Bryan Hanley

Natural and Unnatural Product Chemistry

From Molecules to Systems

Springer

Natural and Unnatural Product Chemistry

Bryan Hanley

Natural and Unnatural Product Chemistry

From Molecules to Systems

Springer

Bryan Hanley
Academic Centre for Dentistry
in Amsterdam (ACTA)
Amsterdam, The Netherlands

ISBN 978-3-031-81697-0 ISBN 978-3-031-81698-7 (eBook)
https://doi.org/10.1007/978-3-031-81698-7

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

This book examines natural product chemistry from the perspective of the success of reductionism at a molecular level and the challenges to adopting such an approach at a system level. The development of natural product chemistry has been built upon the twin achievements of the growth in expertise in isolation and characterisation of molecules and the intellectual examination of the links between molecular entities. This has enabled chemists to create a picture of how individual molecules are created (their biosynthesis) and, to an extent, what their role is both in the context of the living systems responsible for their creation and other organisms that come into contact with them. While there have been significant advances that can be attributed to the application of reductionism, it is, both philosophically and intellectually, incomplete. There are significant barriers to trying to reduce living systems to the sum of their individual parts, and this has been recognised by significant figures including Descartes. The mechanical, reductionist view of living systems can only provide a partial answer.

Two of the aspects of natural product chemistry that are dealt with are the perspectives of the creation of molecules and their interactions with living systems. In the case of the former, the pivotal role of enzymes as macromolecular 'creators' of natural products and the importance of biosynthetic pathways and gene manipulation has led to advances in bioengineering and systems biology. For the latter, the interaction of natural products with living systems is both more complex and more challenging to disentangle. At one level, molecules interact with biological macromolecules (particularly receptors and enzymes) and these interactions have consequences for the biological system which, in some cases, can be predicted but in others cannot. Biological macromolecules (and, indeed biological entities) do not exist in isolation but are part of complete systems. One of the most significant examples of this is the existence of microbial communities in an environment that includes microbes, the host that accommodates them and the general environment (microbiomes). This gives rise to new paradigms in which the role of natural products is less easy to define in terms of individual interactions but also more significant in terms of potential impact.

The future of natural product chemistry will include both a better and more complete understanding of molecular interactions and the consequences of the exposure of biological systems to natural products. This book seeks to demonstrate examples that illustrate progress along these pathways. It will be of interest to chemists and biologists whose view is not limited by, at the one extreme, severe reductionism and, at the other, by a purely observational approach to understanding the complexity of living systems.

Amsterdam, The Netherlands Bryan Hanley

Contents

1 Natural Product Chemistry. Reductionism and Biosynthesis 1
1.1 Introduction to Natural Product Chemistry 1
1.2 Non-medicinal Natural Products 6
1.3 The Alchemical Basis of Natural Product Isolation 14
1.4 The Isolation and Purification of Bioactive Natural Products 17
1.5 Chromatographic Techniques in the Purification of Natural Products 22
1.6 The Isolation and Purification of Sensorially Active Natural Products—Colours and Flavour/Odour Compounds 29
1.7 Structure Determination of Natural Products 37
1.8 Synthesis and Biomimetic Synthesis 38
1.9 Biosynthesis—Techniques and Applications 43
1.10 Biological Engineering of Biosynthetic Pathways 45
References 46

2 Improbable Structures 53
2.1 Basis of Improbability and Nature 53
2.2 Improbable Structures 56
2.3 Proof of Structure Natural Product Isolation and Synthesis 59
2.4 Beta Lactamases and Anti-microbial Resistance 63
2.5 Approaches to Anti-microbial Resistance (AMR) 67
2.6 The Microbiome and Quorum Sensing 68
References 72

3 Naturally-Occurring Sulphur Compounds 77
3.1 Sulphur-Containing Primary Metabolites 78
3.2 Mustard Oil Glycosides (Glucosinolates)—Chemical Properties 80
3.3 Sulphur Compounds of the Genus Allium—Chemical and Biological Properties 84
3.4 The Biological Consequences of Natural Products of Plants of the Genus Brassica 88

3.5 Biological Effects of Glucosinolates in Humans 91
3.6 Natural Product Bioactivity 94
References 98

4 The Biosynthesis of Natural Products 105
4.1 The Principles of Natural Product Biosynthesis 106
4.2 The Development of Natural Products—From Molecules to Genes 107
4.3 The Biosynthesis of Polyketides. A Paradigm for Progress of Natural Product Chemistry 111
4.4 Polyketides in Nature 115
4.5 Mutations and Expression Systems to Effect Biosynthetic Modifications 117
4.6 Synthetic Biology and Biosynthesis 118
References 125

5 Genetic Manipulation in Agriculture 131
5.1 Traditional Plant Breeding—Opportunities and Impacts 132
5.2 Molecular Biology of Plants 136
5.3 Manipulation and Expression of Non-homologous Genes 137
5.4 Gene Editing 140
5.5 The Nature of Biosynthesis—Natural and Unnatural Products 143
5.6 Regulatory Aspects of Genomic Modification 144
5.7 Long-Term Options for Genetic Technologies in Natural Products 147
References 149

6 Bioengineering 155
6.1 Living Systems as Machines 156
6.2 Principles of Bioengineering 159
6.3 Bioengineering of Natural Products—The Example of Polyketides 162
6.4 Gene Manipulation and Heterologous Gene Expression Systems 165
6.5 Polyketide Synthases. Systems Biology and Bioengineering 169
6.6 Natural Products and the World of Utility 170
References 173

7 Enzymes and Natural Products 177
7.1 Enzymes and Natural Products 178
7.2 Enzymes in Novel Environments 183
7.3 Enzyme Engineering and Directed Evolution—Successes and Failures 188
7.4 Abzymes and Molecular Imprinting 192

7.5 Conclusion 194
References 195

8 Receptors and Natural Products 199
8.1 Receptors and Cell Signalling—A Cascade of Events 201
8.2 Natural Products and Enzyme Inhibition 203
8.3 Taste Receptors and Natural Products 207
8.4 Natural Products as Agonists and Antagonists 211
8.5 Summary and Conclusions 215
References 216

9 Natural Products and Microbiomes 223
9.1 Definition and Characteristics of a Microbiome 224
9.2 Natural Products and Microbiomes 225
9.3 Quorum Sensing and Natural Products 228
9.4 Modification of Pathogenicity and Biofilm Formation by Natural Products 230
9.5 Natural Products—A Changed Paradigm 239
References 243

10 The Future of Natural Product Research. New Horizons 249
10.1 Reductionism and Complexity 249
10.2 Heterogeneity of Biosynthesis of Natural Products 252
10.3 Natural Products and Biological Systems 254
10.4 The Future Directions of Natural Products 255
10.5 Synthetic Genes 257
10.6 Extremophiles 257
10.7 New Techniques and Methods 258
10.8 Conclusion 259
References 259

List of Figures

Fig. 1.1 Proposed synthesis of quinine by Perkin 2
Fig. 1.2 The beginning and end of the Stork synthesis of quinine (16 steps) 3
Fig. 1.3 6,6′-Dibromo indole (Tyrian Purple) 7
Fig. 1.4 Indigo dye 8
Fig. 1.5 Modern synthetic pigments 10
Fig. 1.6 Structures of PROP and goitrin 11
Fig. 1.7 Enzymic Hydrolysis of glucosinolates—Principal Products 16
Fig. 1.8 Synthesis of quinine from quinotoxine 20
Fig. 1.9 Silanol group 23
Fig. 1.10 Synthesis of a cyclobutane to use as an imprinting molecule . . . 25
Fig. 1.11 Removal of malodorous thiols by reaction with alpha angelica lactone 34
Fig. 1.12 Structures of menthol and menthones 35
Fig. 1.13 Enzymic Hydrolysis of glucosinolates—Mechanistic Detail 36
Fig. 1.14 Formation of goitrin from progoitrin 37
Fig. 1.15 Formation of 6-amino penicillanic acid from Penicillin G 41
Fig. 1.16 The Robinson synthesis of tropinone 42
Fig. 2.1 The Walden inversion 54
Fig. 2.2 Formation of an isothiocyanate by enzymic Hydrolysis of a glucosinolate 55
Fig. 2.3 Structure of a beta lactam ring 56
Fig. 2.4 Thiazolidine and oxazolone rings 56
Fig. 2.5 Binding of penicillin to bacterial cell wall transpeptidase 57
Fig. 2.6 The biosynthesis of glucosinolates 59
Fig. 2.7 Lactam structures 64
Fig. 3.1 The structure of NABAM (ethylene bis dithiocarbamate) 79
Fig. 3.2 Production of isothiocyanates from glucosinolates 81
Fig. 3.3 Enzymic Hydrolysis of glucobrassicin 82
Fig. 3.4 Disulphide from attempted trapping of indolyl-3-methyl isothiocyanate 83

Fig. 3.5 Representative S compounds formed in allium species 85
Fig. 3.6 Formation of lachrymatory factor 86
Fig. 3.7 Formation of goitrin 92
Fig. 3.8 Structures of goitrin, PTC and PROP 93
Fig. 3.9 Rate limiting step in cholesterol formation 96
Fig. 4.1 Structure of chloroquine 109
Fig. 4.2 Summary of polyketide biosynthesis 112
Fig. 4.3 Structure of 6,6′-dibromoindigo 116
Fig. 4.4 Manipulation of indigo for dying 120
Fig. 4.5 Formation of indigo from precursors 121
Fig. 5.1 Structures of docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA) 139
Fig. 5.2 Structure of ajoene 148
Fig. 5.3 Structure of sulphoraphane 148
Fig. 6.1 Structure of quinine 166
Fig. 7.1 Structures of NSAIDs 179
Fig. 7.2 Representative bitter compounds 180
Fig. 7.3 Examples of glycosides 181
Fig. 7.4 Hydrolysis of 1-methoxy indole glucosinolate in a low water system 187
Fig. 7.5 Putative transition state and stable transition state analogue 193
Fig. 8.1 Structure of 3-nitropropanoic acid 205
Fig. 8.2 Structure of galantamine 206
Fig. 8.3 A Dithiocarbamate as a ‘pro-drug’ for hydrogen sulphide 207
Fig. 8.4 Goitrin with NH–C=S feature highlighted in bold 208
Fig. 9.1 N-acyl homoserine lactone structure 228
Fig. 9.2 Pyocyanin 232
Fig. 9.3 Some representative mycotoxins 233
Fig. 9.4 Structure of triptolide 238

Chapter 1
Natural Product Chemistry. Reductionism and Biosynthesis

Abstract The key concepts that lie within the realm of natural product chemistry are considered. This includes a consideration of isolation, synthesis and biosynthesis of natural products and the importance of reductionism and complexity. Kant's view of empiricism is mentioned. Particular examples are given from the area of bioactive natural products (medicinal chemistry), colour and flavour compounds and interactions of natural products with living systems. Themes are introduced which are subject to more detailed consideration in succeeding chapters. Particular attention is given to isolative techniques and specific examples examined (for example, the use of chromatography) and to bioactivity guided isolation. The concept of the 'Magic Bullet' as described by Ehrlich is discussed and its ramifications noted. Finally two important aspects of natural product chemistry—the interaction of compounds with complex biological systems and the idea of engineering pathways to optimise natural product isolation is introduced.

Keywords Natural products · Isolation · Synthesis · Biosynthesis

1.1 Introduction to Natural Product Chemistry

Chemistry can be considered to be the ultimate empirical scientific discipline. At its root (and as a crucial component of the original aims of alchemy and even to pre-empirical science), it is an endeavour that is grounded in the need to carry out experiments and to then explain what is observed. Recording in detail what you did, what the result was and the logic behind any changes you might make to the experimental protocol is essential. The trail of successes and failures is both a record of what works and, simultaneously, a window into how the world was thought of at the time. To take two examples, Perkin's attempts to synthesise quinine and the later successful procedure developed by Woodward and even later by Stork represent different ways of working and different levels of understanding. In Perkin's time, even the tetrahedral nature of carbon bonds was not yet known and transformations were carried out on the basis of being able to determine the numbers of different atoms

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_1

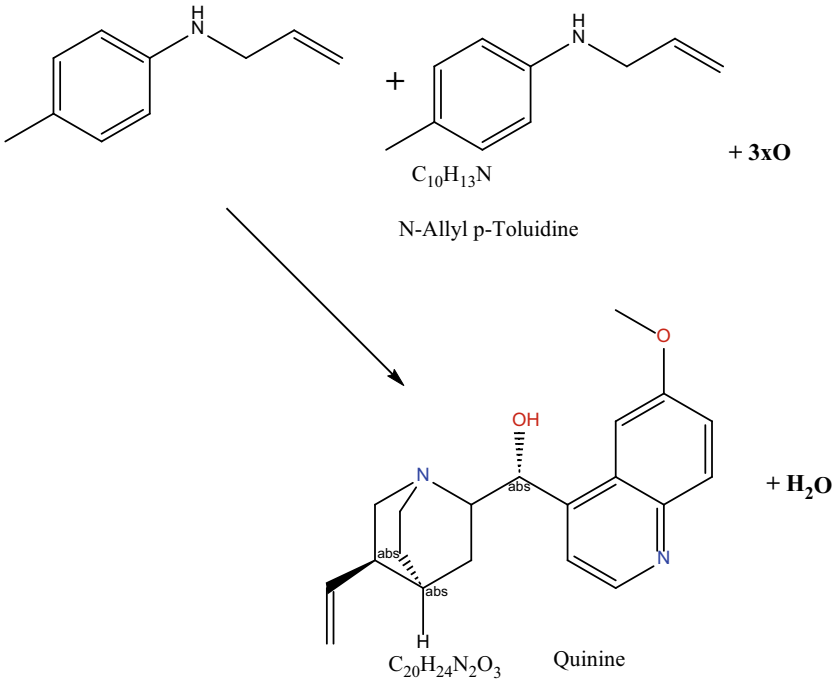


Fig. 1.1 Proposed synthesis of quinine by Perkin

in a desired product and these were used to combine plausible starting materials with the addition of primitive reagents and the application of heat to carry out a 'metamorphosis'. The basis of this reaction using 2 molecules of N-allyl p-toluidine and the addition of oxygen in the form of an oxidising agent is shown in Fig. 1.1.

By the time Woodward and, later, Stork, carried out the chemical synthesis of quinine (or in the case of Woodward, a significant precursor) the understanding of structures, reagents and analogous reactions were very much better understood. This is clear from the reaction scheme followed by Stork (Fig. 1.2). This synthesis starts with 4-Vinyl butyrolactone and comprises no less than sixteen steps.

The understanding of reactions and transformations travels beyond classification of the type carried out for plants and animals since it also encompasses change and transition. In addition to observing, chemists also need to explain and carrying out experiments is a crucial aspect of that. There is also a philosophical basis for chemistry which travels from the known properties of materials to their molecular characteristics. Observing complex molecules as they react and interact in real time is beyond the experimental scope of chemistry. What is within grasp are the predictable consequences of such interactions once the reactions are completed. Based on our knowledge of the molecular properties of chemicals, we can engineer conditions such that they respond in a way that we are able to predict and which, in turn, increases our confidence in our molecular models.

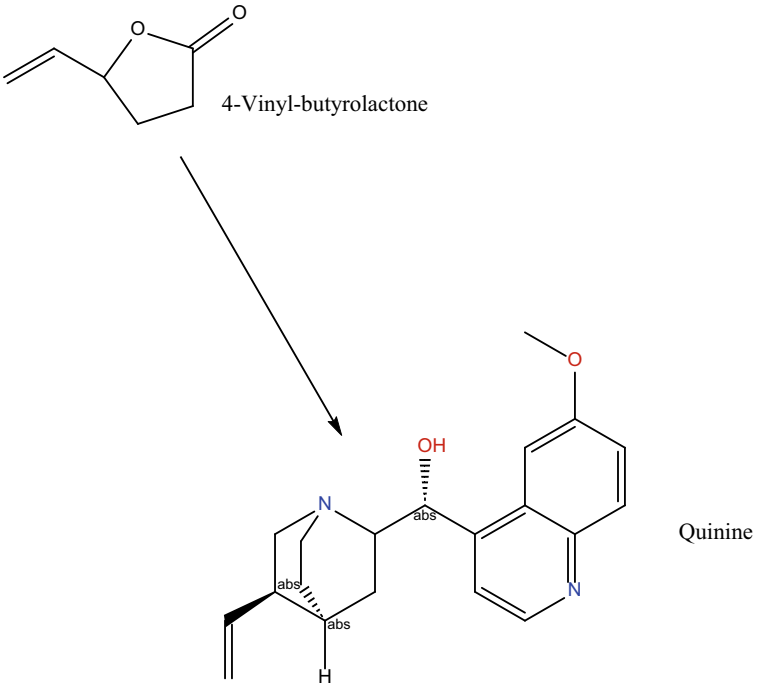


Fig. 1.2 The beginning and end of the Stork synthesis of quinine (16 steps)

In the case of natural products, the molecular structures do not exist in isolation but are part of a larger and much more complex living system by which they are created and with which they interact. The known and examinable part of natural product chemistry is dependent on our ability to measure properties and to extrapolate from our instrumental methods to the molecular structures. Previously whole living systems were difficult to interrogate due to their complexity and entangled interrelationships. Isolative procedures on living systems can additionally have the unfortunate consequence of destroying the integrity of the living system itself. The history of natural product chemistry contains many stories about the application of reductionist techniques to break apart complex systems so that they can be examined at a level that is comprehendible. From these molecular fragments, chemists have tried to reimagine the complex interrelationships that govern the whole system. There are, then, two extremes—a molecular structure that cannot be examined directly but whose properties can only be inferred and a complex living system whose precise structure and function can only be extrapolated from an examination of its constituent parts. Natural (and unnatural) products sit at this juncture looking towards the interpretation of molecular interactions to explain complex systems.

There have been (and still are) several drivers for the study of natural products. Many investigations were initially stimulated by the opportunity to isolate medicinally beneficial compounds and to regulate the amounts of material administered.

This type of activity has its origin in the production of herbal and other remedies. These became part of the distinctive medical tradition of different cultures which were, at times, intermingled and expressed as combinations of different ways of thinking. The benefits of plant-based remedies were known to the ancient Greeks. These included oregano, marjoram, dill, fennel, mint, rosemary, siderites, sage, camomile, thyme, basil and St Johns Wort (Hypericum). Traditional Chinese medicine promotes a holistic approach to health by balancing what is described as the 'vital life force' or qi. Herbal remedies were developed and considered to be an important aspect of Chinese medicine. One of the key principles that underlie Chinese medicine is the Confucian principle of Eight Principles that are organised into opposing pairs. A crucial aspect of this thinking is that of balance. This is analogous, in some ways, to the idea of homeostasis however, as we will see, simple concepts of healthy homeostasis and resilience provide an incomplete picture. Another approach is that of the five elements and it has been suggested that this can be applied on a cellular as well as an organotypic level [1]. The corresponding Indian system of medicine is called Ayurveda. It also employs specific herbs and plants including cumin, turmeric, cardamom and others. While traditional herbal remedies have a long history and significant numbers of adherents who will attest to their effectiveness, their use remains somewhat controversial. Much of the European interest in herbalism derives from the Greek and Roman traditions and is called Galenism. The essence of Galenism has something in common with both Chinese and Ayurvedic principles. Galenism (which was codified by Hippocrates and Galen) suggested that the body contains 'humours' and that keeping these in balance was the key to good health. The four humours were blood, phlegm, and black and yellow bile. Each humour was associated with a different aspect of life—simplistically described as earth, fire, air and water. These attributes also described particular personality types. The Galenic approach was supplanted by the thinking of Paracelsus (Theophrastus von Hohenheim). Perhaps the main difference between the approach advocated by Paracelsus and the Galenic paradigm was the derivation of ill health. In the Galenic philosophy, the body was composed of humours and it was an imbalance between these internal characteristics that led to ill health. Paracelsus proposed that illness arose from external substances that the body came into contact with rather than the disruption of an internal balance. This led him to propose treating 'like with like' and to consider the role of external agents rather than internal balance. This led him towards the use of chemicals (including inorganic salts) and to emphasise the importance of using the correct dosage. In this sense, Paracelsus was an early exponent of the concept of the Therapeutic Index (TI)—the ratio of the lethal dose to the effective dose. The difference between the efficacious dose and the toxic dose is the therapeutic window. Another aspect developed by Paracelsus was the concept that a disease could target a specific organ in the body. This is essential to any concept of selective toxicity. Some of the ideas that Paracelsus suggested were somewhat less recognisable as being scientifically underpinned. For example, if you had a malady in your head then the best cure was to consume some of the head of a healthy individual who had died in their prime (those who died in battle or who were executed were suitable candidates). This interpretation of treating 'like with like' led to a number of remedies that would

not be recognised as such. In late Medieval times the 'mosses of dead men's skulls' were a recognised therapy [2]. The wound-healing properties of skull lichens were well known. More recently it has been noted that lichens are a source of antibiotics so this may be a very early use of natural products as curative agents [3]. There is a tradition of using *Usnea longissimi* and other members of the class of lichen to treat wounds and as a source of antibiotics [4]. The active principles are thought to include usnic acid (2,6-Diacetyl-7,9-dihydroxy-8,9b-dimethyldibenzo[b,d]furan-1,3(2H,9bH)-dione), a polyketide-derived molecule which, while exhibiting a range of biological activities, is also toxic [5].

