover, such classification of genes is done at different levels. First, there is a

---

[10] For the importance of functional similarity of products in genes classification, see Fogle 2010. Moreover, despite the importance of the functional aspects of the genes, genes are not mere functional kinds, because the material component they present as union of genomic sequences is also relevant in determining their identity.

broad classification of genes into two functional subtypes: i) genes that encode regulatory RNAs that play different functions within cellular processes; ii) genes that encode an RNA for the amino acid sequence of a polypeptide (Perini 2011). Then, we can find more specific classifications of genes tokens into gene types according to the given molecular product of the gene under consideration. Tokens of the same type of genes are clustered together if they encode the same molecular products or molecular products that are considered similar enough (Gerstain et al. 2007; Fogle 2010; Griffiths, Stotz 2013). In line with the importance of the link between genes and products, protein-coding genes (those encoding an RNA for a polypeptide) are normally named after the protein they encode. In this classificatory system, the locus is indicated often in italic capital letters while the name of the gene in capital letters and the protein made in normal characters. An instance of gene classification based on DNA sequences and function is the one of the genome of *Saccharomyces cerevisiae*, sequenced in 1996 and constituted of about 6,275 genes, organised on 16 chromosomes. Of these genes, about 5,800 are identified by their function. For example, the gene *DCS2* encodes the protein dcs2, which takes part in the biological processes that regulate the response of the cell to heat (Liu et al. 2017).[11] Generalising, the link between genes and products, together with the consideration of the relevant union of genomic sequences, is what normally allows gene individuation and gene-talk.

As far as other taxonomies or nomenclatures are considered, genes are classified in a variety of ways often referred to as "gene ontologies". In order to aid scientific practice, these ways of classifying the genes have been grouped within the project The *Gene Ontology*, a resource that provides a computational representation of our current scientific knowledge about the functions of genes according to the functions of the product they encode. Overall, this project considers three aspects when classifying genes: i) molecular functions performed by gene products, ii) cellular components in which the gene product performs a function, iii) biological processes in which gene products are involved. For example, the gene for the product "cytochrome c" can be classified according to the molecular function of the product (oxidoreductase activity), the biological process (oxidative phosphorylation), and the cellular component (mitochondrial matrix)

---

[11] For the database on *S. cerevisiae* genome and the relevant articles see `www.yeastgenome.org`.

(Gene Ontology).[12]

Lastly, it is possible to find different ways to organise genes in taxonomies contingent on specific contexts and given scientific goals. For example, in evolutionary developmental biology and comparative genomics, genes are often classified in terms of evolutionary history. However, these taxonomies tend to be highly-context relative and presuppose the characterisation of the gene mentioned above. First, scientific practice clusters tokens of genes across species into the same type according to a given union of sequences and a given function in encoding a product that takes part in molecular and cellular processes. Then, thanks to the similarity of the sequence and the RNA encoded, they can be further classified in terms of evolutionary similarities. An example is constituted by genes that can be classified as homologous if they are inherited by two different species from a common ancestor and such classification is done in terms of sequences and RNA encoded.

To conclude, genes can be classified and organised in taxonomies according to different practices and in a variety of ways. Nevertheless, this classification presupposes an initial identification of the gene as a given union of genomic sequences that encodes an amino-acid sequence. Once the gene is individuated, then it can be classified for further purposes. This is possible thanks to a general characterisation of the gene concept, as that union of genomic sequences that are transcribable in a target molecule (Bellazzi 2022a). Such characterisation of the gene is the one relevant for the purpose of the chapter.

### 4.4.2 The gene *is* a natural kind

Having summarised how genes can be organised in taxonomies, I will now consider whether the category gene can be considered a natural kind. To do so, I will start considering whether the properties associated with the natural kind category respect the requirements for naturalness presented Chapter 2 and in section §3.

In section §2, I have pointed out the concept of gene relevant to our analysis. Genes are the union of sequences transcribable in a target molecule and play their function only in a wider

---

[12] For the Gene Ontology `www.geneontology.org`.

system of interactions and environments (Griffiths, Stotz 2013). According to this definition, the gene presents a two-fold identity: in terms of structure or composition and terms of function. First, the gene has a structural component as a region or union of regions of nucleic acids, while not every region of the given nucleic acid is a gene (Fogle 2010; Griffiths, Stotz 2013). Specifically, the gene is composed of those regions that can be considered images of the target molecule and are actively involved in transcription. This leads us to the second aspect of the gene. The union of sequences has a given function in encoding the primary structure of a polypeptide or a functional RNA molecule. The functional component of the identity of the gene is evident during transcription, a reactive process that can be up-regulated and down-regulated thanks to specific intra-, inter- and extra-cellular interactions. This makes the gene a context-dependent entity and a multiply-composable one (Germain et al. 2014; Bellazzi 2022a) [Chapter 3]. Specifically, a given instance of the gene kind can be composed of different stretches of nucleic acids while the function characterising the gene is maintained as the final transcribed molecule is.[13] This feature is important to account for the complexity of genetic phenomena. Accordingly, to understand what a gene is, one needs to consider: first, the identification of those sequences that are transcribable in a target molecule and, second, the process of transcription where the gene plays its function. The identification of the main features of the gene category is the starting point in the analysis of whether the gene can be considered a natural kind.

Now, we need to explore whether these properties respect the three main requirements for naturalness: i) the presence of this category in our best scientific theories; ii) projectibility, whether the category has explanatory power and can be projected from one instance to another; iii) whether the gene is a node within a causal network.

Let me start with the first, being a category postulated and discussed in our scientific theories. Despite their complicated history and the current challenges raised by postgenomic analysis, genes still retain an important role across a different variety of life sciences. Specifically, they

---

[13] In order to appreciate how the definition allows for multiple realisability and composition, it is important to consider that the gene type is the one that can be multiply composed, while each gene token is going to be composed by specific nucleic acids sequences. Accordingly, for the gene definition, we just need to identify which specific token nucleic acids could compose it while maintaining that the relevant function is realised. For further analysis of the complexity of the relations between the properties of the gene, see Chapter 3, Bellazzi 2022a.

are often invoked in the study of protein synthesis and the impact that alterations in protein structure can play at the cellular and organismal levels [as presented in Chapter 3]. Moreover, genes still retain a role in developmental and evolutionary biology. The gene is a category that stably figures across different scientific disciplines, thus meeting the first requirement for naturalness. Genes also present projectibility and are nodes in causal networks. First, we can project the properties of the genes from one instance to another successfully. This is particularly evident in protein synthesis, one of the phenomena in which genes play an important role. In this case, scientists project the two main properties of the genes from one instance of protein synthesis to another in order to be able to explain and predict general patterns. At the same time, protein synthesis shows how genes are nodes in causal networks: they have a causal role in the synthesis of proteins and the two properties are causally related for this to happen.

To further support my argument and appreciate why genes respect the requirements of the natural kindhood, I will consider a specific case study: the gene mentioned above *DCS2* in *Saccharomyces cerevisiae.* This gene encodes the protein dcs2, which takes part in the biological processes that regulate the cell's response to heat (among other processes) (Liu et al. 2017). Furthermore, it presents the two properties of the category "gene" - being composed of nucleic acids and having a function -, making it a good case study to see how these properties support the naturalness of the kind.[14] Following the characterisation of the gene presented here, this gene is transcribable in a given target molecule, the mRNA for the protein dcs2, and i) it has a specific basis on chromosome XV of the cell; ii) it has a specific function, i.e. the one of encoding the protein dcs2.

First, we can assess the projectibility of this gene, seeing if it allows better explanations and predictions based on its properties and across different instances. The gene *DCS2* encodes the mRNA for the protein dcs2, a regulatory protein that acts as a pyrophosphatase regulator in many cellular processes and specifically in the heat response of the cell. The instances of this

---

[14] The purpose of this section is not to show that "*DCS2*" is a gene, but rather that the properties that genes display (and are present in *DCS2*) respect the requirements for the naturalness of kinds. Moreover, the causal component required by the naturalness claim supports using an instance of the kind to explore whether it is natural. I thank an anonymous reviewer for the volume *Life&Mind* for insisting on this clarificatory point.

gene have a common (type) nucleic acid base, which makes the relevant gene identifiable on a precise locus on chromosome XV of the cell, and they encode a given protein within the right circumstance. Also, instances of this gene take part in the process of transcription, which leads to the presence of the protein within the cell, thereby taking part in a causal process. The same role is played by distinct instances of this gene across different cell individuals and the identification of this causal pattern allows us to predict better and explain heat regulation processes. Moreover, this gene can be deemed homologous to the gene *DCS2* in other species such as *Homo sapiens* or *Mus musculus* (house mouse), allowing this category to be projectible not only across instances of *Saccaromyce cerevisiae* but also to other species. Accordingly, the gene *DCS2* is projectible because it has the properties of the kind "gene". This example illustrates how it is possible to project the category "gene" from one instance to another to make reliable explanations and predictions about genetic phenomena.

Second, we can explore whether genes can be nodes in causal networks. First, we need to see if the core properties of the gene can be considered causally related. Second, we need to consider whether the gene plays a causal role within a broader context. I will do so with the help of the previous example. The two properties that seem "core" for the gene *DCS2* are i) being composed of a given union of sequences and ii) encoding for the protein dcs2. The two seem to be at least partially causally related as the sequence allows for the interactions necessary for the transcription of the DNA sequence into the mRNA that then encodes the protein dcs2. The "cross-talk" between the sequence, RNA polymerases and the actors of transcription is what causes the second definitional property of the gene to be present. The sequence has a causal role in the manifestation of the functional property together with the relevant factors, allowing the encoding for a specific protein. This supports the presence of a causal link between the two main properties of the gene and allows a hierarchical causal characterisation of the other properties that might be an effect of those two. Furthermore, the gene under consideration plays a role within the general causal network. For instance, the encoding of the protein dcs2 contributes causally to heat regulation within the cell. Lastly, having considered the features of a specific gene to support the argument does not compromise the generality of the conclusion because the discussed properties of the gene *DCS2* are those typical of the kind "gene" - being composed of nucleic acids and being transcribable in a target molecule and not those specific of *DCS2*. Accordingly, similar considerations can be made for other instances - even more complex ones in virtue of the two core properties of the kind.[15]

As follows, the gene category captures a natural kind according to the requirements for the naturalness of a category. First, this category has an important role within our scientific theories, respecting a naturalist approach. Second, it is possible to identify two main properties that are important for gene definition. These properties allow its identification and can be used as a basis for the analysis of naturalness. The gene category is projectible, as its definitional properties have a predictive and explanatory role, and scientists can project results about one instance to other instances based on its core properties. Moreover, these properties can be deemed to be causally related, and the gene can be considered a node in a causal network in different processes.

## 4.5 Conclusion

In this chapter, I have argued that the gene can still be retained as a natural kind within the postgenomic context, following the view of natural kinds presented in Chapter 2 and by Khalidi (2013, 2018, 2020, 2021). The discoveries that followed the Human Genome Project and ENCODE project have shown that it is difficult to identify the gene with a precise stretch of DNA and that it is important to consider a full system of interactions to understand the complexity of genetic phenomena. A possible reaction to this controversy is to take a nominalist view of the gene and consider it a conventional or instrumental category. However, the success of the genes category across different sciences and its wide applicability can be considered a hint in support of the naturalness of the kind.

In order to argue in favour of the genes as a natural kind, I have illustrated in section §3

---

[15] I thank the reviewer for the volume *Life&Mind* for the suggestion to clarify the relation between the example of gene considered in the paper and the general kind "gene". The definition of gene defended in the paper allows the generalisation from simple to more complex case studies, as the definition presented here is compatible with both multiple realisation and composition. Accordingly, we can generalise it to cases such as those involving alternative splicing because it is possible to identify the two core properties of the kind "gene" also in those cases as presented in Chapter 3, Bellazzi (2022a).

what makes a given category natural. To do so, I have used Khalidi's account as a starting point for the analysis. Specifically, a kind is said to be natural when it respects three requirements i) it figures in our scientific theories; ii) it is projectible; iii) it is a node in the causal network. Then, in section §4, I have illustrated why the gene is a natural kind. This has also bought in necessary considerations on genes taxonomies and classifications into genes type and tokens. The definitional properties of the gene are individuated mostly synchronically and respect the criteria proposed for the naturalness of the kind. First, the gene category figures in scientific theories, respecting the naturalist requirement. Then, the individuated properties of the gene allow the members of the kind to be part of generalisations and projections concerning the category. This is possible because the properties are causally related to each other's and allow the instances of the kind to enter into causal interactions that are uniform and identifiable.

A realist approach to genes as natural kinds has some benefits. First, if we accept that one of the aims of science is to discover what kinds of things are in the world, then an argument in favour of the naturalness of the gene category supports the success of genetics. Second, having a realist or instrumentalist view towards a scientific category might have an impact on scientific practice and scientific discoveries. An example of this can be found in the history of the gene itself, summarised at the beginning of this chapter. The realist understanding of the gene supported by Muller offered a theoretical framework for which scientists started looking for the gene as a material entity rather than a mere theoretical instrument.[16] This contributed to the discovery and identification of the gene as specific stretches of DNA or at least located on the DNA. A realist understanding of the gene category seemed to have impacted the direction of research (Fox-Keller 2000). Accordingly, considering the gene as a natural kind can impact how scientists think about this category and offer a broader framework for future research, even if it might not change scientific practice on a daily basis. Lastly, applying a single account of natural kinds to a category complex as the molecular gene can be taken as further support for Natural Kind Monism, allowing for a form of unity [as presented in Chapter 2, as will be further discussed in Chapter 7, Chapter 8].

---

[16] A reference is the short article Bellazzi 2022c, "Genes: from instruments of thought to the base of life" published on *Jargonium*.

Concluding, Part II focused on some philosophical and metaphysical issues related to the molecular gene. Molecular genetics is an interesting interplay of chemical and biological phenomena. Specifically, the molecular gene represents a good starting point for the present research because it has a crucial role in protein synthesis and other biological processes, while at the same time it is composed of chemical compounds and nucleic acids. This detailed case study aimed at deepening the understanding of which properties are at play at the interface between chemistry and biology and at offering a first step to support the general thesis of this research. In Chapter 3, I argued in favour of the weak emergence of the molecular gene, which offers a view of the relation between the function of the molecular gene and the molecular structure of the nucleic acids. In Chapter 4, I considered the status of the category "molecular gene" and argued in favour of its natural status. This part has shown that two main properties are at play when considering kinds at the interface between chemistry and biology, those related to molecular structure and biochemical functions. This result will be the starting point for Part III.

Moreover, this case study supports that there is something specific about biochemical kinds and their properties. The existence and specificity of the molecular gene has been motivated by the two main theoretical steps that underpin this research [as presented in Chapter 1]: the characterisation of the relation between biochemical functions and chemical structure in terms of weak emergence and an account of natural kindhood, for which the molecular gene is a natural category.

# Part III

# The Metaphysics of Biochemical Kinds

# An introduction to biochemical kinds

> *"The biochemical, for example, is not just the chemical aggregated.* Biochemical entities may be individuated functionally in terms of their effects on the organism or the role that they play in a particular system."
>
> (Khalidi 2013, 200; emph. added)

Part II has presented an analysis of the molecular gene. Together with the results mentioned in the previous chapters, this case study showed that the gene has an important functional component: a gene is a stretch of nucleic acids that has the particular function of being transcribed. This is fundamental for both genes' emergence and the status of the molecular gene as a natural kind [Chapter 3, Chapter 4]. But what does it mean that the gene has a particular function, and, more generally, i) what is a function when ascribed to a biochemical molecule? The answer to this question opens two others: ii) What is the relation between the molecular structure of a biochemical kind and its biochemical function? iii) How can we identify a biochemical kind, and is this category natural?

Part III considers these questions and try to offer some more general metaphysical considerations of the biochemical world. The results of this part are important because they generalise the themes and topics discussed in Chapter 3 and Chapter 4, in particular concerning the weak emergence of biochemical properties and phenomena and the natural kind status of biochemical kinds. Moreover, these chapters also offer a novel metaphysical analysis of biochemical kinds. These kinds are, as previously underlined, interesting and important case studies for philosophers working on inter-level relations and questions concerning the unity of science. Specifically, this is so (as a result also of the analysis of the molecular gene) because biochemical kinds are often characterised by both chemical (offered by a structural characterisation) and biological (offered by a functional one) features. The nature and the relation between these two characterisations have been recently discussed in the literature (Slater 2009; Tobin 2010; Goodwin 2011; Bartol 2016; Havstad 2016, 2018, 2021; Kistler 2018; Tahko 2020, 2021), but some of the metaphysical questions that concern these kinds still need to be answered. This part aims to cover some of them and will use mostly vitamin B12 as a case study.

Part III is structured as follows. Chapter 5 represents the bridge chapter between Part II and Part III and considers in detail the theme of biochemical functions with the usage of vitamin B12 and the molecular gene as case studies. As we have seen, the ascription of functions is crucial for molecular genes. Moreover, such ascription has recently been the object of controversy. One of the published results of ENCODE is that the 80% of the human genome has been assigned a *biochemical function*, as discussed by Griffiths (2009) and Germain et al. (2014). However, some evolutionary biologists do not accept such attribution as genuine because it does not use an evolutionary notion of function (Doolittle 2013). This opens questions on the nature of the biochemical functions ascribed to the molecular gene. The problem of functional ascription and characterisation does not concern only the molecular gene. The ascription of functions and their consideration is well spread and important for other biochemical kinds: biochemists commonly ascribe functions to biomolecules and use such attribution in their classifications. In this regard, proteins, vitamins and other biochemical kinds are classified in terms of chemical structure and biochemical function. The relevance of functional attribution offers further support to explore and consider what a biochemical function is. This will be the purpose of Chapter 5. As will result from this analysis, the biochemical function and the chemical structure of a given biochemical kind are related. This opens questions concerning the identification of the exact relation between the chemical structural properties and the biochemical functional ones and whether a form of unity in biochemistry can be maintained given the complexity of biochemical molecules and interactions. Chapter 6 focuses on exploring the relation between biochemical functions and chemical structure. It argues that biochemical functions are weakly emergent[17] from the chemical structure and that this is compatible with a form of unity, further elaborating some of the considerations presented in Chapter 3. The analysis of biochemical functions and their relation to chemical structure calls for further clarifications on the nature of biochemical kinds *per se*. Chapter 7 considers biochemical kinds directly. This chapter argues that biochemical kinds are natural according to Khalidi's account and that biochemical molecules can be considered instances of a *sui generis* natural kind.

In conclusion, Part III argues that there is something *sui generis* about the biochemical realm, characterised by weakly emergent biochemical functions and natural kinds. This result is supported thanks to the two theoretical steps explored in Chapter 5, Chapter 6, Chapter 7 [with the same structure of Part II]. Once biochemical functions are defined, it is possible to identify the relation between biochemical functions and chemical structure as a form of weak emergence [Chapter 6]. Moreover, biochemical kinds can be considered natural kinds [Chapter 7]. The results expressed here also have two implications concerning unity of science [Chapter 1, Chapter 8]. As previously mentioned, the disunity of science is often defended via the irreducibility of the different entities and by a pluralism concerning natural kinds [Chapter 1, Duprè 1983, 1993]. The results of Part III will try to counter-argue these two aspect. Firstly, I will argue that a form of unity can be preserved in identifying weak emergence as the relation between biochemical functions and chemical structure [Chapter 6]. Secondly, I will consider a form of unity via Natural Kind Monism (Tahko 2021). Biochemical molecules will be seen as instances of a natural biochemical kind that is identified as natural in the same way chemical and biological ones are [Chapter 7]. This form of Natural Kind Monism can offer further support for the unity of science.

---

[17] The relevant detailed definition of weak emergence will be presented in Chapter 6.

# Chapter 5

# Biochemical functions

## 5.1 Introduction

Function talk common to within different life sciences.[1] From macroevolution to genetics, functions are mentioned everywhere, e.g. a limb's function is to allow movement and RNA polymerases' function is to transcribe DNA. Biochemistry is not immune from such characterisation. Specifically, biochemists commonly ascribe functions to biomolecules and can classify them accordingly. This has been noticed in the recent philosophical literature on biochemical kinds. For instance, Slater (2009), Tobin (2010), Goodwin (2011), Bartol (2016), Havstad (2016, 2018), Kistler (2018) and Tahko (2020, 2021) mention that biochemical kinds are defined in terms of chemical composition and the function their instances exhibit in biological systems. Functional characterisation is well spread in biochemistry and involves genes (Germain et al. 2014) and kinds as vitamins (Combs 2012, 377; Carr, Maggini 2017; Fang et al. 2017). Proteins offer a common case study: they are defined in terms of the amino-acid chain that composes them and the function they play. The attribution of a functional property is not secondary as proteins are classified into two varieties: fibrous proteins, which mostly have a structural function; and globular proteins, such as enzymes or antibodies, which have a physiological function (Bartol 2016,

---

[1] A version of this chapter is forthcoming in Bellazzi F., "Biochemical functions", *British Journal for Philosophy of Science.*

534). But, while a lot has been written on biological and psychological functions, little has been said explicitly about biochemical ones. It is often assumed that biochemical functions are just biological functions (as mentioned in Slater 2009; Bartol 2016), as those that can be identified following the principles of natural selection. Nevertheless, the important chemical compositional nature of biochemical kinds requires a deeper analysis of what biochemical functions are. The biochemical world seems to be a chemical world embedded within biological processes, and this impacts the nature of biochemical functions as well.

This chapter explores the ontological consequences of accepting functional attribution to biochemical molecules, granted that scientific practice and relevant literature seem to accept that this is the case.[2] Such analysis is important because biochemical kinds have recently caught the attention of philosophers interested in the unity and disunity of science due to their representing “kinds at the borders” (Kistler 2018; Tahko 2020, 2021). On the one hand, biochemical kinds have a structural chemical nature, on the other, they also have a biological functional one. Accordingly, understanding the nature of biochemical functions can inform such debates. Moreover, it can provide further clarity when considering functions in the biochemical context for different purposes. For instance, the ascription and understanding of biochemical functions can improve the explanation of the biosynthesis of some biochemical kinds and their necessary introduction via nutrition. Here, I will argue that if we accept such attribution as genuine, biochemical functions consist of causal dispositions to contribute to biological processes and manifested in specific circumstances because of the evolutionary history of the systems in which the molecule interacts. The proposed view has the benefit of being in dialogue with contemporary characterisations of functions while considering the peculiarity of biochemical functions, which cannot be simply identified as biological or chemical functions.

The structure of the chapter is the following. Section §2 will illustrate the status of the controversy over the characterisation of biochemical functions and the relevance of this debate. Section §3 will consider whether a biological or chemical view of function alone can be enough for biochemical characterisation and exemplify my account with vitamin B12. The choice of this

---

[2] The conclusions are based on the conditional acceptance of genuine attribution of biochemical functions.

case study, in addition to the molecular gene of the previous chapters, allows us to infer more general information about biochemical kinds and functions. As will result from the analysis, none of the theories mentioned seems adequate on its own for biochemical functions. Evolutionary accounts of function are not satisfactory from a semantic and from an ontological point of view (Griffiths 2009; Germain et al. 2014; Garson 2018), while the chemical approach does not capture the specificity of the contribution made within biological processes.[3] Section §4 will present the most sensible way to understand biochemical functions. These functions consist of chemical dispositional properties that causally contribute to selected processes or traits within an organism. Their manifestation conditions are explicated in terms of actual chemical reactions, while the identification of the relevant chemical properties in the considered context should be given in terms of evolutionary history. Moreover, once an account of biochemical functions is put forward, I will also apply it to genetic functions. Section §5 will assess the proposal, considering whether it respects some standard requirements for a theory of function. In conclusion, biochemical functions should be understood by considering a causal dispositional view of functions and an evolutionary perspective jointly.

This chapter represents a bridge between Part II and Part III as it discusses biochemical functions, a feature of the molecular gene and biochemical processes and molecules. Together with having implications for a better understanding of the molecular gene, the present account will then allow us to explore the relations between structure and functions in relation to the unity of science [Chapter 6] and biochemical kinds [Chapter 7].

## 5.2 Biochemical kinds and biochemical functions

Biochemical kinds are characterised in terms of structure and function: they are some macromolecules that play some role or function within given contexts. A structural characterisation of

---

[3] I refer to "evolutionary accounts" of function to mean both backwards-looking and forward-looking accounts of biological functions. I count as evolutionary all of those accounts that make a reference to maximising Darwinian fitness of those bearing the traits with the relevant function either in the past or in the future. Accordingly, also Bigelow, Pargetter's account counts as evolutionary as they make explicit reference to maximising Darwinian fitness in the future (1989, 189-191).

a biochemical kind will allow us to identify its chemical structure and composition (Slater 2009), while the biological side is given by the functional aspect of these molecules. Nevertheless, what biochemical functions are is open to debate. In the philosophical literature on biochemical kinds, functions are presented in relation to *"biological functions"*, while not expanding on which account of biological function is relevant and what the implications of such attribution could be. For instance, in his paper about monism or pluralism concerning macromolecules, Slater writes that we can read these functions within the framework of the biological theories of function and with a holistic enterprise, even if he does not further specify how this could be done (2009, 858-859). A similar suggestion is presented by Bartol when he argues for the duality of biochemical kinds (2016). He considers these kinds as presenting "two joints": chemical in terms of structure and biological in terms of function. This is supported by the fact that, for Bartol, biochemical functions are *biological functions* and so should be interpreted within a historical and evolutionary framework (2016).[4] However, the historical and evolutionary characterisation does not pertain to the chemical, so biochemical kinds present a dual nature (according to Bartol). The interplay (and potential tension) between the chemical structural nature of biochemical kinds and the biological functional one is also present in Havstad (2018). Instead of contrasting the two approaches, she underlines that protein characterisation is complex and involves the consideration of microstructural properties, biological functional properties and etiological ones. She reports that "somehow, scientists are routinely and consistently individuating proteins according to some combination of their microstructural and aetiological properties" (2016, 739). Finally, in his work on biochemical kinds, also Tahko discusses functions in the biochemical context as *biological evolutionary functions* that can be interpreted either with the etiological theory or the with goal-directed dispositions one (2020, 2021). While his account considers functions as dispositional properties, biochemical functions are still identified as biological ones.

An intense debate concerning what biochemical functions are and which account of functions is relevant at the biochemical level has been present also in the philosophy of genetics. Specifically, the publication of the ENCODE results has opened some questions concerning which

---

[4] This hypothesis is presented, even if not discussed in detail, also in Slater (2009), and Tahko (2020).

account of function should be used to determine genetic functions. A result of the ENCODE project is that the 80% of the human genome has been assigned a *biochemical function* (as discussed by Griffiths (2009) and Germain et al. (2014)). However, evolutionary biologists do not accept this attribution as genuine and claim that a too broad and permissive notion of function has been used (Doolittle 2013). They argue that a selected effect view of function is the one that allows us to identify relevant functions, and this has not been respected by the attribution of biochemical functions at the genetic level. Such improper usage, some claim, compromises the informativeness and efficaciousness of the notion of function (Doolittle 2013). However, the success of the ENCODE project challenges this claim and opens questions on the relevant sense of biochemical functions or functions at the molecular level and whether a biological account of function is appropriate for biochemical and molecular analysis.

Leaving aside the concerns over what properties are relevant for kinds to be biochemical for the time being [this will be considered in Chapter 7], I will focus here on what biochemical functions are and how their attribution works, considering how little is said explicitly on this in the literature. The answer to these questions is important because of the role biochemical functions have in the nature and characterisation of biochemical kinds.[5]

### 5.2.1 Vitamin B12 and its function

In order to evaluate how to interpret biochemical functions, granted that we accept their ascription as legitimate, I will focus on vitamins' function, not yet analysed in the philosophical literature.[6] This case is helpful because vitamins are complex chemical compounds characterised in terms of their chemical composition and the *metabolic function* they play within the organisms

---

[5] For instance, if we want to discuss functional monism or the prevalence of structural characterisations in biochemistry, as in Tobin (2010) and Goodwin (2011), the importance of individuating such functions is evident.

[6] As pointed out by one of the anonymous reviewers for the BJPS version of this chapter, one could legitimately ask why not to focus on a simpler case study, such as oxygen O2. The choice of vitamin B12 is based on the fact that its complexity makes it a good case to characterise the biochemical properties. Specifically, vitamin B12's functions represent a case of multiple realisation that allows us to appreciate the important functional characterisation of its definition and identification: the different vitamers compounds are classified together because they have the same function. Furthermore, given the interest in considering etiological characterisation, this case study has features relevant to the purposes of this chapter.