The ideas of Paracelsus were an important aspect of the foundation of the medicinal basis of the development of natural product chemistry. Another major advance was the concept of the 'Magic Bullet'. The idea of a magic bullet, a medicine or other remedy that is wonderfully specific and efficacious, was used by Ehrlich in 1906 when he used the German word zauberkugel. Ehrlich would also have been aware of the German myth related to 'magic bullets'. A "Freischütz" ("freeshooter"), in German folklore, is a marksman who, by a contract with the devil, has obtained a certain number of bullets destined to hit without fail whatever object he wishes. The tale became widely circulated in 1811 when Johann August Apel included it as the first tale in the first volume of the "Gespensterbuch" or "Book of Ghosts" and was later adapted into opera. Der Freischütz, Op. 77, J. 277, (usually translated as The Marksman) is a German opera with spoken dialogue in three acts by Carl Maria von Weber with a libretto by Friedrich Kind. It premiered on 18 June 1821 at the Schauspielhaus Berlin. It is considered the first important German Romantic opera and was hugely popular throughout the nineteenth century in Germany and across Europe. Ehrlich was born in 1854 in Upper Silesia and studied both in Leipzig (where Apel came from) and in Berlin and so was likely aware of the folk tale and/or the opera. In Ehrlich's view, a magic bullet was a chemical agent that would destroy an infective microbe without harming the host. The first such compound was an organo-arsenic compound which was called Salversan. It was one of a batch of compounds synthesised by the German chemist Albert Bertheim and tested in Ehrlich's laboratory. This followed Ehrlich's discovery that some of the dyes used to stain cells, developed by Koch, had some anti-infective properties however these were limited and resistance developed. This most celebrated example of a selectively toxic agent was responsible for a significant decrease in the prevalence of syphilis and was the medication of choice for treatment of syphilis until the development of natural product antibiotics like penicillins and cephalosporins. While the arsenical nature of Salvarsan was not in any doubt, the original proposed structure (with an As=As bond) was later amended and the nature of the active agent reassessed.

1.2 Non-medicinal Natural Products

While the discovery of bioactive agents was a considerable spur to the development of natural product chemistry, other uses of plant, animal (including insects) and microbially-derived compounds was also a significant contributor to chemistry understanding and the development of how natural products were constituted and reacted.

One of the most important aspects of chemistry is that it is both an empirical and a reductionist exercise. The role of abstract thought in chemistry is something that occupied Kant and other philosophers. It has been suggested that Kant believed in a hierarchy of sciences and he thereby classified the various sciences according to the respective importance they each accord to abstract thought. Kant believed that the laws of physics were able to be derived independently of experience. Chemical laws, by contrast, must be validated through empirical observation and experiment. This raises the paradox of whether the laws of chemistry are actual laws or are just the summation of empirical observation. Kant's view, at its most simplistic, is that science is divided into proper science (physics) and improper science (everything else). However, within that second category there are sciences like chemistry whose laws are conditional upon observation and are not a property of the objects that the science is applied to. However chemistry is empirical and rational unlike, for example, psychology whose laws are not purely abstract thought (like physics) nor able to be derived from empirical observation and experimentation (like chemistry). In the Kantian universe, then, the rational feature of chemistry is that it can be inspected and its laws derived from rational observation. Another important aspect of chemistry relates to scale. Mathematics deals with the world be explaining it in terms of equations—that was one of the attractions of Newtonian thought. Physics considers basic systems on a sub-atomic level and includes the quantum world. Chemistry (at least in its connection to living systems) is primarily preoccupied with molecular interactions and explaining he world in these terms. Biology deals with whole systems from cells to whole and complete living entities. This simplistic, scaled view of different scientific disciplines has been challenged—not least by the development of ideas related to quantum biology [6] and biological Dark Matter [7].

In practice, chemistry actually has a dual function and can thereby unite abstract thought and empiricism through the development and testing of 'proofs'. In the development of natural product chemistry, a primary driver was the isolation of 'useful' chemical entities. These were important per se but were also important in providing a way of validating ideas about molecules and their interactions. As we have seen, uniting abstract thought and empiricism was an important aspect of the growth and popularity of alchemy. It also, as we will, see impacted on the use of natural products for other purposes such as the production of dyes and colourings, food and flavourings and others. The other driver is the use of natural products as a way of probing how living systems operate in all of their complexity. The meaning of living systems and why they operate the way they do has both empirical (observational) and abstract aspects. The contribution of natural products to the development of dyes

and to the study of flavour provide examples of this dichotomy between seeing/doing and thinking.

The history of dying and dyestuffs was initially a result of the need to develop signifiers of rank and entitlement through the wearing of clothes of a particular colour. The Stockholm papyrus (Papyrus Graecus Holmiensis) contains 154 recipes for the manufacture of dyes and colours to create artificial precious stones and gems. It was originally written in Greek and is one of the earliest descriptions of ancient chemical methodology. An English translation was made [8] and this increased awareness of the recipes. A similar manuscript in the Rijksmuseum van Oudheden in Leiden, also contains different recipes for the manufacture of materials.

Natural dyes came from three sources—plants, animals (including shellfish and fish) and minerals. The dyes that were used were of more or less availability and their use was, therefore reflective of the status or wealth of the individual—there was a distinct hierarchy in the use of dyes. One of the rarest (and hence most expensive) dyes was Tyrian Purple which was derived from the mucus derived from Mediterranean sea snails. Three different species of snail were used—Hexaplex trunculus (bluish purple), Bolinus brandaris (reddish purple), and Stramonita haemastoma (red). The secretions of the snails is colourless and must be processed to create the red/purple colour of Tyrian Purple. Up to 10,000 snails were required to make just one gram of the dye and it was more expensive than gold. Unusually for a natural dye, Tyrian Purple is both bright and colourfast. The main chemical component which creates the colour is indigo dye (see below) and the corresponding 6-bromo indigo and 6,6′-dibromoindigo (Fig. 1.3).

A range of other colours including red/purple, blue and violet dyes can be produces from different marine molluscs. The variations in colour are a consequence of the presence, in varying amounts, of a range of indigo-related compounds.

Blue colours are also obtained from woad (Isatis tinctoria) and the indigo plant itself (Indigofera tinctoria). Woad is a member of the Brassicaceae family and the blue colour produced by processing of woad leaves was used for colouring clothes, bodies and as an ink for manuscripts. The principal blue colouring compound is indigo dye—[2(2′)E]-[2,2′-Biindolylidene]-3,3′ (1H,1′H)-dione (Fig. 1.4).

While extraction from plant material was the main source of indigo until the twentieth century, an effective synthesis was carried out on an industrial scale. The intermediate in any synthesis of indigo is indoxyl which oxidises in air to give indigo. The industrial method which came to be adopted was developed by BASF and relies

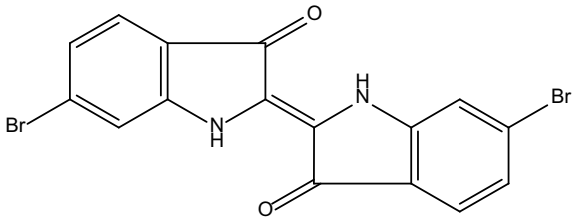
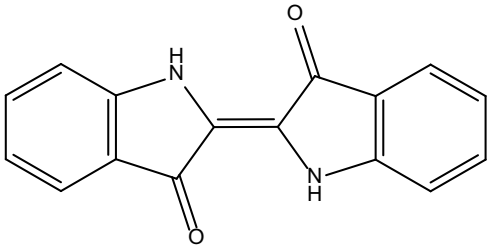


Fig. 1.3 6,6′-Dibromo indole (Tyrian Purple)

Fig. 1.4 Indigo dye



upon the synthesis of N-(2-carboxyphenyl)glycine from the easy to obtain starting material, aniline. This process was in use by 1897, at which time 19,000 tons of indigo were being produced annually from plant sources. While the synthetic route to indigo (and other common colours) largely replaced extraction as the preferred source, recently interest has grown in the use of natural colours. This is in part a move away from petrochemicals and towards more sustainable approach. Further impetus was given to the use of natural colours by a study carried out in 2007 in which 6 synthetic food colours and/or the preservative sodium benzoate were possibly implicated in hyperactivity in children [9]. The six colours were tartrazine (E102), quinoline yellow (E104), sunset yellow (E110), carmoisine (E122), ponceau 4R (E124) and allura red (E129). While the study suggested a link between the specific colours tested, sodium benzoate and hyperactivity the UK' Committee on Toxicity noted that "increases in mean levels of hyperactivity observed in this study were small relative to normal inter-individual variation and that changes in behaviour were not evident in all children in any one group and were not consistent across age groups or across the different mixtures used in the study" (COT statement 2007/04. September 2007). It should also be noted that a range of other dietary constituents have been associated with an effect on ADHD in children. Nonetheless, it appears that natural products have undergone something of a renaissance in relation to their use in foods as colouring agents [10].

The use of natural colours in art has similarly been a progression from natural products to synthetic analogues and then back to naturally-derived materials. While it is not possible to definitively date the first example of art, the beginnings of figurative art occur in caves and are thousands of years old. The earliest pigments were iron oxides (red) and charcoal (black). In addition other minerals were ground up and combined (sometimes after heating to alter their chemical composition and colour) and suspended in water before use. The principal mineral-derived colours used in cave art are black (often from charcoal, soot, or manganese oxide), yellow ochre (often from limonite—hydrated iron oxides), red ochre (haematite, or baked limonite), and white (kaolin clay, burnt shells, calcite, powdered gypsum, or powdered calcium carbonate). The collective term 'ochre' is used to describe the iron-based (generally red) materials used in cave painting. These earth pigments were the first to be used since they were persistent (did not fade) and required minimal treatment before use. The Egyptians introduced the idea of further stabilising pigments by using a process to produce 'lakes'. This involved fixing colours onto an inert binder. It can be used

with all types of dyes including those derived from plants. By adding various salts, the colours of lakes can be adjusted.

Mineral dyes fall outside the area of natural products. Dyes used for art can be produced from a number of plants and these produce a range of colours, as we have seen previously, including blue (woad and indigo plant), rose madder (madder plant) and carmine red (cochineal). The nomenclature of colours that occur in nature was important both for artists and naturalists but was largely subjective and this resulted in confusion. The American naturalist, Robert Ridgeway became frustrated at the lack of any systematic nomenclature for the colours that occurred in nature and in 1912 he published a book titled "Color Standards and Color Nomenclature". He mounted a whole range of colour blocks and named 1115 colours. This huge undertaking arose from his earlier publication in 1886—"A Nomenclature of Colors for Naturalists and Compendium of Useful Knowledge for Ornithologists". The fascination with colour in nature (and the development of classification) underpinned Darwin's work that led to the publication of his great work 'On the Origin of Species by Means of Natural Selection'. However it is noteworthy that the first edition of Darwin's tome preceded the Ridgeway classification by some 30 years so both can be considered to be part of a general move towards classification of nature but the process was perhaps most widely applied by Linnaeus in his classification publishing in 1735. There is, therefore, an implied reductionist theme. First comes classification of plants and animals (Linnaeus) then an explanation of how those classifications arise (Darwin) and finally a classification of one of the key components of nature, colour (Ridgeway).

In art, the development of portable colours (particularly the invention of the metal tube for pigments) allowed painting en plein air and, ultimately, facilitated the growth of the Impressionist Movement. By the time tubes (with caps) were introduced in the early part of the twentieth century, both synthetic and natural colours were available. By 1905, Monet, in answer to a question about which colours he used, said "As for the colours I use, what's so interesting about that? I don't think one could paint better or more brightly with another palette. The most important thing is to know how to use the colours. Their choice is a matter of habit. In short, I use white lead, cadmium yellow, vermilion, madder, cobalt blue, chrome green. That's all." The Impressionists as a group used a range of colours derived from cobalt, copper, chromium and other Transition metals. The main colours that are commonly used today which they did not have access to are phthalocyanine, quinacridone, azo colours, and titanium white (Fig. 1.5).

In addition, French Ultramarine had been synthesised by heating together kaolin, sodium carbonate and sulphur and was a much cheaper and more available alternative to lapis lazuli (a rare mineral from South Asia).

While many colours used by artists are now synthetic and generally have improved colour fastness and are more readily available in large amounts (as well as being cheaper), a number of manufacturers still produce paints derived from extracts containing natural products.

One of the most interesting aspects of natural colours is the appearance of iridescence. These colours (also termed 'structural colours') change when viewed from different angles. It has been suggested that they owe their properties to the wave

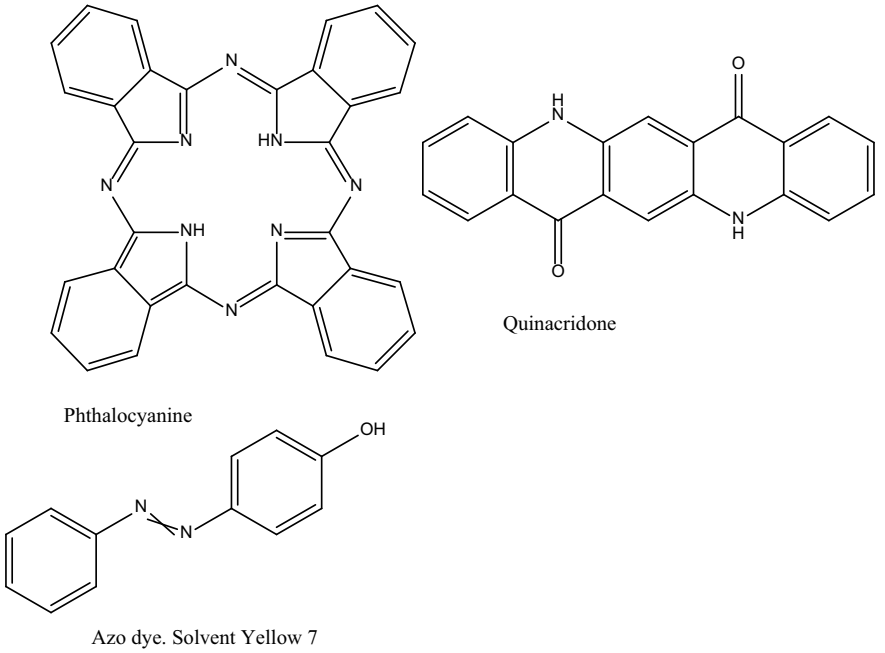


Fig. 1.5 Modern synthetic pigments

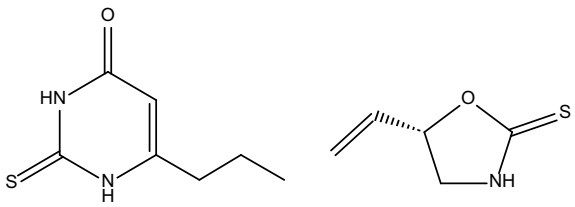
nature of light. Pigments create colour by absorbing certain wavelengths of light and reflecting others. This affects the way that colours are formed. Mixing coloured light (as Newton did in his investigations into prisms) will result in additive mixing and therefore different wavelengths are added together. Eventually, if all the visible wavelengths are added then white light is obtained. In the case of mixing paints and pigments, the opposite effect (subtractive mixing) is the underlying principle. Each pigment absorbs light of particular wavelength and only reflects back a proportion of the light. As more and more pigments are added together, more and more wavelengths are absorbed and less and less of the visible spectrum is reflected until eventually the mixture will be perceived to be black. In practical terms, this means that mixing pigments to get different colours will make the resultant mixture less 'colourful' and so single pigment colours will be stronger and more luminous. Iridescence is a consequence of light being reflected and also passing through an object. Some of the light that passes through is subsequently reflected. The two types of reflected light combine and, due to the wave nature of light, this produces an interference pattern and the colour varies depending on the position of the observer. Iridescent colours combine in the way that light combines rather than the way that traditional colours combine. Examples of iridescence are common in nature and include shells, fish, the colours of some flowers and the wings of insects including butterflies. Such effects were created by, for example, applying iridescent materials from nature directly to the surface of paintings on paper. Such appropriation increased the popularity and

use of materials such as Mother-of-Pearl. The use of substances such as mica flakes have helped in the production of iridescent artist paints, and these are commonly available.

The second example of the use of natural products for a non-medicinal purpose is the study of flavours. These will be discussed in more detail later however some examples will illustrate their potential significance.

Flavour chemistry and its understanding has something in common with some types of medicinal chemistry. Just as one of the mechanisms whereby bioactive compounds exhibit their activity by binding to receptors, so the perception of taste is governed by the binding of compounds to taste receptors. The role of taste receptors and taste buds in the appreciation of ingested components has been extensively studied and reviewed [11] Taste receptors have different characteristics—some are G-protein coupled receptors (GPCRs) while others are ion channels. Sweet receptors are classified as belonging to the T1R family and are of 3 types. Most mammalian species contain genes for up to 20 bitter receptors (T2R type). Some bitter receptors respond to a small number of specific molecular types while others are considerably more catholic in their affinity for bitter compounds. The large number of bitter receptors was postulated to be a response to the need to detect a range of toxins—many of which have a bitter taste profile, however further studies suggest that many bitter taste receptors have an extra-oral role. Within some of the bitter taste receptors there are a number of polymorphisms that affect the binding and subsequent signal transduction of bitter compounds. One bitter receptor in particular T2R38, seems to be involved in a number of biological effects (not just taste) and these may have alternative physiological effects. Polymorphisms in T2R38 have been associated with vegetable intake [12]. The perceived bitterness of (in particular) Brassica vegetables is associated with polymorphisms in T2R38. One particular allele (called PAV) is associated with sensitivity to the bitter test compound 6-n-propylthiouracil (PROP) and also with an aversion to Brassica vegetables. This may be a consequence of the structural similarity between PROP and goitrin (a compound commonly present in Brassica vegetables) (Fig. 1.6).

Other members of the T2R family of bitter receptors are found in gut epithelial cells and in pancreatic cells and their presence is associated with the release of hormones involved in appetite regulation, such as peptide YY, glucagon-like peptide-1 and ghrelin [13].



6-n-Propylthiouracil (PROP)

Vinyl oxazolidine thione (Goitrin)

Fig. 1.6 Structures of PROP and goitrin

In addition to their presence in the GI tract, T2Rs are also found in immune cells and have a role in the innate immune response [14]. Bitter taste receptors occur in a number of extra-oral sites including the brain, bladder and lower and upper respiratory tracts. Of particular interest is the impact of T2Rs on quorum sensing. Quorum sensing is the regulation of gene expression in response to fluctuations in cell-population density. Bacteria which make use of quorum sensing produce and release chemical signal molecules called autoinducers that increase in concentration as a function of cell density. These in turn have an effect on virulence and on the formation of biofilms (see below) One particular class of quorum sensing molecules are homoserine lactones and some members of the class have been found to interact with T2Rs and to induce a range of responses including phagocytosis and inhibition of biofilm formation [15]. This suggests that a primary purpose of T2Rs outside of the oral cavity may not be solely associated with detection of bitter compounds.