(as Combs 2012, 69). Many of them come in *vitamers or vitamin families*: different compounds that exhibit the same biochemical function. Often, vitamers are classified under the same kind, given that they share the same function and some common chemical composition and structure. Specifically, I will focus on the case of vitamin B12 because the ascription of function to these vitamins plays an important role in characterising what these kinds are.[7]

Vitamin B12 can be defined as "the generic descriptor for all corrinoids (compounds containing the cobalt-centered corrin nucleus) exhibiting qualitatively the biological activity of cyanocobalamin" and it comes in four vitamers: cyanocobalamin, methylcobalamin, hydroxycobalaim, adenosylcobalamin (Combs 2012, 377; Fang et al. 2017). The handbook *The Vitamins: Fundamental Aspects in Nutrition and Health* identifies the "biochemical functions of vitamin B12 as a coenzyme in the metabolism of propionate and the biosynthesis of methionine" (Combs 2012, 377). B12 vitamers are all characterised by a cobalt-corrin complex and by having a coenzyme function in humans for various biochemical processes such as hematopoiesis, DNA and RNA production, neural metabolism, and carbohydrate, fat, and protein metabolism.[8] These chemical compounds are classified under the same category, B12 vitamin, because they display a combination of stable microstructure, a cobalt-corrin complex, and physiological functions.[9] This might represent an instance of multiple constitution of the kind B12, where this kind is constituted by different chemical compounds that share some functional properties (Kistler 2018).[10] For the purposes of this chapter, the most relevant implication of the multiple constitution of B12 vitamin is that different cobalamin compounds are classified as B12 also because of the function they play in organisms. The chemical properties, as being a cobalt-corrin complex, are

---

[7] Function ascription in the biochemistry of vitamins is common. Another example is the ascription of function to vitamin C: "Vitamin C is actively accumulated into the epidermal and dermal cells via the two sodium-dependent vitamin C transporter (SVCT) isoforms 1 and 2 [27], suggesting that the vitamin has crucial *functions* within the skin." (emph. added Carr, Maggini 2017). More on functions and vitamins in Combs (2012).

[8] Reference for chemical structure and function of vitamin B12 `https://pubchem.ncbi.nlm.nih.gov/compound/Cobalamin`. Also chapter 17 "B12 Vitamin" in Combs' *The Vitamins : Fundamental Aspects in Nutrition and Health* (2012).

[9] The chapter does not aim to provide an analysis of biochemical kinds and of which properties are necessary and sufficient for a biochemical kind to be considered as such. Rather, it is relevant to consider that a functional characterisation of such kind is often present, combined with a microstructural one.

[10] In Kistler, a kind is multiply constituted when there are two or more microscopic structures that obtain it (2018, 18). See also Gillett (2013).

not enough to classify B12 as such and classification practices consider the functional component as well, such as the role they have in different physiological processes.

Moreover, B12 vitamins are considered an essential nutrient because they play a variety of different important functions in human physiology. However, an interesting feature of vitamin B12 is that its biosynthesis is confined to a few prokaryotic species, making its absorption possible only via nutrition (Combs 2012; Rowley, Kendall 2019). This implies that despite the necessary role that B12 vitamins have in human physiology, they are not synthesised or produced directly by humans but need to be introduced via dietary intake. Among the important functions of B12, one that has captured the attention of scientists is the function of vitamin B12 to contribute to the proliferation of erythroblasts during their differentiation (Koury, Ponka 2004). But what does it mean that B12 has the function to contribute to hematopoiesis processes? In order to explore these issues, we need to understand what biochemical functions are.

## 5.3 Biological functions and chemical functions

Some suggest that biochemical functions are biological functions as those conserved in evolution or those that will be conserved in evolution (Bartol 2016; Havstad 2018; Tahko 2020, 2021). However, since biochemical kinds instances are complex chemical molecules or compounds, functional attribution within an evolutionary framework might not be as straightforward as it seems.[11] In this section, I will explore whether we should interpret biochemical functions as biological evolutionary functions. Having shown that such attribution is not straightforward, I will consider whether they can be considered chemical functions. As I will argue, neither the assumed evolutionary biological account nor the chemical one can be deemed satisfactory for biochemical functions.

---

[11] Slater (2009) points out the difficulties of defining biochemical functions in an evolutionary sense.

### 5.3.1 Biological evolutionary functions

The philosophical literature on biochemical function often refers to "biological theories" of function when considering the functional nature of biochemical molecules, specifically those that consider an evolutionary perspective on the topic (Bartol 2016; Havstad 2018; Tahko 2020, 2021).[12] However, we can explore whether the role of biochemical functions can be interpreted or identified evolutionarily.[13] The discussion about biological functions in an evolutionary framework identifies three broad possible approaches: etiological theories, the theory of biological advantage and the theory of generalised selected effect. Here, I will explore if functional attribution to biochemical molecules can work following any of the theories mentioned above and assuming that function ascription to B12 vitamins is genuine.

First, let us start by considering etiological theories of function.[14] What is referred to as the "strong etiological theory" states that a trait's function is a difference-maker effect of such trait that contributed to the fitness of the ancestors and has been retained via natural selection because of these benefits (Neander 1991; Davini 2021). A trait can be defined as "any detectable phenotypic property of an organism" (Valles 2013) and so a biological function according to this theory is the function of a phenotypic property of the organism. However, B12 vitamins are not properly phenotypic properties of human beings but rather chemical compounds that can only be absorbed via nutrition. B12 cannot be considered traits of organisms if we follow this definition of traits. Accordingly, the very identity of B12 vitamins rules out the strong etiological theory so formulated in the first place, given that this theory assigns functions to traits and B12 are not traits. Let me also consider Mitchell's definition (1993, 2003), for which a function is what

---

[12] Bartol explicitly mentions evolutionary theories of function about the function of proteins, also suggesting that proteins could be considered historical entities (2016). Also, Tahko (2020) discusses two evolutionary approaches to biological functions, the backwards-looking and the forward-looking one.

[13] The literature on biological functions also presents another approach, the causal theory of function. This will be discussed with the chemical view of functions and in section §4.

[14] It is possible to distinguish between the strong and the weak etiological theory. In this section, I discuss the strong etiological theory, but a similar argument can be made for the weak version. This theory states that a trait T has the function of producing an effect of type E in an organism O only if T satisfies the following: i) T must have contributed to the fitness of O's ancestors by producing effects of type E; ii) T must be hereditary (Buller 1998). However, in humans, B12 vitamins are not hereditary traits nor are produced by hereditary traits. As a consequence, also this theory would not accept B12 functions as such.

allowed its bearer to be produced or reproduced in the ancestors of the population bearing that trait. B12 vitamins are bio-synthesised in nature only by prokaryotic single-celled organisms, as some bacteria and archaea. These organisms synthesise cobalamin compounds because they have cobalamin-dependent enzymes that need them as cofactors for various biological processes[15]. In these prokaryotic cells, the retention and production of B12 can be accounted for by natural selection, and so the action of B12 can be considered a proper function, according to Mitchell's account. However, things are different if we consider the function that B12 vitamins play within human (or animal) hematopoiesis. Given their chemical nature as compounds and the fact that they can only be introduced in the body by nutrition, they can hardly be considered produced or reproduced as a direct result of the selection process *in virtue of* their hematopoietic role. Indeed, the reasons why mechanisms of production of B12 in prokaryotic organisms have been selected are not directly linked to the hematopoietic function B12 vitamins play in human beings.

If B12 do have functions, the etiological theories are not the right place to look into. This does not come as an attack on the theories themselves though. Probably, the proponents of the etiological theories would not ascribe the relevant function to B12 themselves, but rather to the biochemical processes in which B12 participate and for which B12 might even be necessary. Or they would not ascribe biological functions to biochemical molecules. Nevertheless, if we operate under the assumption that we are genuinely ascribing functions to biochemical molecules and compounds and that biochemical functions are evolutionary, then these theories cannot provide what is needed.

Another theory of function within the evolutionary framework is the theory of biological advantage, according to which a function is a disposition to maximise the organisms' fitness (Bigelow, Pargetter 1987; Wouters 2003). Also, this theory fails to capture B12 functions. B12 vitamins are not traits of an organism that can maximise an organism's reproductive fitness in the future. And, even more, it seems that B12 vitamins would still be difference makers in hematopoiesis even if this did not maximise the organism's fitness. This scenario is possible in cases of dramatic environmental changes or, more simply, in cases where there are anomalies in

---

[15] "Among the most prominent vitamin B12-dependent enzymes in bacteria and archaea are the methionine synthase isozyme MetH from enteric bacteria" (Rodionov et al. 2003)

the quantity of B12 assumed. In the lack or excess of B12, these macromolecules would still play a role within hematopoiesis, even if this might have no impact on fitness or even a detrimental one. The difference vitamins are making is not interpreted by reference to maximising fitness in the future, and this rules out the theory of biological advantage.

Lastly, there is the theory of generalised selected effect. This theory does not rely on evolutionary selection mechanisms but simply claims that a function is what allows its bearer to be retained within a process of selection. Specifically, “the function of a trait consists in the activity that contributed to its bearer’s differential reproduction, or differential retention, within a population” (Garson 2017, 534). If we allow B12 vitamins to have undergone a process of selection, as general retention compared to other macromolecules or compounds, then we might be tempted to accept the generalised selected effect theory. However, there is one main criticism that plagues the position: it allows too many things to count as function, it is too liberal. An example is the hard-rocks on the beach case proposed by Kingsbury (2008). Let me consider a rocky beach at the seaside. Due to hydraulic action and abrasion by water, a rock survives on the beach longer if it is harder because the environmental conditions will cause the lighter rocks to become sand sooner. In this case, we have a process of selection, erosion done by the sea, and a series of objects surviving selection, the hard-rocks. Following the generalised theory of selected effects, hard rocks have the function “hardness” that allows them to survive erosion. But is hardness *a function* of rocks? It seems not. To deal with these concerns, Garson adds the requirement that the entities being selected need to be part of a population, where a population has members that “must engage in fitness-relevant interactions, whether competitive or cooperative” (Garson 2017, 536; 2019b, 42) and rocks are not part of a *population*. This implies that rocks cannot bear a function in principle because they are mere aggregates, and the relative success of one rock does not depend on the chances of persistence of the others (Garson 2017, 537). What about B12 vitamins? Taken as chemical compounds themselves, B12 vitamins do not seem to constitute a population that engages in fitness-relevant interactions: one B12 compound is not in competition with other ones. This rules out the generalised theory as well. However, it is true that they do interact within organisms that undergo a process of selection and that have been (and are) in fitness competition with each other. This case illustrates the tension concerning biochemical functions that I will elucidate in §4.

In conclusion, while we are intuitively inclined to think that the relevant activity of vitamin B12 is a function (as scientific practice considers it so), none of the aforementioned accounts of function allows us to say as much. Surely, the assimilation and usage of B12 molecules by organisms can be accounted for by evolutionary history, but this does not give us a full story about what the function of B12 is. When we are at the micro-scale, functions do not seem to be identified in terms of evolutionary contributions but rather as answers to the "what-does-it-do questions" (Slater 2009; Garson 2018).

### 5.3.2 Two problems for biological evolutionary functions

The B12 case has allowed me to explore how to interpret (or not) the function ascription to some biochemical molecules if we follow an evolutionary account of functions. Trying to generalise, there are two main problems that we face if we consider biochemical functions from an evolutionary perspective, as identified by Garson (2018). These are the "socio-linguistic argument" and the "ontological argument".

The "socio-linguistic argument" is based on actual scientific and biological practice. This argument says that biologists do not attribute functions referring to selection, specially in "proximal" questions in biology such as in molecular biology, physiology and biochemistry (see also Wouters 2003, 658; Griffiths 2009). In these disciplines, functions of parts and processes are not identified in evolutionary terms, but rather they refer to causal contributions to given processes. As Griffiths writes: "the identification of function in these biosciences seems to be a straightforward experimental matter. Ascriptions of function are confidently made by biologists who take organisms apart and examine their workings but do not test hypotheses about their evolution" (2009, 15). In the case of B12 vitamin, scientists are appealing to the current behaviour and causal effect of these molecules in the production of erythroblasts. Taken at face value, in biochemistry practitioners do not talk in terms of evolutionary functions. This might be a first source of problems in applying an evolutionary theory of functions in biochemistry (Germain et al. 2014, Garson 2018). However, despite the importance of analysing scientific practice when asking ontological questions, this argument cannot be considered conclusive. For instance, it is possible that an historical analysis is implicit or deemed irrelevant for practical reasons. And even if some biologists are committed to an evolutionary theory of function, they might still find more useful for their purposes to consider function as current behaviour.

This argument can then be complemented with an "ontological argument" that focuses on the kind of entities that can or cannot bear an evolutionary function. For instance, functions can be attributed, as in the biochemical case, to items that do not take part in the evolutionary selection mechanism or do not have the relevant history. Accordingly, the evolutionary approaches to functions discussed are not appropriate to characterise biochemical functions because they wouldn't admit as functions those that are considered as such. This appears to be the case of B12 vitamins. It is at least debatable whether they had the *relevant* history of selection in virtue of their hematopoietic function in humans. First, they are not traits because they are not properly phenotypic properties of human beings (Valles 2013). Second, their production is not a direct effect of evolutionary selection processes because of the role they play in human beings, but rather because they work as co-enzyme in some prokaryotes. Lastly, B12 vitamins do not constitute a population because, as non self-reproducing chemical compounds, they do not interact in a competitive or cooperative manner. An evolutionary explanation of the mechanisms synthesising B12 in prokaryotes can be offered, however these are not related to the effects that these molecules have in human hematopoiesis.

### 5.3.3 Chemical functions

Despite functions in biochemistry often being associated with the biological component of their nature, there is a form of function talk also within organic chemistry. Here, I will explore whether chemical functions can help us understand biochemical ones.

Chemical functions are related to functional classifications of compounds in terms of functional groups. Functional groups are groups of atoms or bonds in a molecule that are responsible for some typical chemical reactions. Also, the presence of the same functional group is informative about the possible reactions that different compounds can undergo. These groups are called functional because they provide information about the specific role of compounds within a chemical reaction and the contributions provided to particular processes. Chemical functions, so interpreted, are then contributions to chemical reactions. This is in line with the causal theory of function, for which a function is the causal role of an entity within a performed activity in a complex system or "what an item does or is capable of doing" in a complex system (Wouters 2003, 636; also Cummins 1975; Sterelny, Griffiths 1992; Germain et al. 2014). According to this theory, the ascription of a function amounts specifically to the identification of a property or a set of causal properties that contribute to given phenomena. This property or this special set of properties can be understood in terms of dispositional causal powers that are manifested within the right set of conditions and the right processes (Cummins 1975; Tahko 2020). Chemical functions are then dispositional properties that allow the compounds to contribute to given reactions. Moreover, the strict relation between functional groups and structural features, specifically the presence of specific atoms or bonds, seems to make them reducible: chemical functions correspond to a specific structure within the compound molecule. This analysis gives us a sense of chemical function, as the role that a given compound can play in virtue of the presence of a well-identifiable functional group.

Let us now consider the B12 case and biochemical functions. Can such a view of function be adequate? The attribution of these functions is often based on relevant activities of the entities under consideration, i.e. "those likely to make a relevant difference" for a given phenomenon (Germain et al. 2014, 817). This is close to the discussed above view of function. When we ascribe a function to cobalamin molecules, classifying them as B12 vitamins, we associate them with a special set of causal properties (Griffiths 2009). These causal properties are interpreted as dispositions that can be manifested or not in different circumstances (Tahko 2021). Accordingly, the disposition is explicated in terms of specific chemical bonds between the molecules in the cobalamin compound and other molecules involved in erythropoiesis, and its manifestation can be expressed in terms of specific chemical reactions and given environmental conditions. As in §2.1, B12 vitamins have the disposition to contribute to the proliferation of erythroblasts during their differentiation (Koury, Ponka 2004). This happens because vitamin B12 acts as a coenzyme in the reaction involved in regenerating methionine, as required in normal erythropoiesis. Specifically, methylcobalamin (a vitamer of B12) takes part in "the transfer of a methyl group from 5-methyl-THF to homocysteine via methylcobalamin, thereby regenerating methionine" (Koury, Ponka 2004, 109). In this sense, when we say that B12 vitamin has an hematopoietic function in humans we are referring to its dispositional property to react in a specific way during the regeneration of methionine required in erythropoiesis. Furthermore, this theory is compatible with the semantic and ontological considerations presented earlier in the chapter. First, it seems that scientists identify and classify biochemical functions in terms of causal contributions to a given environment and to given phenomena. In these "micro-contexts", evolutionary considerations are often not present, while the consideration of the causal action of these molecules is. Second, it satisfies the ontological requirement as well. Compounds, such as vitamins, can have chemical dispositional properties, so we are ascribing to them properties that are apt for their kind. Moreover, we can also provide a detailed story of the chemical reactions happening for the manifestation of such properties. Thus, we might be content in saying that we have found a functional framework for biochemical functions and their ascription.

Nevertheless, some dissatisfaction can still be present: following Bartol's suggestion (2016), biochemical kinds seem to have a dual nature, a chemical and a biological one, and this affects the nature of biochemical functions as well. This is what makes them so interesting. At face value, biochemical kinds are not just chemical compounds but those that have *specific functions within biological systems*, while a function is not taken as a definitional component of every chemical kind. Moreover, a mere chemical view of biochemical functions might be too restrictive. Biochemical functions refer not only to the different functional groups the compounds have or to the presence of specific atoms and bonds, but also to a role within a given biological process or system. The hematopoietic function of B12 does not simply refer to the presence of a particular functional group or to the contribution to chemical processes or reactions, but to those chemical reactions that play a role within biological processes.[16] *Prima facie*, this biochemical function

---

[16] It is also important to keep in mind that this function is multiply realisable.

ascription refers to a relevant activity or role of the molecule as one that is "likely to make a relevant difference" for the given biological process under consideration (Germain et al. 2014, 817): B12 "makes a difference" for hematopoiesis. However, this difference-making is not within a chemical reaction only but within those chemical reactions that operate in a biological process. B12 is not only a coenzyme in the reaction mentioned but also a vitamin and essential nutrient because of the role played in human physiology. The biological environment in which biochemical kinds operate provides a specific context to the phenomenon in consideration. This has made some philosophers, among which Slater (2009), Bartol (2016), Havstad (2018) and Tahko (2020, 2021), to interpret biochemical functions as something more than chemical functions, and often they have seen them within an evolutionary or historical framework.

In the next section, I will argue that we can still retain biochemical functions as functions by proposing a combined account that embraces the dual nature of such kinds, the chemical and the biological one. The proposal is based on a causal account of functions *à la* Cummins (1975), as similar to the one used in cases of chemical functions, combined with an evolutionary perspective.

## 5.4 Biochemical functions: a unified approach

In this section, I will first present an account of biochemical functions, exemplifying it via vitamin B12. Then, I will apply it to the molecular gene.

Although the chemical view of function is not adequate as it is, a causal theory of function still remains a valid starting point in our understanding of biochemical functions (Cummins 1975; Wouters 2003; Griffiths 2009; Germain et al. 2014). Together with the advantages mentioned above, this theory is invoked for cases like the ones considered here or for cases of "proximal biology" (as in Wouters 2003). A causal consideration of function is present in the mechanistic view and seems to fit the case of molecular biology or biochemistry.[17] For instance, Griffiths (2009,

---

[17] As suggested by an anonymous reviewer for the BJPS version of this chapter, the causal view of function has also been discussed in the mechanistic philosophy of science. According to these approaches, a component of a mechanism has a function when it makes a causal contribution to a system-level behaviour. This approach has been applied to biological functions as well and is in line with the view defended here.

15) and Germain et al. (2014) state that genetic functions do not refer to evolutionary functions but rather to causal functions within biological systems. Here, I elaborate on their suggestion by extending it to other biochemical cases and combining it with an evolutionary perspective. To do so, my analysis will be in two steps. First, I will endorse a specified version of the causal analysis which states that a biochemical function is a dispositional causal property of an entity involved in biological processes, that is those involved in the life of an organism and that have been evolutionary selected (Wouters 2003). The view defended operates in the broad framework of the causal account of function, with the addition of the context and the details relevant to the biochemical context. Second, I will unify this account with an evolutionary approach. The proposed view has some novelty because it considers an application of the causal account of function to the biochemical domain precisely and dialogues with the previous literature on biochemical kinds. Moreover, it combines it with evolutionary considerations that allow the identification of those causal contributions relevant to the biological processes considered.

In the study of life, biologists divide the organism into different systems and subsystems that have specific capacities in maintaining the general phenomenon of life. In each of these systems and subsystems, the parts play a causal role, which allows them to have functions. Such functions are causal dispositions, but, as stressed by Cummins (1975) and Germain et al. (2014), functions as dispositions are always contributions to something in a given context, and this context is the phenomenon of life of an organism. Consequently, the relevant biochemical dispositions are those that causally contribute in *biological processes* and not any causal dispositional property of a given chemical molecule or compound. They are dispositions to bring out a specific effect within physiological processes, and their causal contribution can happen independently from its being beneficial for the final survival or fitness of the organism.

The context of life makes it necessary to address a second point: why these molecules contribute to these specific processes in the way they do (Garson 2018, 2019). An answer can be

---

However, this approach is conditional on a mechanistic view, a commitment that goes beyond the scope of the chapter. Moreover, the account defended here gives a stronger role to the evolutionary component that we need to consider for biochemical functions. Further references are Cummins (1975), and Craver (2001).

given considering a modified etiological account of functions.[18] In a nutshell, some entities play a specific causal role within an organism because they interact with parts or traits that have been directly selected. In the biochemical case, some specific molecules or compounds manifest some specific causal powers because they contribute to processes involving traits or components of an organism that have been evolutionarily selected in the ancestors. These functions are not straightforwardly evolutionary, as chemical molecules such as B12 do not have the right characteristics to be deemed as properly selected nor to be a trait or parts of a biological population themselves. Nevertheless, their action is within organisms that have undergone a process of selection and we can still consider an evolutionary perspective about them. This is because a given causal disposition counts as a function if the traits or processes that interact with it have been selected in the past when going into specific physiological processes (and not others) that involve this molecule. With Garson (2019) and Griffiths (2009), I agree that there are no entirely ahistorical functions within biological organisms and in the life sciences: the context in which biochemical functions work is embedded in evolution. For functional attribution, one has to consider the evolutionary history of the organism under consideration.

This allows identifying not only the manifestation conditions of the dispositions (as in a standard causal theory of function) but also the relevant and adequate chemical dispositional properties that contribute to the processes within the biological organisms. Accordingly, the consideration of the evolutionary history of the processes to which the biochemical function contributes permits having an account that selects only those contributions that are appropriate for the processes [Figure 6]. From this, two main benefits follow i) it allows us to appreciate how, in this account, functions can be relational, ii) it allows the identification of those causal contributions in a non-arbitrary way (*contra* Tahko 2020).[19] First, different functions can be played by the same biochemical kind in different contexts and accordingly, the functions are

---

[18] A further reference can be Kitcher's indirect theory of selection as presented by Wouters: "Functions are indirectly linked to selection if it is assumed that the function helps the organism to respond to selection pressures, but if it is left open whether or not the presence of the functional item is completely explicable in terms of selection" (2003, 655).

[19] I thank an anonymous reviewer for the BJPS version of this chapter for pointing out this issue.

relative to given organisms.[20] The relativity of functions of biochemical kinds has been pointed out by Tahko (2021, 33) as well and seems to be a feature characterising these phenomena. The different evolutionary histories of the different processes identify the relevant chemical dispositions. Moreover, considering the different evolutionary histories of the relevant biological processes to which the functions contribute allows us to answer a potential concern presented by Tahko (2020). Specifically, he is asking whether the various biochemical functions and, accordingly, the relevant chemical dispositions, are just pragmatically selected according to the discipline that is inquiring into the function. If we consider the evolutionary history, then the individuation of the relevant dispositional properties is not discipline relative but relative to the different evolutionary histories.

In conclusion, an adequate view of biochemical functions needs to consider both the causal dispositional account to understand what these functions are and the evolutionary perspective to identify the adequate set of dispositional properties. This perspective is also important because it allows distinguishing biochemical kinds with positive biochemical functions from poisons or other dispositional properties that can be manifested within biological processes.[21]

Let me recall the B12 vitamin case. These vitamins have a series of functional dispositional properties. Such dispositional properties are manifested thanks to identifiable and specific biochemical reactions and the function, as explored in the previous section, is not directly evolutionary. Nevertheless, the hematopoietic function operated by B12 might be considered indirectly evolutionary or indirectly selected. Even if B12 is not selected for its hematopoietic role in humans, some traits interacting with them have been evolutionarily selected for their efficiency

---

[20] The same biochemical molecule can have different biochemical functions in different biochemical processes and/or in a different organism. These phenomena of multiple determination have been discussed in the literature by Tobin 2010, Goodwin 2011, Bartol 2016 and Tahko 2020.

[21] Specifically, we can identify poisons and distinguish them from vitamins or nutrients because the chemical dispositional properties of the first one do not interact with the process in the way it has been selected for. We can still allow poisons to have a biochemical function as far as a given chemical causal interaction with the organism is possible and happening, but then we might not deem them as nutrients or vitamins but as poisons. While both poisons and biochemical kinds can have biochemical functions, whether the causal chemical contributions are beneficial or not can be established only within an evolutionary framework. Moreover, we can include metabolic responses to poisons as biological processes to which biochemical kinds causally contribute. These processes have been selected in the past to interact in specific ways with the chemical molecules - that is to mitigate their poisoning effect. I thank an anonymous reviewer for the BJPS version of this chapter for suggesting the discussion on poisons.

within humans' ancestors. Accordingly, the biochemical function of B12 is a set of dispositions to bring in a specific causal effect in the processes of hematopoiesis and erythroblasts proliferation that *are a result of the process of selection*. An approach that considers the evolutionary history permits the ascription of different functions to B12 vitamins in different organisms, as the same set of cobalamin compounds can have different dispositional properties that are manifested in different circumstances[22]. In bacteria, B12 manifests the set of dispositional properties relevant to a specific co-enzymatic action, while in human beings, it manifests the set of dispositional properties relevant to other phenomena, such as erythropoiesis. The identification of these dispositional properties is not interest-relative, but rather is given by the different evolution of the processes in humans and bacteria to which vitamin B12 contributes [Figure 6].

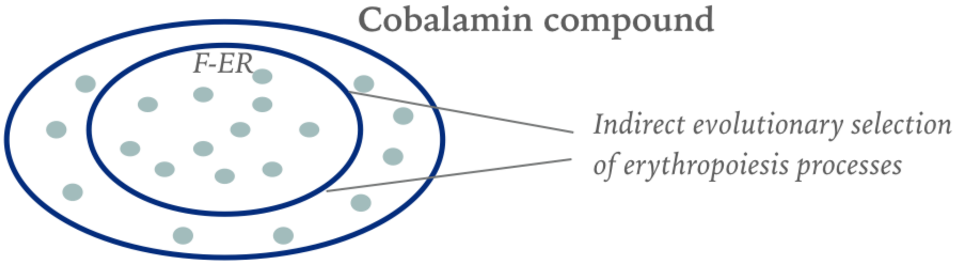

Figure 6. The evolutionary selection of the relevant dispositional properties or chemical powers for biochemical functions. In the example, F-ER is the function to contribute to erythropoiesis for vitamin B12.

In conclusion, to understand biochemical functions, we need to consider jointly a causal dispositional view of functions and an evolutionary perspective.[23] This allows us to formulate the following definition of biochemical function:

---

[22] This can be accepted if biochemical kind ascription is given in terms of both composition and a set of functional properties, as the action of B12 compounds as co-enzymes. However, this chapter does not consider biochemical kind ascription itself.

[23] As suggested by an anonymous reviewer for the BJPS version of this chapter, this account is different from forms of pluralism in the literature. Pluralism usually refers to the usage of different accounts of function in different contexts (Allen, Neal 2020). In this case, instead, I am defending a view for which, if we want to understand the nature of biochemical functions properly, we need both approaches combined.

**BC-function**: A biochemical function consists in a specific set of chemical dispositional properties (or chemical powers) that manifest in specific circumstances because of the evolutionary history of the traits or processes in which the biochemical molecule is involved and, as such, the relevant chemical powers are indirectly evolutionary selected.[24]

Moreover, the manifestation of these dispositional properties can be identified with specific chemical reactions. The reason why such reactions happen in the way they do can be provided via the evolutionary history of the traits or features involved in the relevant subsystems. This approach captures the specificity of biochemical functions that are both chemical and biological and is explanatory adequate. This has some advantages compared to the two views in isolation. On the one hand, the causal account on its own does not identify the dispositional properties relative to the specific processes or might do that in a discipline-relative way (as pointed out by Tahko 2020). On the other, a biological evolutionary account on its own is not adequate for the reasons mentioned in the previous sections. Moreover, the account presented here is compatible with the relativity of functions in terms of causal dispositional properties that are present but manifested in different environments. This allows the same molecules or compounds to have different functions within different biological organisms (as in Tahko 2020, 2021) and to identify these properties in a non-arbitrary way.