The link between bitter compounds and cancer has been noted. One potential side effect of activation of bitter receptors by pharmaceuticals arises from the activation of the ATP-binding cassette B1 (ABCB1). Stimulation of ABCB1 increases the efflux of bitter compounds from extrauterine cells [16]. This can occur in response to specific bitter tastants binding to T2R38. While a link between consumption of Brassica vegetables such as cabbage and the prevention of cancer has been the subject of considerable interest [17] only a small number have considered the effect of Brassica vegetables on existing cancers. In one study it was found that the incidence of N-nitrosobis-(2-oxopropyl)amine (BOP)-induced skin and pancreatic adenocarcinoma was elevated in cabbage-fed hamsters [18]. Feeding the animals cabbage in a high fat diet also elevated the yield of BOP-induced pancreatic ductular carcinoma and may have limited the effectiveness of anti-cancer medication. It was found that TAS2R38 was expressed in pancreatic ductal adenocarcinoma cells Both certain bitter compounds and quorum sensing molecules are able to act as agonists and this leads to activation of ABCB1 [19]. The upregulation of ABCB1 could increase the efflux of anti-cancer drugs and other chemoprotective agents from the cell. The authors suggested that TAS2R38 may take part in pancreatic cancer progression via the induction of chemoresistance mechanisms. Given the known facility of goitrin to bind to the TAS2R38 bitter receptor [20], it is reasonable to suppose that the effect of Brassica vegetables on recovery from cancer may be worth investigating. Interestingly it has been found that a proportion of the population are ‘super bitter tasters and have a polymorphism in their TAS2R38 gene that makes them particularly sensitive to bitter compounds such as those found in Brassica vegetables. In a nutritional epidemiological study carried out, “Super-tasters had about a 58% higher risk of cancer incidence, and the tasters had about a 40% higher risk of developing cancer, compared to women who were classified as non-tasters,” however the study did not examine cancer endpoints (mortality rates) and there was a further complication in that “cancer risk was higher among women classified as tasters or super-tasters, but we didn’t see any differences in bitter-tasting vegetable consumption.” [21]. It is not clear if super bitter tasters have a higher induced level of expression of ABCB1 than normal tasters. The principle is clearly observed—a receptor that was first defined in terms of the chemical entities and physiological/behavioural observation was later

found to have other properties. The existence of a family of bitter taste receptors was predicted by a geneticist (Lush) who studied mouse strain differences in bitter taste avoidance, and who suggested that a cluster of bitterness-tasting genes "have evolved from one original bitterness gene by a process of local duplication and differentiation" [22]. The genes themselves were identified by genomic analysis as being linked to bitter taste responsiveness in humans and mice. Interestingly the number of structurally distinct compounds perceived by humans as bitter is much larger than the number of bitter taste receptor genes, implying that each bitter receptor responds to more than one bitter ligand. In addition, human bitter receptor genes have substantial diversity of coding sequence which suggests that receptor polymorphisms are responsible for the genetic component of individual differences in bitter taste. The question outlined above for super bitter tasters remains, what is the physiological purpose of these polymorphisms apart from an increased aversion to certain bitter compounds. Of course many bitter compounds (e.g. naturally occurring alkaloids) are bitter so an enhanced aversion to such compounds may confer an advantage on those with particular polymorphisms. It would seem likely, given what we already know about bitter taste receptors in the gut, that other physiological mechanisms (e.g. enhanced efflux of toxins) are also relevant. The use of certain types of plants containing bitter, anti-parasitic compounds by primates infected with parasites has been observed in the wild [23]. In addition parasitized deer and lambs will select and consume the bitter and astringent Puna chicory, thereby reducing their parasite load [24]. In medieval times worms were treated using bitter, purgative plants such as wormwood or gentian, which would both kill the worms and expel them by provoking a bout of diarrhoea.

The relationship of living systems with natural products is, as we have seen from the examples of colour and taste, complex and multifactorial. In both cases the natural products themselves provided the impetus for scientific investigation which yielded further information about the biological system. The biological systems from which the natural products were isolated (or interacted with) have properties which are not necessarily a function of the natural products themselves. For example, while colours can be defined in terms of our (and natures') reaction to them, the property of iridescence requires that we think in terms of the wave:particle duality of light and of interference patterns. This is not simply a consequence of an additive or subtractive approach to colours but is an important consequence of wave and particle theories. The properties of taste are, as we have seen, closely linked to ingestion and consumption but the receptors to which taste molecules are agonists for are much more than part of a taste transduction pathway. They have a range of other physiological functions and may, as we have seen, be involved in holistic aspects of physiological health. While attention has been drawn to one particular member of the class of bitter receptors (T2R38), other members of the class have extra-oral significance also.

The application of reductionist principles has facilitated the study of simpler aspects of naturally occurring molecules and provided a starting point for probing more complex systems. Reductionism implies isolation and purification and these are important aspects of the history and current reality of natural product chemistry.

Some of the techniques that have been developed by chemists are part of, and have their roots in, the alchemical tradition and involve breaking down complex systems (including molecules) until a recognisable fragment can be identified while others are designed to preserve the particular property of the substance of interest so that it can be further studied and used. An overview of these approaches will now be examined.

1.3 The Alchemical Basis of Natural Product Isolation

While alchemy has a reputation of, by current thinking, being unscientific and closely linked with the transmutation of metals into gold, it also overlaps with folk medicine and the use of natural products. In addition, many techniques used by more 'conventional' chemists in the eighteenth and nineteenth centuries have their roots in the investigations and methods used by alchemists. Although alchemy is rooted in misunderstanding about the nature of substances, it was not simply a series of random explorations and, in their records of the details of their investigations, alchemists were true to the traditions of experimentation and empiricism. For example, one of the guiding principles of alchemy was that, under certain conditions, substances reacted together to give something that was different to either of the constituents. This underpinned the principle of chemical reactivity and chemical reactions as being distinct from the result of simply mixing substances together.

Aristotle developed the hypothesis that humans can either observe and mimic nature or attempt to improve and perfect nature. Alchemy was part of this latter tradition of improvement and perfection and, as we will see later, such a distinction was part of the underpinning for the subsequent synthesis of natural products (and analogues of natural products) by chemists. The basis of alchemy involved experimentation and intervention in nature—not simply observation and discovery. As part of this, alchemists sought to determine the nature of matter through experimentation. This quest went beyond metallurgy and technical modification of metals (metalworking) and probed the philosophical nature of substances. Much of alchemy was concerned with the nature of metals and transmutation. However, there was also a tradition of curing disease and finding the means towards eternal life. This was more pronounced in some traditions than in others.

The tradition of alchemy and herbal medicine had a focus upon the preparation of extracts containing active ingredients. The attempted isolation and separation of active principles was a later concern. This distinction between extraction to provide a crude, active mixture and separation in order to optimise the active ingredient (and to understand more about the constituents of an extract) was a key aspect in the progression of the post-alchemical philosophy of science. While alchemical extracts could vary depending on how they were extracted, the idea of separating the constituents of an extract was a paradigm change. The development of chromatography as a technique to separate pigments from plants was first described by Mikhail Semenovich Tsvet in 1905. He filtered an extract of leaves through powdered chalk packed into a

glass tube and achieved the separation of carotenoids and chlorophylls. The more efficient method for extraction by repeated circulation of the same solvent through a solid material was developed by Franz Ritter von Soxhlet, some 25 years earlier in 1879 in Munich. Both techniques were a significant advance on the methods employed by alchemists, were repeatable and led to the greater used of natural products in a defined way with predictable properties.

The extraction of natural products (in particular the solvent used) is dependent upon the nature of the products that are required to be isolated. Many natural products contain sugar residues (glycosides) or have ionic groups (e.g. sulphates and phosphates) and are easily transported in the aqueous-based systems of living systems. These properties facilitate their extraction into polar solvents such as water or water/ethanol mixtures. Others are considerably less polar and may not be available for simple extraction. The simplest method for extraction of natural products, which is still the basis of methods currently used, is suspension of the material containing the product of interest (usually dried then ground up) in a solvent of the appropriate polarity. In the case of alchemy, extraction was a process of trial and error. In addition, the range of solvents used later by chemists was not available and aqueous extractions were most common. The suspension would be agitated for some period and the supernatant isolated by filtration to remove the residual solid matter. If yield is important or if the product is only fractionally soluble in the solvent then the extraction procedure can be repeated. This type of crude extract was the basis of further alchemical procedures. While this methodology is gentle and allows for the extraction of thermolabile compounds, it is generally an inefficient process and is solvent intensive. A further problem arises if the compounds of interest are subject to autolysis by co-occurring enzymes—as is often the case for plant-derived natural products. Hydrolytic enzymes can be released if the cellular structure of the source material is disrupted by grinding it up and suspending it in water. For example the whole class of mustard oil glycosides (glucosinolates) are broken down by a co-occurring thioglucosidase enzyme (myrosinase) to give, after a spontaneous reaction involving a molecular rearrangement, a range of organic compounds including isothiocyanates and nitriles (Fig. 1.7). These compounds cannot be extracted into water but may be distilled and therefore purified in that way.

The second approach to the isolation of crude extracts of natural products involves heating the solvent which may simultaneously break down the living material and improve extraction efficiency. If the solvent is heated to (or near to) its boiling point then the technique is termed alchemical infusion. The extraction can either be carried out by filtering the hot liquid or allowing it to cool first. The third method commonly used in alchemical studies is the process of percolation. In essence, percolation is akin to a single pass Soxhlet extraction. The ground plant material is packed into a suitable vessel and the solvent is passed through it slowly, gradually extracting the compounds of interest. Fresh solvent is added to maintain a flow of extract. The extracted plant material is generally compressed to remove any residual solvent and the total extract can then be used as a source of the compounds of interest.

In summary, alchemical isolation of natural products was typically carried out by one of three methods.

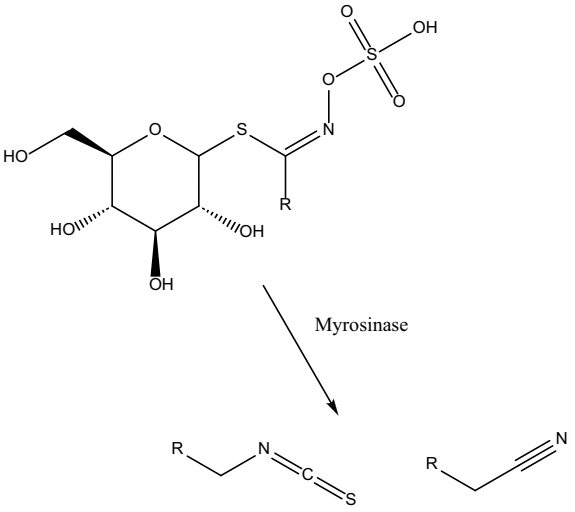


Fig. 1.7 Enzymic Hydrolysis of glucosinolates—Principal Products

1. Cold (generally aqueous) extraction.
2. Hot (infusive) extraction.
3. Repeated (percolative) extraction.

While extraction using water or alcohol (or water alcohol mixtures) were among the earliest and most common methods for obtaining active principles, the use of distillation (as noted above) was also a frequently used technique for volatile compounds. Chemical distillation is the heating, evaporation and condensing of substances. It is one of the simplest methods of purification and is a key step in obtaining oils and similar volatile compounds. In alchemical terms, distillation is the sixth stage of the alchemical process. The substance is heated until boiling and the vapours collected and condensed back into a liquid form. These vapours represent the purification of the essence of a substance. The artist David Teniers the Younger painted a number of representations of alchemists and alchemical workshops. In many of these, alchemical instruments and apparatus can be seen including the three key components of the alchemists distillation practice—curcubits (gourd shaped vessels for heating substances), alembics (a type of condenser that sat on top of the curcubit) and lute (a sticky mixture that sealed the joint between the curcubit and the alembic). While distillation was known in the ancient world (a terracotta distillation apparatus dating from 3000BC was found in the Indus Valley, Pakistan) [25], distillation of alcoholic beverages was also known in the twelfth century. A comprehensive review of various aspects of distillation is available [26] and is a useful reference for this technique.

1.4 The Isolation and Purification of Bioactive Natural Products

The isolative techniques developed during the practice of alchemy form the basis of current extraction technology for natural products. As noted previously, the subsequent development of chromatographic techniques represented a paradigm shift in the purification of natural products. Ensuring the maximum effectiveness of both steps will enable the natural product to be obtained in high yield and purity. Optimisation of extraction is a necessary pre-requisite for the development of subsequent effective and efficient purification procedures for individual natural products. The first consideration is that the solvent (or solvent mixture) should be inert with respect to the compounds of interest and of the correct polarity to effectively extract the compounds of interest—preferably at ambient temperature. Organic solvents, while being the most traditionally used, have certain hazards associated with them including toxicity and flammability. There is also the question of disposal of the excess solvent afterwards. In addition, many natural products are thermolabile and extraction using heated or boiling solvents—which is commonly used to increase the solubility of the extractable chemicals—may result in the breakdown of the target compounds and therefore, reduced yields. A range of preparation techniques have been developed including ultrasound-assisted extraction (UAE), pressurized liquid extraction (PLE) and microwave-assisted extraction (MAE) and these have allowed more efficient extraction to be carried out using more sustainable conditions—particularly less solvent.

A number of alternatives have been developed based around modification of the extracting solvent either to make the extraction process greener or to reduce the amount of solvent required by increasing the solvation power of the solvent. These include the use of ionic liquids, eutectic solvents and supercritical extraction methods. In some cases the solvent can be 'tuned' to be more specific for the extraction of the products of interest. Deep eutectic solvents (DES) are solutions of acids and bases that can form a eutectic mixture. The particular advantage of DES is that they have a low vapour pressure and are non-flammable and relatively non-toxic. While they have application as tuneable solvents, there remains significant unknown details about how to optimise them and their use [27, 28]. The use of natural DES specifically for the extraction of bioactives from plant material has been reviewed. These are eutectic mixtures of two or more natural compounds forming a liquid phase at room temperature based in hydrogen bond acceptors (quaternary amines, alcohols) and hydrogen bond donors (sugars, organic acids) [29]. Ionic liquids are an organic cation paired with an inorganic or organic anion. They have been adopted as an alternative to organic solvents for natural product extraction and a range of uses have been reported [30]. The use of supercritical fluids for natural product extraction has been proposed however the initial capital cost and the specialised equipment required tends to dissuade researchers from the use of supercritical techniques as a first choice. The most commonly used fluid is supercritical carbon dioxide and the use of that (and other solvents) has been reviewed [31]. One of the particularly

attractive properties of supercritical fluids is that they are able to permeate material with an efficiency greater than that of a normal solvent (although less than that of a typical gas). This leads to enhanced extraction yield. In addition, one of the perceived drawbacks of supercritical carbon dioxide—its relatively low polarity—means that by adjusting the pressure in the extraction vessel, selective extraction can be carried out. In addition, co-solvents can be used to optimise selective extraction conditions.

The use of technology and the development of new solvent systems has taken the extraction of natural products beyond the capabilities and knowledge of alchemy. This has improved the efficiency of extraction and active materials that are only present in small amounts are able to be isolated. In addition the synthesis (or semi synthesis) of natural products and analogues of them has facilitated the development of medicinal chemistry. The basis of natural products lies in the concept of the uniqueness of living systems and their components. Modification of natural products—or even the use of natural systems to synthesise chemicals from novel starting materials is considered to give rise to 'unnatural products'. Such a definition is predicated upon the source (derivation) of the product rather than its chemical structure. Even isotopically labelled materials—which are identical to the natural product except in the ratio of different isotopes present—are, by this definition, unnatural. Such thinking underpins the concept of the unique nature of natural products and was expressed through the concept of vitalism. This idea, that materials from living systems contain some 'essence' as a function of their derivation was an influence from Galenic medicine to Ayurvedic and ancient Chinese medicine and to Descartes. Indeed, the principle was only fatally undermined by Wohler's synthesis of urea (a compound that occurs naturally) from inorganic starting materials (lead cyanate and aqueous ammonia) [32]. It is noteworthy, however, that Wohler did not set out initially to disprove vitalism and his teacher Berzelius, remained unconvinced. Indeed one of the starting materials that Wohler used derived from the tanning process and therefore, for an adherent to vitalism, the link between the synthetic urea and a living system was maintained. However, to all intents and purposes, following the experiment by Wohler, the concept of vitalism was a discredited paradigm and this paved the way for the simultaneous growth of natural product and organic chemistry as two facets of the chemistry of carbon-containing compounds.

The growth in interest in bioactive constituents provided much of the impetus for the isolation and purification of natural products in the nineteenth century. Indeed, until the end of the nineteenth century, natural products (and minerals) were the source of all drugs since living systems were considerably more advanced in their capability to synthesise compounds than the chemist in the laboratory. The earliest 'modern' pharmaceutical preparation was morphine which was prepared and marketed by Merck in 1829. The crude opium material had been in use for thousands of years prior to its more recent manifestation as a marketable drug. Crystalline morphine was isolated from opium resin by purification by Serturner in 1805 by adding ammonia to the dried latex from the opium poppy and the process repeated in 1817 [33]. While the empirical formula of morphine was not known until it was deduced by Robinson in the twentieth century, the purification by Serturner allowed for the preparation of safe dosages of morphine and to the purification of a range of other alkaloids [34].

In addition to their molecular characteristics which facilitated their isolation, since the same set of techniques could be used for all of them, the early alkaloids also had a significant biological activity and this facilitated the use of activity-guided separations. Bioactivity guided fractionation has made a significant contribution to the isolation and characterisation of natural products. The approach is, in essence, one of trial and error. A mixture is split into different fractions and each fraction is tested for biological activity. The fractions containing the bioactive principle(s) can be combined and all other fractions discarded. Repetition of these steps (often using different fractionation techniques) gradually removes all the inactive components leaving just the active ingredient(s) which become increasingly more concentrated since it is no longer 'diluted' by inactive components. The development of methods to isolate and purify natural products was a crucial step in the development of safe drugs and this in turn had other consequences. For example, the relevance of quinine as an anti-malarial agent was not only medicinally important but also had far reaching social and commercial implications. In cities where malaria was endemic, wealthy citizens moved to cooler regions—particularly in the summer. The name of the disease implied that it was caused (or exacerbated) by 'bad air' so moving away from the city was a sensible and practical step. Those who could not afford to 'flee to the hills' remained in the city and suffered parasitisation. The prevalence of malaria also inhibited European settlement in Africa. The spread of malaria to South America (by Europeans) led to the realisation that the powdered bark of the Cinchona tree contained a potent anti-malarial substance (quinine). This treatment was already used by the indigenous population to treat a number of fevers and the extension to a new fever disease (malaria) was a logical step. Quinine is also a non-narcotic analgesic. The levels of quinine in cinchona bark extract can vary and, because the therapeutic window is relatively narrow (i.e. overdosing with deleterious side effects or ingestion of an insufficient dose, were not uncommon) and quinine itself is toxic, development of a standardised therapeutic dose was important. In addition, cinchona bark grown under different conditions contains different amounts of the different cinchona, quinine alkaloids (quinine, quinidine, cinchonine and cinchonidine). Quinine was therefore extracted from Cinchona bark by Pelletier and Caventou in 1820 [35] and this allowed standardised efficacy trials to be carried out. The other important aspect of quinine purification is that, while it only comprises some 10% of the weight of cinchona bark, it can be purified as a sulphate by addition of sulphuric acid to an extract. The isolation is facilitated by the quinine sulphate being the least soluble of the cinchona alkaloid sulphates and therefore the easiest to isolate as a crystalline solid. Unlike many pharmaceuticals, the preferred method for producing quinine remains isolation from the bark of the Cinchona tree although a complete synthesis has been carried out using modern synthetic techniques [36]. Previously, Woodward and Doering claimed to have achieved the first total synthesis of quinine in 1944 [37] however Rabe and Kindler had previously reported the synthesis of quinine via the important intermediate quinotoxine (Fig. 1.8). Stork, in a later paper, cast doubt on this route and, therefore, on Woodward's total synthesis since Rabe and Kindler never published the details of how they arrived at quinine from quinotoxine [38]. The controversy was finally laid to rest with the synthesis of quinine

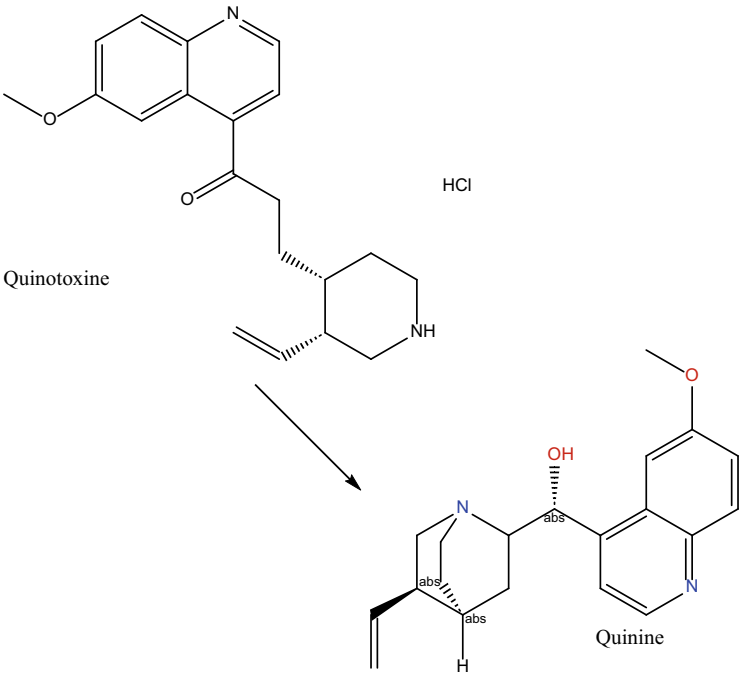


Fig. 1.8 Synthesis of quinine from quinotoxine

from quinotoxine using methods available to Rabe and Kindler [39] but this was not reported until 2008. This is the type of 'angels dancing on the head of a pin' argument that occupies synthetic chemists.

To take a purely empirical and practical stand, it is, in fact, of little consequence since the preferred route to obtaining quinine remains natural product isolation. Probably the most consequential 'failed' synthesis of quinine was that carried out by Perkin in 1856 when he prepared the first synthetic dye, mauveine when attempting to make quinine [40]. This is described in more detail in Sect. 1.6.

While the isolation of natural products with common chemical characteristics (like alkaloids) can be achieved, in some cases, without intermediate separation—if the amounts of the product are small or comprise part of a complex mixture, isolation can be more challenging or may require chromatographic separation. Once extraction has been optimised, other purification techniques become important. The combination of extraction and purification can, as noted previously, be a matter of trial and error and while the chemical (and biological) properties of the potential target compounds precludes the development of a single all-encompassing methodology, there are some common aspects which may guide the natural product chemist in their quest to obtain a pure product.

Initial approaches to the separation of natural products are simpler if the chemical characteristics of the compound(s) of interest are known or can be deduced. For

example, alkaloids are basic and normally contain at least one ionisable nitrogen atom. This means that they can often be ionised and extracted into an aqueous solution. The residual neutral (and sometimes the acidic) molecules can be removed by shaking with an organic solvent into which they will partition. Separation of the two phases is generally carried out in a separating funnel or similar piece of glassware and the solution can then be basified and the un-ionised alkaloids extracted from the basic solution into an organic solvent. It may seem obvious but it is worth remembering that the organic solvents used should be immiscible with water so alcohols are a poor choice. The second point is that some molecules do not separate simply into acid or base aqueous soluble forms so it is generally a good idea to retain all fractions until you are sure you have the material of interest in the expected fraction. It is also notable that the use of strong acids and bases can be detrimental to alkaloids (and other natural products) and should be avoided where possible. The other issue is that if fatty acids are present then strong alkali will cause saponification and lead to the formation of emulsions which may be difficult to resolve into an organic and an aqueous fraction. The steps in isolating a crude alkaloid mixture are as follows

1. Grind up plant or other material (lyophilisation or freeze drying can be a convenient method).
2. Extract the mixture with dilute aqueous acid solution. In most cases it is not necessary to optimise the yield.
3. Filter to remove any residual solid material. Wash the aqueous extract with an immiscible solvent (hexane, ethyl acetate).
4. Basify using a mild base (ammonium hydroxide is generally a good choice). Avoid strong bases such as sodium or potassium hydroxide.
5. Extract into an organic solvent. The choice of solvent depends on the polarity of the compounds and successive extraction with hexane, ethyl acetate and finally chloroform may be suitable. Dry the separate extracts by addition of anhydrous magnesium or sodium sulphate and filtration.
6. Evaporate the solvent from each extract and take up the extracts in the minimum volume of solvent (ethanol can be used as an alternative to the extraction solvent). Store sealed in a freezer prior to analysis.

In some cases, the alkaloid mixture may crystallise from the acidified aqueous solution either spontaneously (after concentration) or by addition of ethanol or isopropyl alcohol and cooling in an ice bath. The case of quinine (mentioned above) is a case in point. The commercial production of known and biologically assayable quinine had to be achieved. Cinchona bark was mixed in water containing slaked lime and it was then extracted directly into a hot organic solvent. The organic layer was extracted with diluted sulfuric acid and upon concentration of the aqueous phase quinine sulfate crystallized.

In the case of target-driven isolation, the development of suitable bioassays to monitor biological activity is necessary. This will be dealt with at greater length in Sect. 1.5.

From a practical standpoint, the isolation of large amounts of single compounds can be challenging both because of the presence of mixtures of closely structurally

related compounds and/or the low level of occurrence of the compound of interest. And in this case a semi-synthetic route may be a better option than pure isolation. In this approach (which will be discussed further in Chapter Two) a core molecule which is produced in relatively large amounts is isolated from a natural source and this is then modified using synthetic chemistry. Examples of this approach are the use of 6-amino penicillanic acid from a fungal culture as a core molecule for functionalised penicillins with improved pharmacological properties. A second example is the isoprenoid compound paclitaxel (PTX), a potent anti-cancer drug which only occurs in small amounts in nature. While PTX is isolated from plants of the Taxol family (yew trees), the yield is small (ranging from 0.001 to 0.06% of the dried bark's weight) and isolation destroys the plant host. For this reason, a precursor (baccatin III) is used as a precursor for the semi-synthetic preparation of PTX [41]. As with quinine, while a formal synthesis of PTX has been carried out, it is far easier (and cheaper) to make use of a partially biological route.

At the other end of the scale (both in terms of complexity and physicochemical characteristics) is the isolation of the toxin, 3-nitropropionic acid. This small (3 carbon) compound is found in several plants and fungi [42]. It is a potent inhibitor of the important metabolic enzyme, succinate dehydrogenase. Physiologically it is a neurotoxin and administration has been used to create an animal model of Huntington's Disease [43]. It is biosynthesised in the fungus, *Penicillium atrovenetum*, from L-aspartic acid [44]. It is simply isolated from the fungus in which, after 2 days, it is present in the medium at levels up to 2 g/litre [45]. The toxin can be routinely isolated from the fungal culture medium by acidification of the culture supernatant, extraction of 3-nitropropionic acid into diethyl ether, evaporation of the ether and then sublimation of the product and recrystallisation from chloroform.

While compounds containing ionisable features (e.g. alkaloids and acids) can sometimes be sufficiently purified by selective extraction (as demonstrated above), chromatographic methods are often required for the isolation of compounds that are either present in small amounts or are present alongside functionally similar compounds. There is considerable complexity in the use of chromatography however the basic principles can be illustrated with a few examples.

1.5 Chromatographic Techniques in the Purification of Natural Products

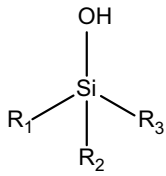
While alchemists developed a range of extraction and concentration techniques and these were applied by herbalists and others interested in treating human disease, it was the development of chromatographic techniques that probably had the most significant impact on the supply of purified natural products for medicinal (and other) uses. The use of analytical and preparative chromatographic methods has underpinned much research on phytochemicals and natural products.

The first example of the use of chromatography in a recognisable form (as noted previously) was reported by the botanist Mikhail Tsvet in 1903. He used a column of calcium carbonate in a glass tube to separate plant pigments. Unlike the elution chromatography we are more used to, Tsvet realised that different pigments travelled different distances through the calcium carbonate and so could be retrieved as separate fractions by sampling the solid adsorbent from each of the colour zones. Some years later Lederer and Kuhn at the University of Heidelberg also used chromatography to separate plant pigments. A range of other advances then followed as chromatography was applied to increasing numbers of types of natural product purifications. These all had the same basic principle—a mobile phase would carry a mixture of compounds through a stationary phase and the different compounds would be separated based upon their interactions with the stationary and mobile phases. A range of techniques were then developed using different stationary and mobile phases to facilitate the separation of molecules based on size, charge, volatility etc. and these led to liquid:liquid chromatography, ion exchange and gas chromatography and high performance liquid chromatography (HPLC). A range of more precise techniques including reverse phase chromatography, affinity chromatography, chiral chromatography and flash chromatography were developed. In some cases, the chromatographic method could only be developed as a result of advances in other areas (e.g. the development of microparticulate silica) and a range of packing advances has led to the development of separation methods for some of the most recalcitrant molecules including DNA, RNA, peptides and polypeptides. Once again, the importance of scale is seen. Analytical chromatography is small scale and is used for the identification of molecules. Larger scale chromatography is used for preparative or semi-preparative purification of compounds.

The simplest (and most commonly used) technique for separating and purifying large amounts of natural products is column chromatography using a silica or microporous resin solid phase and a suitable solvent (or solvent mixture). Silica is a polar stationary phase which contains silanol groups (Fig. 1.9).

The polar groups of natural products bind to the silica through Van Der Waals or hydrogen bonding and are eluted by the solvent in order of increasing polarity. The most polar compounds are eluted last. Isocratic elution requires a constant mobile phase composition (but not necessarily a single component) throughout complete procedure. The elution characteristics of the mobile phase remain constant. This type of chromatography is commonly used when the compounds to be separated have similar polarities. For gradient elution involves the composition of the mobile phase changes. This type of elution is commonly employed to separate compounds of

Fig. 1.9 Silanol group



very different polarities. In either case a range of solvent mixtures may be employed however normally a two solvent system comprising solvents with different polarities is used. Some of the most common solvent systems are chloroform:methanol and hexane:ethyl acetate. Particularly in the case of isocratic based separation, the separation is not, however, purely guesswork or based on theoretical calculations of molecular polarity and size. Prior to setting up a preparative column, the crude sample is run on a silica thin layer chromatography (tlc) plate with varying solvent mixtures to establish the optimal solvent mixture. Further discussion of the principles and importance of tlc is given below. Visualisation of the compounds on the tlc plate either using a general stain (iodine) or a compound-specific stain (e.g. p-dimethylamino cinnamaldehyde) will ensure that the compounds of interest are able to be separated. Samples that are collected from the column are also analysed by tlc and those containing the compound(s) of interest are combined, the solvent removed and then, if necessary re-chromatographed. For some samples (often those containing highly ionisable groups) a small amount of an additional solvent (e.g. tetrahydrofuran, triethylamine) can be added to sharpen the elution profile and prevent 'tailing'. Tailing is the overlap of one component into the next component in a column chromatograph. Preparative HPLC is the method of choice for compounds that are chromatographically similar. HPLC can also be used as an analytical tool to check the purity and homogeneity of samples that have been fractionated.

The use of functionalised stationary phases can be advantageous—particularly when dealing with highly polarisable molecules. For example alumina is able to absorb highly polar compounds such as alkaloids. Acidic alumina was successfully used in the isolation of glucosinolates (a group of natural products containing a sulphonated oxime moiety) however fractionation of the individual glucosinolates was not attempted [46]. In some cases—particularly when attempting to fractionate highly polar natural products such as glucosinolates, polyamine functionalised resins such as diethylaminoethyl Sephadex can be used. While organic solvent mixtures can be used, elution with ammonium salts can give fractions from which much of the solute can be removed by lyophilisation prior to a final desalting stage by passage through a size exclusion gel.

While the above techniques can be used (when coupled to a suitable bioassay system) to determine completely uncharacterised compounds, they are more often used for the isolation and purification of known individual or classes of compounds. For novel compounds, the isolative system is often coupled to a suitable bioassay and full structural characterisation using a range of spectroscopic techniques can be carried out. A more detailed, illustrative description of these techniques is outlined below (Sect. 1.10). A range of other techniques including counter-current chromatography, centrifugal thin layer chromatography, gel filtration chromatography and various ionic-based techniques have also been developed often for specific applications. The use of molecular imprinting technology has increased in popularity [47]. In essence a polymer is prepared and is used to enclose a template molecule. The template is removed and the shape of the imprinting molecule left in the polymer. The polymer can then be used to selectively adsorb molecules with structural characteristics like those of the template molecule. Molecular imprinting can be thought of as

a biomimetic technique with similarities to the production of molecular recognition-specific antibodies. While the use of polymer-based imprinting molecules together with a suitable template and a crosslinking molecule is the usual form that this technology takes, the idea of using proteins has been suggested since proteins are multifunctional polymers whose constituent amino acids contain all the required functionality to absorb a range of molecular types—if the amino acids can be fixed in a suitable relative orientation. Selective absorption of substituted phenyl propanoids was achieved using denatured bovine serum albumin as a protein, substituted cinnamyl diesters as imprinting agents and glutaraldehyde as a cross-linking agent (Hanley A B and Russell W R—unpublished). Using the same principle, the 2 + 2 cycloaddition of cinnamic acids was carried out to give phenyl cyclobutanes which can be used as imprinting templates (Fig. 1.10) to bind free cinnamic acids which can then be cross linked by uv irradiation. In the absence of an imprinted protein (i.e. in solution) no reaction takes place. In this case, the molecularly imprinted protein is acting in a similar way to an antibody-enzyme (abzyme) [48].

One technique which has greatly improved the capability to isolate and purify natural products is Flash Chromatography. This procedure is briefly described below.

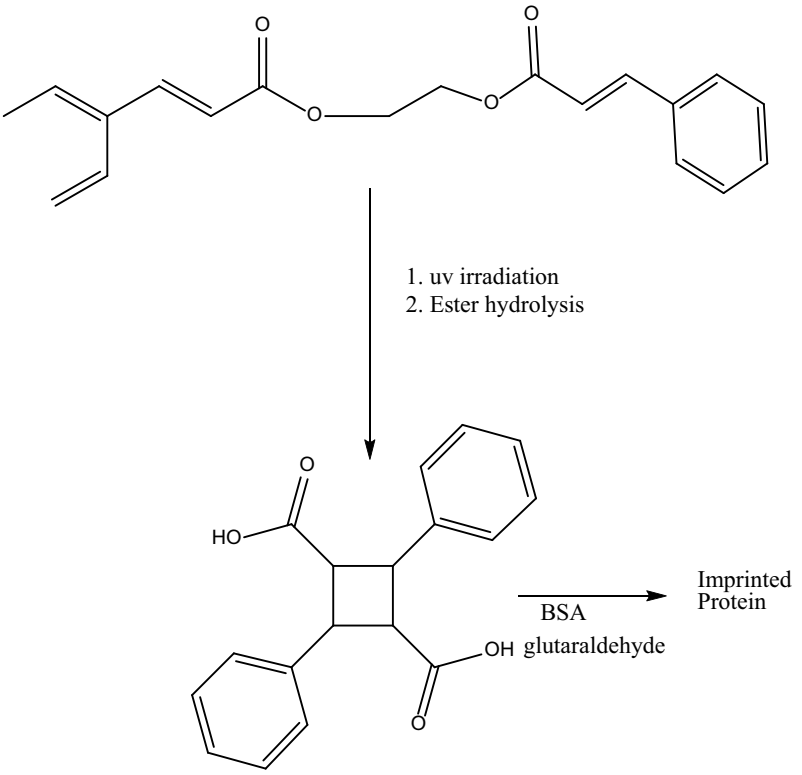


Fig. 1.10 Synthesis of a cyclobutane to use as an imprinting molecule

The first stage in chromatographic separation, as noted above, is determining the best solid phase (silica is commonly used) and the most appropriate solvent system. The simplest way to do this is to run the extract containing the compound(s) to be separated (if possible alongside authentic standards) by tlc and to use a range of solvents mixtures (binary solvents in an isocratic system are, as noted above, most convenient). Retention on tlc is generally measured as the Rf (Retention factor) which is defined as the distance travelled by the compound of interest divided by the distance travelled by the solvent. The solvent front distance is standardised to 1.0 so a compound that travels half way up a plate will have an Rf of 0.5. In general solvent systems are chosen in which the compound of interest has an Rf of about 0.25. Compounds can be visualised by staining with iodine. As noted previously, a range of specific stains are also available which can help in the identification of specific classes of compounds (https://www.epfl.ch/labs/lcso/wp-content/uploads/2018/06/TLC_Stains.pdf).

The essence of flash chromatography is that the silica column is eluted under a small amount of pressure. This can be a complicated procedure however a number of manuals are available (see, for example, https://chemistnotes.com/organic/flash-chromatography-principle/).

An alternative method, which was first published in Aldrichimica Acta in 1985, makes use of a sintered glass funnel and tlc grade silica [49]. The funnel is filled with silica and placed on a Buchner flask and a water pump suction applied. The silica will be compressed and it should be pressed down with a suitable flat object. The column is pre-eluted with the solvent of choice and this fraction is discarded. The sample is loaded onto the surface of the silica in solution, absorbed under vacuum and the solvent is then added in aliquots and the column sucked dry after each addition. The fractions are collected separately and analysed by tlc to identify which fractions contain the target compound. Fractions containing the compound of interest are combined, evaporated, re-chromatographed if necessary and the purified compound thereby isolated. The elution profile can be isocratic or a gradient can be incorporated since the column is eluted to dryness after the addition of each solvent fraction. (https://www.erowid.org/archive/rhodium/chemistry/equipment/flashpad.html). The use of this abbreviated technique does not require the use of either pressure from a gas cylinder or specialised equipment. It is also rapid and simple to carry out.

While isolating known compounds or even unknown members of a class of compounds can be carried out relatively easily, by comparing the chromatographic and/or staining characteristics of the compound with the characterised compound or with an analogue, the success of natural product medicinal chemistry relies heavily on the development of activity guided separation methods. In this case the separation and subsequent characterisation of a compound relies upon its biological interactions with a living system rather than purely its structural similarity to a known compound. While some of these are complex, the same principles apply in most cases.

The first stage is deciding on the biological property that is desired and the development of a suitable assay and its use to screen extracts for active principles. The driver for such assays can be a desired pharmacological property which can be measured or

the use of data analysis to predict the bioactivity of an isolated compound or extract [50]. The particular advantages of this latter approach are that it is not necessary to have the biological material (e.g. receptor or model system) to conduct the assay and it is not even necessary to, in the first instance, isolate and fractionate the compounds of interest. The 'best' compound can be constructed and then potential sources of this can be determined based on known phytochemical principles. The drawback to a purely in silico approach is that it does not directly provide a measure of biological response and context and both false positive and false negative results are possible. These will only be revealed when the compounds are isolated and tested in a suitable biological assay system. Despite this drawback, in silico approaches can save considerable time and may enable some structures to be eliminated. In this way they can facilitate rapid throughput assays.

A simple assay which can be repeated on many fractions in a reproducible manner is the preferred option for a medicinal chemistry approach however biological context is important. Rather than try to examine the generic rules that apply to activity-guided assays, an illustrative example will serve to show some of the principles. For example, in the process of identifying and assaying allergens in vitro, the use of an immunoglobulin E (IgE) screen is well established [51], however a wholly IgE-based test has limitations (discussed below). In the testing of unknown (or suspected) allergens, the individual response to a skin prick allergen test is considered to be the preferred method. There are limitations including the possibility of severe and deleterious systemic reactions in the host (these are rare events but are not unknown) and variation in the interpretation of results. The second possibility (noted above) is to measure the binding of a specific IgE to an allergen. This test relies upon the allergen being tested provoking the same molecular response as the known allergen that the IgE molecule recognises. This forms the basis of a number of analytical tests including the Enzyme Linked Immunosorbent Assay (ELISA) and the RadioImmunosorbent Assay (RIA). The problem with these assays is that a full allergic response requires the linking of IgE receptors on the surface of Mast cells to provoke degranulation and release of a cascade of biological mediators. IgE binding assays will not detect that capacity of an allergen to cause degranulation and it is, therefore, insufficient to reflect fully a biological response. It detects binding of a molecule to an antibody and not the full biological consequences of that binding. For this reason the use of human tissue which is presensitised to the target allergen has been carried out [52]. In this case a whole biological response, including the release of biological factors associated with degranulation, can be assayed. The major drawback to a human tissue-based approach is that a supply of tissue is required and each tissue sample must act as its own control. The assay is also more complex than an IgE-binding assay and cannot easily be automated.