### 5.4.1 Genetic functions

Let us go back to genetic functions and the attribution of biochemical functions to molecules. As presented in Chapter 3 and Chapter 4, genes have a functional component that is relevant for both the weak emergence of their instances and their natural kind status. However, the interpretation and the attribution of functions at the genome level have been the object of controversy. Specifically, the ENCODE project has attributed a function to the 80% of the genome, but this is incompatible with an evolutionary account of function (Doolittle 2013). Evolutionary

---

[24] In this thesis, I take powers and dispositional properties to mean causally efficacious properties that manifest themselves in specific circumstances. I am aware of the extensive discussion on the relation between powers and dispositions, but for the present purposes, I take a minimal commitment, and I define them as above.

biologists have pointed out that a form of selection or natural selection is an important part of biological functions, but normally estimates are such that “only the 3-8% of the genome has been undergoing purifying selection” (Germain et al. 2014, 808; also in Doolittle 2013). This would compromise most of the biochemical functions attributions made to genome sequences and genes within the ENCODE project. In this regard, Griffiths does an analysis of the function of TATAAA sequences in binding transcription factors that can allow us to illuminate the issues and controversies, as we did with B12 vitamins (2009). If we take an etiological approach, when scientists ascribe a function to the nucleotides TATAAA they just mean that this “sequence evolved by natural selection because it had the effect of binding transcription factors in ancestral organisms” (Griffiths 2009, 14). Or, following a theory of biological advantage, function ascription means that the role of TATAAA will maximise the reproductive fitness of its bearers. However, following Griffiths (2009), these evolutionary interpretations do not seem straightforwardly apt when considering how the notion of function is used when referring to the micro-scale of biological organisms. In these contexts, it is questionable whether scientists are referring to evolutionary history when they speak of the function of TATAAA or whether the TATAAA counts as an evolved sequence at all. At the micro-scale, functions do not seem to be identified in terms of evolutionary contributions but rather as answers to the “what-does-it-do questions” (Garson 2018). This has motivated Griffiths (2009) and Germain and colleagues (2014) to argue for a causal theory of function when considering genetic phenomena, compared to an evolutionary one.

Nevertheless, it remains true that the TATAAA sequences play their role within selected processes and within organisms that have undergone a selection process. The same can be said for genes. Their function is understood as, for instance, a contribution to protein synthesis because they are sequences transcribed into mRNA. Such contribution is not necessarily understood in evolutionary terms (as selected effects) but rather, as the ENCODE analysis shows, as causal contributions to given processes. Nevertheless, the simplicity of DNA sequences, on the one hand, and the complexity of genetic phenomena, on the other, require more than a causal analysis to understand what genetic functions are. The fact that specific sequences are expressed in specific ways during transcription can offer an evolutionary explanation, if not directly, at least indirectly, via the processes in which genes take part [more will be said in Chapter 7].

The application of the account of biochemical functions presented here can help to understand genetic functions within the postgenomic project world. Genes' functions are dispositional properties manifested thanks to identifiable and specific biochemical reactions during transcription. While an evolutionary explanation can be given of why, for instance, protein synthesis happens in different ways considering different organisms. This approach retains the usage of biochemical functions applied within the ENCODE project, as this is close to a causal and dispositional view of function (Griffiths 2009; Germain et al. 2014). Furthermore, this allows the multiple realisability of the genetic functions and the multiple-composition of the genes to be present [Chapter 3]. The same sequence can play different functions because different dispositional properties are existent but manifested in different environments because of the evolutionary history of the process in which they are taking part.

A last important point is worth being mentioned. The fact that genes can be attributed a biochemical function does not exclude the fact that some might also be assigned a biological function. When there is a clearer correspondence between genes and sequences and when such sequences have undergone a clear process of selection, there is a sense in which the gene can have a biological function: the one it has been selected for due to the benefits it had in the past/previous ancestors (as in Doolittle 2013). However, not all genes might be attributed a biological function, and some might just have a biochemical one, where a dispositional functional property contributes to selected processes.

## 5.5 Assessing the view

Having provided a possible way to retain biochemical functions, I will consider in this section whether this proposal is adequate and we can maintain biochemical functions as real functions. To do so, I will first assess if and how the theory satisfies three common desiderata for a theory of function (Nanay 2010; Garson 2017; Davini 2021). Then, I will face a possible objection:

whether the theory allows too many entities to have a biochemical function.

### 5.5.1 Requirements

In the philosophical literature on functions, three requirements have been identified for a theory of function: 1) it should distinguish between function and lucky accident; 2) function ascription has to be explanatory; 3) functions ought to be normative (Nanay 2010; Garson 2017; Davini 2021). Let me consider them in order.

The first requirement states that a function should not be any role that an entity can play but a given specific one. One example can be that the function of the limbs is to walk and not to allow us to wear skirts or trousers. A good theory of function should be able to discriminate the real functional role played by a given entity in a system from spurious or accidental functions. This requirement is satisfied because the theory constrains the function effects of a molecule or compound to those that contribute to the biological processes of an organism. Biochemical functions are these specific dispositions and not any dispositional property of the considered biochemical kind: the organismal context constrains biochemical functions to a specific set of dispositional properties. For instance, B12 vitamins can take part in many chemical reactions without these having a causal impact on the relevant biological processes; while the hematopoietic function of B12 is so because it causally contributes to hematopoiesis, a biological process that has been selected to be so. The evolutionary component of the considered theory constrains all the possible dispositional causal properties to those that are manifested in selected environments and the organisms under consideration. Furthermore, the same biochemical kind can have different functions in different organisms because of the different histories of selection of the processes in which they are involved. The causal dispositional account can help us to understand what biochemical functions are. At the same time, the evolutionary component allows us to identify which of these properties are relevant in different organisms and environments.

Second, ascribing function should be explanatory. The explanatory power of function has two meanings. First, it helps to understand the behaviour or the contribution of a given trait or entity to processes or systems in the organism; second, function ascription can explain why a given trait, entity or feature exists. For instance, the function of sight for eyes explains the behaviour of eyes in a given organism and sight also explains why eyes have been selected.[25] Function ascription to biochemical molecules can be explanatory in the same way within the theory presented here. First, it illuminates some biochemical processes, such as vitamin functioning and vitamin biosynthesis. For instance, the function of B12 vitamins during hematopoiesis is explanatory because it improves the understanding of the proliferation of erythroblasts during their differentiation (Koury, Ponka 2004). Second, the ascription of functions to B12 allows us to explain why it is synthesised by some prokaryotes and why it is necessary to introduce B12 in the human diet. B12 is produced by some bacteria and archaea because the dispositional properties of B12 played a causal role as co-enzymes in the ancestors of these populations. Moreover, the theory presented allows explaining why B12 can be considered an essential nutrient: it has dispositional causal properties that play a role in some selected processes and are manifested in selected processes.

Let me now move to the last requirement, the normativity of function. This represents a more complicated issue, as normativity means that “it is possible for something to have a function that it cannot perform” (Garson 2017, 1111). If the theory of function is normative, then the theory can account for a trait or entity to have a function that it is not capable of performing (see also Nanay 2010). This condition is normally explicated with a theory of dysfunction, for which a trait is dysfunctioning when the trait fails to perform the role for which it has been retained in evolutionary processes. As Garson (2017) points out, it is more difficult to account for a theory of dysfunction if one accepts a causal theory of function, as this theory considers a function just any causal contribution to the process under consideration. What can we say about the amended theory defended here?

The first answer comes from the fact that it is not appropriate to ask how to make sense of biochemical “dysfunctioning”. Biochemical molecules or compounds do not present what is normally called “dysfunction”, as they are not proper traits of an organism, but rather entities

---

[25] This is normally done with an etiological theory of function, as sight is the function of eyes because it is the behaviour of eyes in the ancestors of the organisms under consideration and eyes exist because sight caused them to be selected.

that contribute to biological processes, whether this is beneficial or not.[26] The specific dispositional properties are manifested whenever the chemical conditions are adequate. This can also happen when the action is not (or was not) beneficial for the organism. Thus, the scientific community speaks of *biochemical alterations* rather than dysfunction and they consist in chemical alterations of the molecules or when there is a lack or an excess of them. This does not exclude forms of dysfunctioning when considering the biological processes impacted by the biochemical alterations. However, such dysfunctioning does not lie at the level of the biochemical alterations, but at the level of the processes, cells or organs where the molecules act. Specifically, biochemical alterations can lead to the dysfunction of the relevant biological processes in a "standard" evolutionary biological sense: the processes (or relevant biological entity) do not perform the role for which they have been selected in the past. For instance, we can have a dysfunction of erythroblasts due to an alteration, such as lack, of B12 vitamins (Koury, Ponka 2004), but B12 vitamins do not display a dysfunction themselves.[27]

Nevertheless, recalling the dispositional and evolutionary components of the presented theory of function, the theory still allows for a form of normativity. There are cases in which the molecule or compound can have a function without manifesting or performing it in three ways. First, the manifestation conditions might be challenged in a way that the dispositional property is not manifested. For instance, there might be changes in the environment that prevent B12 from contributing to erythroblasts proliferation. Second, the trait or the cells with which the biochemical molecule or compound normally interacts can be damaged or are dysfunctioning for

---

[26] Relevant to this see the debate on finks and masks of dispositions as in Bird, Handfield 2007; Choi 2012.

[27] As pointed out by an anonymous reviewer for the BJPS version of this chapter, misfolded proteins represent an interesting case of biochemical alterations. Misfolded proteins are those folded in a non-standard structure and are associated with different diseases, such as degenerative ones (Reynaud, 2010). In these cases, different chemical configurations can lead to the manifestation of different chemical dispositions and impact the processes to which the molecules contribute. The chemical contribution remains relevant to processes in a causal sense, and the possibility of such interaction is given by evolutionary history, but the process is then dysfunctioning in that it does not perform what it has been selected for. Specifically, in the case of misfolded proteins and degenerative diseases, it is the accumulation of such proteins (so an excess in quantity) that can lead to a detrimental chemical contribution to the relevant cellular process and prevent the functioning (of the process) in the way it has been evolutionary selected for. In this case, the "dysfunction" is not of the protein *per se*, but of the process to which the proteins contribute due to an alteration in quantity of proteins.

unrelated reasons. This latter dysfunction is understood within an etiological theory of function, where the trait or cell is not performing the activities that were beneficial in the ancestors of the relevant population. Third, there might be some structural damage to the chemical structure of the macromolecule that prevents the manifestation of the dispositional functional property.

In conclusion, the theory of functions presented here can allow for normativity thanks to the combination of a causal theory of function together with an evolutionary approach. This latter etiological approach allows us to constrain the possible dispositional causal properties of a given biochemical kind to the only relevant ones for a given organism. It also gives the context for the manifestation or not of these properties. All things considered, we can retain this approach to biochemical functions as a valid account of functions.

### 5.5.2 Biochemical functions and biochemical kinds

Let me now move to consider a possible objection: the theory allows too much. Here we have defined biochemical function as a specific set of dispositional properties that manifest themselves in particular circumstances because of the evolutionary history of the traits or processes in which the biochemical molecule is involved. These dispositional properties are those that causally contribute to biological processes, where the causal contribution is neutral as it can be beneficial or not. However, this position risks allowing for too many biochemical functions as "causal contribution to biological processes" can imply that many different kinds have a biochemical function. For instance, zinc (Zn) is a chemical element, and thus a chemical kind, with atomic number 30 and a series of properties that have little to do with biological properties, such as being a brittle metal at room temperature. This element, however, is very important for many different biological processes. Zinc is defined as "an essential trace element" necessary for "the normal growth and the reproduction of all higher plants and animals, including humans. In addition, it plays a key role during physiological growth and fulfils an immune function. It is vital for the functionality of more than 300 enzymes, for the stabilization of DNA, and for gene expression" (Frassinetti et al. 2006). When considered in the biological environment, it seems justified to accept that zinc has biochemical functions because this element has a series of dispositional properties that contribute to biological processes. Furthermore, some traits or processes interact with zinc and have been evolutionarily selected. This makes the function of zinc a biochemical function according to the present account. However, scepticism might rise in accepting zinc as having a biochemical function, as this can imply that any chemical element that satisfies the requirements can be considered a biochemical kind as well.

To answer this objection, I think it is important to distinguish between biochemical functions and biochemical kinds as much as it is possible. Specifically, even if biochemical kinds need to have a biochemical function, having a biochemical function is not enough for a kind to be a biochemical one. As the zinc example shows, also pure chemical kinds can have a biochemical function, even if this does not make zinc necessarily a biochemical kind. The answer to this objection is a "bite the bullet" strategy: many kinds can have a biochemical function if they satisfy the requirements. However, this does not necessarily imply that they are biochemical kinds.[28] The purpose chapter is just to consider biochemical functions and a more detailed theory of biochemical kinds will be presented in Chapter 7. As will be explored, one way to develop an account of biochemical kinds can involve different structural and biological conditions as pointed out by Slater (2009).[29] For instance, a biochemical kind can be a chemical molecule or compound that displays a set of biochemical functions, and its existence is a direct or indirect effect of evolutionary pressure because it is bio-synthesised by organisms that have undergone a process of selection.[30] If we follow this possible definition, then zinc would not be a biochemical kind, even if it has a biochemical function. Alternatively, B12 vitamin would be a biochemical

---

[28] As pointed an anonymous reviewer for the BJPS version of this chapter, we can ask if entities such as colours could have a biochemical function. I think that this would not be the case because a crucial component of the provided definition of biochemical functions is that they are constituted by chemical dispositional properties of chemical molecules or compounds: such chemical dispositions need to be identifiable. The status of colours is metaphysically controversial, but even within a physicalist account of colours, they refer to physical properties rather than chemical ones. Some physical properties, such as electromagnetic radiation, can have a function within biological processes, but this would require a specific analysis that goes beyond the scope of the chapter.

[29] This definition is a possible attempt to identify biochemical kinds and the author does not commit to it as a set of necessary and sufficient conditions. The purpose for this chapter is only to identify some requirements that can allow distinguishing molecules or compounds playing a biochemical function from biochemical kinds.

[30] This does not imply that these kinds are *only* bio-synthesised as a result of selection, as many biochemical kinds can also be synthesised artificially [Chapter 7].

kind with different biochemical functions, because it is bio-synthesised by organisms, bacteria or archaea, and it plays different roles in relation to cells or traits that have been selected in these organisms or others. In conclusion, if some molecules play a role in biological processes we should not be afraid to accept them as having a biochemical function, even if we might not accept them as instances of biochemical kinds.

## 5.6 Conclusion

In this chapter, I have argued that if we accept that function ascription to biochemical molecules is genuine, then we should interpret biochemical function as composed of chemical dispositional properties to causally contribute to biological processes within an organism. A biochemical function consists of a specific set of chemical dispositional properties that are relevant to chemical reactions in physiological processes. These dispositions are manifested in specific conditions because of the evolutionary history of the traits or processes in which the biochemical molecule is involved. This account is of relevance for the discussion on biochemical kinds and the philosophy of biochemistry because it discusses biochemical functions in their peculiarity and makes reference to the positions already present in the literature.

First, while the account presented is not entirely novel, as it is built in dialogue with the already existent debate, the application to the biochemical domain represents a novel contribution. Function ascription in biochemistry is important both in the current philosophical and scientific domain, and an explicit clarification of what this entails is of value. Second, this account stresses the importance of historical evolutionary considerations in identifying which chemical dispositional properties constitute the different biochemical functions. This allows the identification of the relevant functions, which is not easily possible with a simple causal contribution approach that might include all possible chemical contributions of a molecule. Moreover, the combination of a causal and historical approach is in line with the recent work in the literature that has underlined the role of historical theories of function (see Garson 2019). Third, this chapter represents a bridge between Part II and more general metaphysical considerations of biochemical kinds.

A clear understanding of biochemical functions supports the results presented on the molecular gene and will allow us to explore the relation between biochemical functions and chemical structures and biochemical kinds in the next Chapter 6 and Chapter 7.

To reach a clear account of what biochemical functions are, I have first presented the status of the controversy and what biochemical functions are taken to be in the debate. Then, I have considered whether the biochemical functions can be interpreted as either biological functions or chemical functions using the case of vitamin B12. Given the inadequacy of these views on their own, I have proposed a positive account for which we should consider both a causal theory of function together with an evolutionary approach. A causal theory of function offers us an account of what functions are, while an evolutionary perspective answers the question of why they play a particular role within the organism. Moreover, an evolutionary approach aids the identification of which chemical dispositional properties are relevant. The present theory can also be applied to genetic functions, contributing to our understanding of genetic phenomena and the ENCODE controversy. Lastly, I have considered some requirements for which this approach to biochemical functions can be considered a genuine account of functions and a possible objection. The account presented respects the requirements, allowing us to retain biochemical functions as functions. Moreover, I have explored whether the theory might allow too much, and I have concluded that we can answer this concern by separating biochemical kinds from biochemical functions.

In conclusion, biochemical functions seem to be a crucial property for our understanding of the biochemical world and genetics, but little has been said in detail about what biochemical functions amount to. Here, I have tried to fill this gap and I have offered a theory of biochemical function that can be a starting point to explore the connection between the chemical world and the biological one. Together with the already discussed implications of this account, biochemical functions are often seen as a possible source of disunity for the biochemical world. A clear account of biochemical functions is a necessary step for the topic of the next Chapter 6, where I will consider more in detail the relation between structure and function for biochemical kinds.

# Chapter 6

# Biochemical functions and unity in biochemistry

## 6.1 Introduction

This chapter considers the relation between biochemical functions and chemical structure, elaborating further on the themes presented in Chapter 5.[1] The discussion of the relation between biochemical function and chemical structure is relevant to the debate concerning the unity of science, together with being a foundational topic for biochemistry (Santos et al. 2020). Specifically, it explores how the presented notion of biochemical functions can be used to approach questions concerning an ontological sense of unity of science and argue in favour of a form of non-reductive unity in biochemistry. This will be defended by proposing that biochemical functions, often considered a root for disunity, are weakly emergent from the chemical features of the molecule, following the account of weak emergence presented by Wilson (2011, 2015, 2021). Weak emergence allows for retaining a form of ontological autonomy from the chemical basis while at the same time maintaining a form of dependence that saves a form of unity within biochemistry (Wilson 2011, 2015, 2021). Throughout the chapter, the argument is supported by using vitamin B12 as the main case study [Chapter 5].

---

[1] I thank Jessica Wilson for all the discussions and helpful comments on this chapter.

The results of this chapter have various implications. First, one of the main research questions of biochemistry, as presented by the Biochemical Society, is "to understand how the structure of a molecule relates to its function, allowing them to predict how molecules will interact"[2] (as also in Santos et al. 2020). While understanding the precise relation in the various individual case studies is a matter of empirical consideration, the present chapter presents a philosophical and metaphysical framework compatible with these results and offers a more general context to understand the relations between structure and function. Second, this chapter provides a novel application of the account of Wilson's account of weak emergence, enriching the case studies that can fit with her framework and offering new insights into the understanding of weak emergence in not-yet-considered cases. Third, the weak emergence of biochemical functions supports the main thesis of the present research, which is that there is something *sui generis* about biochemical kinds and their functions. Biochemical functions can be seen to preserve a form of autonomy, given that they present a distinct causal profile. Nevertheless, this form of autonomy does not prevent us from maintaining a form of unity in biochemistry [as further discussed in Chapter 8]. In detail, it argues that unity can be maintained first because biochemical functions are materially dependent on chemical features, as will be spelt out throughout the chapter, because of the identity of tokens condition that characterises the defended form of weak emergence. Secondly, unity is achieved because weak emergence is a way to identify the relations that hold between biochemical functions and chemical structure: the two are not disentangled but rather linked and connected.

The structure of the chapter is the following. Section §2 briefly introduces the unity of science debate and its relation to the one concerning biochemical kinds and biochemical functions. Section §3 discusses why biochemical functions are often taken as a source of disunity between the properties held by biochemical kind and summarise the view of biochemical functions presented in the previous chapter. As in Chapter 5, biochemical functions can be seen as sets of chemical dispositional properties that contribute to biological processes. Section §4 explores how, given this definition of biochemical functions, we can interpret the relation between chemical structure

---

[2] From the website of the *Biochemical Society*: `https://biochemistry.org/`.

and biochemical structure via weak emergence. This will also allow exploring how unity can be preserved, given the characterisation of the relation presented. Section §5 presents some challenges and clarificatory remarks. I also relates the account of weak emergence presented here with the one in Chapter 3.

## 6.2 (Dis-)Unity of science

Biochemical kinds represent an interesting case to look at for those philosophers working on the unity or disunity of science. They are so because they seem to be a bridge between the chemical and the biological domain, and, as such, they can inform the debate on the relations between the two (Tahko 2021). Before exploring how unity can be achieved in the biochemical realm, I will briefly recall which sense of unity of science we are concerned with for this chapter [previously discussed in Chapter 1].

As presented in Chapter 1, the idea of the unity of science has been connected with a form of reduction: if we are able to reduce one entity to another, then the two things are unified (Oppenheim, Putnam 1958; Cat 2017; Tahko 2021).[3] Reduction itself is a broad relation that expresses a form of priority of something over something else and can be more or less strong. If x reduces to y, this can mean, for instance, that y is more basic than x, that x fully depends on y or is constituted by y, or, in an even stronger sense, x is nothing over and above y (v. Riel, v. Gulick 2019). For example, given that an organ is composed of its cells, it can be seen as reducible to them and nothing over and above them. Here, we will be focusing on reduction and unity in the sciences, and this can take two forms: an epistemic form and an ontological one. The first explores whether some scientific theories can be theoretically or semantically reduced to or unified with others in any of the senses mentioned above.[4] For instance, an epistemic sense of (dis-)unity

---

[3] For a historical overview on the unity of science see Chapter 1.

[4] Epistemic reduction and epistemic unity can take a theoretical or semantic form, as in Tahko 2021. The first comprises that a certain set of distinct phenomena captured by a higher-level science are described in some ways in terms of a unified lower-level formal mathematical framework. The second refers to an identity relation between all predicates of higher-level sciences and the lower-level ones in a way that all higher-level predicates and explanations are replaceable by lower-level ones. In this thesis, we are not concerned with epistemic reduction or unity but rather with an ontological sense of unity.

concerns whether biochemistry and chemistry can be theoretically or semantically unified. The second form of (dis-)unity regards the ontological relations between the two considered levels, that is, whether ontologically some lower-level entities can be considered more basic than a higher-level class of entities. In the example, we can ask whether biochemical kinds or properties are in some ways reduced to or unified with chemical ones. This ontological form of unity is supported by a general scientific realism approach for which science is taken to be one of our best disciplines to know the world and, accordingly, can inform us about the existent entities in the world. The debate concerning the unity and disunity of science itself has taken different perspectives and approaches. Often, a disunity approach is achieved by embracing different forms of pluralism, supported by the variety of successful special sciences and the causal autonomy that higher-level entities seem to have (as Duprè 1983, 1993). The support for unity instead comes from the successes of inter-disciplinarily approaches and the level of interactions between different higher and lower-level entities (Tahko 2021).

Leaving aside for the present purposes the general debate on unity of science, this chapter focuses on ontological unity and will explore whether a form of it can be found in biochemistry. This domain is the interface between chemistry and biology, but the two disciplines inquire mostly into different kinds of properties. Chemistry is often taken to be the domain of chemical structure and kinds characterised in microstructural terms, such as constituent atomic properties.[5] For example, chemistry considers the microstructure of a molecule and how such structure remains stable. Biology, instead, is the domain of evolutionary functions, etiological classifications and pluralism (Slater 2009; Bartol 2016). For instance, the study of the function of the eye involves etiological characterisations of how the eye has evolved. Biochemical kinds are characterised in terms of both properties, structural and functional. Proteins, for example, are characterised both in terms of microstructure and in terms of the functional roles that they play within biological systems.

In the literature, inter-level questions concerning biochemical kinds have been explored by Bartol (2016), who has argued in favour of a dualist view: some molecules are instances of both

---

[5] Even if this has been challenged as in Tobin 2010, Havstad 2018.

chemical and biological kinds. In his account, chemical properties and biological ones are not linked in a straightforward way, and instead of favouring the chemical or the biological approach, we should just embrace the duality of biochemistry. Within the biochemical domain, we find chemical kinds, identified by their microstructural properties and biological kinds, identified by their biological evolutionary functions. This gives us a disunity view of biochemistry: there is the chemical world and there is the biological one and the two are disconnected. This form of disunity can be seen particularly if one considers Duprè's arguments for the disunity of science: the first regards the plurality of classifications into various kinds of kinds, and the second considers the irreducibility of the various entities (1983, 1993). If one follows such arguments, then the disunity of biochemistry might seem evident. First, biochemical molecules can be classified and can be seen as instances of two different kinds of kinds, the chemical and the biological, as will be further considered in Chapter 7. Second, the various instances of multiple determinations and multiple realisations in biochemical phenomena might support the irreducibility and the disunity of biochemical functions and chemical structural properties. In this chapter, I will focus mostly on the latter sense of disunity, and I will move in the next section to the consideration of these issues attending to some examples.

### 6.2.1 Multiple determinability and multiple realisability

One of the arguments for which there seems to be a duality or disunity in biochemical level kind and phenomena is based on the complexity of the relations between the chemical properties of a given biological relevant molecule and its functional role. These relations often take the form of multiple realisability and multiple determinability. Multiple realisability (MR) refers to a phenomenon in which the same entity or property can be realised by different ones.[6] For example, the property of being an eye can be realised by different organs in different animals. Multiple determinability (MD) refers to the opposite phenomenon: when the same entity can determine different properties or other entities. For example, the same chemical compound can enter into

---

[6] Realisation can be defined as a "synchronic ontological dependence relation, distinct from identity, and that transmits physical legitimacy from physical realisers to what is realised" (Polger, Shapiro 2016, II, 4).

different chemical reactions, realising different properties. Both these relations can be taken to support a form of disunity of science or failure of reductionism, as extensively explored in the metaphysics of science and philosophy of mind literature (see Fodor 1974; Duprè 1983, 1993; Polger, Shapiro 2016; Shapiro 2020; Tahko 2020).

In the biochemical case, MR and MD are particularly relevant because the same biochemical function can be realised by multiple microstructures and the same microstructure can realise multiple biochemical functions (Tahko 2020). Let me present a couple of examples: haemoglobin for MR and the crystalline proteins for MD. As discussed and presented by Tahko (2020, 2021), haemoglobin is a protein with the function of binding and releasing oxygen and can be constituted by at least two different polypeptide chains (or more). The biochemical function of haemoglobin can be considered an instance of MR, as the function of binding and releasing oxygen is realised by at least two distinct macromolecules (chains of polypeptides) that present some microstructural differences. This can lead to a disunity approach when considering haemoglobin: it is difficult to map an identity reductive relation between the chemical structural properties and the functional ones.[7] Multifunctional proteins or "moonlighting" proteins such as crystallines represent instances of MD instead. Crystallines are structural proteins present in all vertebrates' eye lenses, having a function in allowing sight, but they can also have an enzymatic role in digestive processes. In these cases, we notice a form of MD, as the same chemical structure can lead to very different functions in sight and digestion mechanisms (Tobin 2010; Bartol 2016; Tahko 2020). This again challenges forms of reductionism and unity, as a strict relation between function and structure does not hold. Moreover, both MR and MD generate issues of taxonomy or classification. If we follow a microstructuralist approach, then we should favour structure over function and have either many kinds that have the same function (in the case of MR) or one unique kind that has different functions (in the case of MD). If we follow a functional approach, then we have two or three - or as many as the functions - different kinds (in the case of MD) or one kind (in the case of MR).[8]

---

[7] MR represents a challenge to identity reductionism because if tokens of H are realised by $L_1$ and by $L_2$ (and the two are different), then there is no identity between H and $L_1$ or $L_2$ (Shapiro 2020). Later in the chapter, I will consider the disjunctive strategy to multiple realisation.

[8] More on the classification and identification of biochemical kinds will be presented in Chapter 7.

As a reaction to these tensions, Bartol argues that we should bite the bullet and simply embrace the duality of biochemical kinds: nature has cut two joints, the chemical structural one and the biological functional one (2016). Nevertheless, it might be that arguing for disunity is giving up too soon (as Tahko 2020, 2021). First, disunity does not really do justice to the features of biochemical macromolecules that display *both* chemical structure and biological function. These two features are strongly entangled, as supported by some more complex relations between the functions and the chemical structure (see also Goodwin 2011). For instance, Tahko suggests that some cases of MD can be explained or derived from the amphoteric nature[9] of some microstructures (2020). In the cases of some moonlighting proteins for instance, their dual-functions nature can be seen as rooted in some chemical properties of the molecule (Goodwin 2011; Tahko 2020), or at least this can be an option to be analysed in detail.[10] The scientific successes of biochemistry in predicting, manipulating and explaining phenomena support a form of unity as well, as this discipline combines chemical and physical model systems to explain and predict biological phenomena.[11] Second, a more detailed account of biochemical functions is needed if we take these special properties as the root of disunity. These properties are often taken either as irreducible or as unifiable, but little in the literature has been said about them. If, as Tahko (2019, 2021) argues, we can offer an explanation of the link between biochemical functions and chemical structure, this can support unity. In the next section, I will further explore this topic and consider why biochemical functions can be both a challenge and a solution for forms of unity in the biochemical domain.

---

[9] An amphoteric chemical substance is one that can react both as a base or as an acid.

[10] The reducibility of the dual nature of moonlighting proteins has been challenged by Santos et al. (2020). This article stresses the importance of analysing the "dynamical interplay between the micro-level of the parts and the macro-level of the relational structures of their systems" in order to understand these proteins (2020,1). Here I am not supporting the reducibility of biochemical functions to chemical structural properties but rather the relation between functional and structural properties.

[11] The Biochemical Society defines biochemistry as "the branch of science that explores the chemical processes within and related to living organisms." (`https://biochemistry.org/education/careers/becoming-a-bioscientist/what-is-biochemistry/`).

## 6.3 The double problem of biochemical function

While it is not easy to provide a set of necessary and sufficient conditions for a kind to be biochemical, the literature on the topic agrees that biochemical kinds need to exhibit at least two kinds of properties: structural ones and functional ones (Slater 2009; Khalidi 2013; Bartol 2016; Kistler 2018; Tahko 2020). Prima facie, this definition or the combination of these two sets of properties might not be particularly problematic. However, as summarised in the previous section, here lies the tension of the biochemical unity/disunity debate: biochemical functions are often multiply realised and biochemical molecules are often multiply determinable. The argument for a form of disunity or dual approach is often based on these kinds of phenomena. Nevertheless, the topic of biochemical functions needs to be explored in further detail before assessing whether a form of unity can be found or not.

Functional properties in the biochemical context generate what we can call the double problem of biochemical function: the "relation problem" and the "function problem". The "relation problem" asks about the relationship between the chemical structure and the function of a biochemical molecule: how a chemical structure can realise a given biochemical function.[12] As briefly introduced in the previous section, the relation problem is generated by the fact that functional properties in the biochemical domain are often multiply realised, and because biochemical molecules manifest multiple determinability (see Slater 2009; Bartol 2016; Tahko 2020). Furthermore, it is difficult to understand which of the two components, the functional or the structural, has ontological priority in the taxonomy and identification of the biochemical kinds (Slater 2009; Bartol 2016; Tahko 2020). The "function problem" instead asks what biochemical functions are. As discussed in Chapter 5, it is assumed that functional biochemical properties should be interpreted in an evolutionary sense (see Bartol 2016; Tahko 2020). However, this interpretation is not straightforward and needs to be expanded further. Let us consider these problems in more detail with the main case study of this chapter, vitamin B12, already introduced.[13]

---

[12] This problem is presented in Tahko 2020 as well as the "Bio-chem" problem. However, the difference between the approach defended here and his is that we distinguish between biochemical and biological functions.

[13] In Chapter 3 and Chapter 5, I have also discussed how the molecular gene presents forms of MD and MR and how to interpret genetic functions with the account previously presented.

### 6.3.1 Vitamin B12

Vitamins B12 are cobalamin chemical compounds that can act as coenzymes in specific biological processes - specifically, propionate metabolism and methionine biosynthesis. This vitamin comes in four forms - or vitamers - that display similar but different chemical structures: cyanocobalamin, methylcobalamin, hydroxocobalamin, adenosylcobalamin (Combs 2012, 377; Fang et al. 2017).[14] They share a cobalt-corrin complex and the coenzyme function in humans for various biochemical processes such as hematopoiesis, DNA and RNA production, neural metabolism, and carbohydrate, fat, and protein metabolism.[15] Accordingly, these chemical compounds are classified under the same category, 'B12 vitamin', because they display a combination of stable microstructure, a cobalt-corrin complex, and physiological functions.[16]

Vitamin B12 represents an interesting case study relevant to discussing unity and disunity because it presents both MR and MD. First, it presents a form of MR in that the biochemical functions of vitamin B12 can be realised by each of the four vitamers recognised in scientific practice.[17] Second, vitamin B12 plays various roles in human physiology, acting in different biological processes, from DNA and RNA production to hematopoiesis, displaying a form of MD too. The combination of MR and MD challenges forms of identity-based reduction, in which the functions of vitamin B12 would be identical to some of the properties of the microstructure (Tahko 2020). For the sake of the example, let me focus on the function "being a coenzyme in hematopoiesis (the production of blood cells)" (**Coenz-Blood**). Specifically, B12 vitamers have a biochemical function in the proliferation of erythroblasts (red blood cells) during their differ-

---

[14] A more detailed description is the following: vitamin B12 is "the generic descriptor for all corrinoids (compounds containing the cobalt-centred corrin nucleus) exhibiting qualitatively the biological activity of cyanocobalamin".

[15] Reference for chemical structure and function of vitamin B12 `https://pubchem.ncbi.nlm.nih.gov/compound/Cobalamin`. Also, chapter 17 "B12 Vitamin" in Combs' *The Vitamins: Fundamental Aspects in Nutrition and Health* (2012).

[16] The purpose of the chapter is not to provide an analysis of biochemical kinds or of which properties are necessary and sufficient for a biochemical kind to be considered as such. For the present purposes, it is sufficient to consider that a functional characterisation of such kind is often present, combined with a microstructural one.

[17] This might represent an instance of multiple constitution of the kind B12, where this kind can be constituted by different chemical compounds that share some functional properties (Kistler 2018). In Kistler, a kind is multiply constituted when it can be constituted by two or more microscopic structures (2018, 18). See also Gillett (2013).

entiation (Koury, Ponka 2004). This happens because vitamin B12 acts as a coenzyme in the reaction involved in regenerating methionine, which is required in normal erythropoiesis. This function is a definitionally important part of the four vitamers of B12: it distinguishes generic cobalt-corrin complexes from B12 vitamers, and this shows that, even if it might not be necessary and sufficient on its own to define B12, the functional component is nevertheless important.

Let us go back to the double problem of biochemical functions and elucidate them with the example. First, the "relation problem": **Coenz-Blood** is realised in four different ways via the four vitamers of vitamin B12 and, as such, the relation between the chemical properties of the vitamin B12 and one of its functions should be further explored. The MR of **Coenz-Blood** means that it is at least challenging or not straightforward to map a 1:1 correspondence between it and the possible underlying physicochemical properties. The realisation of this function should be further explored. Second, the "problem of function": what does it mean that vitamin B12 has **Coenz-Blood** as a *biochemical function*? As discussed in Chapter 5, an answer is often offered within the etiological selected-effects approach. Vitamin B12 has **Coenz-Blood** because playing this causal role resulted from natural selection and benefitted its retention in the past. This would make this function of B12 a biological property dependent on evolutionary history rather than a chemical one.

The combination of these two problems of biochemical functions might support a theory of disunity between the two domains. The realisation problem challenges the unification or reduction between the biochemical functions of B12 and its chemical structure. The function problem supports such disunity because the nature of biochemical functions will be subsumed under some biological characteristics, which do not relate straightforwardly to the chemical. Disunity might seem to be the way to go. However, the successes of biochemistry itself seem to provide reasons for the opposite: if we can explain, predict and manipulate biochemical kinds in terms of their function and composition, the two aspects need to be related and, to some extent, ontologically unified. Here, I will recall the account biochemical functions previously presented and see how this can support the unity of the two domains, contrasting the fact that these properties could be taken to be the roots of disunity.

### 6.3.2 Biochemical functions

The two problems of biochemical function do not seem to be easily settled by a disunity approach. As previously mentioned, there are two main reasons for that. First, as pointed out by Tahko (2020, 2021), some biochemical functions can be related to the structure of the molecule under consideration in a way allowing for forms of unity or reduction. Second, biochemical functions do not seem to fit into an evolutionary account of function (and so be straightforwardly biological) [as in Chapter 5]. In order to explore whether a form of unity can be achieved and how, let us briefly summarise what biochemical functions are. This will allow us to inquire into how such properties are realised and whether a unified view can be provided.

Often, as in Bartol (2016) and Tahko (2020), an answer to the function problem is given by the fact that biochemical functions are *biological* functions, and these are often identified with evolutionary functions, either backwards-looking or forward-looking.[18] However, this might be inadequate for two reasons, as discussed in detail in Chapter 5. Let us recall briefly the argument previously mentioned (Garson 2018). First, biochemical functions are not evolutionary as follows from a "socio-linguistic argument" based on actual scientific and biological practice, for which biologists do not attribute functions referring to natural selection, especially in "proximal" questions in biology such as in molecular biology, physiology and biochemistry (see also Wouters 2003, 658; Griffiths 2009). In these disciplines, functions of parts and processes are not identified in evolutionary terms but rather in terms of causal contributions to given processes. Second, evolutionary attribution might not be adequate for an "ontological argument", as biochemical molecules do not take part directly in the evolutionary selection mechanisms or do not have the relevant history. According to evolutionary theories of functions, a function is a property of a trait, where a trait is any "detectable phenotypic property of an organism" (Valles 2013). However, it seems questionable whether B12 vitamins are a trait of an organism. B12 vitamins are not phenotypic properties of human beings but rather chemical compounds absorbed via

---

[18] According to a backwards-looking theory of function, a trait's function is a difference-maker effect of such trait that contributed to the fitness of the ancestors and has been retained via natural selection because of these benefits. A forward-looking theory of function, instead, sees the trait's function as a difference maker for the future fitness of the population having the trait [more in Chapter 5].

nutrition. As such, they cannot be considered traits of organisms and be assigned evolutionary functions. Accordingly, if we accept that biochemical functions are ascribed to vitamin B12, biochemical functions do not seem to be straightforwardly evolutionary [as discussed in Chapter 5].

The functional aspect of B12 remains important. These vitamins are identified as chemical compounds with some given biochemical functions, such as contributing to hematopoiesis. As mentioned, this function does not seem to refer to an etiological function for semantic and ontological considerations. What are those functions then? In genetics and at the micro-level, functional attribution seems to be done in terms of the basis of relevant activities of the entities under consideration, i.e. “those likely to make a relevant difference” for a given phenomenon (Griffiths 2009; Germain et al. 2014, 817), rather than in terms of history. In the case of B12 vitamins, **Coenz-Blood** makes a difference for hematopoiesis, jointly with other enzymes and actors. This functional attribution is in line with a causal theory of function, the main competitor to the etiological accounts (Slater 2009: Griffiths 2009; Germain et al. 2014). According to this theory, a function is the causal role of an entity within a performed activity in a complex system (Cummins 1975; Sterelny, Griffiths 1992; Germain et al. 2014). The ascription of a function amounts specifically to the identification of a property or a set of causal properties (from now on *powers*) that contribute to given phenomena.[19] The considered function is associated with chemical causal powers manifested within the right set of conditions and processes [Chapter 5].

In more detailed terms, the biochemical function of vitamin B12 **Coenz-Blood** consists of a set of chemical powers manifested in “the transfer of a methyl group from 5-methyl-THF to homocysteine via methylcobalamin, thereby regenerating methionine” (Koury, Ponka 2004, 109). In this sense, when we say that B12 vitamin has a haematopoietic function in humans, we are referring to its dispositional property to react in a specific way during the regeneration of methionine required in erythropoiesis, the production of blood cells. This, however, is not enough to pin down the nature of biochemical functions, as they are not any chemical powers of

---

[19] As mentioned in Chapter 5, I take powers and dispositional properties to be causally efficacious properties that manifest themselves in specific circumstances. I am aware of the extensive discussion on the relation between powers and dispositions, but for the present purposes, I take a minimal commitment and I define them as above (Wilson 2021).

the molecule but very specific ones (those relevant to the process).[20] The action of the chemical powers of **Coenz-Blood** depends on the right biological context for their contribution and on the presence of the right process to which the molecules can contribute. Specifically, as stressed by Cummins (1975) and Germain et al. (2014), functions are always contributions to something in a given context, and this context is, in the case of **Coenz-Blood**, hematopoiesis. For **Coenz-Blood** of vitamin B12 to operate, the process of hematopoiesis needs to be happening thanks to cell regulatory mechanisms and the action of various enzymes and co-factors.[21] **Coenz-Blood** is associated with those chemical powers that contribute to specific biological processes. The biological processes to which **Coenz-Blood** contributes are evolutionarily selected, and, as such, a form of indirect evolutionary selection is present for the function too: the relevant powers are those that contribute to processes that are a result of evolution. As such, evolution indirectly selects those chemical powers relevant to the biochemical function. Nevertheless, it is important to underline that the causal contribution can happen independently from its being beneficial for the actual/ final survival or fitness of the organism, as what is needed for their contribution is the presence of the biological processes to which they contribute. Combining the two discussed aspects (being related to chemical powers and being dependent on a biological context), we can offer the following characterisation of biochemical functions [as presented in Chapter 5]:

**BC-function** Biochemical functions are associated with a set of chemical powers to bring about a specific effect within biological processes. These biological processes are a product of evolution and, as such, the relevant chemical powers are indirectly evolutionary selected.[22]

This account of biochemical functions is in line with the general characterisation of biochemistry as the science that considers the behaviour and effects of chemical processes in biological systems (Santos et al. 2020). Moreover, this approach to biochemical functions allows us to answer the function problem, telling us what these properties are. It provides a starting point to explore the unity in biochemistry, as the biochemical functions are identical or associated with some

---

[20] More on this will be said in §5.

[21] The importance of the context for the manifestation of biochemical functions is also underlined in Santos et al. (2020).

[22] In Chapter 5, the definition is also presented in terms of dispositional properties to make the connection between the chemical account of functions and the biochemical one stronger.

specific chemical powers manifested during biological processes. The challenges of MD and MR still remain, and in the next section I will explore a more detailed account of how to answer the realisation problem and achieve unity.

## 6.4 Which unity?

In the previous section, I considered an answer to the function problem for which biochemical functions are causal contributions to biological processes. In this section, I will focus instead on the relation problem and how unity can be achieved despite MR/MD.

At a first level of dependencies, a basis for unity can be seen in the nature of biochemical functions. Biochemical functions are associated with a set of chemical powers that contribute to biological processes. This can provide a basis for a sense of unity: there is a way in which biochemical functions are in some way reducible to chemical properties. However, as mentioned in §2, a straightforward form of identity reduction is challenged by the widespread cases of MD and MR in the biochemical domain. Moreover, the set of dispositions relevant to biochemical functions are not any arbitrary chemical power of the considered molecule or compound but some very specific ones. The relevant powers are those contributing to biological processes and have undergone at least an indirect selection process [Chapter 5]. In some ways, the reductionist approach does not tell us enough about which chemical powers are relevant to the given process. The consideration of the biological process to which they contribute and of their evolutionary history is necessary to understand the relevant set of powers (a similar suggestion is in Santos et al. 2020) [Chapter 5]. Moreover, the causal efficacy of biochemical molecules is distinctive in that it should bring about specific effects within biological processes.

Different ways have been explored to maintain a form of unity without committing to forms of identity reduction.[23] Here, I will consider weak emergence via the proper subset strategy and non-reductive physicalism, as in Wilson (2011, 2015, 2021) and as suggested by Tahko

---

[23] As summarised by Wilson (2021, 55), some ways to retain a form of unity without identity reduction are functional realisation, constitutive mechanisms, mereological realisation, determinable/determinate relations, supervenience, and forms of emergence.

(2020). This account, I will argue, provides an answer to the relation problem and allows for the specificity of biochemical functions while maintaining a unified account.

### 6.4.1 Weak emergence and the "proper subset strategy"

Weak emergence is a form of emergence compatible with non-reductive physicalism and, accordingly, with forms of unity: there is only one broader kind of properties, physical properties. According to non-reductive physicalism, higher-level entities are real and constitute a novel level of reality, being distinctively causally efficacious; at the same time, their causal actions operate in a way respecting physical causal closure and hence in line with physicalism.[24] This combination of distinctiveness and causal efficacy, together with a sense of dependence, can be maintained by defending a form of weak emergence based on the "Proper Subset of Powers strategy" (Wilson 2011, 2021; Tahko 2020).[25] This strategy comprises two steps: i) accepting the *Token Identity of Powers Condition*; ii) accepting the *Proper Subset of Powers Condition.* The first states that every token power of a given token feature H on an occasion *t* is identical with a token power of the token feature L on which H co-temporally materially depends at *t*.[26] The second states that the token feature H has at *t* a non-empty proper subset of the token powers of the token feature L on which H co-temporally materially depends on at *t* (as formulated in Wilson 2021, 57-58). The combination of these two conditions constitutes the basis for a weak emergence relation between the higher and the lower-level entities or features:

**WE**: "What is it for token feature H to be **Weakly metaphysically emergent** from token feature L on a given occasion is for it to be the case, on that occasion, i) that H co-temporally materially depends on L, and ii) that H has a non-empty proper subset of the token powers had by L" (Wilson 2021, 75; variables modified, emph. added).

---

[24] The principle of causal closure is often taken as a condition for forms of physicalism and claims that "all physical effects have sufficient physical causes", avoiding cases of problematic overdetermination.

[25] This strategy presupposes a very simple ontology of objects, properties, and powers. Properties are instantiated by objects and are identified by a range of causal powers (Shapiro 2020). In this case, a biochemical molecule instantiates the property "having a given biochemical function", individuated by a specific set of causal powers.

[26] Material dependence implies a form of substance monism, in line with physicalism, and a form of minimal nomological supervenience of the emergent features *type* H on the base features *type* L (Wilson 2021, 73). This means that supervenience should happen with at least nomological necessity.

The first condition i) allows for a form of dependence as there is a token identity of the powers associated with the two features; the second condition ii) allows for a form of distinctiveness. In particular, this account allows for a form of unity because the token powers of the realised feature H are nothing more than a subset of the token powers of a realising feature L, and the two features can be unified as the two sets of powers are both physically acceptable and the token powers of both sets are identical.[27] At the same time, H is ontologically autonomous from L because H has a *proper subset* of the token powers of L and by Leibniz's laws and via set-theory principle, a proper subset of token powers is different from its set of token powers. This permits to maintain the type difference between H and L. The proper subset strategy also allows for a form of causal autonomy, as discussed by Wilson (2011, 2021). Specifically, H has a distinctive causal profile compared to L because it possesses a distinctive set of causal powers or distinctive causal profile compared to L. H's causal autonomy is based on the fact that H has a distinctive set of powers compared to the feature from which it emerges.

Summarising unity can be maintained because the higher-level properties are "physically acceptable", as they materially co-temporally depend on lower-level properties, but they can still be maintained as ontologically autonomous (Wilson 2011). Moreover, as I will explore in section §5, the proper subset strategy and weak emergence are able to deal with MR and MD. In the case of MR, it can be possible to identify more than one district token power subset of the lower-level L that can be associated with the higher-level feature H. While in the case of MD, the token set of powers of a given lower-level feature L could present different proper subsets of token powers associated with different higher-level emergent feature H. More on this will be said once we have seen how biochemical functions can be weakly emergent according to the presented account.

### 6.4.2 Biochemical functions are weakly emergent

Let us now consider the interface between biochemical functions and chemical properties and the answer to the relation problem. As we discussed in the previous section (and more in detail in

---

[27] See also Shapiro 2020.

Chapter 5), a biochemical function is associated with a set of chemical token powers to bring in a given effect within biological processes. More precisely, the relation between the token powers associated with BC and the correspondent chemical powers can be interpreted with the proper subset view. BC has in a given $t$ a proper subset of token powers of the set of chemical token powers of the chemical molecule. This proper token subset is individuated via the evolutionary history of the biological process to which BC contributes. Accordingly, following the account considered, we can state the weak emergence of BC:

**$\mathbf{WE_{BC}}$**: A biochemical function BC weakly emerges from the chemical compound C under consideration at a given $t$ because: i) BC co-temporally materially depends on C at $t$; ii) BC has an identifiable and non-empty proper subset of token powers of C at $t$.

The powers associated with BC are token identical at $t$ to powers in C, but we can identify the biochemical functions as being associated with a proper subset of the chemical powers (as in Figure 7). This makes the biochemical function BC *type* different from C, while it also allows us to maintain that the biochemical functions are co-temporally materially dependent on the chemical ones. Biochemical functions can then be considered weakly emergent from the chemical powers of the molecule and this provides an answer to the relation problem: the relation between the chemical properties of a biochemical kind and the functions is weak emergence. This also allows having a form of unity given by the token identity of the instances of the biochemical functions and the chemical properties, while at the same time maintaining a type difference and the related causal efficacy.

In the case of vitamin B12, **Coenz-Blood** has a specific proper subset of the chemical powers of cobalamin, the ones relevant to the regeneration of erythroblasts in hematopoiesis. Those powers are the ones involved in the relevant co-enzymatic action that the vitamin plays: the token of the powers of **Coenz-Blood** are the same toke powers of the cobalamin compound involved in the process, however the causal contribution is distinctive. The function **Coenz-Blood** emerges from the chemical compound in that it has a proper specific subset of causal powers. Specifically, in this specific case, it amounts to those chemical properties that allow for the regeneration of methionine via “the transfer of a methyl group from 5-methyl-THF

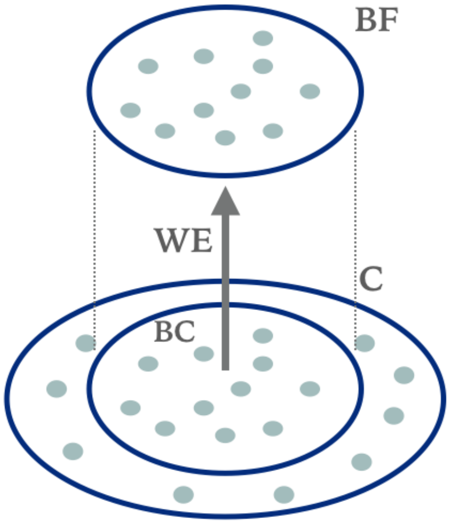

Figure 7. The weak emergence relation (WE), for which a token biochemical function BC is associated with a set of token powers that is token identical to a proper subset of token powers of the token chemical set C.

to homocysteine via methylcobalamin" (Koury, Ponka 2004, 109). This set is not arbitrarily chosen, but it is identifiable thanks to the evolutionary history of the different biological processes in which B12 acts as a co-enzyme [see Figure 6 in Chapter 5]. The causal contributions are those relevant to the given environment and the given process. The biochemical functions of B12 vitamins can be considered weakly emergent from the chemical dispositional properties of cobalamin compounds at a given time *t*.[28] This makes the causal profile of vitamin B12 distinctive, as recognised in scientific practice and in the functional characterisation of B12. At the same time, this emergence is only weak as it does not presupposes any stronger forms of ontological novelty, as the one of a strong form of emergence of a physically unacceptable variety.[29] The identity of the token powers associated with both the emergent feature and the

---

[28] In Chapter 3, I have presented the weak emergence of the molecular gene. Also in this case, it is possible to identify a proper subset of powers held by the nucleic acids molecule associated with the function. As will be discussed later in this chapter, the two views of emergence are compatible. Moreover, the definition of weak emergence presented here is temporal, being defined for every token instance in a given *t*, as in the case of the molecular gene. Accordingly, weak emergence, as defined here, is compatible with the existence of the biochemical function or entity at a particular time and in a given context.

[29] The contrast between weak and strong emergence is often expressed in terms of whether the token

lower basis allows us to maintain a form of ontological unity.

### 6.4.3 Unity within autonomy

Given the importance of this aspect for the general goals of the thesis, let us summarise how weak emergence so formulated allows a form of unity that combines the autonomy of biochemical functions with the dependence they have on chemical powers and answers the relation problem.

Firstly, unity is maintained because this account presupposes a form of material dependence given by the identity of tokens condition. This gives a dependence relation between the chemical and the biochemical features as, at a given $t$, the instances of the two kinds of properties are the same: the token powers of the biochemical functions are identifiable among a broader set of chemical powers. This is further supported by the fact that this kind of emergence is *weak*: the causal powers of the biochemical functions are not novel (or an addition) compared to the chemical causal powers, but a proper subset of them. Second, the emergence discussed here and obtained via the proper subset view offers us a view of the connection between biochemical functions and chemical structure: the two are not disentangled but rather connected. At last, weak emergence, as proposed, is compatible with a broader form of substance monism or physicalism for which there is only one kind of substance (Wilson 2021). Once we grant the weak emergence relation to hold, biochemical functions are nothing over and above physical properties. In conclusion, unity via weak emergence accounts for the biochemical features to be dependent on the chemical features in some important respects, as the latter is ontologically prior and there is a level of material dependence, while at the same time, the causal autonomy and distinctiveness of the biochemical functions can be maintained (Tahko 2021).

---

powers of the emergent feature are novel or not compared to the lower-level. In the case of weak emergence, such token powers are not novel, while in the case of strong emergence there is an addition of powers (for more see Wilson 2015, 2021). Strong emergence can be considered compatible with forms of unity, but more needs to be said for such compatibility to hold.

## 6.5 Challenges and clarificatory remarks

Before concluding, some clarifications and answers to some objections should be offered. As previously mentioned, multiple determinability and multiple realisability represent a standard challenge to forms of unity. Accordingly, it is important to say more and expand on how the proper subset strategy can deal with these phenomena. Secondly, Tahko, in his 2020 paper, presents an objection to the proper subset view and how it can deal with unity in biochemistry. Here, I am going to provide an answer to this objection. Lastly, I will briefly present how this view is compatible with the account of weak emergence introduced in Chapter 3.

### 6.5.1 Multiple realisability and multiple determination *again*

Multiple realisability and multiple determination seem to undermine forms of straightforward reduction and unity, as presented in section §2.1. This is so because, in MR, the same higher-level feature seems to be realised by different lower-level ones. While in MD, the same lower-level feature can be determined or can realise different higher-level entities. In both cases, we cannot identify a straightforward reductive relation between the two features. This challenges forms of unity or reductions. In the cases discussed here, biochemical functions are multiply realisable, and in some biochemical cases, such as in the crystallin protein, the same chemical features can be determined into many biochemical functions. Here, we have presented the proper subset view and weak emergence as a way to maintain the autonomy and ontological relevance of the biochemical functions while preserving a form of unity. However, more must be said on how this view can be compatible with MR and MD cases.

To start with, MR and MD are what we can call "type issues": it is the realised *type* that can be multiple realisable or be one of the determinations of a given lower-level feature.[30] Starting with MR, it is the type function **Coenz-Blood** that is multiply realisable by the four vitamers of B12. However, in a given moment, such as during a specific instance of hematopoiesis, a token instance of **Coenz-Blood** will be realised by a specific token instance of the four vitamers

---

[30] More on biochemical kinds and the identification of the various types of biochemical molecules will be discussed in Chapter 7.

of B12. At the time $t$, *only* the token powers of a proper subset of the lower-level entity are identical to the token powers of the emergent feature **Coenz-Blood**. This implies that despite MR at the type level, at $t$ the token entity is realised by one token lower-level set of features. In the case of MD instead, there is only one token subset of powers that in a given time $t$ realises the biochemical functions under discussion [Figure 8]. A token biochemical function is emergent in that it has a proper subset of the token powers of chemical features. This makes the proper subset view straightforwardly compatible with MR and MD, as discussed by Tahko (2020, 2021). Let us consider these two in more detail.

For MR, there may be several distinct token proper subsets of powers of the chemical features that can be associated with the biochemical function. In the case of **Coenz-Blood**, there are several distinct token proper subsets of the B12 vitamers that can be associated with the function and, as such, can realise the biochemical function under consideration. This is possible because while the type is multiply realised, the token is always realised by a specific subset of token powers. For MD, two aspects can be considered. From the perspective of the token realised feature, one identifiable proper subset of chemical powers is associated with the higher-level feature, and, as such, MD is not problematic. From the multiply determinable feature perspective, instead, the token set of powers of a given chemical feature could present different proper subsets of token powers associated with different biochemical functions [Figure 8]. Or, as suggested by Tahko 2020, there could be one proper subset of powers associated with two distinct type features, bringing in different effects in the relevant biological context. Accordingly, the token powers of the functional properties are a subset of those of a single chemical kind.