The specific example outlined above illustrates some of the salient points about activity-guided fractionation using the detection and assay of allergens as a specific example. More generally, the use of activity-guided isolation of natural products is conceptually simple and can be summarised in the following steps (extraction followed by fractionation), once a suitable bioactivity assay has been created and validated:

1. Extraction of compounds from the organism of interest. The aim of the extraction is to ensure that all bioactive constituents are present in the extract. At this stage it can be helpful to carry out a series of extractions with solvents of gradually increasing polarity ending with aqueous-based solvents. Each sample can then be evaporated and assayed separately for the activity. If an expected level of activity is not found at this stage this may be due to two possible factors
    i. The active principle may not have been extracted (unlikely) or may have been decomposed by the extraction procedure (more likely)
    ii. The activity resides in more than one fraction or is a result of synergy between different components. If this is the case then samples of different fractions can be combined in order to try to 'rediscover' the activity. It is worth noting that if activity is a result of the interaction of a number of components, this may be a 'go/no go' stage gate since it is much more difficult to get a drug approved if it consists of more than one entity.
2. Once the extraction (and pre-fractionation, if carried out) has been completed, the extract containing the active material can be further fractionated using a range of (generally) chromatographic techniques as described previously. Each fraction should be assayed and the fractionation procedure continued until a single active component has been isolated and can then be characterised (Sect. 1.7). Once again, loss of activity may occur and a similar procedure of combining fractions, as before, should be carried out in order to 'rediscover' the activity. If the activity cannot be detected then it may have been lost due to the nature of the fractionation protocol or decomposition of the active principle. If this appears to be the case then an alternative fractionation procedure may prove more successful.

In addition to traditional fractionation/screening, a number of advanced data-based techniques are available for the natural product chemist. Phenotypic assays (those which produce phenotypic changes to the test system rather than measuring the interaction with a specific molecular target) can used in conjunction with advanced analytical techniques such as liquid chromatography coupled to high resolution mass spectrometry (LC-HRMS) to cluster groups of natural products according to chemical and biological characteristics. The use of metabolomic data can help to determine the molecular level cause of phenotypic variation in cell-based assays. Metabolomics can also yield useful information regarding the profile of natural products in an extract and enable prioritisation of novel bioactive natural products.

Nuclear Magnetic Resonance (NMR) spectroscopy is one of the most important tools for natural product identification. In addition, by using stable isotopes to trace biosynthetic pathways, it can also be a key tool in the manipulation of pathways. It is important to realise, however, that NMR is a relatively insensitive technique compared to HRMS and is therefore of greatest value in structure determination and in the analysis of major metabolites. HRMS is far more sensitive and can even be used to determine the spatial distribution of natural products within an organism [53, 54].

1.6 The Isolation and Purification of Sensorially Active Natural Products—Colours and Flavour/Odour Compounds

While the development of methods to use, extract and then purify biologically active natural products has been the subject of many investigations, both colours and flavour compounds have also attracted considerable attention. Approaches to these have been noted previously however they will now be considered and contextualised in greater detail. The isolation and purification of colour and flavour compounds from natural sources can be carried out with relative ease and assays for such compounds which conform to the principles of 'activity-guided' isolation can be developed. In addition, the purification of the 'active principle' to single component homogeneity may not be necessary.

Flavours are compounds which elicit a biological response by binding to a flavour receptor. Five flavours (and five classes of flavour receptors) are found in humans. These are salt, sour, sweet, umami and bitter. Fatty acids have been associated with a range of physiological responses and the presence of free fatty acid receptors has been postulated. Indeed it has been suggested that fat could be a sixth taste modality [55]. This small number of taste receptors, nonetheless, is in marked contrast to fragrance compounds and fragrance (olfactory) receptors. In the case of olfactory receptors, it is thought that about 400 different types of olfactory receptor exist. These are members of the class of G protein-coupled receptors (GCPRs)—this is in contrast to the much smaller number of taste receptors. Both taste and odour receptors can occur in physiological centres that are distant from their primary site. In the case of olfactory receptors, one member of the class OR51E2 is found in the gut and in other organs. This, of course raises the question of what the function is of these receptors in a range of other organs. As we will see, taste receptors (particularly bitter receptors) seem to be responsible for a range of extra-oral physiological responses. Studies have demonstrated that binding of odorant molecules induces a conformational change in the odour receptors [56, 57] which may affect functionality. In terms of extra-odorant activity, the odour receptor OR51E2 binds propionate and appears to modulate inflammation caused by colitis, in a mouse model [58]. In addition the link between odour receptors in the olfactory bulb and a range of neurological outcomes has been noted [59]. It is likely that other biological consequences related to odour receptors will be discovered and this could form the basis for the manipulation of physiological endpoints through specific interactions with members of the class. It is clear from these simple and isolated examples (and, as we shall see with taste receptors) that biological responses may be far more complex than anticipated and natural products may have a range of activities beyond the simplistic explanation suggested by initial studies.

In contrast to the very complex situation with odour receptors, taste receptors, as noted previously, are relatively simpler both in number and apparent function. They are organised into a limited number of types. Sweet taste is detected by a receptor composed of two different GCPRs (in other words, a heterodimer). These are

classified as taste receptor type 1 member 2 (T1R2) and member 3 (T1R3). Different sweet molecules (sugars, sweet peptides/proteins and small molecule sweeteners) bind to different parts of the receptor dimer. Molecular details of the binding of sweet compounds to the sweet receptor is outside the scope of this chapter however a number of mutation studies have suggested, among other details, the mechanism whereby sweet taste is potentiated by sodium ions and an increase in plasma insulin levels as a result of the taste receptor response to sugar, before sugar can travel to the gut. It has also been suggested that the use of artificial sweeteners may lead to an increase in insulin resistance [60], however causality in the link between artificial sweeteners and obesity or type 2 diabetes has not been established [61]. The umami receptor is a heterodimer of T1R1 and T1R3. Saltiness is mediated by an epithelial sodium channel [62]. Sour taste has been suggested to be a consequence of the presence of ion channels that are sensitive to organic acids.

Bitter taste is more complex and is a consequence of the interaction of multiple taste receptors of the class of family type 2 (Tas 2R or T2R). The T2R family are a group of GCPRs. Over 20 genes have been found in humans and these are responsible for the bitter tastes experienced. While a limited number of compounds have been found in nature that have a sweet, sour, umami or taste profile, a large number of bitter compounds with a range of structural characteristics have been found. It has been estimated that the number of bitter compounds is in the order of thousands or tens of thousands [63]. The use of binding assays to human bitter receptors has been used to identify bitter compounds and these have been collected into a database (http://bitterdb.agri.huji.ac.il/dbbitter.php) Indeed, the automation of bitter taste studies has been approached using machine learning [64]. It has been suggested that the development of bitter taste receptors was an evolutionary advance to prevent (or at least, discourage) the consumption of toxins—many of which have a bitter taste profile. This is not invariably the case, however and it has been suggested that about 60% of known bitter compounds are significantly toxic and a little over 50% of toxic compounds are bitter. This, together with the extra-oral physiological location of bitter receptors, suggests that they may have other significant roles apart from being a detection system for bitter compounds and a discouragement to ingestion [65].

While bitter taste receptors are found in the oral cavity as a crucial component of taste buds, as noted previously, they are also present elsewhere and have different physiological purposes [66]. A definitive theory that may explain the multi-biological role of taste receptors remains elusive however it seems clear that they have a number of important roles to play in the interaction of living systems with externally-generated natural products. For example, it has been found that bacterial quorum sensing molecules can act as agonists for certain T2Rs. This property can activate the human innate immune response (including stimulation of the secretion of cytokines and IL-6 in epithelial cells) and thereby elicit a response to bacterial invasion. One particular quorum sensing molecule (a homoserine lactone termed AHL-12) was found to act as a ligand for a particular bitter receptor, T2R38, which in turn stimulates phagocytosis [67]. Bitter receptors of the T2R family are found in various locations including in the brain. It has been suggested that bitter receptors

may have a role to play in neuroprotection [68]. The mechanism of any such interaction is not clear however if the inflammation hypothesis of Alzheimer's and other neurological diseases is correct [69] then the use of small molecules that are agonists for T2R38 may induce phagocytosis and protect the brain from infection and inflammation. Other biological roles for bitter tasting natural products and their interaction with specific receptors remains an area of active investigation. While biological systems can demonstrate exquisite specificity, it is also the case that multifunctional, organ-specific responses are found and their delineation may help to further elucidate the biological roles of natural products. The study of flavour from a commercial and natural product standpoint will be dealt with later.

The use of natural products as colouring compounds has, as mentioned previously, a long and significant history. The first colours were derived from minerals which, although they exist in the natural world, are not 'natural products' as we understand the term. Nonetheless, many of the techniques used (grinding to a fine powder, mixing, suspending in a suitable carrier (solvent), the use of binding materials and adhesives), are similar to those used in the extraction of natural products from living material. The concept of suspending a colouring agent in a suitable medium is the origin of the term 'tempering' and the resulting mixture was referred to as tempura. Mixtures of milk-derived (casein) protein and natural oils such as linseed were also used in the production of material for tempura painting. The attempts to create a mobile paint that can be applied to a surface is just one aspect of the use of natural products. Natural colours can be composed of single pigments but are more often a mixture. The concept of additive and subtractive mixtures has been discussed previously. Because many colours are mixtures, they can separate into their component parts. Such a process has in the past been used to visually demonstrate the principle of chromatography however, for artists and others who work with paint it can be problematical. Mixtures of different pigments may also be differentially sensitive to light or to oxidation by air. These properties of mixtures of pigments encouraged the development and use of single component, synthetic colours rather than natural extracts. A second challenge is the degree of light fastness associated with different colours. Light fastness is the degree to which a colourant can resist fading when exposed to light. The simplest, qualitative way of measuring light fastness is to apply a colour to a surface, to split the material in two and to expose one part to light while keeping the other part in the dark. A number of protocols have been developed, depending on the nature of the paint and, in addition to visual inspection, samples can be compared spectrophotometrically. The standard technique is referred to as the British Standard Blue Wool Scale (BWS). The scale is calibrated from 0 to 8 with 8 being completely lightfast and 0 denoting extremely poor colour fastness. There are a number of uv-initiated processes that result in colour fading in light. These include photolysis, reduction, oxidation and photoexcitation. It is not just the applied colours that can fade as a result of exposure to light—the colouring materials themselves may also lose vibrancy and colour even before they are applied.

Another aspect of the use of natural products as colouring materials is the dyeing of textiles, wood, pottery and other solid materials. One of the most important properties of dyestuffs is that they must be absorbed by the material being dyed and not wash

out. Various fixing agents (called mordants) have been used. These include tannic acid-based fixing agents, anionic and cationic fixing agents and cross-linking agents. Choosing the best mordant is an important aspect of the use of dyes.

In terms of molecular structure, organic colouring materials generally contain a conjugated system together with a functional group that can enhance or stabilise the conjugation. The relative orientation of conjugation and functional groups can result in the production of a darker (bathochromic shift) or lighter (hypsochromic shift) colour. The effect of different groups can be calculated and predicted according to the Woodward-Fieser Rules and these have been calculated for a number of chemical moieties [70]. Such methods can be used to try to predict the colour spectrum of synthetic compounds with various conjugated groups attached.

The history of the use of colouring materials both in textiles and in painting is long and varied and has been referred to earlier. Certain types of natural colours (e.g. Tyrian Purple, ultramarine) were expensive and highly sought after both because of the scarcity of the natural sources themselves and/or the complicated or involved preparation methods. As a result certain colours were considered to be a status symbol and their use was restricted to certain members of the societal hierarchy. For example, as noted previously, Tyrian Purple is produced from the extraction of a gland from the sea snail. It has been estimated that 12,000 snails are required to give 1 g of the purple dye [71]. Some natural colours (e.g. saffron) are still used and harvesting the plant or animal source remains an important process. Naturally derived colours can command a premium price. In other cases (e.g. indigo), the production of the dye is predominantly from synthetic sources. The advantages of the synthetic route lie both in the consistency of the product and the price. In addition, extraction and processing can leave residual plant material behind which has to be disposed of. For example, in the case of extraction of indigo from Indigo tinctoria, it has been estimated that over 99% of the plant material must be disposed of afterwards [72]. The possibility of using manipulated bacteria as a source of biotechnologically produced indigo has been noted [73].

Significant advances were made in the production of synthetic dyes and colours in the eighteenth century. The first synthetic dye produced was mauve which was an unintended consequence of an attempt by W H Perkin to prepare the remedy for malaria, quinine. Perkin was tasked by his supervisor, Hoffman (the first Director of the Royal College of Chemistry in London, a private venture supported by donors), with attempting to synthesise the anti-malarial drug and thought to achieve this via an oxidation reaction of aniline with dichromate. Some of the details of this reaction have previously been noted (Fig. 1.1) and it relied upon calculation of the mass equivalence of starting materials and products. What he produced instead was the first stable synthetic dye mixture, mauveine. This was the first of the azo dyes and has subsequently been found to be composed of a mixture of compounds. In addition to being relatively simple to prepare on a large scale, mauveine had the advantage that it was light fast and could easily be used to dye cloth. Interestingly Perkin tried the original reaction while Hoffman was on holiday and carried out the reaction in his ‘home laboratory’. Following his success, Perkin left Hoffman’s laboratory and set up a factory at Greenford Green in west London (next to the Grand Union canal).

While Perkin's use of his home laboratory has been suggested to be a result of his dedication and work ethic, avoiding working at the Royal College of Chemistry had the advantage that the patent was assigned to Perkin alone [74] and therefore he was the sole beneficiary. This enabled him to make a considerable fortune and to later sell his factory and retire at the age of 36 years. The question of the precise method used by Perkin and the identity of the dyestuff itself has been the subject of some speculation [40]. Much of the subsequent progress in synthetic colours was a result of efforts in Germany where the chemical tradition was more rapidly established. Interestingly W H Perkin founded the Yorkshire Dyeware and Chemical Company in 1900. This company, now called Yorkshire Farben GmbH, is headquartered in Krefeld (a city in North Rhine-Westphalia).

Within a relatively short period, synthetic colours had moved beyond mimicking nature and completely new, colour fast and stable colours were being developed. Much of the impetus for this came from the automobile industry. General Motors, in particular, pioneered the use of new colour combinations to enable them to compete with the all black Model T Ford (black was chosen because it was the fastest drying colour for cars coming off the assembly line) and GM set up an "Art and Colour Section" in 1927.

While the development of synthetic colours greatly increased the numbers, stability and use of colours—both for dyeing/colouring and art painting—flavour chemistry still has a significant natural product component. Previously the biological importance of flavour has been noted and the significance of bitter receptors, in particular, commented upon. The practical importance of flavour compounds is an important aspect of the commercialisation and use of flavourings. While synthetic flavours have been developed (analogously to the development of synthetic colours) the complexity of natural flavour compounds (and mixtures of them) has encouraged the continued use of natural (and nature identical) flavours. Flavours can also evolve depending on what other components are present and the individual response. Two areas—the development of chewing gum and the perception of bitter taste—will be used to provide examples of the use of natural products in the consequences and development of flavour.

The first commercial chewing gum was introduced by Thomas Adams in 1848. It used the natural product chicle as its main ingredient. Although chicle-derived gum was used in America by ancient civilisations, there is a suggestion that Northern Europeans chewed birch bark. Dry distillation of birch bark produces birch bark tar which can then be chewed. Birch bark tar has a number of uses including as an antiseptic and a mollusc repellent. Apart from medicinal uses, chewing gum has been of particular interest to people to freshen breath and to clean teeth. This led to the development of flavoured gums. One of the key companies in this early development was the Wm Wrigley Company of Chicago who introduced Spearmint and Juicy Fruit flavoured gums in 1893. Wrigley was not the only gum manufacturer but he recognised the need to publicise his brand such that it became synonymous with chewing gum. The idea of using gum as a breath freshener—and even to replace tobacco—began in the 1850s. The patent for a tobacco substitute gum was granted in 1870. However it was not until 1880 that the addition of sugar and corn starch to

chicle allowed gum that retained its flavour to be produced. Wrigley's Doublemint was introduced in 1914. While the nature of the flavouring in Doublemint (and other gums) is a trade secret, peppermint (a cross between water mint and spearmint) is the predominant mint variety used. A number of plants and seeds have traditionally been used to freshen breath and mint was a significant contributor to these however, it was not until 1790 that peppermint lozenges (Altoids) were first commercially produced to freshen breath. Both breath freshening mints and chewing gum are designed to remove food debris (by stimulation of saliva production) and mask the odours of bad breath. Apart from specific, food-borne compounds, the main compounds that are known to cause bad breath are the bacterial compounds methane thiol and hydrogen sulphide. It has been suggested that both compounds may undergo a Michael type reaction which will remove their odiferous properties. This property has been used as a rationale for using certain lactones (alpha angelica lactone and related compounds) (Fig. 1.11). An alternative approach was to use a combination of allyl isothiocyanate (the active principle of mustard oil) and a zinc salt. This was shown to be effective in the removal of oral malodour compounds for an extended period of time [75].

The flavour of mint was a major impetus for the cultivation, growth and isolation of active principles however mint (*Menthol*) also contains a number of other bioactive principles [76]. In addition, mint oil contains potent anti-microbial activity. The major mint flavouring compound is menthol and the oxidised derivative, menthone (Fig. 1.12).

Synthetic routes to menthol have been developed by BASF, Symrise and Takasago [77]. Menthol is a terpene alcohol (2-isopropyl-5-methylcyclohexanol) and binds to a neuronal receptor, TRPM8. This then allows calcium ions to enter the neurone to give the impression of coldness. The active (menthol and related compounds) in mint are extracted using a range of possible extraction systems noted previously. Steam

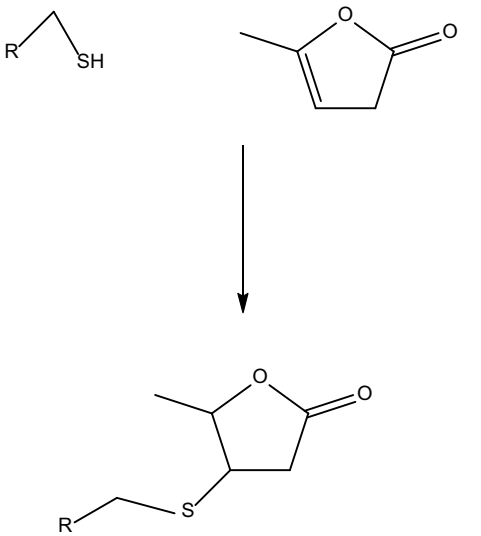


Fig. 1.11 Removal of malodorous thiols by reaction with alpha angelica lactone

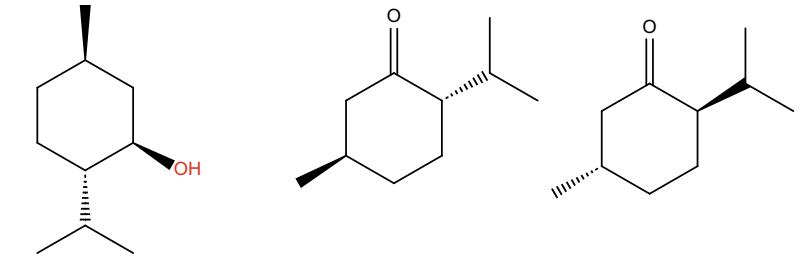


Fig. 1.12 Structures of menthol and menthones

distillation, fractional distillation and chromatography can also be used to separate the various components of mint oil [78]. It is interesting to note that TRP receptors in the oral cavity are also involved in pain signals (and perhaps analogously to bitter receptors) in the avoidance of the ingestion of harmful compounds [79]. It has been noted that some menthol compounds impart bitterness and this is a challenge in the formulation of chewing gum. In addition, the presence of S compounds such as dimethyl sulphide can impact upon the flavour properties of mint oil [80].