A last clarification might be useful for the present purposes. A common reductionist answer to MR is that it does not represent a serious problem for reductionism because it can be explicated in terms of a closed disjunction. Let us consider this debate in more detail. MR represents a challenge to reductionism because there are instances of the higher-level H that are not instances of the lower-level $L_1$ to which it should be reduced, but are instead instances of $L_2$, which also realise H in given circumstances. However, some have argued, MR is not a problem

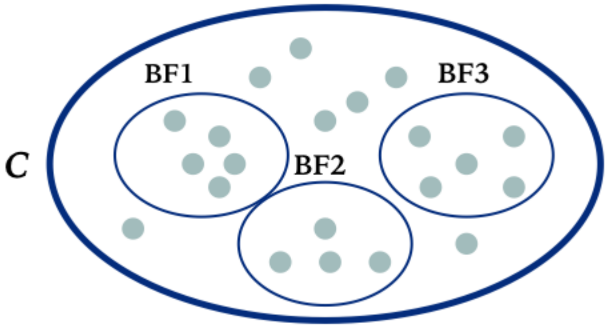

Figure 8. Multiple determinability of the cobalamin molecule, for which only one subset of powers is realised at a given *t*.

for reductionism if H is reducible to the disjunction of $L_1$ and $L_2$ such that $H=(L_1$ *or* $L_2)$. This would imply that H is reducible to either one of the Ls, and accordingly, reductionism is still safe. To this, one can answer in three main ways: the first two are metaphysical and concerns the nature of genuine kinds and the proper subset view, and the third is contextual to the kind of properties biochemical functions are. The main answer to the disjunctive strategy is famously discussed by Fodor (1974), Kim (1992) and is based on the fact that disjunctive properties cannot represent a strong enough basis for natural kinds, as they would represent a category too heterogeneous, especially when such disjunctions are open-ended. Wilson also discusses how the proper subset view can ensure a form of ontological autonomy *contra* the disjunctive strategy (2021). In the case of MR, when the entity H is weakly emergent (as in the cases discussed here), the token powers of H are a proper subset of the token powers of either $L_1$ or $L_2$. This makes H type different from the disjunction of Ls because of Leibniz's law: there are some powers of L that are not of H. Accordingly, the function **Coenz-Blood** is not reducible to the disjunction of the chemical powers of the four vitamers of B12 because such function has a proper subset of causal powers of the vitamers.

To these metaphysical considerations, I am going to add a last contribution that considers more in detail the nature of biochemical functions. These properties are associated with a set of powers whose selection is at least indirectly a result of evolution, and their causal efficacy is embedded in biological systems that are currently evolving. This has an impact on the fact that the types of realisers of the biochemical functions can change or increase in time. Moreover, there could be a biologically possible world in which the biochemical function is realised by another chemical molecule which we do not know, or that does not play the function in current systems, but could have the function. This would make the disjunction an open disjunction, and, as such, challenges a straightforward reductionist approach as there could always be a case in which H is not realised by $L_1$ or $L_2$. This, as reported before, is particularly challenging given the counterintuitive existence of genuine natural kinds or features with disjunctive properties.

In conclusion, the proper subset view and an account of weak emergence seem to be compatible with accounting for forms of MR and MD.

### 6.5.2 Biological context to the rescue

MR and MD are not the only aspects that need to be taken into account from the perspective of the proper subset view. As briefly mentioned above (and as will be discussed in more detail in Chapter 7), one of the tensions concerning biochemical kinds is that they can be interpreted and classified according to the biological or the chemical perspective and the two perspectives can be incompatible. For example, haemoglobin is multiply realisable: the same set of biochemical functions is realised by different chemical structures. If we follow a functional-based approach, then we have one kind - haemoglobin - that is realised by at least two different chemical structures. The same can be said of the biochemical function "binding oxygen" of haemoglobin. However, if we follow a reductionist chemical approach, then we would have two different kinds characterised in terms of different chemical features, and so the functions would be different as they would be reduced or identified with different chemical properties.

Here, we have introduced a more detailed account of biochemical functions for which a biochemical function has a proper subset of token powers of the chemical set of powers. This can help us to identify which biochemical functions are relevant to the biochemical kind: those contributing to specific biological processes. Accordingly, discussed in more detail in Chapter 5, this might provide us with a way to identify biochemical functions as those that contribute to biological processes. Moreover, the account proposed is compatible with unity, but different from reductionism as the proper subset of powers of the feature is not identical to the set of chemical powers. There is something special about being a biochemical function compared to being any chemical causal contribution to a process.

Debating these types of issues, Tahko puts forward an objection to the proper subset view that could come from ontological reductionism (2020). Specifically, he is asking whether the various biochemical functions and, accordingly, the relevant proper subsets, are just pragmatically selected according to the discipline that is inquiring into the function. If the selection is pragmatic or discipline relative, the ontological reductionist could argue that the subsets of powers can be reducible to larger subsets of powers contained in the molecular structure, but not manifested. If this is right, the relevant sets could be interpreted according to the chemical perspective with a larger proper subset as well (making the view compatible with reductionism). It is the approach or the discipline we are considering that gives the relevant proper subset, and this might be - from the reductionist perspective - the chemical one. In a nutshell, Tahko objects that the proper subset view does not give us what we need to identify the specific proper subset relevant to biochemical functions.

An answer to this issue can be found in the more precise account of biochemical functions, presented in Chapter 5 and here. If we want to consider biochemical functions and not biological or chemical functions, then the selection of which proper subset is relevant is not pragmatically determined by the two disciplines, but by the nature of such properties. Let us recall the account of biochemical functions again:

**BC**: “Biochemical functions are associated with a set of chemical powers to bring out a specific effect within biological processes. These biological processes are a product of evolution and, as such, the relevant chemical powers are indirectly evolutionary selected”.

Here, it is important that the set of chemical powers associated with a biochemical function is not just any collection of powers, but the one indirectly (or directly in some very specific cases) evolutionary selected. It is the evolutionary selected biological process to which they contribute that identifies which chemical powers are associated with a biochemical function.[31]. Accordingly, the identification is not discipline-relative or pragmatic, but, in biochemical phenomena, it is given by the context in which the molecules operate. A function is always a contribution to something in a context (Germain et al. 2014), and this context is given by biological processes.[32] This answers to the reductionist, as the set is not any chemical set, but a specific context-relative one where the context is the specific biological process to which they contribute.[33]

The presented definition of weak emergence also helps here. According to this, there is a token identity of the powers in a given moment, but it is at the moment in which the biological process is happening that the powers are identified and, then, manifested. These powers are such because of the evolutionary history of the biological process considered. Nevertheless, this does not imply that one could not consider different aspects of these molecules and study them independently for pragmatic or discipline-relevant interests. One could still approach the functions of haemoglobin or vitamin B12 from a chemical perspective or a biological only one. However, if the interest is not only epistemic, but ontological, one needs to consider the nature of biochemical functions *per se* and this offers a selection of powers, that are not any, but specific ones. In conclusion, the objection based on the fact that the selection of powers is pragmatic and, as such, compatible with also the reductionist approach can be answered by considering a more detailed definition of biochemical functions. The biological evolutionary context comes to the rescue and offers us which powers are those relevant in a non-pragmatic or arbitrary way.

### 6.5.3 Two accounts of weak emergence

A last clarificatory note is needed before concluding the chapter. Weak emergence can be deemed a way to obtain unity when considering the relation between chemical structure and biochemical

---

[31] More on the role and the nature of the evolutionary properties of vitamin B12 will be discussed and presented in Chapter 7. For other references on this, see Chapter 5 and Bellazzi *forthcoming*.

[32] The importance of the context for the comprehension of biochemical functions, in particular concerning proteins, has been presented in Santos et al. 2020.

[33] A similar suggestion is presented by Goodman concerning protein folding: "The tertiary structure of a protein, when there is one, represents the most energetically stable conformation available to the protein in the relevant biological circumstances. "(Goodwin 2011, 543)

functions. This view combines a form of material dependence with the ontological autonomy and relevance of the emergent entity. We can maintain a proper sense of biochemical functions while keeping them distinct from the chemical features from which they are realised. In this chapter, we have proposed weak emergence via the proper subset view. However, the concept of weak emergence has already been discussed in Chapter 3. In Chapter 3, weak emergence is defined as the combination of novelty and robustness: a given phenomenon is weakly emergent when it is considered novel and robust. How can we combine this approach with the proper subset view? In this section, I am going to do that in two steps. First, I will consider the relation between novelty and robustness and weak emergence. Second, I am going to expand on the relation between robustness and multiple realisability.

In Chapter 3, I have introduced emergence as the combination of novelty and robustness. Novelty implies that it is possible to identify the emergent property in a distinctive way from the properties of the lower-level entities, and the consideration of such a property improves explanations, leading to new ones. Robustness implies that the emergent property displays stability within a certain range of perturbations of the lower-level entities, interpreted in Chapter 3 in terms of multiple realisability. Moreover, I have clarified that these two features are *the ratio cognoscendi*, i.e. how to know that something is emergent, while, from the ontological point of view, it is the fact that the phenomenon is emergent that represents the *ratio essendi* of the features, novelty and robustness. We can know that something is emergent when it displays novelty and robustness. But something displays novelty and robustness because it is emergent. Emergence allows us to postulate the existence of something and its knowledgeability. But we have epistemic access to its emergence thanks to its being novel and robust. This makes the approach presented in chapter 3 compatible with what has been in this chapter, where more details about the ontological aspects of emergence have been presented. Something is novel and robust because it has a *proper subset* of the token powers of the lower-level: being weakly emergent means holding such a proper subset of causal powers.

Let us expand on this compatibility more. Novelty is defined as the capacity to lead to new explanations. This criterion is mostly epistemic and as such can be seen as compatible with the proper subset view. The entity holding the proper subset of powers of the lower-level should provide more explanatory power than simply postulating the lower-level entity from which it emerges. Let us now turn to robustness. Robustness has been defined in Chapter 3 as a form of multiple realisation. More precisely, this multiple realisation should be seen as having the same type of entity, in the examples discussed a given gene or B12 vitamin, realised by different type entities of the lower-level, either DNA strands or different cobalamin compounds at the given time $t$. Accordingly, robustness implies that the emergent type entity can be realised by different lower-level types, but every token emergent will be realised at a given $t$ by some specific token features of the lower-level, as discussed in the previous section. For example, a given gene can display novelty and robustness because it improves a particular set of explanations and is multiply realisable. These features are so because there is a proper subset of the powers held by the DNA that is token identical with the powers of the gene at a given moment and the selection of such powers is given by the process of transcription. Accordingly, the gene is weakly emergent as it has this proper subset of causal powers, and we know so because it displays novelty and robustness. A similar form of reasoning can be provided for the emergence of the function **Coenz-Blood**.

In conclusion, it seems that the two views can be seen as compatible and complementary. Novelty and robustness can be seen as the indicators for emergence and our epistemic tools to access it.[34] While the proper subset view offers a way to understand what is happening at the ontological level between the powers associated to the emergent entities and those of the lower-level.

## 6.6 Conclusion

This chapter considered how biochemical functions can be linked to chemical structure and how this informs the unity of science debate. First, it briefly presented the theme of unity of science and why biochemistry is an interesting field to explore unity and disunity issues. It pointed out

---

[34] This is also discussed in Chapter 3, in particular in §3.3 and footnote 18.

that functional properties challenge senses of unity or reducibility in biochemistry. Specifically, the double problem of biochemical function, the "function problem" and the "relation problem", can support forms of disunity. Here, the two problems have been explored by recalling the definition of biochemical functions presented in Chapter 5. The answer to this problem is needed to deal with the relation problem and the ontological dependence relations between chemical structure and biochemical function. Then, it considered the relation problem and argued that biochemical functions could be considered weakly emergent from the molecular structure. The end of this chapter considered three further aspects or challenges of the proposed view. It showed how this is compatible with both MD and MR and I have considered a reductionist objection, as in Tahko 2020. It then discussed briefly how this view of emergence can be linked to novelty and robustness and how it is compatible with the results in Chapter §3. In conclusion, the weak emergence of biochemical functions and the related autonomy can be seen in support of the main thesis of the present research, that is that there is something *sui generis* about biochemical kinds and their functions.

This chapter has a series of interesting results. First, it relates to one of the main research topics of biochemistry, the relation between biochemical functions and chemical structure. The account presented allows us to maintain a form of autonomy for biochemical functions while being compatible with forms of unity. Specifically, a form of unity in biochemistry between the structural and functional features can be achieved once we have defined such functions and we have identified the relation between them. A similar conclusion has been put forward by Tahko (2020, 2021), as he discusses that there is some space for ontological unity if we can establish a relation between the biological and chemical aspects of biochemical kind. The definition of biochemical functions can be explicated in terms of chemical powers that contribute to biological processes. However, a straightforward reductive unity view does not seem available due to the various forms of multiple realisation and multiple determination present in biochemical phenomena. A way to achieve unity that remains is a form of weak emergence that allows us to retain both dependence and autonomy. This also provides a solution for the relation problem.

Concluding, unity is firstly obtained because biochemical functions are materially dependent on chemical features, thanks to the identity of tokens condition. Secondly, unity is achieved because we can spell out the clear relations between the biochemical and chemical features - they are not disentangled but linked and connected. The consideration of biochemical functions and their relation to chemical structures opens up questions on what biochemical kinds are and whether they can be considered natural. This will be explored in the following Chapter 7.

# Chapter 7

# Biochemical kinds

## 7.1 Introduction

Is there something specific about being a biochemical kind, or do biochemical molecules instantiate chemical or biological kinds? This question stems naturally from the discussions in the previous chapters that consider biochemical functions and their weak emergence [Chapter 5, Chapter 6]. This chapter takes a step forward and considers the topic of biochemical kinds *per se*.

Biochemical kinds are mentioned in the scientific and philosophical literature, but whether biochemical molecules instantiated chemical, biological or biochemical kinds still needs to be established. Here, we explore this issue and the chapter argues that biochemical kinds can be considered a natural category because they respect the requirements for natural kindhood presented in Chapter 2 and Chapter 4 (and as in Khalidi 2013).[1] While the argument is conditional upon the acceptance of the assumed account, considering the natural status of biochemical kinds can open the debate about the naturalness of these kinds - something that has not yet been fully developed in the relevant philosophical literature. Moreover, the naturalness of biochemical kinds supports a form of autonomy to biochemistry as a discipline that refers to categories

---

[1] This chapter relates to the naturalness question mentioned in Chapter 2 and Chapter 4. This question considers how we can distinguish natural from non natural categories.

characterised by an adequate cluster of properties. This is in favour of the general result of this thesis for which there is something *sui generis* about biochemical phenomena.

The status of biochemical kinds as both distinct from chemical or biological ones and natural is important because it can inform a series of debates. The first concerns the taxonomy of biochemical kinds, as discussed in the philosophy of biochemistry (see Slater 2009; Goodwin 2011; Tobin 2010; Havstad 2016, 2018). Biochemical kinds present chemical structural properties, functional properties and etiological ones[2], and there is tension regarding which of the features is more important for biochemical classification. Accordingly, establishing whether biochemical kinds are a natural category and can be classified accordingly might the relevant taxonomic practices. The second debate that can benefit from assessing the naturalness of biochemical kinds is the one concerning unity in biochemistry and the relation between the biochemical, chemical and biological domains. As already discussed, Bartol (2016) argues that there are no biochemical kinds, but only chemical and biological kinds: nature has cut two joints. Some molecules can instantiate both chemical kinds and biological kinds. The chemical kind membership is bestowed by their structure, while the biological kind membership is given by their functional component and being evolutionarily conserved units in virtue of these functions. However, as in Chapter 6, biochemistry as a discipline and the characterisation of biochemical kinds seem to indicate that there might be something specific about the biochemical domain as a link between the chemical and the biological. In this respect, biochemical kinds might then result in being *sui generis* kinds that are what they are exactly in virtue of being chemical molecules that play a role in biological processes. In the previous chapter, I considered how unity could be maintained once the relation between the features is maintained. In this chapter, I will argue for a unity within kinds classifications, following Natural Kind Monism. This can be seen in contrast of Duprè's analysis for which a plurality of kinds or a duality of kinds, in the case of biochemical molecules, can be seen as a source of disunity (1983, 1993). Here, instead, the chapter argues that biochemical kinds are natural kinds, and this can be seen as compatible with forms of unity of science.

The structure of the chapter is the following. Section §2 presents the status of the controversy

---

[2] Etiological considerations are put forward for proteins, genes and other biochemical kinds, as suggested previously in this thesis [Chapter 4, Chapter 5, Chapter 6.].

concerning biochemical kinds. Section §3 summarises again the considered account of natural kind [already presented in Chapter 2, Chapter 4]. Specifically, natural kinds are projectible categories in our best scientific theories and they represent nodes in the causal network of the world (as in Khalidi 2013, 2018, 2020). It then considers which accounts of natural kindhood are mostly applied in chemistry and biology and why they do not seem suitable for the characterisation of biochemical kinds. Having set down the main criteria, section §4 identifies some of the core properties that most instances of biochemical kinds have. These are: presenting a chemical structure, having a given biochemical function in a biological context and a form of evolutionary history in being or having been bio-synthetised.[3] It shows that these three properties are related and respect the criterion of kindhood mentioned above, supporting my reasoning with two case studies and putative instances of biochemical kinds: vitamin B12 and the molecular gene. Section §5 considers how this approach can inform debates in taxonomy. Section §6 concludes by exploring how the naturalness of biochemical kinds can be seen as compatible with a form of unity of science via Natural Kind Monism.

## 7.2 The problem

As presented in Chapter 1 and throughout the thesis, biochemical kinds can be broadly identified as those that figure within biochemical explanations and predictions and are studied by biochemistry. They come in a variety of different instances studied by disciplines that go from molecular biology to genetics. Standard examples of biochemical kinds include proteins, nucleic acids, carbohydrates, lipids, and vitamins. In order to understand more about what these kinds amount, let us consider the definition of biochemistry as presented by the *Biochemical Society*:

> “The branch of science that explores the chemical processes within and related to living organisms. It is a laboratory based science that brings together biology and chemistry. By using chemical knowledge and techniques, biochemists can understand and solve biological problems. Biochemistry focuses on processes happening at a molecular level. [...] Biochemists

---

[3] The issue of biochemical molecules synthetised in laboratories will be considered in detail in section §4.1.

> need to understand how the structure of a molecule relates to its function, allowing them to predict how molecules will interact."[4]

Following this definition, while it is not easy to provide a set of necessary and sufficient conditions for a kind to be biochemical, the literature on the topic agrees that biochemical kinds need to exhibit at least two kinds of properties: structural chemical and properties related to the functional role they have in biological processes (Slater 2009; Goodwin 2011; Khalidi 2013; Havstad 2016, 2018, 2021; Bartol 2016; Kistler 2018; Tahko 2020). Moreover, as Havstad (2016, 2018) reported, some biochemical kinds - such as proteins - can also be characterised in terms of evolutionary or etiological properties. This adds a third cluster of properties that should be considered in relation to biochemical kinds, as already mentioned in Chapter 5, Chapter 6. Chemical structural properties are mostly those shared by organic compounds (which are often identified with biochemical kinds), and they include being composed of carbon atoms covalently linked to other elements. Functional properties refer instead to having a role within biological processes and causally contributing to them [Chapter 5, Chapter 6]. While biochemical kinds seem to display this combination of properties, the relation between them is far from simple, and this challenges their understanding and status as natural kinds. On the one hand, they have properties typical of chemical kinds; on the other, they have properties typical of biological ones. Moreover, the structural and functional properties display multiple realisability and multiple determinability, as explored throughout Chapter 6. This is summarised in the previously discussed "relation problem" (or Bio-chem problem, following Tahko 2020) that questions the relationship between the biochemical function of a biochemical molecule and its microstructure.

In the thesis, I have explored the complexity of the relation between function and structure considering the molecular gene, which displays multiple realisability together with the DNA displaying multiple determinability, and vitamin B12, which can be realised by (at least) four different vitamers. In the previous chapter, I have provided an answer to the relation problem with the aid of weak emergence [Chapter 6]. Nevertheless, the complexity of the relation between function and structure still challenges the natural kindhood of biochemical kinds and the way

---

[4] From website of the *Biochemical Society*: `https://biochemistry.org/`.

they can be organised and arranged in taxonomies. Let us consider briefly a couple of examples [already introduced in Chapter 6] to remind the reader about the tensions concerning the properties typically ascribed to biochemical kinds.[5]

The first example, extensively discussed by Tahko (2020) and presented in Chapter 6, is the one of haemoglobin. Haemoglobin is a kind of protein with the function to bind and release oxygen. This protein can be formed by at least two different amino acid chains (and actually many more) that fold in similar ways so that they can play the same function. The biochemical functions of haemoglobin are multiply realisable, as the function of binding and releasing oxygen is realised by at least two distinct macromolecules that present some microstructural differences. Another interesting example is the case of multifunctional proteins or moonlighting proteins such as crystallines discussed by Tobin (2010) and Bartol (2016). Crystallines are structural proteins that are present in all vertebrates' eye lenses, but they can also have an enzymatic role in digestive processes. This makes crystallines a case of multiple determination, as we have the same series of amino acid chain that can lead to very different functions. These phenomena generate the tension between the chemical and the biological approach to biochemical kinds and affect how we organise our biochemical knowledge and classification.

What is the kind of these biochemical molecules? If we were to favour the structural component of the kind, then we would deem biochemical molecules such as haemoglobin as chemical kinds - but, in this case, we would have two different kinds characterised by different structures; conversely, if we were to favour a functional or biological approach, we would get one kind in the case of haemoglobin, but two in the case of crystallines or moonlighting proteins. None of these approaches seems to be satisfactory - also because the two properties, as already discussed, are related to each other in an important and relevant way and this is acknowledged also by scientific practice (Havstad 2016, 2018).[6] Moreover, a satisfactory answer to these problems is made more complicated by the fact that the inquiry into the kind of biochemical molecules can be seen as comprised of two different, even if related, philosophical issues. The first is what I call

---

[5] These examples are discussed also in Bellazzi (2022b): "The superpowers of proteins", `https://www.jargonium.com/post/the-superpowers-of-proteins`.

[6] The analysis of why these accounts are not satisfactory will be presented in more detail in section §3.3.

the "problem of definition". This asks about the definition and identification of a biochemical kind as such. The second is what I call the "problem of taxonomy". This asks about which approach should be favoured in taxonomic systems for biochemical molecules. Let me present and discuss these in more detail.

### 7.2.1 The problem of definition

The first problem concerning biochemical kinds is the one that inquires into whether there is something such as a *biochemical kind*. As mentioned, biochemical molecules display two families of properties: chemical structural and biological ones (functionally related and etiological). And while it seems simple to identify the type of given molecule, for instance haemoglobin, there are difficulties in deciding whether these molecules should be considered chemical or biological or what it takes for a kind to be biochemical and how this can be defined.[7] Specifically, we can frame the "problem of definition" in terms of the following questions:

**Problem of definition**: Can we define something such as a biochemical kind?

From this, two questions can follow:

- Are biochemical molecules just instances of chemical or biological kinds?
- What does it take for a kind to be biochemical?

The answer to these questions is made difficult by the complexity of the relation between the two types of properties. At least four options can be explored to define biochemical kinds, as I will discuss later in §3.3. The first is the "dual option", presented by Bartol (2016): in every biochemical compound there are two kinds instantiated, a chemical kind and a biological one. This gives us a dual view of kinds, and neither is biochemical. The chemical kind is defined by the chemical structure, while the biological one by the function. For instance, in the case of every haemoglobin molecule, there are two kinds of kinds: the chemical kinds (given by the

---

[7] Havstad defines three classificatory practices concerning biochemical kinds: classificatory characterisation, individuation and organisation (2021, 7670). The first - which is the one we are concerned with -focuses on the definition of the kind. The second focuses on identifying which tokens belong to a given kind. The last focuses on organising taxonomies.

different amino acid sequences) and the biological kind (given by the function). The second is a reductionist approach, suggested by Goodwin (2011) and discussed further by Tahko (2020). According to a reductionist account, biochemical kinds are just chemical ones because their function can be reduced to chemical properties. In this case, haemoglobin is not a biochemical kind but a chemical one (or more different chemical ones), defined by given chemical structural proteins. The third one is to take the opposite direction and argue that biochemical kinds are just biological kinds in that they display a given biological function. According to this view, haemoglobin would be one biological kind identified by functional properties. The latter option is to consider a form of pluralism and assess each kind discussed by biochemistry on a case-by-case basis deciding which properties are relevant. In all these cases, the relation between chemical structure and biochemical functions discussed in the previous chapters make these answers more complicated. The answer to the problem of definition will be the main concern of this chapter and will be inquired by asking whether the category "biochemical kinds" can be considered a natural kind and upon which properties.

### 7.2.2 The problem of taxonomy

The second problem that stems from an inquiry into biochemical kinds is the one that regards their organisation and classification. This is what we can call the "problem of taxonomy":

**Problem of taxonomy**: How can we organise biochemical compounds in taxonomic systems?

From this, two questions can follow:

- Which taxonomies should we prefer, those that favour chemical or functional properties?
- How do we classify these kinds?

The problem of taxonomy can take even more complex forms as it relates to distinct classification practices. As summarised by Havstad (2016; 2021, 7670), it is important to distinguish three different aspects of taxonomic classification: a) classificatory characterisation or definition,

b) individuation and c) organisation [as in Chapter 4]. The first focuses on the definition of a specific (biochemical) kind, for example, what it means for something to be a gene or a vitamin B12 - as discussed in Chapter 4 regarding the molecular gene. This implies offering a definition for which something can be ascribed to the given category considered. The second focuses on identifying which tokens belong to a given kind or type - for instance, we can individuate whether a given biochemical compound is an instance of vitamin B or B12. The last focuses on organising the various kinds in distinct taxonomies and arranging different biochemical kinds into hierarchical classification schemes or taxonomies. While the problem of taxonomy is important and interesting for scientific practices, I will not be focusing on this for most of the present chapter. This is because taxonomic practices also depend on pragmatic considerations, and accounts that favour different properties can be maintained as equally valid. Nevertheless, taxonomies can be informed by the answer to the problem of definition, as will be briefly discussed in section §5.

## 7.3 Kinds of kinds

Despite its complex nature, the biochemical realm retains great scientific utility and underpins the understanding of biological beings. There is a consistent talk of biochemical kinds and usage of this category. This supports inquiring into the problem of definition and asking whether biochemical molecules are instances of chemical or biological natural kinds or, instead, they instantiate biochemical kinds, which could be considered a natural category. To answer these concerns, I will explore whether biochemical molecules are instance of which kind of *natural* kinds.[8] In this section, I will first summarise the account of natural kinds already presented in Chapter 2 and Chapter 4 (and as in Khalidi 2013, 2018, 2020). Then, I will introduce chemical and biological kinds and consider different options for the kindhood status of biochemical kinds. As will follow, these approaches to the naturalness of the biochemical world do not capture its complexity. This will motivate the consideration of biochemical kinds as natural kinds *per se*.

---

[8] This relates to what in Chapter 2 and Chapter 3 has been presented as the naturalness question, which inquires into how a category can be considered natural.

### 7.3.1 Natural kinds

As already summarised in Chapter 2 and Chapter 4, a starting point for discussing the naturalness of a category is a form of naturalism: to understand which kinds are natural, we should start by looking at those kinds postulated by different sciences.[9] This, however, is not enough on its own to capture natural kinds. Even within the best scientific theories, the history and philosophy of science have shown that not all categories present in the discipline can be considered those carving reality at the joints. For instance, a category like hysteria, which has been used as a scientific category in the past, has proved not to be a natural kind of disease and was abandoned as a kind (Khalidi 2013, 59). Other examples are phlogiston or the medical humors. Moreover, some categories can have an instrumental role or cannot be considered stable or robust enough to be really informative about the world. Accordingly, Khalidi has introduced two further requirements for naturalness: i) genuine projectibility and ii) being a node in a causal network (Khalidi 2013, 2018, 2020, 2021). Let us summarise them briefly again.

First, natural kinds should be projectible, in the sense that they can be projected successfully from one instance to another, having an explanatory and predictive power across contexts (Griffiths 2004; Tobin 2013; Khalidi 2013; Crane 2021). This means that it is possible to predict which properties will be present in different instances and project them from one instance to the other. Projectibility leads us to consider a second aspect of the theory: these successes appear to be a "reflection" of the causal network in which instances of the kinds are involved. The success of some kinds categories is so because the properties of the natural kinds are causally clustered (Khalidi 2013, 2018, 2020). The joints that natural kinds carve so successfully are those that can be found in the causal structure of the world. The relation between natural kinds and the causal structure of the world is dual: i) these kinds present a set of projectible properties that are hierarchically ordered as "causes and effects in recurrent causal processes" (Khalidi 2018); ii)

---

[9] Often, this is combined with a form of weak or moderate realism (r) for which kinds are objective features of reality but do not necessarily correspond to distinct metaphysical categories such as universals. This means that nature has some joints - objective features of the world (whether we know them or not) - and our best theorising should aim at carving these joints. However, this should be considered as an answer to the ontological question rather than the natural one, and, as such, it will not be considered in this chapter [more on this in Chapter 2].

these kinds themselves play a crucial role in the causal structure of the world: they are causally efficacious on other kinds. More specifically, it is possible to identify a "cluster of core causal properties" that cause the other properties associated with the kind. These properties represent some criteria to distinguish natural from conventional kinds. Following this account, natural kinds can be defined as those categories with a set of properties discoverable by science and whose co-instantiation causes the instantiation of other properties together with being nodes within causal networks (Khalidi 2013, 2018).

More can be said about the properties that these kinds share, following a distinction always presented by Khalidi [Chapter 2]. For some kinds, we can identify and define clearly some necessary and sufficient properties for their membership - these kinds are those that can be defined "monothetic kinds". In more detail, a monothetic kind is one "associated with a property or set of properties of which is singly necessary for membership in the kind and all of which are jointly sufficient" (Khalidi 2013, 16). Other kinds do not satisfy this condition because we cannot identify a clear set of necessary and sufficient conditions in every instance and, as such, they are called "cluster kinds" or "polythetic kinds". As a result of this distinction, instances of monothetic kinds will possess all and only the same properties in quality of being members of a kind. In contrast, instances of a polythetic kind might possess different properties despite being instances of the same kind. Examples of monothetic kinds normally come from fundamental physics; for example, only members of the kind protons have the properties "*positive charge 1.6 x* $10^{-19}$ *C, mass of 1.7x* $10^{-27}$*kg, and spin 1/2*". On the side of polythetic kinds, biological species can represent a good example, as it is difficult to identify a well-established set of necessary and sufficient properties that all and only the instances of these kinds share.

Once we have stated the conditions for natural kinds in general, we need to consider what are the kinds that can be instantiated in biochemical molecules. The debate has seen them as either chemical or biological, according to which properties we favour (Slater 2009; Tobin 2010; Goodwin 2011; Havstad 2016, 2018). Bartol has favoured a dual theory of kinds for which biochemical molecules instantiated both kinds, chemical and biological (2016). While these different authors subscribe to different views of natural kinds (and these are different from Khalidi's), the properties that are associated with chemical or biological kinds remain constant in the debate, and there is a tendency to reduce or consider biochemical kinds in these two frameworks. But why is this the case? What is normally taken to confer chemical kind membership and biological one? In order to answer these questions, let us consider in detail kind ascription in the chemical and biological case.

### 7.3.2 Chemical kinds and biological kinds

Chemical kinds have been considered paradigmatic cases of natural kinds (Ellis 2002; Serri 2020). These kinds display a series of stable properties across instances and allow their successful usage in different sciences. Chemical kind membership is often conferred by various forms of microstructuralism for which a given chemical kind is identified by the microstructural properties that its instances display (Hendry 2006; Slater 2009; Goodwin 2011; Tobin 2010; Havstad 2018).[10] These can include all the microlevel properties deemed satisfactory for a given molecule to be chemical. Chemical kinds can also display a level of "messiness" in terms of the relevant properties needed to identify them (to say it with Havstad, 2018), but even in more complex cases - such as those of compounds - microstructural properties are often deemed sufficient to define and identify them. Moreover, microstructural properties are those that allow the consideration of chemical kinds as natural, according to the definition presented above.

First, the characterisation of chemical kinds in terms of microstructural properties is present in contemporary chemistry, respecting the naturalistic criterion. Second, microstructural properties allow the projectibility of these kinds: we can infer the properties of one instance from another in virtue of the presence of some core microstructural ones. Chemical kinds are also nodes in causal networks as they have a causal impact on the world around them: a particular microstructure allows them to enter into chemical reactions and provides them with the properties characteristic of the various kinds. A simple example to further understand the naturalness

---

[10] An alternative account of chemical kinds can be seen in the work of Needham (2000, 2010) and LaPorte (2003) that suggest considering also macroscopical properties for the identification of chemical kinds, such as thermodynamic properties (Serri 2020). Here, I focus on microstructuralist accounts because these are the ones used to ascribe chemical kindhood to biochemical molecules.

of chemical kinds are chemical elements, such as Lithium. Lithium is a chemical element with precise microstructural properties, such as having atomic number 3. This core property allows making precise projections about instances of Lithium, expecting particular properties and behaviours in virtue of such structure. Moreover, having the microstructure characterised by the specific atomic number causes the instantiation of other properties crucial for the kind and its interaction with the surrounding world.[11]

Conversely, biological kinds are mostly individuated by evolutionary history (Khalidi 2013,2021; Bartol 2016). Specifically, the members of a given biological kind are often considered so because they share some diachronic properties (causal, phylogenetic descent or causal history) and not because they possess the same synchronic causal properties or microstructural properties. Accordingly, these kinds are also referred to as etiological or historical kinds: their evolutionary or causal history makes the instances of the kind they are.[12] For example, the species *Panthera tigris* is identified in virtue of a specific evolutionary history of its members as they result from similar selection processes. Also, in the case of biological kinds, we can easily see that these kinds are discussed by science, respecting naturalism. Moreover, the historical or etiological properties of these kinds make them projectible and nodes in the world's causal structure: the instances of biological kinds have given properties and causal powers thanks to their evolutionary history. For instance, *Panthera tigris* instances have properties that can be inferred from one instance to another, thanks to a specific causal evolutionary history. We can project the properties of one tiger to another once we consider their evolutionary history, and the evolutionary history can be seen as the core causal property for which other properties of the kind are instantiated.[13]

---

[11] For more on why some chemical kinds are natural kinds, see chapter 5.2 on *Lithium* in Khalidi 2013. Moreover, it is important to notice that for compounds, microstructuralism can be combined with forms of macrostructuralism, and these kinds can present different types of properties, as in Needham (2000, 2010) and Havstad (2018).

[12] Whether *all* biological kinds are historical or etiological kinds is an issue that is open for discussion; for example, Khalidi 2013 discusses the example of eyes and points out that sometimes there is an interest also in the present causal powers of a given biological kind and not only in evolutionary history. Nevertheless, at least for the majority of kinds, there is a consensus that biological kinds are individuated mostly evolutionarily. Moreover, it is important to point out that biological kinds are not the only etiological or historical kinds, as we find instances of such kinds, for example, in geology.

[13] For an extensive analysis of why historical and etiological kinds can be considered natural kinds, see Chapter 4 of Khalidi 2013 and Khalidi 2021.

While chemical elements and biological species can be seen more easily as members of either chemical or biological kinds, it is more difficult to understand and identify the kind of biochemical molecules. As Havstad (2016) points out, in the case of macromolecules such as proteins, scientists use a combination of functional, structural and historical properties in biochemical taxonomies. Are then biochemical molecules instances of chemical or biological kinds? Or can they be considered instances of both? Or instances of neither, that is they are their own category?[14]

## 7.4 Some kinds options for biochemical molecules

In order to answer the problem of definition about biochemical molecules, we can explore whether they can be defined in terms of the instantiation of chemical or biological kinds. I do so by considering the previously introduced case of crystallines, a moonlighting protein that has a functional role in both sight and digestive processes (Tobin 2010; Bartol 2016).

### 7.4.1 Biochemical kinds as chemical kinds

Let us start by exploring whether biochemical molecules can instantiate chemical kinds. Some biochemical molecules, such as proteins, could be considered chemical kinds because of the importance of some microstructural properties. This would lead to a reductionist approach for which biochemical properties can be reduced to structural properties or for which structural properties are the most stable ones. Crystallines could be deemed chemical kinds because it is in virtue of their molecular structure that they are the kind they are, and - if the reductionist is right - then they would play their functions in virtue of such structure. This position has been defended for instance by Goodwin (2011). He points out that the protein's primary structure provides crucial

---

[14] One could take a reductionist stance here and claim that the only genuine kinds are physical kinds, so, strictly speaking, biochemical kinds would not be chemical nor biological because they should be reducible to physical kinds. While this option should not be ruled out in principle, here I focus on whether biochemical kinds can be chemical or biological kinds because chemical and biological properties are those relevant for their classification. Moreover, it seems that if the goal is to rule out both biochemical and chemical kinds, we should proceed in reduction by considering first whether biochemical kinds are chemical, and then by reducing the chemical to the physical. This still needs to be properly explored in the literature.

information for protein identification and its core properties. To understand his argument, let us summarise the four-levels structure of proteins. Proteins have a "primary structure", which is just the linear sequence of amino acids. They then present a "secondary structure", which refers to stable geometrical patterns localised in some parts of the molecule. Lastly, proteins present a geometric and full structure referred to as "tertiary and quaternary structure". The quaternary structure is linked to the protein's function and provides information about what the protein does.

Given the complexity of the relation between structure and function, Goodwin suggests that the primary structure of a protein (the linear amino acids sequence) can be considered more important than other properties or configurations (2011). This is so because, he argues, the primary structure is easily identifiable, plays a stable role in scientific practice and has an "overall robustness" of primary microstructural properties compared to other macroscopic properties. Primary structure can be seen as displaying a form of ontological robustness in that it remains constant and stable for the protein, while its higher-order structures and functionality might vary. If so, biochemical kinds such as proteins would be chemical kinds rather than biological ones. In the case of crystallines playing different functions, they would be identified as one protein because they share some microstructural similarities. The primacy of the primary structure of a protein compared to the others needs to be further established, but nevertheless, one can still consider the chemical microstructural properties of a biochemical kind as what defines it. A similar argument is put forward by Tahko (2020), who also underlines the importance of the microstructure of proteins for explanatory purposes. Specifically, he argues that some changes in the microstructure can impact other properties of the molecule while leaving its function unchanged. For instance, there could be small changes in the microstructure of crystallines that could impact some of their properties while still allowing for the main function to be played. Accordingly, while he does not take a definitive position, he gives hope to the reductionist or microstructuralist in identifying the properties crucial for biochemical molecules with structural chemical properties. If the chemical structural properties are really the most stable ones both in terms of structure and in terms of causal impact on other properties, then we could consider biochemical molecules to instantiate chemical kinds.

However, the reductionist or chemical approach to biochemical kinds encounters two main problems. Following Havstad (2018), the first one is that microstructural properties are only *some* of the ones relevant for biochemical kinds because functional and historical or etiological criteria come into play. Moreover, such functional properties are biochemical functions and not only chemical functions, as already discussed in Chapter 5. Second, the reductionist has the burden of proof of showing that the other properties of biochemical kinds can be reduced to chemical properties (as also discussed by Bartol 2016). The reduction of biochemical functions to chemical properties is not straightforward and it seems that the consideration of the chemical properties would give a partial story for which crystallines are what they are and play the role they do. As discussed in the previous Chapter 6, the relation between function - crucial for identifying the kind - and structure is better accounted for via weak emergence.[15] In this holds, biochemical molecules do not seem to be instances of chemical kinds - as some of their definitional properties are not only chemical microstructural ones.

### 7.4.2 Biochemical kinds as biological kinds

Let us now consider whether biochemical molecules can be seen as biological kinds. This hypothesis would imply that the functional and the evolutionary properties of biochemical molecules are those most relevant to their natural kind status. This hypothesis has been suggested by Bartol (2016). He argues that given that it is difficult to understand which microstructural properties are the most stable in proteins, one could consider alternatively their evolutionary history and the functions conserved in evolution. This would make biochemical molecules instances of biological kinds. Specifically, proteins within organisms can be seen as displaying a form of evolutionary history because of the functions for which they have been selected. For instance, crystallines can be seen as more than one kind whose functions have been conserved in

---

[15] A further challenge to the reductionist strategy could come from the context sensitivity of some of the processes that are involved in protein folding and how the functionality of the protein can be affected by it. While this might support the thesis that the primary structure is the unchanging property of the protein, it might also show its insufficiency in accounting for all the relevant properties that identify the biochemical kind.

evolution. This evolutionary history would be the most stable property, being the causal root for the different functions of crystallines in both sight and digestive processes. In this framework, the evolutionary properties of a protein are the most stable ones as "through all of this change [different levels of structure and potentially different functions in different contexts], the closest thing to a constant is the biological protein's historical trajectory" (Bartol 2016, 546 - insertion added). In this perspective, the chemical microstructures of the various biochemical molecules and the functions played would be only the result of the "latest stage" of the evolutionary history and dependent upon it. If we grant this as true, then biochemical molecules can be considered biological kinds in that they are *evolutionary kinds*. The biological characterisation might be then the one prevailing. Crystallines would be like the species *Panthera tigris* whose properties depend upon core evolutionary ones.

However, also this approach has some shortcomings. First, the structural and functional properties of biochemical kinds are important for their understanding, they are related, and cannot be considered only in terms of evolutionary history and selection (Havstad 2016, 2018). This is so because many molecules could share a similar evolutionary history but have structural properties that affect their current functions, as underlined by Bartol (2016) and Tahko (2020). Second, the relation between the biochemical functions and the evolutionary history is complex, as biochemical functions do not seem to be straightforwardly evolutionary functions (as discussed in Chapter 5, Chapter 6). In this respect, it is difficult to see why biochemical molecules should be primary evolutionary kinds while they display a complex cluster of different properties, of which none seems prevailing for their identification.

### 7.4.3 Biochemical kinds as dual kinds

Another option for biochemical molecules that combines the two approaches above can be seen in the dual approach proposed by Bartol (2016). He argues that biochemical molecules are both chemical and biological, given that both approaches can be considered satisfactory. These molecules display structure, function and a form of evolutionary history. He considers this as a sign of a *duality* of kinds: biochemical molecules instantiate two kinds - none of them is *biochemical* - the chemical and the biological one. The instantiation of these kinds is disjunctive and can lead to contrasting taxonomies according to the kinds we consider for biochemical characterisation. Crystallines, in the example, would instantiate a chemical kind because of their structural properties and an some evolutionary biological kinds because of their evolutionary history.

While this approach might seem appealing - as it considers, in a way, the complex nature of such molecules - a dual approach to biochemical kinds does not seem satisfactory either. As Havstad points out, biochemical kinds are not individuated simply from a microstructural perspective, or an evolutionary one, or a dual approach (2016, 2018). They are individuated thanks to a *combination* of different approaches which *include* microstructural, functional and evolutionary criteria and these aspects are related to each other. This comes from what she calls "inferential" reasons and "descriptive" reasons. The "inferential" reasons amount to the fact that the same object, in this case, proteins, can do different things in different contexts. Accordingly, the fact that these molecules present different properties should not lead us to postulate two distinct kinds automatically to account for them. In the case of crystalline, the same protein can be inferred to be part of the same kind according to its structural, functional and evolutionary properties without the need to postulate different kinds of kinds. "Descriptive" reasons are based on the fact that in actual scientific practice, biochemical kinds are identified using a series of different criteria that do not prioritise microstructure, function, or evolutionary history. These properties have a similar weight and are related to each other.

### 7.4.4 Pluralism

A latter option is still available: each instance of biochemical molecule can be given a certain kind membership on a case-by-case basis, following a form of pluralism as suggested by Slater (2009). According to pluralism, various classifications are possible following the different inductive and explanatory practices, and we should assess which one is relevant when considering the different cases. While this approach might seem tempting, there are three main considerations that one can put forward.

The first is that the problem of definition (that considers what biochemical kinds are and if they are natural) is different from the problem of taxonomy (that considers how biochemical kinds are organised in taxonomies), as will be discussed in §5. This would allow us to retain a pluralist taxonomy while maintaining a unified or monist account of biochemical kinds. The second is that pluralism does not seem to capture biochemical practice and the considerations above that stress the link between structure, function and evolutionary properties, as underlined by Havstad (2016, 2018) and Bartol (2016) [Chapter 5, Chapter 6]. Havstad, in particular, points out how there are systems to type proteins in a unified way and these support the need for a monist account of biochemical kinds (2016, 2018). The third is that biochemistry is still in its nascence and quickly growing, so there might be cases of kinds discussed by this discipline that can be identified with either chemical or biological kinds, or whose kindhood membership should be considered individually. For instance, one might deem some minerals that play a role in physiological processes, such as zinc, as chemical kind instead of biochemical one. In this respect, it is important to keep open the possibility that the process of identification of the relevant kinds is fallible and can comprise different kinds of kinds. Nevertheless, this does not exclude the fact that there might be some cases, as will be discussed in the next section, in which there are natural biochemical kinds.

In conclusion, biochemical molecules do not seem to instantiate chemical, biological, or both of these kinds. This result is further supported by what has been explored in Chapter 5 and Chapter 6, for which biochemical functions are an important component of biochemical molecules. However, they are neither straightforwardly evolutionary nor chemical functions. In the next section, I will argue that biochemical kinds could be considered natural categories themselves.

## 7.5 Biochemical kinds

In the previous section, I have discussed whether biochemical molecules could be considered instances of either chemical, or biological kinds or an instantiation of both kinds. However, these approaches do not appear adequate to characterise the complexity of the biochemical world. This is because there is a strong relation between the different properties of the biochemical molecules [Chapter 6]. The options are then to a) disregard biochemical molecules as instances of some genuine natural kindhood - or to wait for one of the approaches to be proved satisfactory - or b) to explore whether biochemical molecules could be instanced of *biochemical kinds*. This latter hypothesis would be supported by the moderate naturalism assumed in this thesis: the interest in the biochemical sciences has been growing in recent years, and these disciplines provide successful explanations and predictions. Moreover, looking at how scientists characterise biochemical kinds, it is possible to identify some features that the instances of these kinds share: membership is conferred by an assortment of properties that include microstructure, macromolecular superstructure, functions and etiology (as Havstad 2016, 3; 2018).

Let us consider the properties characterising these molecules more in detail in order to understand whether we can accept the naturalness of biochemical kinds. First, biochemical molecules present a chemical characterisation as organic compounds - chemical compounds that contain carbon and are involved or present within biological processes. Second, as follows from the definition mentioned above, their involvement in biological processes implies that they also display a biochemical function, as already discussed. Third, biochemical molecules seem to have some form of evolutionary properties, whether direct - evolving themselves - or indirect - concerning biological processes.[16] The cluster of these properties has been underlined when considering the molecular gene, vitamin B12 and some of the proteins cases (mostly discussed in the literature). Summarising, biochemical kinds share the following properties: i) their instances display a chemical structure and some chemical properties; ii) they have a biochemical function; iii) there are some etiological considerations that need to be made in relation to such structure, whether they are direct or indirect [as seen in Chapter 5, Chapter 6] (see also Slater 2009; Goodwin 2011; Tobin 2010; Havstad 2016, 2018; Tahko 2020). The next question asks whether the category

---

[16] It is important to notice that evolutionary considerations are relevant, but they are combined with structural and functional ones. Cases of molecular convergent evolution further support this. In convergent molecular evolution, such as the one of the genes codifying for haemoglobin proteins in different mammals, the proteins have a different evolutionary history but share some functional similarities that allow for their common classification (for more on this see “Predictable convergence in hemoglobin function has unpredictable molecular underpinnings” by Natarajan et al., 2016). I thank Prof. Tabler for this suggestion.

"biochemical kind" is *natural* on the basis of these properties.

In order to answer it, let us recall the considered account of biochemical kinds [Chapter 2]. For a kind to be natural its properties should be present in scientific theories and i) be projectible; ii) nodes in causal networks. We can start with the first, which is easily respected. Biochemical kinds are kinds postulated in an efficacious scientific discipline - biochemistry. Moreover, scientific practice, together with the work of some philosophers, tells us that these kinds have the following properties: microstructural and macrostructural properties, functions and a form of etiology or evolutionary history. Granted that biochemical categories respect the naturalness criterion, we should explore whether they respect the other two. The success of the biochemical kind category offers us support for proving the projectability of the category. Specifically, considering that some macromolecules display specific structural properties and functional roles in biological processes improves predictions and explanations of different phenomena. Once we identify molecules such as proteins, vitamins or nucleic acids, we can project some specific chemical structural properties from one instance to another. Furthermore, the knowledge or consideration of the evolutionary history of the molecules or the processes to which the molecules contribute would allow identifying which chemical properties are relevant for the various processes and systems under consideration. This projectability is possible because the three core properties of biochemical molecules cause the other properties they display, such as the causal impact they have on the various organisms or the way they interact with each other. Moreover, it is possible to identify the relation between these properties: the set of powers or chemical dispositions associated with the biochemical function is a proper subset set of those of the chemical structure. This set is identifiable via evolutionary history of the biological processes to which the function contributes [Chapter 5, Chapter 6]. Accordingly, it seems that the properties typical of biochemical kindhood respect the requirements for naturalness.

One might now wonder whether what makes these kinds projectible and nodes in causal network are actually properties coming from their being biological and chemical kinds. This would mean that there is no novelty to their naturalness and they are just a conjunction of the two kinds. In order to answer this concern, it is relevant to consider the status of biochemical functions - presented in the previous two chapters. As argued, biochemical functions are neither chemical nor biological and are weakly-emergent from the chemical structure [Chapter 3, Chapter 5, Chapter 6]. This provides a further ground for the distinction between biochemical kinds and chemical ones. Nevertheless, biochemical functions can be ascribed to kinds we might not want to accept as biochemical [as mentioned in Chapter 5 §5.2]. This supports a more detailed analysis to establish whether the category "biochemical kind" is natural. Let us do so by considering two good candidates for biochemical kinds: vitamin B12 and the molecular gene.

### 7.5.1 Some biochemical kinds: vitamin B12 and the molecular gene

Vitamin B12 and the molecular gene are good candidates for being biochemical kinds because they display the three relevant properties for biochemical kindhood. In this section, I am not going to consider whether these two categories *per se* represent a natural kind (the question regarding the molecular gene has been considered already in Chapter 4), but rather I am going use them as examples in which the category "biochemical kind" can be seen to be natural. As from the previous section, biochemical kinds are projectible categories that play a causal role, and they present three sets of properties that should be linked causally: structural, functional and some evolutionary properties.

As discussed in previous chapters, vitamin B12 is a vitamin that acts as a specific co-enzyme in different biological processes. This vitamin can be composed of four vitamers, all cobalamin compounds able to display the specific biochemical set of biochemical functions relevant to the kind. Vitamin B12 represents a good candidate for being a biochemical kind as we can identify the three properties that characterise biochemical kindhood. First, the chemical component of vitamin 12 is easily identifiable as we can identify at least four varieties (hydroxocobalamin, adenosylcobalamin, cyanocobalamin, methylcobalamin). These compounds all present the biochemically rare element cobalt (chemical symbol Co) positioned in the centre of a corrin ring, a planar tetra-pyrrole ring [Figure 9[17]]. Accordingly, the four vitamers have a similar chemical structure.

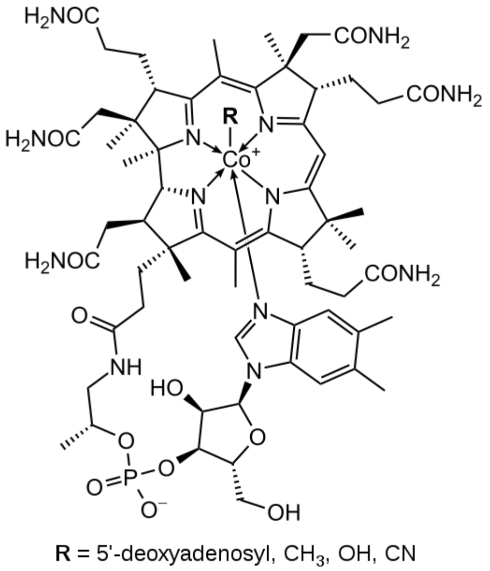

Figure 9. Generic cobalamin molecule, presenting the corring ring in the centre. Image from Wikimedia Commons (Hbf878), Public domain (2017).

Second, vitamin B12 has a series of biochemical functions associated with chemical dispositions to contribute to specific biological processes, such as erythropoiesis. As discussed in Chapter 5 and Chapter 6, the chemical and the functional component of vitamin B12 are related. The chemical structure is the basis from which the biochemical functions of vitamin B12 weakly emerge because these functions can be associated with a proper subset of the chemical powers of the cobalamin compounds. Vitamin B12 displays two of the features of biochemical kinds, and these two features are related. Moreover, the vitamers have a specific chemical structure, being cobalamin compounds, and they have the same function, for instance acting as co-enzymes in hematopoiesis. These two features can also be deemed a node in different causal networks involving the biochemical reactions in which vitamin B12 operates. Third, following biochemistry practice and the considerations mentioned in the previous chapters, etiological parameters should also be acknowledged (Havstad 2018). The identification of relevant etiological parameters is a complicated issue and can be assessed differently for each instance of biochemical kinds, as they

---

[17] Image from Alsosaid1987, Wikimedia Commons (Hbf878), Public domain `https://commons.wikimedia.org/wiki/File:Cobalamin_general.png` (2017).

seem to display evolutionary history in forms that are more or less direct. In the case of vitamin B12, the relevant etiological parameters take two forms. The first is an indirect one: the set of powers associated with the biochemical functions is indirectly evolutionary selected [Chapter 5, Chapter 6]. The biochemical function of vitamin B12 manifests itself because of the evolutionary history of the processes in which it takes part. The second regards whether the instances of the kind considered are or could be bio-synthetised and have a more direct evolutionary history. In the case of vitamin B12, this latter is respected because its vitamers are bio-synthetised by bacterias and archea that use them as co-enzymes in vital processes. Accordingly, one could track the evolutionary history of its synthesis. It also seems possible to individuate an evolutionary history of why this molecule is needed in humans' physiology in relation to its synthesis within bacteria. Whether direct or indirect, it is possible to identify a causal evolutionary history of vitamin B12. Accordingly, this kind seems to be a *biochemical kind* because it has the relevant properties.

Moreover, vitamin B12 as a biochemical kind shows also how the properties of biochemical kinds can be projectible and nodes in causal networks. Specifically, the fact that the four vitamers have a similar structure and function makes it possible to infer and project the properties from one instance of vitamin B12 to another. Considering the relevant evolutionary properties allows for inferring which functions will be played by the vitamin in different organisms and processes and identifying possible dysfunctions - as discussed in Chapter 5. This projectibility is made possible by the causal relations between these core properties of the kind and other properties that the kind display. The relation between the chemical structure and biochemical function is the one mentioned above - functions can be associated with a set of chemical dispositional properties of the cobalamin compounds. These properties are also related to evolutionary historical properties: the set of chemical powers associated with the biochemical functions is indirectly evolutionary selected. These properties determine the causal impact of these kinds on the surroundings: these features make biochemical molecules nodes in the causal network, e.g., co-enzyme in different processes such as erythropoiesis.[18]

---

[18] More will be said on this in the next section.

The case of the molecular gene and how these properties are projectible and causally related has already been discussed in detail in Chapter 4. Nevertheless, I will stress here that the properties definitional of the molecular gene are also suitable to make it a good candidate for being a biochemical kind and showing that biochemical kinds are natural. To start with, the molecular gene is discussed in different scientific disciplines and in molecular biology, a part of biochemistry and accordingly respects the naturalist requirements. Moreover, it has two main properties that can be identified i) it is composed of nucleic acids (DNA or RNA) ii) it has the function of being transcribed, which is associated with a specific subset of causal powers of the nucleic acids (Bellazzi 2022a). The etiological or evolutionary properties of the genes are more easily identifiable compared to those of vitamin B12. The evolution of nucleic acids within biological organisms and, specifically, of genes has been discussed in the literature, to the point that genes have been considered primary units of selection (see Lloyd 2017 for further debates). While the primacy of genes in evolution is not relevant for present purposes, we can still accept that it is possible to identify an evolutionary history of the different genes. This evolutionary history has a role taken into account by the different gene taxonomies and genomic comparative studies [as in Chapter 4]. Accordingly, the molecular gene can be considered a good candidate for being a biochemical kind. In the gene case, one can also notice how these three sets of properties respect the requirement for natural kindhood. Both a molecular structure and a functional characterisation allow projecting the properties from one instance to another, and the evolutionary considerations permit us to do so in a relevant and specific way. Moreover, the relation between structure, function and evolutionary history supports the causal role that the molecular gene, as a biochemical kind, has in different biological processes, such as protein synthesis.