In the case of the bitter components of Brassica vegetables, the properties of the compound are a direct consequence of binding to a specific bitter receptor. In addition, as we have seen, genetic polymorphisms mean that individual taste perception may vary. Brassicas contain a number of amino-acid derived compounds called glucosinolates. When plant material is crushed, the glucosinolates are brought into contact with a co-occurring thioglucosidase enzyme myrosinase) and, following cleavage of the thioglucose moiety, the aglucone undergoes a rearrangement reaction to give a range of (mainly) volatile compounds including isothiocyanates and nitriles (Fig. 1.13).

These compounds are responsible for the typical bite and flavour of raw brassicas and mustards. Indeed, the release of these pungent molecules has been called the ‘mustard oil bomb’. While some insects are able to disarm the bomb by secreting a sulphatase which converts a glucosinolate to a non-reactive desuphoglucosinolate [81], others consume glucosinolates and, with an insect produced thioglucosidase, make themselves very unpalatable to predators [82]. Another glucosinolate (progoitrin—Fig. 1.14) is responsible for much of the bitter taste associated with brassica vegetables. The initially formed isothiocyanate spontaneously cyclises to give L-5-vinyloxazolidine-2-thione (Goitrin) (Fig. 1.14).

This compound contains an NH–C=S moiety which is also found in and is responsible for the taste of the bitter test compounds phenylthiocarbamide (PTC) and 6-n-propyl thiouracil (PROP). It has been shown that the ability to taste these types of compounds has a genetic determinant—in particular a polymorphism in the T2R38 bitter receptor gene [83]. While other compounds may also contribute to the bitter taste, goitrin appears—based on taste testing and chemical similarity arguments, to be a major contributor [84]. The T2R38 genotype has been shown to have a significant influence on bitter taste sensitivity to brassicas [85]. The genetic background of

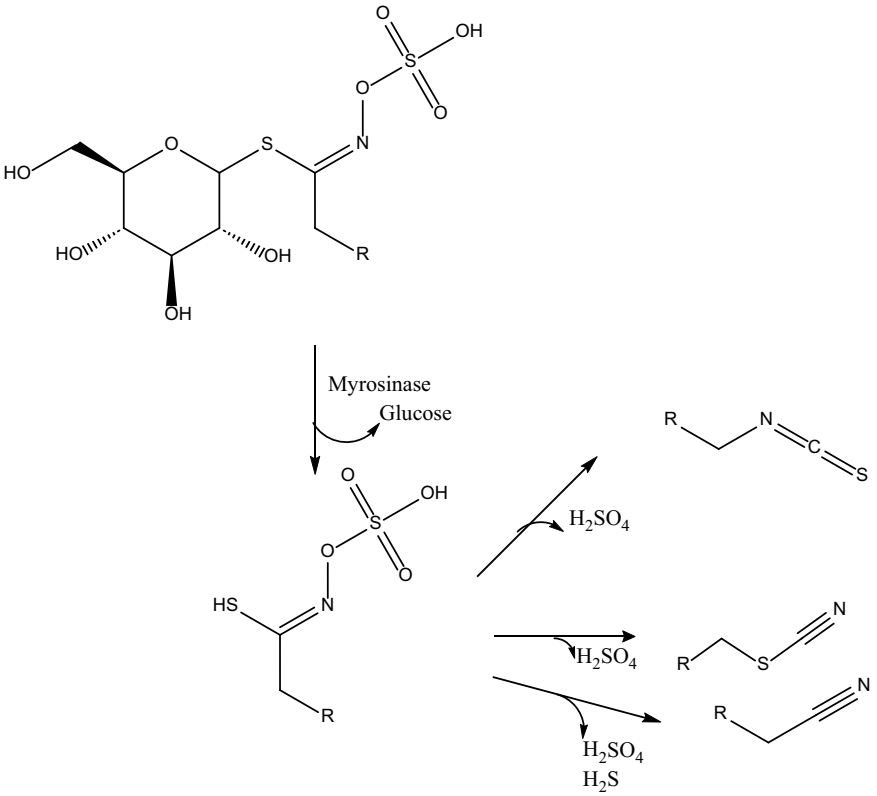


Fig. 1.13 Enzymic Hydrolysis of glucosinolates—Mechanistic Detail

taste perception has been reviewed and a number of conclusions drawn [86]. There was found to be a significant correlation between bitter and sweet taste perception and polymorphisms in the T2R38 receptor. Another gene (coding for a fatty acid transporter) called CD36 appears to exist in a number of polymorphic forms that affect fat taste perception [87]. It seems that, as is the case for some bitter receptors, the gene for fatty taste perception (CD36) has at least one other function—fatty acid transport. These interactions might suggest that bitter natural products are more polyfunctional than realised from a simplistic interpretation of their interactions with the taste receptors that comprise their primary targets. In which case, other natural products may also have numerous targets and physiological activities.

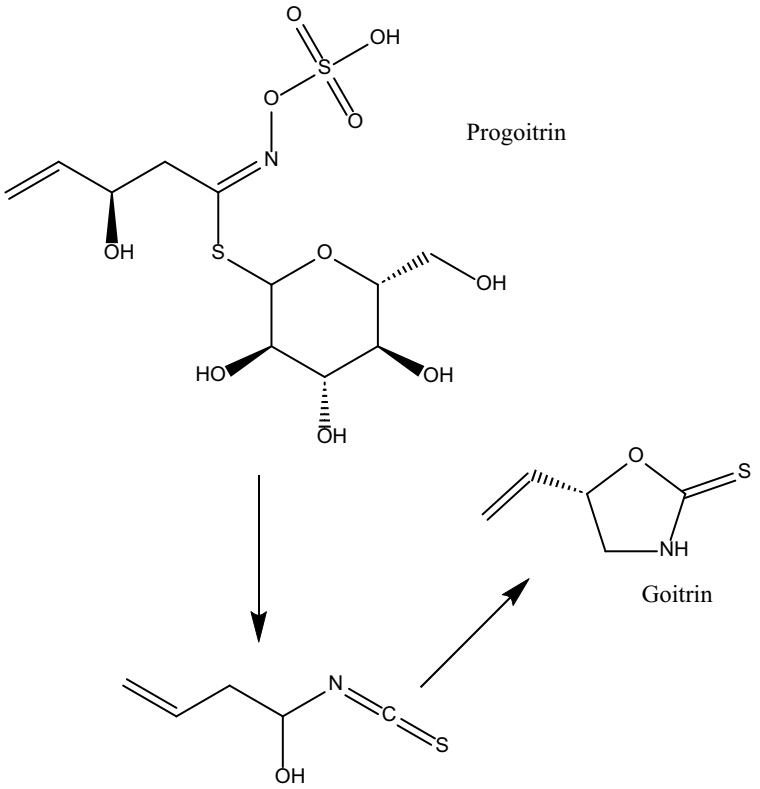


Fig. 1.14 Formation of goitrin from progoitrin

1.7 Structure Determination of Natural Products

Structure determination has been one of the most significant areas of progress for natural product chemistry. Natural product chemistry is, above everything, the science of atoms and molecules and the interaction of these with large systems. At one end of the scale, the structure of molecules is inferred from their properties and reactions and, at the other end of the scale the hugely complex living systems that molecules interact with are probed using highly specific tools that provide evidence for how systems work and interact with molecules. In both cases, while absolute knowledge of the structural properties may not be available, the knowledge we have allows us to construct hypotheses and, perhaps more importantly given the importance of empiricism to chemistry, to test them. In most cases, absolute proof is not attainable—we have an increasing weight of evidence that supports the validity of our hypotheses however, in a Popperian world, it only takes one piece of contradictory evidence to bring a whole edifice crashing down. Inference of chemical structures is one of the

most important and significant parts of the edifice that has been built around natural products.

While a range of techniques exist including, as we shall see, spectroscopic and spectrometric techniques, X-ray crystallography provides probably the clearest and most definitive evidence for a particular structural assignment. While a number of small molecules have been subjected to crystallographic analysis, one of the most interesting areas, from the standpoint of natural product chemistry, is the determination of the absolute configuration of chiral natural products. The separation of chiral components is possible using a range of chromatographic methods and, while nuclear magnetic resonance (NMR) spectroscopy is a very useful tool for structure determination, its use in an achiral solvent will not generally allow for the determination of the absolute conformation of enantionmers. In this case crystallography can be the method of choice. While a number of molecules can be crystallised and analysed, a significant number do not form crystals that are suitable for analysis and, in that case, co-crystallisation is a preferred technique. The technique of co-crystallisation involves the entrapment of a small natural product in an inclusion complex. Such an approach was described for the inclusion compound derived from tetra aryladamantanes (TAAs) [88]. This is interestingly analogous to the use of chaperone molecules to help in the determination of protein structures. The Cambridge Structural Database is a database of validated and curated small molecules and contains both current and historical data. Another technique that can make use of very small amounts of material is electron cryo microscopy (cryo EM) and progress in this area has been reviewed [89]. It has been suggested that cryo EM may supplant x-ray crystallography as the preferred structural tool for determining the absolute conformation of small molecules. In addition cryo-EM has been used for the structural determination of biological macromolecules [90].

The principal spectroscopic method used in natural product chemistry, as mentioned previously, is NMR spectroscopy. In addition, however, ultraviolet/visible (UV–VIS) infrared (IR) spectroscopy and mass spectrometry have been used [91]. The use of machine learning and library structure matching has grown substantially as computing power and the number of techniques available has increased [92]. In the case of NMR, techniques have also been developed that allow the chemist to investigate the natural product both in solution and bound to a target ligand. The natural products magnetic resonance database is a freely accessible collection of NMR data on natural products [93].

1.8 Synthesis and Biomimetic Synthesis

The development of isolation methods for natural products has, as has been demonstrated with specific examples, fuelled significant advances in fractionation techniques, however the relative scarcity of some compounds and the investigation of unnatural analogues has been responsible for the development of significant synthetic methodologies. While chemists have devised ingenious and elegant strategies for the

synthesis of natural products (and their analogues), the development of approaches that mimic how biological systems synthesise compounds has been an attractive goal, not only because it can demonstrate the chemical plausibility of a natural reaction pathway but also because it can illustrate the optimal way of producing natural products which has been subjected to incredibly long periods of evolution to find the most favourable routes. From a philosophical standpoint, the isolation of natural products is attractive since the chemist can be certain that the product isolated is precisely identical in all respects to that which is found in nature. (This assumes that care is taken to prevent the degradation of the natural product). This is particularly the case with chiral molecules. If a synthetic route produces a mixture of natural and unnatural enantiomers, not only must the natural product enantiomer be separated, it is also an inherently inefficient process since up to half the product is of the 'wrong' type. Two approaches to this problem are to start with a chiral molecule which can then be modified while retaining the chirality at the key centre of the molecule. This often involves a combination of biological and chemical synthesis. The second way is to introduce chirality at a late stage in the synthetic procedure either by using a specific chemical agent (often a catalyst) or by using an enzyme to carry out the reaction. In the case of the latter approach, the development of synthetic biology approaches where a whole biosynthetic pathway can be imported into a suitable host which acts as a multi-stage reactor, has become more important. While isolation of a natural product, as has been noted, has the advantage that the isolated product is in the precise form it is found in nature, it is possible, on occasion, to subvert the biosynthetic machinery by introducing a novel starting material which then enters the biosynthetic cycle and the resulting product will possess the molecular characteristics of the analogous natural product (e.g. chirality) but will be 'unnatural' since it was created from a different starting material from that found in nature. An example of this will be presented.

One aspect of natural products that attracted considerable attention and was mentioned previously, was the concept of *vitalism.* The essence of vitalism is that living systems (and by implication, anything derived from them) have some property which distinguishes them from non-living matter. In the case of chemistry, the implication of vitalism is that natural products, since they derive from living systems, cannot be created ex vivo and organic compounds can only come from living systems. The first synthesis of a natural product (urea) from inorganic starting materials by Wohler was noted previously and this seemed to discredit the concept of vitalism—certainly as applied to chemistry. Despite the proof presented by Wohler that organic material can arise from inorganic precursors without the intervention of any living system, the concept was not immediately adopted wholesale and was still a matter of some debate many years after Wohler's key experiment. In 1898 Professor Francis Japp of the Department of Chemistry, the University of Aberdeen, delivered a lecture to the British Association for the Advancement of Science titled 'Stereochemistry and Vitalism' [94]. Japp was influenced by Pasteur who delivered a series of lectures on the molecular asymmetry of natural products. As Japp points out, living systems can make a single enantiomer whereas chemists in the laboratory generally make two forms of a given compound. It was a lack of understanding of how this situation

might come about that influenced Japp and his way of thinking. The idea that science should not simply observe but should also understand why a particular phenomenon occurs is a perpetual challenge and one to which chemistry—with its emphasis on observation and empiricism—is not always ideally suited to meet. To take an example from Newtonian physics, it might be argued that Newton's mathematical treatment of gravity is a way of describing what we observe but does not impart any knowledge about how gravity actually functions. While vitalism was considered by most scientists in the twentieth century to be a false concept, nonetheless as recently as 1966, Francis Crick could present a lecture in which he posed the question 'Is Vitalism Dead?' Indeed Crick was very aware that vitalism remained attractive to many and might be used to explain the mystery of consciousness [95]. Crick, along with others such as Jacques Monod, was concerned that a 'fourth wave' of vitalism may coalesce around biological phenomena that were not easily explained in physico-chemical terms. Some of these ideas will be mentioned in other chapters. The ideas expressed by Japp are of less concern now since we can understand how chiral biological molecules such as enzymes may in turn be integral to the formation of stereoisomers and how the formation of different enantiomers may be energetically favoured.

Chiral (enantioselective) synthesis is a challenge for synthetic chemist including those who are working on natural products. Methods for the synthesis of a range of chiral centres are reviewed regularly. One of the most common routes is through asymmetric catalysis. One complex example is the synthesis of the alkaloid Incargranine A. The challenges of the synthesis include the presence of 6 stereocentres and the very low abundance of the compound in its plant source. One of the issues that had to be resolved was whether the natural compound was an enantiomer or present in the racemic form. An iridium-catalysed reductive cyclisation was a key step in the synthesis and this enabled the first complete synthesis of (+)-Incargranine A [96].

The use of a key chiral starting material which is modified to create the final product is termed partial or semi-synthesis. One of the most useful techniques is to carry out retrosynthesis which can disclose a core scaffold that can then be functionalised [97]. The identification of a key intermediate that can be used for the synthesis of a range of analogues [98]. Despite the advances both in synthetic chemistry and in natural product isolation, finding compounds that have desired properties or activities remains problematical. Because of this, there has been a substantial increase in the use of bioinformatics to probe structures in silico—particularly in the area of medicinal chemistry. Although some progress has been made in the prediction of biological properties, ultimately potential target molecules have to be synthesised or isolated. Natural products are a useful starting point for analogue synthesis and some of these approaches have been reviewed [99]. One of the first areas of interest in the preparation of semi-synthetic, natural product-derived molecules was the preparation of a range of chemically modified penicillin analogues from the parent molecule, 6-amino penicillanic acid (6-APA). The production of 6-APA is carried out by the cleavage of the side chain of the naturally occurring Penicillin G or V with an acylase (Fig. 1.15). In a similar way, 7-aminocephalosporanic acid (7-ACA) is the starting material for a range of semi-synthetic cephalosporins [100].

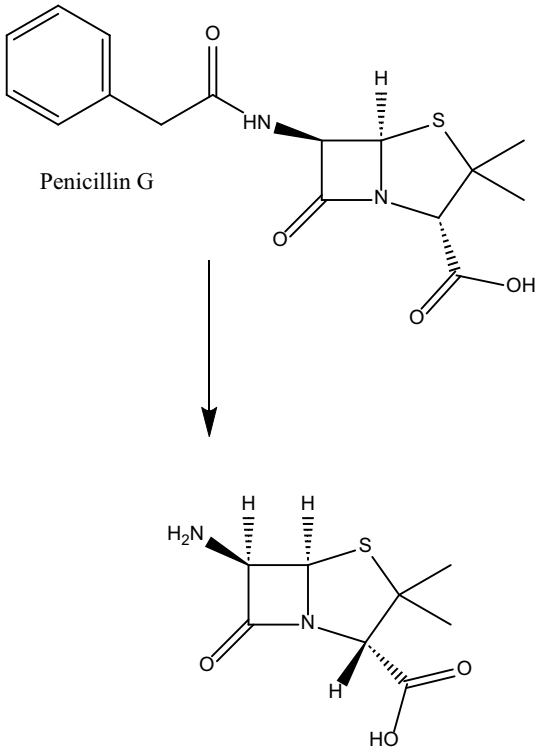


Fig. 1.15 Formation of 6-amino penicillanic acid from Penicillin G

While the use of both chemical catalysis and partial synthesis has had notable successes, another approach to the enantioselective synthesis of natural products entails the use of enzymes to carry out key steps in the synthetic procedure. Directed evolution takes this approach one step further and will be discussed in more detail below.

The idea of biomimetic synthesis is both simple and appealing. In short, by knowing about how a molecule is constructed in living systems, chemists use that sequence as a guide (starting from very basic units) thereby also synthesising, *en route*, a series of molecules that are intermediates in vivo. In addition to revealing a plausible (though not always the best) pathway, biomimetic synthesis allows the chemist to probe a pathway by introducing a labelled atom at a specific point in the molecule (radio or stable isotopes) and to determine if they are incorporated into the final product in the predicted position, thereby confirming the involvement of a particular pathway intermediate and providing further evidence for the validity of the pathway. This aspect of biomimetic synthesis will be discussed further in the section on biosynthesis. In addition to providing insights as to the likelihood of a particular pathway being used in nature, biomimetic synthetic routes can also encourage the

development of new reactions and reagents. One of the earliest biomimetic syntheses was the one pot synthesis of tropinone (a precursor of atropine) by Robinson in 1917 (Fig. 1.16) [101] and this landmark synthesis has been extensively quoted and reviewed [102]. More generally, reviews of biomimetic synthesis are carried out frequently [103] and these emphasise the importance and relevance of this approach.

The ultimate biomimetic synthesis is the use of enzymes and as techniques for the isolation (and stabilisation) of enzymes have grown increasingly sophisticated, the numbers of reactions caried out using biological catalysts has increased. This has even extended to the use of biomimetic systems to stabilise enzymes in nano-structures [104]. Of course enzymes can be used to catalyse reactions in precisely the same manner as they are used in nature however there has also been an increased awareness of the use of enzymes to carry out reactions for which they were not originally intended in 'unnatural' environments [105]. This aspect of natural product chemistry will be dealt with in greater detail in Chap. 7.

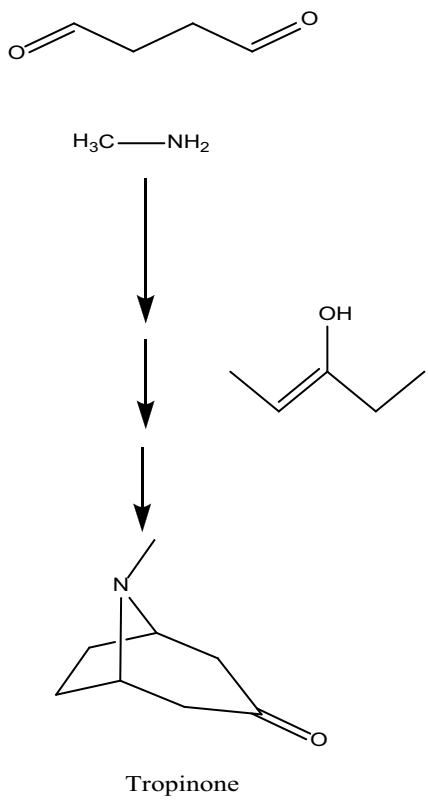


Fig. 1.16 The Robinson synthesis of tropinone

1.9 Biosynthesis—Techniques and Applications

The biosynthesis of natural products has already been touched upon when discussing the derivation of natural products from common precursors and in the section on biomimetic synthesis. Biosynthesis is both an intellectual enterprise (creating knowledge, proving hypotheses) and an intensely practical approach to understanding and then optimising the production and use of natural products. It is, therefore, both a philosophical and an empirical exercise and bridges the divide in science between knowing and doing [106].

It could reasonably be argued that the derivation of any product by a living system is governed by biosynthesis and, in turn, by the enzymes that comprise the biosynthetic pathway. Among the earliest studies were those that led to the elucidation of the Tricarboxlic Acid (TCA) or Kreb's cycle and its constituent enzymes and intermediates. Even before that, in 1833, the first enzyme (diastase—a mixture of amylases) was discovered and its properties elucidated. Much of the early work on enzymes was a result of studies of fermentation by Pasteur and others. In 1926, Sumner' for the first time was able to crystallise an enzyme (urease) and confirmed that it was a protein [107, 108]. In practice, while primary metabolism provided considerable evidence of the importance of pathways for the interconversion of biological molecules catalysed by enzymes, the division of natural products into primary and secondary metabolites was a crucial step in the understanding of biosynthesis. While primary metabolites are those molecules which are considered essential to the survival and growth of an organism, secondary metabolites are substances which are non-essential but provide the living entity with some sort of advantage. These could include toxins and predation-protective agents, quorum sensing and signalling molecules and those which help to optimise growth under different environmental conditions [109]. As is a common habit among scientists, chemists have attempted to classify natural products into different categories based upon structural similarities and biosynthesis. These include alkaloids (identified by a structural similarity) and polyketides (compounds with a common biogenesis—the polyketide pathway) among others. In order to optimise the usefulness of natural products, it is first of all necessary to understand how biosynthetic pathways are assembled. Achieving this makes possible advances in pathway manipulation culminating in synthetic biology [110].

A detailed analysis of the plethora of biosynthetic pathways employed by living organisms is out with the scope of this volume. The example of polyketides will be used to illustrate general and specific principles in Chap. 4. Two of the consequences of the manipulation of natural products by biotechnological methods will, however, be illustrated. The first of these is the production of natural products by taking the genes that code for the enzymes that constitute a biosynthetic pathway and transplanting them into an alternative host. In this case, the reason for carrying out the process is to optimise the production (and ease of isolation) of a given natural product. The second (related) approach is the manipulation of the pathways in order to produce genetically manipulated products that are different from those that occur

naturally. This latter approach will be discussed in more detail in Chap. 5 while the former is the subject of Chap. 6. First of all, as noted previously, it is necessary to understand the biosynthetic pathway and how it operates. This approach began as a largely natural product-led initiative in which the molecular transformations that are carried out in biological systems are studied. By understanding the relationships between different molecules, it was possible to construct hypotheses about the reaction pathways that could effect their transformation. By analogy with known pathways and the plausibility of chemical transformations, it is possible to determine both the nature of the transformation and the biological entity (enzyme) that might carry it out.

The determination of the structure of natural products in the early part of the twentieth century involved purification then controlled degradation to give simpler products whose structures were already known. A significant number of terpenes and steroids were characterised by such methods and some of the most important of these were reviewed [111].

There was a limit to the extrapolations that could be carried out using degradative methods and a number of structures were mis-assigned and later corrected. As the knowledge base of structures grew, there was increased speculation about the biosynthetic origins of groups of natural products and this, in turn, helped to provide plausible structures for related compounds. Equally importantly, such information-guided investigations began to draw attention to compounds whose assumed structure did not fit the biosynthetic rationale.

Prior to the widespread use of NMR spectroscopy (which eventually would be used to solve structures of relatively complex natural products), structures were still assigned on the basis of controlled degradation and the differences in infrared (IR) and ultraviolet/visible (uv/vis) spectroscopy of the starting material and the products. This helped in the identification of specific structural features such as carbonyl groups and their reduction to an alcohol. The other advance was the development of mass spectrometry and an understanding of the rules that governed the fragmentation of molecules. One of the limitations to the use of mass spectrometry was that compounds had to be relatively volatile to be able to be detected. The use of trimethylsilyl (TMS) groups was a major advance. The groups were able to functionalise OH, COOH, NH and SH groups but these derivatives were chemically inert and so did not participate in or enable other reactions to take place. In addition, the functionalisation by TMS increased the volatility of natural products and therefore individual compounds could be separated by gas chromatography prior to mass spectrometric analysis (GC–MS). While the preparation of TMS derivatives can be technically challenging (the reagents are very sensitive to moisture and care must be taken in the derivatisation step), the increased use of automation has enabled the increased application of GC–MS to metabolite profiling in a rapid fashion [112]. In addition, the accumulation of data on mass spectral characteristics has enabled the preparation of libraries of spectra that can be used to rapidly identify compounds. While derivatisation means that GC–MS provides a snapshot of the compounds present at a particular time (and place) the development of what are referred to as ‘soft’ ionisation techniques has enabled less volatile molecules to be analysed directly without prior derivatisation

and has led to the development of HPLC separation coupled to mass spectrometry. The introduction of isotopic labelling techniques coupled to the discovery of gene clusters has enabled the more detailed investigation of biosynthetic pathways [113].

The development of stable isotopic labelling has heralded significant advances in the study of biosynthetic pathways. Initially interest was focussed on the use of C-13 NMR to elucidate the structural characteristics of unknown natural products or to provide information about similarities (and differences) between classes of natural products [114]. In addition, the growth in 2-D correlative spectroscopic techniques greatly increased the strength of the resolving power of NMR. One of the drawbacks of C-13 NMR spectroscopy is the relatively low abundance of the C-13 nucleus (1.1%) however this is an advantage when using C-13 NMR to elucidate biosynthetic pathways. The first way is to directly observe precursorial relationships between molecules by administering compounds with specific labelled atoms and observing if they are incorporated into the final product. The second methos makes use of the fact that certain isotopes (particularly deuterium and O-18) affect the chemical shift of C-13 in adjacent parts of the molecule. Such studies can, by interpretation, provide information about the sequence of bond breaking and bond formation in the biosynthesis of a natural product. Indeed NMR can also be used to investigate the binding of substrates to enzyme active sites. There are numerous papers in the chemistry literature that describe specific examples of the use of NMR and other techniques. The analysis of the progression of a biosynthetic pathway from the standpoint of the natural product, can be seen in the study of the biosynthesis of sialic acid. In this example, one particular group (N-acetylmethyl) whose atoms have a different chemical in the biochemical intermediates, can be followed by Real Time (RT)-NMR. A number of specialised techniques can be used to enable the components of the pathway to be followed [115]. The reaction can be followed both using an isolated enzyme and a rat liver cytosolic extract. The use of C-13 labelling can facilitate the study of complex metabolic events. For example, the use of labelled glucose to study changes in the primary metabolism of E. coli in the presence and absence of sub-lethal amounts of the succinate dehydrogenase inhibitor, 3-nitropropanoic acid, was carried out using C-13 NMR spectroscopy (Hanley, unpublished). Sugar-nucleotide multienzyme cascades were studied using C-13 labelled sugars and C-13 NMR spectroscopy [116]. In both cases chromatographic separation and characterisation of individual intermediates was not necessary and the dynamic processes could be monitored. If such techniques had been available, the study of metabolic pathways such as the TCA cycle could have been completed much more quickly.

1.10 Biological Engineering of Biosynthetic Pathways

The study of biosynthetic pathways by following the chemicals (natural products and their precursors) has been very successful—even to the extent of limited manipulation of pathways as can be seen in the case of penicillins and cephalosporins), however the development of techniques whereby the biological machinery of cells can be

manipulated has greatly increased both understanding of biosynthesis of natural products and ways to manipulate pathways. From the early days of the cloning of genes for individual enzymes in a pathway, synthetic biology allows pathway engineering in a multifunctional way to be carried out.

The route to the synthesis of natural products begins with an appreciation that biosynthesis is generally a multi-stage process and that the individual stages are reactions that are catalysed by enzymes. Once this was realised, the typical chemistry reductionist approach was to attempt to isolate and understand the individual enzymes and reactions in the pathway. The value of understanding complete pathways was appreciated and the discovery of gene clusters was an important step. The transplantation of complete biosynthetic sequences from their original hosts (often plants) into microbes was a key step. Not only did this allow for the much more efficient production of the target natural product, it also enabled biological chemists to manipulate the pathways more easily to make novel, valuable products. Initially synthetic biology was limited to relatively short pathways, the development of new approaches and the refinement of existing methods has facilitated the development of multiple segment pathways to increasingly complex products derived from plants [117] and microbes and combinations of plant and microbial systems [118].

The basic principles of synthetic biology are described by a simple paradigm—Design, Build, Test (DBT). It is underpinned, on a molecular level, by a better understanding of the molecular events that underlie individual biological reactions (what enzymes are involved? is there a need for cofactors? what is the mechanism of the reaction and how can it be manipulated?) and how biological systems as a whole work. The concept of systems biology is, therefore, crucial to synthetic biology. The divergence between systems and synthetic biology is that while the latter is an information and knowledge gathering approach, the latter seeks to create new systems. This can be thought of as being analogous to the dichotomy of chemistry as combining knowing and doing [119].

References

1. Chung S, Cha S, Lee SY, Park JH, Lee S (2017) The five elements of the cell. Integr Med Res 6(4):452–456. https://doi.org/10.1016/j.imr.2017.10.002. Epub 2017 Oct 18. PMID: 29296573; PMCID: PMC5741387
2. Modenesi P (2009) Skull lichens: a curious chapter in the history of phytotherapy. Fitoterapia 80(3):145–148. https://doi.org/10.1016/j.fitote.2008.12.005. Epub 2009 Jan 13. PMID: 19171183
3. Shrestha G, St. Clair LL (2013) Lichens: a promising source of antibiotic and anticancer drugs. Phytochem Rev 12:229–244. https://doi.org/10.1007/s11101-013-9283-7
4. Dela Cruz TEE, Timbreza LP, Sangvichien E, Notarte KIR, Santiago KAA (2023) Comparative study on the antimicrobial activities and metabolic profiles of five Usnea species from the Philippines. J Fungi (Basel) 9(11):1117. https://doi.org/10.3390/jof9111117.PMID:37998922;PMCID:PMC10672636

5. Macedo DCS, Almeida FJF, Wanderley MSO et al (2021) Usnic acid: from an ancient lichen derivative to promising biological and nanotechnology applications. Phytochem Rev 20:609–630. https://doi.org/10.1007/s11101-020-09717-1
6. Lambert N, Chen YN, Cheng YC et al (2013) Quantum biology. Nat Phys 9:10–18. https://doi.org/10.1038/nphys2474
7. MohammadiPeyhani H, Hafner J, Sveshnikova A et al (2022) Expanding biochemical knowledge and illuminating metabolic dark matter with ATLASx. Nat Commun 13:1560. https://doi.org/10.1038/s41467-022-29238-z
8. Caley ER (1927) The Stockholm papyrus: an English translation with brief notes. J Chem Educ IV(8):979–1002
9. McCann D, Barrett A, Cooper A, Crumpler D, Dalen L, Grimshaw K, Kitchin E, Lok K, Porteous L, Prince E, Sonuga-Barke E, Warner JO, Stevenson J (2007) Food additives and hyperactive behaviour in 3-year-old and 8/9-year-old children in the community: a randomised, double-blinded, placebo-controlled trial. Lancet 370(9598):1560–1567. https://doi.org/10.1016/S0140-6736(07)61306-3.Erratum.In:Lancet.2007Nov3;370(9598):1542. PMID: 17825405
10. Sarkar S, Khatun D, Dutta B et al (2022) Journey of natural pigments from ancient antiquity to present: insights on sustainable development. Indian J Hist Sci 57:330–334. https://doi.org/10.1007/s43539-022-00059-9
11. Roper SD, Chaudhari N (2017) Taste buds: cells, signals and synapses. Nat Rev Neurosci 18(8):485–497. https://doi.org/10.1038/nrn.2017.68. Epub 2017 Jun 29. PMID: 28655883; PMCID: PMC5958546
12. Calancie L, Keyserling TC, Taillie LS, Robasky K, Patterson C, Ammerman AS, Schisler JC (2018) TAS2R38 predisposition to bitter taste associated with differential changes in vegetable intake in response to a community-based dietary intervention. G3 (Bethesda). 8(6):2107–2119. https://doi.org/10.1534/g3.118.300547. PMID: 29686110; PMCID: PMC5982837
13. Rozengurt E (2006) Taste receptors in the gastrointestinal tract. I. Bitter taste receptors and alpha-gustducin in the mammalian gut. Am J Physiol Gastrointest Liver Physiol 291(2):G171–G177. https://doi.org/10.1152/ajpgi.00073.2006. Epub 2006 May 18. PMID: 16710053
14. Medapati MR, Singh N, Bhagirath AY et al (2021) Bitter taste receptor T2R14 detects quorum sensing molecules from cariogenic Streptococcus mutans and mediates innate immune responses in gingival epithelial cells. FASEB J 35:e21375. https://doi.org/10.1096/fj.202000208R
15. Tuzim K, Korolczuk A (2021) An update on extra-oral bitter taste receptors. J Transl Med 19:440. https://doi.org/10.1186/s12967-021-03067-y
16. Jeon TI, Seo YK, Osborne TF (2011) Gut bitter taste receptor signalling induces ABCB1 through a mechanism involving CCK. Biochem J 438(1):33–37. https://doi.org/10.1042/BJ20110009.PMID:21592089;PMCID:PMC3319131
17. Higdon JV, Delage B, Williams DE, Dashwood RH (2007) Cruciferous vegetables and human cancer risk: epidemiologic evidence and mechanistic basis. Pharmacol Res 55(3):224–236. https://doi.org/10.1016/j.phrs.2007.01.009. Epub 2007 Jan 25. PMID: 17317210; PMCID: PMC2737735
18. Birt DF, Pelling JC, Pour PM, Tibbels MG, Schweickert L, Bresnick E (1987) Enhanced pancreatic and skin tumorigenesis in cabbage-fed hamsters and mice. Carcinogenesis 8(7):913–917. https://doi.org/10.1093/carcin/8.7.913.PMID:3594724
19. Gaida MM, Mayer C, Dapunt U, Stegmaier S, Schirmacher P, Wabnitz GH, Hänsch GM (2016) Expression of the bitter receptor T2R38 in pancreatic cancer: localization in lipid droplets and activation by a bacteria-derived quorum-sensing molecule. Oncotarget 7(11):12623–12632. https://doi.org/10.18632/oncotarget.7206.PMID:26862855;PMCID:PMC4914309
20. Wooding S, Gunn H, Ramos P, Thalmann S, Xing C, Meyerhof W (2010) Genetics and bitter taste responses to goitrin, a plant toxin found in vegetables. Chem Senses 35(8):685–692. https://doi.org/10.1093/chemse/bjq061
21. Lambert JD, VanDusen SR, Cockroft JE et al (2019) Bitter taste sensitivity, food intake, and risk of malignant cancer in the UK Women's Cohort study. Eur J Nutr 58:2111–2121. https://doi.org/10.1007/s00394-018-1772-4

22. Bachmanov AA, Beauchamp GK (2007) Taste receptor genes. Annu Rev Nutr 27:389–414. https://doi.org/10.1146/annurev.nutr.26.061505.111329.PMID:17444812;PMCID:PMC2721271
23. Wrangham RW, Nishida T (1983) *Aspilia* spp. leaves: a puzzle in the feeding behaviour of wild chimpanzees. Primates 24:276–282
24. Peña-Espinoza M, Valente AH, Thamsborg SM et al (2018) Antiparasitic activity of chicory (*Cichorium intybus*) and its natural bioactive compounds in livestock: a review. Parasites Vectors 11:475. https://doi.org/10.1186/s13071-018-3012-4
25. Allchin FR (1979) India: the ancient home of distillation? Man 14(1):55–63. https://doi.org/10.2307/2801640
26. Górak A, Sorensen E (2014) Distillation: fundamentals and principles. Academic Press. ISBN 978-0-12-386547-2
27. Hansen BB et al (2020) Chem Rev 2021, 121:3, 1232–1285. https://doi.org/10.1021/acs.chemrev.0c00385
28. El Achkar T, Greige-Gerges H, Fourmentin S (2021) Basics and properties of deep eutectic solvents: a review. Environ Chem Lett 19:3397–3408. https://doi.org/10.1007/s10311-021-01225-8
29. Gonzalez-Rodrigo et al (2023) IJFST 58:6660–6673. https://doi.org/10.1111/ijfs.16641
30. Zhang Y, Cao Y, Wang H (2020) Multi-interactions in ionic liquids for natural product extraction. Molecules 26(1):98. https://doi.org/10.3390/molecules26010098.PMID:33379318;PMCID:PMC7796109
31. Uwineza PA, Waśkiewicz A (2020) Recent advances in supercritical fluid extraction of natural bioactive compounds from natural plant materials. Molecules 25(17):3847. https://doi.org/10.3390/molecules25173847.PMID:32847101;PMCID:PMC7504334
32. Wöhler F (1828) Ueber künstliche Bildung des Harnstoffs. Ann Phys Chem 12:253–257. https://doi.org/10.1002/andp.18280870206
33. Sertürner FW (1817) Über das Morphium, eine neue salzfähige Grundlage, und die Mekonsäure, als Hauptbestandteile des Opiums. Annalen der Physik 25(1):56–90. Bibcode:1817AnP....55...56S. https://doi.org/10.1002/andp.18170550104
34. Huxtable RJ, Schwarz SK (2001) The isolation of morphine–first principles in science and ethics. Mol Interv 1(4):189–191 PMID: 14993340
35. Pelletier PJ, Caventou JB (1820) Ann Chim Phys 15:337–365
36. Kaufman TS, Rúveda EA (2005). Die Jagd auf Chinin: Etappenerfolge und Gesamtsiege. Angewandte Chem Int Ed 117(6):876–907 (in German)
37. Woodward R, Doering W (1944) The total synthesis of quinine. J Am Chem Soc 66(849):849
38. Stork G, Niu D, Fujimoto A, Koft ER, Balkovec JM, Tata JR, Duke GR (2001) The first stereoselective total synthesis of quinine. J Am Chem Soc 123:3239–3242
39. Smith AC, Williams RM (2008) Rabe rest in peace: confirmation of the Rabe-Kindler conversion of d-Quinotoxine into quinine: experimental affirmation of the Woodward-Doering formal total synthesis of quinine. Angew Chem Int Ed 47(9):1736–1740
40. Cova TFGG, Pais AACC, Seixas de Melo JS (2017) Reconstructing the historical synthesis of mauveine from Perkin and Caro: procedure and details. Sci Rep 7:6806. https://doi.org/10.1038/s41598-017-07239-z
41. Han F, Kang LZ, Zeng LQ, Lin JF (2014) Bioproduction of baccatin III, an advanced precursor of paclitaxol with XL, Ye ZW, Guo transgenic Flammulina velutipes expressing 10-Deacetylbaccatin III-10-O-acetyl transferase gene. J Sci Food Agric 94(12):2376–2383
42. Francis K, Smitherman C, Nishino SF, Spain JC, Gadda G (2013) The biochemistry of the metabolic poison propionate 3-nitronate and its conjugate acid, 3-nitropropionate. IUBMB Life 65(9):759–768. https://doi.org/10.1002/iub.1195. Epub 2013 Jul 29 PMID: 23893873
43. Upadhayay S, Yedke NG, Rahi V, Singh S, Kumar S, Arora A, Chandolia P, Kaur P, Kumar M, Koshal P, Jamwal S, Kumar P (2023) An overview of the pathophysiological mechanisms of 3-nitropropionic acid (3-NPA) as a neurotoxin in a Huntington's Disease model and its relevance to drug discovery and development. Neurochem Res 48(6):1631–1647. https://doi.org/10.1007/s11064-023-03868-1. Epub 2023 Feb 4 PMID: 36738367