Before concluding, I would like to stress that the fact that vitamin B12 and the molecular gene are biochemical kinds does not exclude the possibility of them being instances of crosscutting kinds.[19] For example, vitamin B12 can also be considered from the chemistry perspective, thus being classified in virtue of its structure as the chemical kind "cobalamin". Nevertheless, if one wants to consider its biochemical nature, then vitamin B12 as a *vitamin* needs to be iden-

---

[19] Crosscutting kinds can be seen as instances of multiple kinds that would correspond to the same hierarchy level (Khalidi 2013, 69). More on this is presented in Chapter 2.

tified as an instance of a biochemical kind with related properties. Similarly, a given molecular gene can also be considered an instance of either a chemical kind from its material composition or a biological kind from its evolutionary properties. Nevertheless, in molecular biology and when considering protein synthesis, it is important to acknowledge its biochemical nature as a biochemical kind that presents properties that are *sui generis*.

### 7.5.2 Chemical kinds, polythetic kinds and evolutionary properties

A satisfactory account of natural kinds does not only need to be consistent with scientific practice but also needs to be able to capture the different instances in a specific enough way. I have already explored in Chapter 5 the objection for which the defended theory of biochemical functions[20] can be too coarse-grained. Let me recall the concern using the discussed example, the one of zinc (Zn). Zinc is a chemical element, and thus a chemical kind, very important for many different processes in life. Zinc is defined as “an essential trace element” necessary for “the normal growth and the reproduction of all higher plants and animals, including humans. In addition, it plays a key role during physiological growth and fulfils an immune function. It is vital for the functionality of more than 300 enzymes, stabilization of DNA, and gene expression” (Frassinetti et al. 2006). When considered in the biological environment, it seems justified to accept that zinc has biochemical functions because this element has a series of dispositional properties that contribute to biological processes. Furthermore, some traits or processes interact with zinc and have been evolutionarily selected. This makes the function of zinc a biochemical function according to the current account. However, one might be sceptical in considering zinc a biochemical kind.

The answer to this possible objection has been to distinguish biochemical functions from biochemical kinds and to allow non-biochemical kinds to have biochemical functions. Specifically, even if biochemical kinds need to have a biochemical function, more than having a biochemical

---

[20] Defined as a specific set of dispositional properties that manifest themselves in particular circumstances because of the evolutionary history of the traits or processes in which the biochemical molecule is involved.

function is needed for a kind to be a biochemical one. Is there a way to distinguish biochemical kinds from other molecules with biochemical functions? As discussed, biochemical kinds comprise an element of evolutionary history that can take a direct and indirect form. Molecular genes, vitamin B12, and proteins share two of their properties with zinc: a chemical composition and a biochemical function. However, these kinds - differently from zinc - also have a more direct evolutionary history.[21] If we follow the account of biochemical kinds proposed here, then zinc would not be a biochemical kind, even if it has a biochemical function. By contrast, vitamin B12 would be a biochemical kind with different biochemical functions because it is bio-synthesised by organisms, bacteria or archaea, and it plays a role in relation to cells or traits that have been selected in these organisms or others. The evolutionary history of genes and the related nucleic acids represent an additional potential source of debate, but their importance in evolutionary phenomena is well established (Lloyd 2017). Accordingly, having a relevant evolutionary history seems to be a further property for biochemical kinds compared to chemical kinds having a biochemical function. It remains that a detailed discussion of the evolutionary properties of biochemical molecules would require a specific analysis that considers what it means for a molecule to display evolutionary properties and a precise consideration of molecular evolution that we cannot explore in detail here [more will be said in Chapter 8]. For instance, one could question whether the evolutionary history of vitamin B12 is that different from the one of zinc. Ultimately, both can have biochemical functions in evolutionary selected processes, and the fact that vitamin B12 also happens to be bio-synethised might seem contingent (differently from the evolutionary history of genes or proteins).

Nevertheless, it remains that procaryotes are synthesising vitamin B12 and studies in ecology and evolutionary biology have underlined the relation between such synthesis and the need of these vitamins for different organisms. It seems possible to track a causal evolutionary explanation of the need for vitamin B12 humans (and other organisms) that considers the relevant connection between the evolution of humans, the relevant evolution of the procaryotes and their interaction. In this respect, the fact that vitamin B12 is bio-synthetised seems to play a rele-

---

[21] This does not imply that these kinds are *only* bio-synthesised as a result of selection, as many biochemical kinds can also be synthesised artificially.

vant difference in evolutionary interactions. The scientific details of this still need to be spelt out. However, it seems a safe assumption that the levels of interactions between interdependent organisms in ecosystems are varied and existent and this can impact our understanding of the biochemical world.[22]

Building upon this concern, one could ask whether a laboratory synthetised instance of a biochemical kind that does not display a direct form of evolutionary history or bio-synthesis (such as a laboratory synthetised cyanocobalamin molecule) should count as a biochemical kind and how strong is the condition. A way to answer this concern is to consider that the requirement of being bio-synthetised or displaying a form of evolutionary history concerns the kind and not every instance.[23] In the specific case, the instances of cyanocobalamin synthetised in laboratories are modelled to have a biochemical function in a precise way - copying those that are biosynthetised in nature by bacteria - accordingly, we could still deem them as instances of biochemical kinds. This is in line with the fact that biochemical kinds could be polythetic kinds - kinds for which we cannot identify a clear set of necessary and sufficient conditions for their membership in every instance. Accordingly, while on average vitamin B12 molecules display a form of evolutionary history, this could be different in some particular instances. In conclusion, this account considers three components that could allow the exclusion of cases that we would not deem as biochemical kinds.

## 7.6 The problem of taxonomy

As mentioned in §2.2, biochemical kinds also encounter what we have called the problem of taxonomy that asks how we can organise our knowledge of biochemical molecules in taxonomic

---

[22] An interesting study in this regard is the one discussed in "Insights into the Evolution of Vitamin B12 Auxotrophy from Sequenced Algal Genomes" by Helliwell and colleagues (2011). In this work, the evolutionary history of the auxotrophy (the inability of an organism to synthesise a particular compound required for its growth) of vitamin B12 in Algae is explored, and so is the importance of taking an evolutionary perspective in this respect.

[23] This could be further expanded if one would want to take a step further from moderate realism and have a stronger commitment on kind universals, as suggested in Chapter 2. While such commitment would allow the kind to have properties not shared in every instance, such requirements could also be accounted for by considering biochemical kinds as polythetic kinds.

systems. While this is not the main aim of this chapter, it is worth mentioning how the account above of natural kinds can aid the discussion. As summarised by Havstad, three different enterprises relate to kindhood and taxonomy in biochemistry (2016; 2021, 7670). The first is how to define specific kinds - such as vitamin B12 or the molecular gene; the second is the individuation of the instances of different kinds; the third is how we can organise the kinds into different taxonomies (Havstad 2016, 2021). The second problem seems scientifically uncontroversial, granted a given standpoint that considers structural, functional and etiological properties. The first and the second still need to be answered, given the complexity of the properties of the kinds (see Havstad 2014, 2016, 2018). In this respect, three approaches remain available, which mirror, in a way, the different answers to the problem of definition.

The first answer to this problem comes from the microstructuralist perspective that claims that the chemical structures of the considered macromolecules should play a leading role in taxonomy. Accordingly, the individuation of specific kinds and their classification should be done regarding chemical structure. A similar position is sustained by Goodwin, for whom the primary structure of the proteins is the most robust and stable criterion of taxonomical systems (2011). However, a first shortcoming of this position is that it does not easily apply to instances of multiple realisation widely spread across the biochemical world, such as the one of haemoglobin or the one of vitamin B12. As already discussed, if one would stick to microstructural considerations, the molecules composing haemoglobin would be organised into different taxonomies. Despite the similarities, there would be enough structural differences to consider these cases as different proteins (see Tahko 2020). The same can be said for vitamin B12. Vitamin B12 is realised by four chemical compounds (or four vitamers) that present a similar but slightly different chemical structure. They all share a cobalt component, but they also present other elements. If the microstructuralist is strict about her characterisation, then she could not cluster these four chemical compounds into the same taxonomy as they would be instances of different kinds - having different microstructures. Instead, the four vitamers of B12 are so because they share a function as co-enzymes of different biochemical processes. Moreover, at a higher taxonomic level, the vitamin B family comprises a family of macromolecules clustered together not only in terms of structure, but also in terms of origin and general biochemical functions. Accordingly, the microstructuralist characterisation does not capture contemporary taxonomic practices.

A second answer could come from the pluralist: given the contrast between the different taxonomies and the fact that they can be organised in different taxonomic systems, we should just take a pluralist approach towards this case, as suggested by Slater (2009). However, for what concerns the problem of definition, there are no competing taxonomies for individuating the different biochemical molecules in terms of either microstructure, macrostructure or function (following Havstad 2014, 2016, 2018). For example, there is a system for typing proteins, and there seems to be a system to type vitamin B12 as well. Accordingly, being pluralist for the problem of definition does not match scientific typing practice. The situation is different for higher-level taxonomies, such as organising biochemical compounds in terms of families or higher-level groups. In this respect, there might be systems in competition, where some might favour macroscopic properties, such as being water-soluble, and others functional characterisations or microscopic similarities.

A third approach could come from a biological characterisation, for which we consider biological properties of the molecule and its evolutionary history. In this case, we would define and classify biochemical compounds in terms of their function. However, there are two problems with this approach. The first is that the structural and microscopic properties of the different molecules are relevant for their understanding and classification - together with the other properties. The second, presented and discussed by Tahko, is that a biological characterisation can be too coarse-grained and does not capture some microstructural differences or modifications that can impact biochemical functions' role (2020). Moreover, the coarse-grained nature of some biological characterisation might then lead to different taxonomies or classifications led by interest relativeness (2020, 16). While interest-relativeness might not represent a problem *per se*, it does not aid an univocal answer for the problem of taxonomy.

Is there then an approach that we can consider more satisfactory than others? Taxonomic organisation is something mostly grounded in scientific practice. It involves a variety of parameters that can weigh different features and that can also be independent of our standpoints on natural kinds. For instance, taxonomies are compatible with nominalist or conventionalist views of kinds. Moreover, taxonomic systems have different utilities in different contexts and circumstances. My suggestion, which follows scientific practice and some philosophical discussions, is to treat the taxonomies of biochemical kinds considering their biochemical nature that includes all the relevant components. These components are not unentangled, as suggested in the previous section, but they are all related and so should be considered. Nevertheless, one could remain pluralist and instrumentalist and accept the best taxonomies for the different scientific practices in the given circumstances and settings.

## 7.7 Conclusion: unity and biochemical kinds

Biochemical molecules and the biochemical world represent a fascinating domain to explore topics concerning the unity of science, intra-level relations and natural kindhood. While in the previous chapters, I have explored the theme of biochemical functions and their weak emergence, here I have been focusing on considering their natural kind status. I have argued for the naturalness of biochemical kinds - which do not enter into the chemical or the biological cluster of kinds. The naturalness of the biochemical category supports the hypothesis that there is something specific about the biochemical interactions and kinds. Moreover, it shows that the same notion of natural kind can be applied to different categories, following Natural Kinds Monism (NKM), which can support unity. As presented in Chapter 2, NKM is "the view that there is a single notion of "natural kind" and anything falling under that notion can be defined in terms of the same general set of identity criteria." and this can ground the unity among sciences (Tahko 2021, 12).

This chapter presents a series of relevant results. First, it allows us to maintain a form of autonomy of biochemistry, offering an account of natural kinds adequate to the properties considered explanatory relevant for this discipline. The consideration of the problem of definition and an answer to it can lead to a better understanding of the biochemical world and can then guide a series of solutions to the problem of taxonomy. Second, this proposal presents a unified view of these kinds - they are not dual in nature - and suggests that structural, functional and evolutionary properties are linked together. Lastly, it shows how a single account of natural kinds can be applied to different instances, which might preserve a sense of unity of science. Here, I have explored how one account of natural kinds, the one presented by Khalidi, can be applied to different kinds of kinds (chemical, biological and biochemical) while retaining the specific features of each kind. This form of unity contrasts with arguments for the disunity of science that have been put forward together with a pluralist account of natural kinds, such as in Duprè (1993). Accordingly, a unified view of natural kinds can support a form of unity of science and unity between the chemical, the biological and the biochemical.

Concluding, Part III has argued that the biochemical realm presents weakly emergent biochemical functions and genuine biochemical natural kinds. This result supports the main thesis that there is something *sui generis* about the biochemical realm. This has been argued with two theoretical steps: defining the relation between biochemical functions and molecular structure in terms of weak emergence and arguing that biochemical kinds are natural kinds.[24] First, Chapter 5 offered a detailed account of biochemical and genetic functions, complementing the results of Part II. This has been the important starting point for assessing the relation between chemical structure and biochemical functions, a crucial topic of discussion in biochemistry. This has been considered in Chapter 6, which argued for the weak emergence of biochemical functions from molecular structure. This provides the link between biochemical and chemical features. Then, Chapter 7 presented an account for which biochemical kinds should be considered natural kinds.

The results expressed here also have two implications in relation to unity of science. Specifically, Duprè has argued that the disunity of science can be defended because the entities of the world are irreducible, and thus unrelated, and because there is pluralism about natural kinds (1983, 1993). Here, I have been exploring how unity can instead be maintained. In Chapter 6, I argued that a form of unity could be preserved thanks to the identification of weak emergence as the relation between biochemical functions and chemical structure. In Chapter 7, I have explored

---

[24] The same argumentative structure is also present in Part II for the molecular gene.

how the view of kinds in biochemistry should not be pluralist and should consider one account of natural kinds as applicable to different instances, following Natural Kind Monism. Biochemical molecules are instances of a natural biochemical kind that is identified as natural in the same way as chemical and biological kinds are. The implications of Part III concerning a more detailed account of unity and the emergence of biochemical kinds will be discussed in the conclusive Part IV of this thesis.

# Part IV

# Biochemical kinds and the unity of science

# Chapter 8

# Conclusion

> "If we have the habit of freedom and the courage to write exactly what we think; if we escape a little from the common sitting room and see human beings not always in relation to each other *but in relation to reality*; and the sky, too, and the trees or whatever it may be in."
>
> Virginia Woolf, A room of one's own

## 8.1 An inquiry into biochemical kinds

This thesis discussed some philosophical questions related to the philosophy and metaphysics of biochemical kinds. This research operated in the context of the inquiry into the relations between domains and entities that present different properties to understand which kind of unity of science can be presented and put forward. While a complete answer to the problem of the unity of science would require a more extended case-by-case analysis, the discussion of different biochemical kinds and their properties can be informative for this debate. The inquiry into the biochemical world and the relations between chemical and biological entities has been made by first considering the molecular gene and then three more general questions concerning biochemical kinds, with vitamin B12 as a second case study. The main result is that there is something specific about biochemical kinds and their properties. This result has been motivated by two general theoretical steps. The first has been to characterise biochemical functions and the molecular gene as weakly emergent from the chemical structure [Chapter 3, Chapter 6]. The second is via an account of natural kindhood, for which biochemical kinds are natural categories [Chapter 4, Chapter 7]. Among other implications, which will be discussed below, the novelty of this research project lies in the fact that it offers a metaphysically informed view of the molecular gene in the postgenomic context and presents a detailed analysis of the metaphysics of biochemical kinds, considering the notion of biochemical functions and their relation with structure.

In detail, this thesis is divided into four parts. Part I was an introductory part with the scope to introduce the aims and the main conceptual tools underpinning this research. Chapter 1 introduced the debate on the unity of science and biochemical kinds, offering this research's context and methodology. Chapter 2 instead presented the account of natural kinds used throughout the thesis for which natural kinds are those scientific projectible categories that capture nodes in causal networks (as in Khalidi 2013).

Part II was devoted to considering the detailed case study of the molecular gene and presents two chapters. The crucial role of the molecular gene in protein synthesis and other biological processes, together with being composed of nucleic acids, makes it a good starting point to discuss the questions this thesis is concerned with. Chapter 3 considered the nature of the molecular gene in the postgenomic context and argued that it is weakly emergent from the DNA (or relevant nucleic acids). Chapter 4 argued in favour of the naturalness of the gene category, defending the thesis that the molecular gene is a natural kind. The detailed exploration of this case study served two main purposes for the general argument of the thesis. The first one has pointed out that there are at least two properties crucial for the interplay between the biological and the chemical realm: those related to biochemical functions and molecular structure. This can also be seen when considering different biochemical kinds, such as proteins or vitamins. The second is that this case study supports the main thesis defended here, for which there is something specific about biochemical kinds and their properties in terms of weak emergence and natural kindhood status. This “mirrors” the two theoretical steps in support of the result of the research.

Part III considered the themes of biochemical functions and how biochemical functions are related to molecular structure and biochemical kinds by employing vitamin B12 as the main case study. Chapter 5 explored the nature of biochemical functions and argued that biochemical functions are neither chemical nor biological functions but can be defined as sets of chemical powers or dispositional properties to bring in a specific effect in biological processes. Chapter 6 considered the relation between biochemical functions and structure, spelt in terms of weak emergence and explored unity in biochemistry. Chapter 7 discussed biochemical kinds in detail, arguing that they can be considered natural categories. Part III presents the same argumentative structure as Part II and the thesis. Firstly, there is a characterisation of the relation between the chemical and the biochemical in terms of weak emergence. Secondly, it considers the naturalness of the biochemical kinds category.

The remaining conclusive Part IV underlines the main outcome and the implications of this research in relation to the unity of science. The main result is that there is something specific and *sui generis* about biochemical kinds and their properties. This has some further implications for future research, as will be spelt out in the remaining part of this conclusive chapter.

## 8.2 Outcome and implications

The nature of biochemical kinds and the properties of the entities in the interplay between the biological and chemical realms have been debated in recent literature (Slater 2009; Tobin 2010; Goodwin 2011; Bartol 2016; Havstad 2016, 2018; Kistler 2018; Tahko 2020, 2021). Specifically, this is so because biochemical molecules can be considered chemical compounds, following a characterisation via microscopic properties, or they can be considered instances of biological kinds or as instances of both. This might lead us to think we have two options for their nature and identity. The first is a view in which all the properties of biochemical kinds are reducible to some chemical structural properties, but this is challenged by the various instances of multiple determination and realisation [Chapter 3, Chapter 6]. The second is a dis-unified view of these molecules in which the chemical structural properties and the biological ones are both present but not related (as in Bartol 2016) [Chapter 6, Chapter 7]. Here, I have taken a different line of argument, arguing that an appropriate and adequate characterisation of biochemical kinds should not be a "dual" approach but rather a *unified* one. Biochemical kinds and their properties are characterised by a unity of properties that comprises their chemical microstructural properties, their biochemical function, an adequate evolutionary characterisation and *the relation* between these properties. This result has been obtained by analysing the molecular gene and considering the main properties of biochemical kinds, such as biochemical functions and the relation between biochemical functions and molecular structure. In detail, the ontological and metaphysical autonomy of biochemical kinds has been spelt out in two ways. The first is the consideration of the weak emergence of the molecular gene and biochemical functions [Chapter 3, Chapter 6]. The second is the relevance of the natural kindhood of biochemical kinds [Chapter 4, Chapter 7]. Moreover, in line with the moderate naturalism used in the thesis, further support for this result can be seen in the growing relevance of biochemistry as a foundational science for the life sciences.

This result has different implications. First, the inquiry into the biochemical world can direct our research to explore further the evolutionary properties of biochemical kinds and the relation between chemical structures and properties and evolutionary ones. This can lead to a further analysis of which kind of emergence can be ascribed to biochemical kinds once the evolutionary properties are considered. Second, this research has implications for the approach to the unity of science that one might want to defend or take. The autonomy of biochemistry and biochemical kinds is compatible with a final picture of the world unified by connections. This form of unity can be seen as an encouragement for inter and cross disciplinarily development in exploring natural phenomena.[1]. Lastly, this thesis has been developed without the commitment to the existence of hierarchically ordered ontological levels corresponding to the domains of the

---

[1] I thank Samuel Kimpton-Nye for noticing this aspect of the present work.

different sciences. Level talk has been present throughout this work, but levels were taken as heuristic and useful tools to explore relations of dependence. The final picture of the world that one can infer from the analysis of the biochemical realm is not a flat ontology but, rather, one in which there are different dependence relations and connections between chemical, biological and biochemical kinds and these dependence relations do not correspond to a hierarchically ordered and neat view of levels [Chapter 1]. These kinds exist in a network of dependence relations. The neutrality towards ontological levels (in favour of relations of dependence) can be considered both an advantage and an implication of this research that can invite the inquiry into how the unity of science can be maintained (together with the exploration of inter-science relations) without the commitment to hierarchically ordered levels (as also suggested by Potochnik 2021).[2]

Let me now expand more on the implications of this research for the emergence of biochemical kinds and the unity of science.

### 8.2.1 Emergence of biochemical kinds

Chapter 3 and Chapter 7 firstly argued in favour of the emergence of the postgenomic gene and secondly for the naturalness of the biochemical kinds category. Upon consideration of these results, the reader may rightly ask herself whether the combination of the two things proves that also biochemical kinds should be emergent. This is further justified by the fact that I have considered that a given kind is emergent when its definitional properties are emergent (as in Chapter 3; Wilson 2015, 2021). Accordingly, given that biochemical functions are emergent from the chemical structure, as discussed in Chapter 6, biochemical kinds could be considered weakly emergent too. Nevertheless, the question of whether weak or strong emergence is more relevant for biochemical kinds remains to be further answered and analysed. This is because, often, biochemical kinds are characterised not only in terms of chemical properties and biochemical functions but also by a form of evolutionary properties. Moreover, given that emergence is a relation that characterises tokens of given kinds on given occasions, the emergence of the different kinds may need to be assessed on a case-by-case basis.

---

[2] I thank Jessica Wilson for underlying this implication of my work.

The general answer then to whether biochemical kinds are emergent should come with the consideration of the nature of the evolutionary properties ascribed to biochemical kinds and the various prospects and details of molecular evolution. The philosophical consequences of molecular evolution for biochemical kinds still need to be explored in the literature. Accordingly, we are not yet in the position to assess with sufficient precision how biochemical kinds are emergent, and this can be the starting point for future studies.

Nevertheless, it is possible to point towards three possibilities to explore the emergence of biochemical kinds. As presented in Chapter 6 (and as in Wilson 2021), a token feature H can be considered *weakly metaphysically emergent* from token feature L on a given occasion when i) H co-temporally materially depends on L, and ii) H has a non-empty proper subset of the token powers had by L. Instead, a given token feature H can be considered *strongly metaphysically emergent* from token feature L on a given occasion when, on that occasion, i) H co-temporally materially depends on L, and ii) H has at least one token power not identical with any token power of L. Starting from these two definitions, one can explore three options for the emergence of the instances of biochemical kinds in given occasions. Let me consider them briefly[3]:

- A given instance of a biochemical kind can be considered *weakly emergent* in a given occasion because its definitional properties are novel and robust, and its biochemical functions correspond to proper subsets of causal powers of the chemical structure, as defined in Chapter 3 and Chapter 6. To what has been previously argued, one could make a further point that considers the role of evolutionary properties. Precisely, one can specify that the manifestation of the proper subset of causal powers associated with the function depends on the given evolutionary properties associated with the kind. This would make the biochemical kind weakly emergent, even when ascribing evolutionary properties to them, as these properties are important for manifesting the relevant function in the given context. This argument is valid upon considering the instance of the given biochemical kind at a particular time. For example, the evolutionary properties ascribable to vitamin B12 allow the function of vitamin B12 to be manifested in the specific context in which the vitamin

---

[3] I thank Jessica Wilson for exploring these with me possibilities and discussing this topic further.

has its role.

- A given instance of a biochemical kind can be considered *weakly emergent* in a given occasion *not only* because the biochemical functions associated to the kind correspond to proper subsets of causal powers of the chemical structure, but *also* because the evolutionary properties of the biochemical kind correspond to proper subsets of causal powers of the chemical structure. This possibility could open a further point of contact between the chemical and the biological domain (as suggested by Pross 2016; Tahko 2020). An example that could support this form of weak emergence is the one of sickle cell anaemia, presented by Tahko 2020. In this case, it seems possible to explain the evolutionary retention of sickle cell anaemia on the African continent in terms of chemical or biochemical properties. Specifically, some chemical reactions resulting from sickle cell anaemia can lead to more carbon monoxide in the blood, which contrasts malaria's plasmodium.[4] However, the possibility of evolutionary properties being associated with chemical behaviours or properties still has to be explored in detail and requires a better understanding of the molecular nature of evolution.

- A given instance of a biochemical kind can be considered *strongly emergent* on a given occasion. This could be so because the evolutionary properties associated with the kind are *novel* and do not correspond to any token power of the chemical structure from which the biochemical kind emerges. This option requires a precise analysis of molecular evolution and of which kind of evolutionary properties are ascribable in each instance to biochemical kinds.

In conclusion, according to the results presented in this thesis, given instances of biochemical kinds could be considered at least weakly emergent according to the first account. This option does not postulate more than what has been already and could be more easily embraced and sustained. Nevertheless, the other two options are worth being explored in more detail, and a

---

[4] Tahko suggests that this supports reductionism rather than weak emergence (2020). However, it could also open the discussion on weak emergence in terms of the proper subset strategy discussed in Chapter 6.

more extended analysis of the philosophical implications of molecular evolution is philosophically considered.

### 8.2.2 A unified view of the world

The thesis focused on providing an account and analysis of biochemical entities and the properties and relations at play in this realm. However, this project has origins in exploring the possibility of the unity of science starting from particular case studies. A detailed account of the unity of science should comprise more than the analysis of one domain of interplay of relations. Still, I would like to consider which kind of unity is compatible with the view of biochemical kinds defended here. As mentioned in Chapter 1, there can be two sources for metaphysical or ontological disunity (as in Duprè 1983, 1993). The first regards the irreducibility and autonomy of the various entities that are seen as having an autonomous status. The second considers natural kinds pluralism: the variety of classifications can lead to a disunified view of the world. As explored in Chapter 6, a form of unity via reduction does not seem feasible for the biochemical world because of the complex relations between the different features that characterise it. Nevertheless, this does not leave us necessarily with a disunified or disconnected world. As explored in Chapter 3 and Chapter 6, it is possible to identify the relations between the biochemical properties in terms of weak emergence. Moreover, as in Chapter 4 and Chapter 7, it is possible to identify a single account of natural kinds that can be applied to different instances (as in Natural Kind Monism). These results support a form of unity. This section will consider more in detail which kind of unity can be preserved given the results discussed in this thesis. The view of the unity of science that can result should be spelt out in terms of identifying the relations between different entities of the world and has three grounds.

The first form of unity that can be maintained follows from the relation of weak emergence spelt out in Chapter 3 and Chapter 6. Specifically, biochemical functions are seen as materially dependent upon chemical features as each token of a biochemical function is associated with a specific proper subset of the token chemical powers from which the biochemical function is emergent. Accordingly, there is a form of dependence between the instance of the biochemical function and its realisation basis. This can be seen equally for the molecular gene and a more general account of biochemical functions, as in the case of vitamin B12. This material dependence can give us the first sense of unity in which the emergent entity is related to the underlying level. Nevertheless, given that the emergent function has a distinct causal profile from its basis, this can allow us to preserve some definite causal profile. Weak emergence then provides a first form of unity between the biochemical and the chemical while maintaining some autonomy the biochemical realm.

The second form of unity that can be maintained is via Natural Kind Monism, for which there is one notion of natural kinds that is applicable across the sciences (Tahko 2021) [Chapter 1, Chapter 4, Chapter 7]. Specifically, we have seen how the account of natural kinds considered can be applied to different kinds, such as chemical, biological and biochemical and specific instances of such kinds, such as the molecular gene. This form of unity seems particularly relevant when contrasted with forms of disunity that are advocated based on autonomy and the differences between kinds of different special sciences (as in Duprè 1983; 1993). As discussed in the thesis, an instance of this can be seen when considering a dual approach to biochemical kinds for which biochemical molecules instantiated two kinds of kinds that are seen as distinct and disunified, as in Bartol (2016). Here, the thesis underlined how biochemical kinds can be considered natural categories by using an account of natural kinds that can be applied across the sciences and different instances of kinds. This form of metaphysical unity via Natural Kind Monism could be a further ground for a unified view of the world compatible with the autonomy of the various sciences and domains.[5]

Finally, here I would want to suggest that a broader sense of unity of science and unity of the world can be preserved when it is possible to identify how the different features of the world are connected and how the entities postulated and discussed by the different sciences are related. Here, we have seen that different entities of the world are related because they are weakly emer-

---

[5] The unity delivered by natural kind monism can be considered even stronger if there is a commitment to a form of kind universals in our ontology, as put forward by Lowe (2005) and as presented in Chapter 2. In this case, the same fundamental category, a *sui generis* kind universal, would be instantiated in the different kinds of kinds. For instance, we would have a kind universal in chemical kinds instances, in biological kinds instances and in biochemical kinds instances.

gent, and there are levels of dependencies that can be spelt out. Moreover, the account of natural kinds used throughout this work underlines the importance of causal relations and networks in the world. These causal relations can provide a further sense of unity and connection between the different entities of the world. It remains important to say that weak emergence and causal connections should not be considered the only relations across different entities of the world. The discussion of these categories and the results of this thesis can be seen as an invitation to inquire how unity in terms of relations can be further spelt out.

Concluding, this thesis puts forward an account for which the biochemical realm comprises emergent entities and properties, and genuine natural kinds without being incompatible with a form of ontological unity of science. This form of metaphysical unity can be seen as supporting a unified view of the world and as a further motivation to promote the interaction between different disciplines when studying complex entities and phenomena such as the biochemical realm and the interplay between biological and chemical entities.

# *Parting thoughts*

> "I think here I will leave you.
> It has come to seem there is no
> perfect ending.
> Indeed, there are infinite endings.
> Or perhaps, once one begins, there
> are only
> endings."
>
> Louise Gluck, The Past

Years ago, I listened to a philosophy lecture where the philosopher tried to capture the difference and the relation between the philosophical and scientific methods. He said that both science and philosophy aim at being precise (from *prae-caedere*, to cut, carve), cutting off the parts of the world they want to explore to focus on them, but philosophy needs to take a step further. Philosophy should aim at being exact (from *ex-agere*, to push out), bringing out from these "precisions" general hypotheses about the world. This thesis tried secretly to implement this method while not having the arrogance to succeed in reaching an *exact* view of the world. At best, it can present some generalisations that I hope are precise in a way. Nevertheless, if the readers allow me, I would like to conclude this thesis with thoughts that, while not yet philosophically rigorous, have been an implicit guidance and an outcome of it. I hope they will take them for what they are, parting thoughts that only have the dream to be exact.

The exploration of the biochemical world has been personally thrilling and mind-blowing. I started the journey with a more or less Aristotelian metaphysics in mind, a world of substances with their own properties, and saw it challenged and changed by the study of molecular genetics first and biochemical interactions after. The world that I started seeing was a world that seemed less divided into objects and more and more related. If there is some general philosophical conclusion that the biochemical world teaches us it is that relations are important. Genes, for many the very core to understand the development of organisms and most of their features and properties, exist during transcription, in the cross-talk between the main actors of it: they exist when *relations* happen. Chemical molecules display biochemical behaviour and instantiate biochemical kinds when they are *in relation* with the biological context that depends on them. A form of unity in the biochemical world can be seen once one considers the *relations* that hold between objects.

This does not seem completely surprising if one considers recent movements in the philosophy and metaphysics of science, such as structuralism or various relational ontologies. They also deliver us a picture of the world that is relational. I still have not done the deep analysis and consideration of these views that should precede the commitment to them. However, it seems that many and various different scientific disciplines are delivering us an image of a world that is complex, more and more interconnected, and where relations are as important as the objects themselves (if not all there are). I have the feeling or the intuition (which is what it is and nothing more) that the world will not be something like a dappled world of independent and various entities nor only a net of fundamental relations. The things that populate it are plenty, rich and surprising, and they are what they are when they are in the relations that allow them to be what they are. These relations constitute a fundamental unity that is a precondition for the very being of every single thing: a unified, interconnected world that gives rise to the special and unique nature of everything.

# Bibliography

[1] Allen, Colin, Jacob Neal. 2020: "Teleological notions in biology." *The Stanford Encyclopaedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), `https://plato.stanford.edu/entries/teleology-biology/`.

[2] Aristotle. *Metaphysics*, Perseus eds. `http://www.perseus.tufts.edu`.

[3] Armstrong, David M. 1993: "A world of a state of affairs". *Philosophical Perspectives, Language and Logic*: 429-440.

[4] Bartol, Jordan. 2016: "Biochemical kinds". *British Journal for Philosophy of Science* 67: 531- 551.

[5] Bellazzi, Francesca. 2020: "Marvellous mysteries and the unity of science". *Arts Matter Blog.* `https://artsmatter.blogs.bristol.ac.uk/`.

[6] Bellazzi, Francesca. 2022a: "The emergence of the postgenomic gene". *European Journal for Philosophy of Science* 12, 17.

[7] Bellazzi, Francesca. 2022b: "The superpowers of proteins" *Jargonium* `https://www.jargonium.com/post/the-superpowers-of-proteins`.

[8] Bellazzi, Francesca. 2022c: "Genes: from instruments of thought to the base of life". *Jargonium* `https://www.jargonium.com/post/genes-from-instruments-of-thought-to-the-base-of-life`.

[9] Bellazzi, Francesca. *forthcoming*: "Biochemical functions". *British Journal for Philosophy of Science.*

[10] Bellazzi, Francesca. *forthcoming*: "The gene as a natural kind". Edited volume *Life & Mind* (Springer Nature) (22 pages).

[11] Beurton, Peter E. 2010: "A unified view of the gene, or how to overcome reductionism". *The Concept of the Gene in Development and Evolution.* eds. Beurton P. J., Falk R., Rheinberger H.J., 286-314. Cambridge (UK): Cambridge University Press.

[12] Bigelow, John, Robert Pargetter. 1987: "Functions". *The Journal of Philosophy* 54(4): 181-96.

[13] Bird, Alexander. 2018: "The metaphysics of natural kinds". *Synthese* 195: 1397-1426.

[14] Bird, Alexander, Emma Tobin. 2022: "Natural Kinds". *The Stanford Encyclopedia of Philosophy* (Spring 2022 Edition), Edward N. Zalta (ed.), `https://plato.stanford.edu/archives/spr2022/entries/natural-kinds/`.

[15] Boiteux, Serge, and Sue Jinks-Robertson. 2013: "DNA repair mechanisms and the bypass of DNA damage in Saccharomyces cerevisiae". *Genetics*, 193 (4):1025 -1064.

[16] Boone, Worth T. 2018: "Multiple realization and robustness". *Biological Robustness: Emerging Perspectives from within the Life Sciences.* eds. Bertolaso M., Caianiello S., Serrelli E. Switzerland: Springer Nature.

[17] Brigandt, Ingo. 2022: "How to philosophically tackle kinds without talking about natural kinds". *Canadian Journal of Philosophy* 52 (3): 356-379.

[18] Brock, Stuart, Edwin Mares. 2014: *Realism and Anti-Realism.* Hoboken, Taylor and Francis.

[19] Brooks, Daniel S., Markus I. Eronen 2018: "The significance of levels of organization for scientific research: A heuristic approach". *Studies in History and Philosophy of Biological and Biomedical Sciences* 68-69: 34-41.

[20] Butterfield, Jeremy. 2011: "Emergence, reduction and supervenience: a varied landscape". *Foundations of Physics* 41: 920-959.

[21] Buller, David J. 1998: “Etiological theories of function: a geographical survey”. *Biology and Philosophy* 13: 505-527.

[22] Burian, Richard M. 2004: “Molecular epigenesis, molecular pleiotropy, and molecular gene definitions”. *History and Philosophy of the Life Sciences* 26 (1-Genes, Genomes, and Genetic Elements): 59-80.

[23] Carr, Anitra C., Silvia Maggini. 2017: “Vitamin C and immune function”. *Nutrients* 9,11: 1211. https://doi.org/10.3390/nu9111211.

[24] Cartwright, Nancy. 1999: *The Dappled World. A study of the boundaries of science*, Cambridge (UK): Cambridge University Press.

[25] Cartwright, Nancy. 2022: *A Philosopher looks at Science*, Cambridge (UK): Cambridge University Press.

[26] Carnap, Rudolf. 1934: *Unity of Science* eds. Classic Studies in the History of Ideas; Bristol (UK): Thoemmes Press (1995).

[27] Cat, Jordi. 2017: “The Unity of Science”, *The Stanford Encyclopaedia of Philosophy* (Spring 2022 Edition), Edward N. Zalta (ed.) `https://plato.stanford.edu/archives/spr2022/entries/scientific-unity/`.

[28] Clark, David, Nanette Pazdernik. 2012. *Molecular Biology.* San Diego (USA): Elsevier Science-Technology, Academic Cell.

[29] Combs, Gerald F. 2012: *The Vitamins: Fundamental Aspects in Nutrition and Health.* San Diego (USA): Elsevier Science-Technology.

[30] Craver, Carl. 2009. “Mechanisms and natural kinds”. *Philosophical Psychology*, 22:5, 575-594.

[31] Crane, Judith K. 2021: “Two approaches to natural kinds”. *Synthese* 199: 12177 -12198.

[32] Craver, Carl. 2001: “Role functions, mechanisms, and hierarchy.” *Philosophy of Science 68: 53-74.*

[33] Craver, Carl. 2009: "Mechanisms and natural kinds". *Philosophical Psychology.* 22, 5: 575-594.

[34] Cummins, Robert. 1975: "Functional analysis". *The Journal of Philosophy* 72 (20): 741-765.

[35] Davini, Claudio. 2021: "Il concetto di funzione in filosofia della biologia" (The concept of function in the philosophy of biology). *APhEx*, 23.

[36] Di Segni, Gianfranco, Serena Gastaldi, Glauco P. Tocchini-Valentini. 2008: "Cis- and trans-splicing of mRNAs mediated by tRNA sequences in eukaryotic cells". *Proceedings of the National Academy of Sciences.* May 105 (19): 6864-6869.

[37] Doolittle, Ford. 2013: "Is junk DNA bunk? A critique of ENCODE" *PNAS*; 110 (14): 5294-5300.

[38] Dupré, John. 1983: "The disunity of science". *Mind* 92 (367): 321-346.

[39] Dupré, John. 1993: *The Disorder of Things: Metaphysical foundations of the Disunity of Science.* Cambridge (MA): Harvard University Press.

[40] Ellis, Brian. 2001: *Scientific Essentialism.* Cambridge (UK): Cambridge University Press.

[41] Ellis, Brian. 2002: *The Philosophy of Nature: a Guide to the New Essentialism.* Routledge (NY, USA): Routledge.

[42] El-Hani, Charbel N. 2007: "Between the cross and the sword: the crisis of the gene concept". *Genetics and Molecular Biology* 30 (2).

[43] El-Hani, Charbel N. 2015: "Mendel in genetics teaching: some contributions from history of science and articles for teachers". *Science & Education* 24: 173-204.

[44] Eronen, Markus I. 2015: "Robustness and reality". *Synthese* 192: 3961-3977.

[45] Ereshefsky, Mark. 2017: "Species". *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition), Edward N. Zalta(ed.) `https://plato.stanford.edu/archives/sum2022/entries/species/`.

[46] Ereshefsky, Mark. 2018: "Natural kinds, mind independence, and defeasibility". *Philosophy of Science* 85, 845 - 856.

[47] Falk, Raphael. 2010: "The gene - A concept in tension". *The Concept of the Gene in Development and Evolution.* eds. Beurton P. J., Falk R., Rheinberger H.J., 317-348. Cambridge (UK): Cambridge University Press.

[48] Fang, Huan, Kang Jie, Dawei Zhang. 2017: "Microbial production of vitamin B12: a review and future perspectives". *Microbial Cell Factories* 16, 15. `https://doi.org/10.1186/s12934-017-0631-y`.

[49] Fletcher, Hugh L. and Ivor G. Hickey. 2012: *BIOS Instant Notes in Genetics.* Fourth Edition Garland Science.

[50] Fogle, Thomas 2010: "The dissolution of protein coding genes in molecular biology". In *The Concept of the Gene in Development and Evolution.* eds. Beurton P. J., Falk R., Rheinberger H.J., 3-25. Cambridge (UK): Cambridge University Press

[51] Fogle, Thomas. 1990: "Are genes units of inheritance?". *Biology and Philosophy* 5: 349-371.

[52] Fodor, Jerry A. 1974. "Special sciences (or: the disunity of science as a working hypothesis)". *Synthese* 28: 97-115.

[53] Fox Keller, Evelyn. 2000: *The Century of the Gene.* Cambridge (MA): Harvard University Press.

[54] Franklin, Alexander, Eleonor Knox. 2018: "Emergence without limits: The case of phonons". *Studies for the History and Philosophy of Modern Physics* 64: 68-78.

[55] Frassinetti Stefania, Bronzetti Giorgio, Caltavuturo Leonardo, Cini Marco and Clara della Croce. 2006: "The role of zinc in life: a review". *Journal Environ Pathol Toxicol Oncol* 25 (3): 597-610.

[56] Garson, Justin. 2017: "A generalised selected effects theory of function". *Philosophy of Science* 84: 523-543.

[57] Garson, Justin. 2018: "How to be a function pluralist". *British Journal for Philosophy of Science* 69: 1101-1122.

[58] Garson, Justin. 2019: "There are no ahistorical theories of function". *Philosophy of science* 86: 1146-1156.

[59] Gerstain, Mark B., Can Bruce, Joel S. Rozowsky, Deyou Zheng, Jiang Du, Jan. O. Korbel, Olof Emanuelsson Zhengdong D. Zhang, Sherman Weissman, Michale Snyder. 2007: "What is a gene, post-ENCODE? History and updated definition". *Genome Research* 17: 669-681.

[60] Gericke, Niklas M., Mariana Hagberg, Vanessa Carvalho dos Santos, Leyla Mariane Joaquim, Charbel N. El-Hani. 2014: "Conceptual variation or Incoherence? Textbooks discourse on genes in six countries." *Science and Education* 23: 381-416.

[61] Germain Pierre-Luc, Emanuele Ratti, Federico Boem. 2014: "Junk or functional DNA? ENCODE and the function controversy". *Biology and Philosophy.* 29, 6: 807-831.

[62] Gillett, Carl. 2013: "Constitution, and multiple constitution, in the sciences: using the neuron to construct a starting framework". *Minds &Machines* 23: 309-337.

[63] Gillett, Carl. 2021: "Using compositional explanations to understand compositional levels: an integrative account". *Levels of Organization in the Biological Sciences* eds. D. Brooks, J. Di Frisco, W. Wimsatt. Cambridge (MA, USA): MIT Press.

[64] Goodman, Nelson. 1954/1979: *Fact, Fiction, and Forecast.* Cambridge (MA): Harvard University Press.

[65] Goodwin, William. 2011: “Structure, function and protein taxonomy”. *Biology and Philosophy* 26: 533-545.

[66] Griffiths, Paul E. 2004: “Emotions as natural and normative kinds” *Philosophy of Science* 71 (5): 901-911.

[67] Griffiths, Paul E. 2009: “In what sense does “Nothing make sense except in the light of evolution”?”. *Acta Biotheoretica* 57: 11-32.

[68] Griffiths, Paul, Karola Stotz. 2006: “Genes in the postgenomic era”. *Theoretical Medicine and Bioethics* 27:499-521

[69] Griffiths, Paul, Karola Stotz. 2013: *Genetics and Philosophy.* Cambridge (UK): Cambridge University Press.

[70] Gu, Wenli, Feng Zhang, James R. Lupsk. 2008. “Mechanisms for human genomic rearrangements”. *PathoGenetics* 1, 4 doi: 10.1186/1755-8417-1.

[71] Hacking, Ian 1999: “The Social Construction of What?”. Cambridge (US): Harvard University Press.

[72] Hall, Brian K. 2001: “The gene is not dead, merely orphaned and seeking a home”. *Evolution and Development* 3 (4): 225-228.

[73] Helliwell Katherine E., Glen L. Wheeler, Kyriacos C. Leptos, Raymond E. Goldstein, Alison G. Smith. 2011: “Insights into the Evolution of Vitamin B12 Auxotrophy from Sequenced Algal Genomes”. *Molecular Biology and Evolution* 28,10: 2921-2933, `https://doi.org/10.1093/molbev/msr124`.

[74] Havstad, Joyce C. 2014: *Biochemical Kinds and Selective Naturalism*, PhD dissertation, UC San Diego.

[75] Havstad Joyce C. 2016: “Protein Tokens, Types, and Taxa”. *Natural Kinds and Classification in Scientific Practice* eds. C. Kendig New York (NY, USA): Routledge.

[76] Havstad, Joyce C. 2018: "Messy Chemical Kinds". *British Journal for Philosophy of Science* 69: 719-743.

[77] Havstad, Joyce C. 2021: "Complexity begets crosscutting, dooms hierarchy (another paper on natural kinds)." *Synthese* 198: 7665-7696.

[78] Heil, John. 2003: "Levels of Reality". *Ratio* 16: 205-221. `https://doi.org/10.1111/1467-9329.00218`.

[79] Juneau, Kara, Corey Nislow and Ronald W. Davis. 2009: "Alternative splicing of PTC7 in Saccharomyces cerevisiae determines protein localization". *Genetics* 183 (1): 185-194.

[80] Kashyap, Luv and Parul Tripathi. 2008: "Alternative splicing: How one gene can make many proteins" *BioScience Explained.*

[81] Khalidi, Muhammad A. 1993: "Carving nature at the joints". *Philosophy of Science* 60 (1): 100-113.

[82] Khalidi, Muhammad A. 1998: "Natural Kinds and Crosscutting Categories". *The Journal of Philosophy* 95, 1: 33-50.

[83] Khalidi, Muhammad A. 2010: "Interactive kinds". The British Journal for the Philosophy of Science, 61 (2), 335-360.

[84] Khalidi, Muhammad A. 2013: *Natural Categories and Human Kinds.* Cambridge (UK): Cambridge University Press.

[85] Khalidi, Muhammad A. 2016: "Mind-dependent kinds" *Journal of Social Ontology* 2 (2): 223-246.

[86] Khalidi, Muhammad A. 2018: "Natural kinds as nodes in causal networks". *Synthese* (2018) 195: 1379-1396.

[87] Khalidi, Muhammad A. 2020: "Are sexes natural kinds," *Current Controversies in Philosophy of Science* eds. Dasgupta S., B. Weslake New York (NY, USA): Routledge.

[88] Khalidi, Muhammad A. 2021: "Etiological kinds". *Philosophy of Science*, 88, 1-21.

[89] Kim, Jaegwon. 1992:"Multiple realization and the metaphysics of reduction". *Philosophy and Phenomenological Research* 52 (1): 1-26.

[90] Kingsbury, Justine. 2008: "Learning and Selection". *Biology and Philosophy* 23: 493-507.

[91] Kistler, Max. 2018: "Natural kinds, causal profile and multiple constitution". *Metaphysica.* I (19) 1 De Gruyter. DOI: https://doi.org/10.1515/mp-2018-0006.

[92] Knox, Eleonor. 2016: "Abstraction and its limits: finding space for novel explanations". *Nous* 50 (1): 41-60.

[93] Kitcher, Peter. 1984: "1953 and all that. A tale of two sciences". *The Philosophical Review* 93 (3): 335-373.

[94] Koury, Mark J., Prem Ponka. 2004: "New insights into erythropoiesis: the roles of folate, vitamin B12, and iron".*Annu. Rev. Nutr.* 24:105-31. doi: 10.1146/annurev.nutr.24.012003.132306.

[95] Ladyman, James, Dan Ross. (with D.Spurrett, J.Collier). 2007: *Every Thing Must Go: Metaphysics Naturalised* Oxford (UK): Oxford University Press.

[96] Ladyman, James. 2012: "Science, metaphysics and method". *Philosophical Studies* 160:31-51.

[97] LaPorte, Joseph, 2003: *Natural kinds and conceptual change*, Cambridge (UK): Cambridge University Press.

[98] Liu, Wei, Li Li, Hua Ye, Haiwei Chen, Weibiao Shen, Yuexian Zhong, Tian Tian, Huaqin He. 2017: "From Saccharomyces cerevisiae to human: The important gene co-expression modules". *Biomedical Reports* 7: 153-158.

[99] LLyod, Elisabeth, 2017: "Units and levels of selection". *Stanford Encyclopaedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), `https://plato.stanford.edu/entries/selection-units/`.

[100] Lowe, E. Jonathan. 2001: *The possibility of Metaphysics: Substance, Identity, and Time* Oxford (UK): Clarendon Press, Oxford University Press.

[101] Lowe, E. Jonathan. 2005: *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford (UK): Clarendon Press, Oxford University Press.

[102] Lowe, E. Jonathan. 2011: "A neo-Aristotelian substance ontology: neither relational nor constituent". *Contemporary Aristotelian Metaphysics* eds. T. E. Tahko. Cambridge (UK): Cambridge University Press.

[103] Magnus P.D. 2015: "John Stuart Mill on Taxonomy and Natural Kinds". *The Journal of the International Society for the History of Philosophy of Science* (5): 269-280.

[104] Magnus P.D. 2018: "Taxonomy, ontology, and natural kinds". *Synthese* (2018) 195: 1427-1439.

[105] Meyer, Lia M. N., Gilberto C. Bomfim G. C.m Charbel N. El-Hani. 2013: "How to understand the gene in the 21st century?". *Science and Education* 22 (2): 345-374.

[106] Mitchell, Sandra D. 1993. "Dispositions or etiologies? A comment on Bigelow and Pargetter". *The Journal of Philosophy* 90 (5): 249-259.

[107] Mitchell, Sandra D. 2003: *Biological complexity and integrative pluralism* Cambridge (UK): Cambridge University Press.

[108] Mitchell, Sandra D. 2008: "Exporting causal knowledge in evolutionary and developmental biology?". *Philosophy of Science* 75 (5): 697-706.

[109] Nagel, Ernest. 1961: "The structure of science: problems in the logic of scientific explanation". *British Journal for Philosophy of Science*, 68: 163-186.

[110] Nanay, Bence. 2010: "A modal theory of function". *Journal of Philosophy* 107 (8): 412-431.

[111] Natarajan, Chandrasekhar, Hoffmann, Federico G., Weber, Roy E., Fago, Angela Witt, Christopher C., Storz, Jay F. 2016: "Predictable convergence in haemoglobin function has unpredictable molecular underpinnings". *Science* 354, 6310: 336-339.

[112] Neander, Karen. 1991: "Functions as Selected Effects: The Conceptual Analyst's Defense". *Philosophy of Science* 58: 168-184.

[113] Neander, Karen. 2006: "The teleological notion of function". *Australasian Journal of Philosophy* 64 (4): 454-468.

[114] Needham, Paul. 2000: "What is water?" *Analysis* 60 (1):13-21.

[115] Needham, Paul. 2011: "Microessentialism: What is the Argument?". *Noûs*, 45 (1): 1-21.

[116] Okasha, Samir. 2019: *Philosophy of Biology: a Very Short Introduction*. Oxford (UK): Oxford University Press.

[117] Perini, Laura. 2011: "Sequence matters: genomic research and the gene concept". *Philosophy of Science* 78 (5): 752-762

[118] Plato. *Sophist*, Perseus eds. `http://www.perseus.tufts.edu`.

[119] Polger, Thomas W., Lawrence A. Shapiro. 2016: *The Multiple Realisation Book*. Oxford (UK): Oxford University Press.

[120] Potochnik, Angela. 2021: "Our world isn't organized into levels". in Brooks D., J. DiFrisco and W. C. Wimsatt (eds.) *Levels of Organization in the Biological Sciences*. Cambridge (MA, USA): MIT Press.

[121] Pross, Addy. 2016: *What is Life? How chemistry becomes biology*. Oxford (UK): Oxford University Press.

[122] Putnam, Hilary. 1975: "What is mathematical truth?". Historia mathematica 2: 529-533.

[123] Oppenheim, Paul, Hilary Putnam. 1958: "Unity of science as a working hypothesis". *Minnesota Studies in the Philosophy of Science* eds. Feigl H., M. Scriven, G. Maxwell Minneapolis (USA): University of Minnesota Press, 2, 3: 36.

[124] Reynaud, Enrique. 2010: "Protein misfolding and degenerative diseases". *Nature Education* 3 (9): 28.

[125] van Riel, Raphael, van Gulick Robert 2019: "Scientific Reduction". *Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), Edward N. Zalta(ed.)`https://plato.stanford.edu/entries/scientific-reduction/`.

[126] Rodionov, Dmitry A., Alexey G. Vitreschak, Andrey A . Mironov, Mikhail S. Gelfand. 2003: "Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes". *Journal of Biological Chemistry.* 278 (42): 41148-59. doi: 10.1074/jbc.M305837200.

[127] Rodriguez-Pereyra, Gonzalo. 1999: "Resemblance nominalism and the imperfect community". *Philosophy and Phenomenological Research* 59(4): 965-982.

[128] Rodriguez-Pereyra, Gonzalo. 2015: "Nominalism in Metaphysics". *Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), Edward N. Zalta(ed.) `https://plato.stanford.edu/entries/nominalism-metaphysics/`.

[129] Rowley, Carol A., Melissa M. Kendall. 2019. "To B12 or not to B12: Five questions on the role of cobalamin in host-microbial interactions". *PLoS Pathog* 15(1): e1007479. doi: 10.1371/journal.ppat.1007479.

[130] Santos, Gil, Gabriel Vallejos, Davide Vecchi. 2020: "A relational-constructionist account of protein macrostructure and function." *Foundations of Chemistry*, 22(3): 363-382.

[131] Schaffer, Kenneth. 1969: "The Watson-Crick model and reductionism". *The British Journal for Philosophy of Science* 20: 325-348.

[132] Scherrer, Klaus, Jurgen Jos. 2007: "The gene and the genon concept: a functional and information-theoretic analysis". *Molecular System Biology* 3: 1-11.

[133] Serri, Eric R. 2020: "On chemical natural kinds". *Journal for General Philosophy of Science* 51: 427-445.

[134] Shaprio, Lawerence A. 2020:"Theories of Multiple Realisation". *American Philosophical Quarterly* 57 (1): 17-30.

[135] Slater, Matthew H. 2013: "Cell types as natural kinds". *Biological Theory*, 7 (2): 170-179.

[136] Slater, Matthew H. 2009: “Macromolecular pluralism”. *Philosophy of Science*, 76: 851-863.

[137] Tahko, Tuomas E. 2020: “Where do you get your protein? Or: biochemical realization”. *British Journal for the Philosophy of Science*, 71 (3): 799-825.

[138] Tahko, Tuomas E. 2021. *Unity of Science*. Cambridge Elements in Philosophy of Science. Cambridge (UK): Cambridge University Press.

[139] Tahko, Tuomas E. 2022: “Natural kinds, mind independence, and unification principles”, *Synthese*, 200: 144.

[140] Tahko, Tuomas E., Matteo Morganti. 2017: “Moderately naturalistic metaphysics”. *Synthese* 194 (7): 2557-2580.

[141] Tobin, Emma. 2010: “Microstructuralism and macromolecules: the case of moonlight proteins”. *Foundations of Chemistry*, 12: 41-54.

[142] Tobin, Emma. 2013: “Are natural kinds and natural properties distinct?”. *Metaphysics and Science*. eds. S. Mumford, Tugby M.

[143] Valles, Sean. A. 2013: “Validity and Utility in Biological Traits.” *Biological Theory* 8: 93-102. doi: 10.1007/s13752-013-0096-0.

[144] Walmacq C, Cheung AC, Kireeva ML, Lubkowska L, Ye C, Gotte D, Strathern JN, Carell T, Cramer P, Kashlev M. 2012: “Mechanism of translesion transcription by RNA polymerase II and its role in cellular resistance to DNA damage”. *Molecular Cell*,13; 46(1):18-29.

[145] Waters, Ken. 2007: “Molecular Genetics”, *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta(ed.), `https://plato.stanford.edu/archives/fall2013/entries/molecular-genetics/`.

[146] Waters, Andrew M., Rachel Bagni, Franklin Portugal, James L. Hartley. 2016: “Single synonymous mutations in KRAS cause transformed phenotypes in NIH3T3 cells”. *PLoS one*. 11(9): e0163272.

[147] Weisberg, Michael. 2013: *Simulation and similarity.* Oxford (UK): Oxford University Press.

[148] Wilson, Mark. 1985: "What is this thing called "pain"? - The philosophy of science contemporary debate." *Pacific Philosophical Quarterly* 66, 3-4: 227-267.

[149] Wilson, Jessica. 2011: "Non-reductive realization and the powers-based subset strategy." *The Monist* 94.1: 121-154.

[150] Wilson, Jessica. 2015: "Metaphysical emergence: weak and strong".*Metaphysics in contemporary physics: Poznan studies in the philosophy of the sciences and the humanities.* eds. T. Bigaj, C. Wuthrich: 251-306. .

[151] Wilson, Jessica. 2019: "Between scientism and abstractionism in the metaphysics of emergence". *The Routledge Handbook of Emergence* eds. Gibb S., R. F. Hendry, T. Lancaster (1st Ed.). New York (NY, USA): Routledge.

[152] Wilson, Jessica. 2021: *Metaphysical Emergence.* Oxford (UK): Oxford University Press.

[153] Wouters, Arno G. 2003: "Four notions of biological function". *Studies in the History and Philosophy of Biological and Biomedical Sciences*, 34: 633-668.