44. Baxter RL, Hanley AB, Chan HW-S, Greenwood SL, Abbott EM, McFarlane IJ, Milne K (1992) Fungal biosynthesis of 3-nitropropanoic acid. J Chem Soc Perkin Trans I:2495
45. Porter DJ, Bright HJ (1987) Propionate-3-nitronate oxidase from Penicillium atrovenetum is a flavoprotein which initiates the autoxidation of its substrate by O_2. J Biol Chem 262(30):14428–14434 PMID: 3667582
46. Hanley AB, Heaney RK, Fenwick GR (1983) Improved isolation of glucobrassicin and other glucosinolates. J Sci Food Agric 34:869
47. Refaat D, Aggour MG, Farghali AA, Mahajan R, Wiklander JG, Nicholls IA, Piletsky SA (2019) Strategies for molecular imprinting and the evolution of MIP nanoparticles as plastic antibodies—synthesis and applications. Int J Mol Sci 20:6304. https://doi.org/10.3390/ijms20246304
48. Pollack SJ, Jacobs JW, Schultz PG (1986) Selective chemical catalysis by an antibody. Science 234:1570–1573
49. Harwood LM (1985) Aldrichimica Acta 18:25
50. Moshkov N, Becker T, Yang K et al (2023) Predicting compound activity from phenotypic profiles and chemical structures. Nat Commun 14:1967. https://doi.org/10.1038/s41467-023-37570-1
51. Casas ML, Esteban Á, González-Muñoz M, Labrador-Horrillo M, Pascal M, Teniente-Serra A (2020) VALIDA project: validation of allergy in vitro diagnostics assays (Tools and recommendations for the assessment of in vitro tests in the diagnosis of allergy). Adv Lab Med 1(4):20200051. https://doi.org/10.1515/almed-2020-0051.PMID:37360620;PMCID:PMC10197418
52. Opara EI, Oehlschlager SL, Hanley AB Immunoglobulin E (1998) mediated food allergy. Modelling and application of diagnostic and predictive tests for existing and novel foods. Biomarkers 3:1
53. Covington BC, McLean JA, Bachmann BO (2017) Comparative mass spectrometry-based metabolomics strategies for the investigation of microbial secondary metabolites. Nat Prod Rep 34:6–24
54. Hou J, Zhang Z, Wu W et al (2022) Mass spectrometry imaging: new eyes on natural products for drug research and development. Acta Pharmacol Sin 43:3096–3111
55. Besnard P, Passilly-Degrace P, Khan NA (2016) Taste of fat: a sixth taste modality? Physiol Rev 96:151–176
56. Choi C, Bae J, Kim S et al (2023) Understanding the molecular mechanisms of odorant binding and activation of the human OR52 family. Nat Commun 14:8105
57. Billesbølle CB, de March CA, van der Velden WJC et al (2023) Structural basis of odorant recognition by a human odorant receptor. Nature 615:742–749. https://doi.org/10.1038/s41586-023-05798-y
58. Kotlo K, Anbazhagan AN, Priyamvada S, Jayawardena D, Kumar A, Chen Y, Xia Y, Finn PW, Perkins DL, Dudeja PK, Layden BT (2020) The olfactory G protein-coupled receptor (Olfr-78/OR51E2) modulates the intestinal response to colitis. Am J Physiol Cell Physiol 318(3):C502–C513. https://doi.org/10.1152/ajpcell.00454.2019. Epub 2020 Jan 8. PMID: 31913697; PMCID: PMC7099522
59. Olofsson JK, Ekström I, Larsson M, Nordin S (2021) Olfaction and aging: a review of the current state of research and future directions. Iperception 12(3):20416695211020332. https://doi.org/10.1177/20416695211020331.PMID:34249327;PMCID:PMC8239976
60. Mathur K, Agrawal RK, Nagpure S, Deshpande D (2020) Effect of artificial sweeteners on insulin resistance among type-2 diabetes mellitus patients. J Family Med Prim Care 9(1):69–71. https://doi.org/10.4103/jfmpc.jfmpc_329_19. PMID: 32110567; PMCID: PMC7014832
61. Alsunni AA (2020) Effects of artificial sweetener consumption on glucose homeostasis and its association with Type 2 diabetes and obesity. Int J Gen Med 6(13):775–785. https://doi.org/10.2147/IJGM.S274760.PMID:33116769;PMCID:PMC7547772
62. Nomura K, Nakanishi M, Ishidate F, Iwata K, Taruno A (2020) All-electrical Ca2+-independent signal transduction mediates attractive sodium taste in taste buds. Neuron 106:816–829

63. Meyerhof W et al (2010) The molecular receptive ranges of human TAS2R bitter taste receptors. Chem Senses 35:157–170. https://doi.org/10.1093/chemse/bjp092
64. Schreurs M, Piampongsant S, Roncoroni M et al (2024) Predicting and improving complex beer flavor through machine learning. Nat Commun 15:2368. https://doi.org/10.1038/s41467-024-46346-0
65. Nissim I, Dagan-Wiener A, Niv MY (2017) The taste of toxicity: a quantitative analysis of bitter and toxic molecules. IUBMB Life 69(12):938–946. https://doi.org/10.1002/iub.1694. Epub 2017 Nov 11 PMID: 29130618
66. Tuzim K, Korolczuk A (2021) An update on extra-oral bitter taste receptors. J Transl Med 19(1):440. https://doi.org/10.1186/s12967-021-03067-y. Erratum in: J Transl Med. 2021 Nov 26;19(1):478. PMID: 34674725; PMCID: PMC8529754
67. Gaida MM, Dapunt U, Hänsch GM (2016) Sensing developing biofilms: the bitter receptor T2R38 on myeloid cells. Pathog Dis 74(3):ftw004. https://doi.org/10.1093/femspd/ftw004. Epub 2016 Jan 17. PMID: 26782143; PMCID: PMC4748947
68. Duarte AC, Costa AR, Gonçalves I, Quintela T, Preissner R, Santos CRA (2022) The druggability of bitter taste receptors for the treatment of neurodegenerative disorders. Biochem Pharmacol 197
69. Seaks CE, Wilcock DM (2020) Infectious hypothesis of Alzheimer disease. PLoS Pathog 16(11):e1008596. https://doi.org/10.1371/journal.ppat.1008596.PMID:33180879; PMCID:PMC7660461
70. vlab.amrita.edu (2012) Calculation of λmax of organic compounds using Woodward Fieser Rules. Retrieved 13 May 2024. from vlab.amrita.edu/?sub=2&brch=191&sim=850&cnt=1
71. Silk economics and cross-cultural artistic interaction: Byzantium, the Muslim world, and the Christian west. Dumbarton Oaks Pap 58:197–240. 210
72. Bechtold T, Mussak R (2009) Handbook of natural colorants. Wiley, Hoboken
73. Linke JA, Rayat A, Ward JM (2023) Production of indigo by recombinant bacteria. Bioresour Bioprocess 10:20. https://doi.org/10.1186/s40643-023-00626-7
74. Perkin WH (1856) Producing a new coloring matter for dyeing with a lilac or purple color stuffs of silk, cotton, wool, or other materials. William Henry Perkin, United Kingdom, BP1984
75. Tian M, Hanley AB, Dodds MW (2013) Allyl isothiocyanate from mustard seed is effective in reducing the levels of volatile sulfur compounds responsible for intrinsic oral malodor. J Breath Res 7(2):026001. https://doi.org/10.1088/1752-7155/7/2/026001. Epub 2013 Mar 8. PMID: 23470258
76. Hejna M, Kovanda L, Rossi L, Liu Y (2021) Mint oils: in vitro ability to perform anti-inflammatory, antioxidant, and antimicrobial activities and to enhance intestinal barrier integrity. Antioxidants (Basel) 10(7):1004. https://doi.org/10.3390/antiox10071004. PMID: 34201645; PMCID: PMC8300686
77. Dylong D, Hausoul PJC, Palkovits R, Eisenacher M (2022) Synthesis of (−)-menthol: industrial synthesis routes and recent development. Flavour Fragrance J. https://doi.org/10.1002/ffj.3699
78. Zhao H, Ren S, Yang H, Tang S, Guo C, Liu M, Tao Q, Ming T, Xu H (2022) Peppermint essential oil: its phytochemistry, biological activity, pharmacological effect and application. Biomed Pharmacother 154:113559. https://doi.org/10.1016/j.biopha.2022.113559. Epub 2022 Aug 19 PMID: 35994817
79. Aroke EN, Powell-Roach KL, Jaime-Lara RB, Tesfaye M, Roy A, Jackson P, Joseph PV (2020) Taste the pain: the role of TRP channels in pain and taste perception. Int J Mol Sci 21(16):5929. https://doi.org/10.3390/ijms21165929.PMID:32824721;PMCID:PMC7460556
80. Johnson SS, Greenberg MJ (1995) Method for refining mint oils and chewing gum made from same, US Patent 5,425,962
81. Ratzka A, Vogel H, Kliebenstein DJ, Mitchell-Olds T, Kroymann J (2002) Disarming the mustard oil bomb. Proc Natl Acad Sci U S A 99(17):11223–11228. https://doi.org/10.1073/pnas.172112899. Epub 2002 Aug 2. PMID: 12161563; PMCID: PMC123237
82. Kazana E, Pope TW, Tibbles L, Bridges M, Pickett JA, Bones AM, Powell G, Rossiter JT (2007) The cabbage aphid: a walking mustard oil bomb. Proc Biol Sci 274(1623):2271–2277. https://doi.org/10.1098/rspb.2007.0237.PMID:17623639;PMCID:PMC2288485

83. Tepper BJ, Banni S, Melis M, Crnjar R, Tomassini BI (2014) Genetic sensitivity to the bitter taste of 6-n-propylthiouracil (PROP) and its association with physiological mechanisms controlling body mass index (BMI). Nutrients 6(9):3363–3381. https://doi.org/10.3390/nu6093363.PMID:25166026;PMCID:PMC4179166
84. Fenwick GR, Griffiths NM (1981) The identification of the goitrogen, (−)5-vinyloxazolidine-2-thione (goitrin), as a bitter principle of cooked brussels sprouts (Brassica oleracea L. var. gemmifera). Z Lebensm Unters Forch 172:90–92. https://doi.org/10.1007/BF01042410
85. Nor NDM, Mullick H, Zhou X, Oloyede O, Houston-Price C, Harvey K, Methven L (2023) Consumer liking of turnip cooked by different methods: the influence of sensory profile and consumer bitter taste genotype. Foods 12:3188. https://doi.org/10.3390/foods12173188
86. Diószegi J, Llanaj E, Ádány R (2019) Genetic background of taste perception, taste preferences, and its nutritional implications: a systematic review. Front Genet 19(10):1272. https://doi.org/10.3389/fgene.2019.01272.PMID:31921309;PMCID:PMC6930899
87. Franzago M, Borrelli P, Di Nicola M, Stuppia L, Vitacolonna E (2023) Genetic variants in CD36 involved in fat taste perception: association with anthropometric and clinical parameters in overweight and obese subjects affected by type 2 diabetes or dysglycemia—a pilot study. Nutrients 15(21):4656. https://doi.org/10.3390/nu15214656.PMID:37960309;PMCID:PMC10647499
88. Krupp F, Frey W, Richert C (2020) Absolute configuration of small molecules by co-crystallization. Angew Chem Int Ed Engl 59(37):15875–15879. https://doi.org/10.1002/anie.202004992. Epub 2020 Jun 29. PMID: 32441841; PMCID: PMC7540501
89. Danelius E, Halaby S, van der Donk WA, Gonen T (2021) MicroED in natural product and small molecule research. Nat Prod Rep 38(3):423–431. https://doi.org/10.1039/d0np00035c. Epub 2020 Sep 17. PMID: 32939523; PMCID: PMC7965795
90. Strack R (2020) Cryo-EM goes atomic. Nat Methods 17:1175. https://doi.org/10.1038/s41592-020-01014-1
91. Rodríguez J, Crews P, Jaspars M (2012) Contemporary strategies in natural products structure elucidation. In: Fattorusso E, Gerwick W, Taglialatela-Scafati O (eds) Handbook of marine natural products. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-3834-0_7
92. Semenov VA, Krivdin LB (2022) Computational NMR of natural products. Russ Chem Rev 91:RCR5027 https://doi.org/10.1070/RCR5027
93. Wishart DS, Sayeeda Z, Budinski Z, Guo A, Lee BL, Berjanskii M, Rout M, Peters H, Dizon R, Mah R, Torres-Calzada C, Hiebert-Giesbrecht M, Varshavi D, Varshavi D, Oler E, Allen D, Cao X, Gautam V, Maras A, Poynton EF, Tavangar P, Yang V, van Santen JA, Ghosh R, Sarma S, Knutson E, Sullivan V, Jystad AM, Renslow R, Sumner LW, Linington RG, Cort JR (2022) NP-MRD: the natural products magnetic resonance database. Nucleic Acids Res 50(D1):D665–D677. https://doi.org/10.1093/nar/gkab1052.PMID:34791429;PMCID:PMC8728158
94. Japp FR (1898) Stereochemistry and vitalism. Report of the 68th meeting of the British Association for the Advancement of Science. John Murray, London, pp 813–828
95. Aicardi C (2017) Francis Crick, cross-worlds influencer: a narrative model to historicize big bioscience. Stud Hist Philos Biol Biomed Sci 55:83–95
96. Miller AAM, Biallas P, Shennan BDA, Dixon DJ (2024) Angew Chem Int Ed 63:e202314308. https://doi.org/10.1002/anie.202314308
97. Hudlicky JR, Sulikowski GA (2019) Simple start for complex products. Nat Chem 11:297–298. https://doi.org/10.1038/s41557-019-0241-x
98. Fay N, Kouklovsky C, de la Torre A (2023) Natural product synthesis: the endless quest for unreachable perfection. ACS Org Inorg Au. 3(6):350–363. https://doi.org/10.1021/acsorginorgau.3c00040.PMID:38075446;PMCID:PMC10704578
99. Maier ME (2015) Design and synthesis of analogues of natural products. Org Biomol Chem 13(19):5302–5343. https://doi.org/10.1039/c5ob00169b. PMID: 25829247
100. Ozcengiz G, Demain AL (2013) Recent advances in the biosynthesis of penicillins, cephalosporins and clavams and its regulation. Biotechnol Adv 31(2):287–311
101. Robinson R (1917) J Chem Soc 111:762

102. Medley JW, Movassaghi M (2013) Robinson's landmark synthesis of tropinone. Chem Commun (Camb) 49(92):10775–10777. https://doi.org/10.1039/c3cc44461a. PMID: 24116374
103. Bao R, Zhang H, Tang Y (2021) Biomimetic synthesis of natural products: a journey to learn, to mimic, and to be better. Acc Chem Res 54(19):3720–3733. https://doi.org/10.1021/acs.accounts.1c00459. Epub 2021 Sep 22 PMID: 34549936
104. Ortiz C, Jackson E, Betancor L (2022) Immobilization and stabilization of enzymes using biomimetic silicification reactions. J Sol-Gel Sci Technol 102:86–95. https://doi.org/10.1007/s10971-022-05751-x
105. Koeller K, Wong CH (2001) Enzymes for chemical synthesis. Nature 409:232–240. https://doi.org/10.1038/35051706
106. Dear P (2008) The intelligibility of nature. University of Chicago Press
107. Sumner JB (1926) The isolation and crystallization of the enzyme urease. Preliminary Pap J Biol Chem 69:435–441
108. Sumner JB, Kirk JS, Howell SF (1932) The digestion and inactivation of crystalline urease by pepsin and by papain. J Biol Chem 98:543–552
109. Breitling R, Ceniceros A, Jankevics A, Takano E (2013) Metabolomics for secondary metabolite research. Metabolites 3(4):1076–1083. https://doi.org/10.3390/metabo3041076.PMID:24958266;PMCID:PMC3937839
110. Garner KL (2021) Principles of synthetic biology. Essays Biochem 65(5):791–811. https://doi.org/10.1042/EBC20200059.PMID:34693448;PMCID:PMC8578974.)
111. Hanson JR (2017) One hundred years in the elucidation of the structures of natural products. www.scienceprogress.co.uk Structures of natural products. Sci Prog 100(1):63–79 Paper 1700221 https://doi.org/10.3184/003685017X14876775256129
112. Zarate E, Boyle V, Rupprecht U, Green S, Villas-Boas SG, Baker P, Pinu FR (2016) Fully automated trimethylsilyl (TMS) derivatisation protocol for metabolite profiling by GC-MS. Metabolites 7(1):1. https://doi.org/10.3390/metabo7010001.PMID:28036063;PMCID:PMC5372204
113. McCaughey CS, van Santen JA, van der Hooft JJJ et al (2022) An isotopic labeling approach linking natural products with biosynthetic gene clusters. Nat Chem Biol 18:295–304. https://doi.org/10.1038/s41589-021-00949-6
114. Cox IJ, Hanley AB, Belton PS (1984) Fenwick G R Nmr spectra (1H, 13C) of glucosinolates. Carbohydrate Res 132:323
115. Gorenflos López JL, Schmieder P, Kemnitz-Hassanin K, Asikoglu HC, Celik A, Stieger CE, Fiedler D, Hinderlich S, Hackenberger CPR (2023) Real-time monitoring of the sialic acid biosynthesis pathway by NMR. Chem Sci 14(13):3482–3492. https://doi.org/10.1039/d2sc06986e.PMID:37006695;PMCID:PMC10055903
116. Singh S, Peltier-Pain P, Tonelli M, Thorson JS (2014) A general NMR-based strategy for the in situ characterization of sugar-nucleotide-dependent biosynthetic pathways. Org Lett 16(12):3220–3223. https://doi.org/10.1021/ol501241a. Epub 2014 Jun 9. PMID: 24911465; PMCID: PMC4075999
117. Cravens A, Payne J, Smolke CD (2019) Synthetic biology strategies for microbial biosynthesis of plant natural products. Nat Commun 10:2142. https://doi.org/10.1038/s41467-019-09848-w
118. Skellam E, Rajendran S, Li L (2024) Combinatorial biosynthesis for the engineering of novel fungal natural products. Commun Chem 7:89. https://doi.org/10.1038/s42004-024-01172-9
119. Garner KL (2021) Principles of synthetic biology. Essays Biochem 65(5):791–811. https://doi.org/10.1042/EBC20200059.PMID:34693448;PMCID:PMC8578974

Chapter 2
Improbable Structures

Abstract A consideration of improbable structures in nature and the link between improbable and novel structures. Creation of molecules by synthesis and by using biological systems. The history of beta lactams as improbable structures and their eventual resolution. Hermeneutics and natural products. The example of glucosinolates as a misassigned structure and the eventual unveiling of the correct structure by synthesis and X-ray crystallography. The development of combinatorial chemistry and natural product libraries. Beta lactams and beta lactamases in the study of adaptive responses. Anti-microbial resistance (AMR) and its inevitability. The concept of quorum sensing and its use as a tool to overcome AMR. The complexity of microbiomes and cell–cell communication.

Keywords Structures · Beta lactam · AMR · Quorum sensing

2.1 Basis of Improbability and Nature

Improbable natural products can be considered to be of several types and these will be examined in the first part of the chapter. The following are the results of or consequences arising from improbable structures:

a. Structures that were either misassigned (i.e. were wrong) or which are in fact correct but seem to be improbable either on the grounds of the structures themselves possessing unlikely characteristics or their biosynthetic origin seems anomalous.
b. The proof of structures by synthesis or degradation studies (including enzymic degradation). Until the development of advanced spectroscopic techniques proof of structure by synthesis or degradation was the most useful of techniques. This approach is also inherent in the application of reductionism to natural products.
c. The creation of novel molecules from the modification of biosynthetic pathways. This has already been alluded to in the previous chapter when discussing synthetic biology and will be a recurring theme. In some cases the pathway can be altered

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_2

by changing the characteristics of the product of a single gene. In other cases, the order of genes in a pathway can be manipulated giving rise to novel products.

The second part of the chapter will deal with biological responses to natural and unnatural molecules using beta lactamases as an example. There will then be a discussion of anti-microbial resistance and ways to change the paradigm of the introduction of new anti-microbials and the gradual growth in resistant strains of microbes. In this 'anti-microbial arms race' evolution favours microbes rather than the chemist and therefore an alternative approach is needed rather than simply attempting to devise novel structures that are biologically active but to which resistance will not initially be present. The history of anti-microbial resistance suggests that the development of resistance is inevitable.

One important aspect of how natural (and unnatural) products should be defined depends upon the frame of reference that is adopted. Living systems produce natural products using the biosynthetic/enzymic tools that are available to them and which have evolved over time. Certain mechanistic attributes are commonly encountered and these provide the biosynthetic rationale for the accumulation of some natural products. While the tools may be significantly different to those that the synthetic chemist uses, there is a sense in which they are similar. Single reactions can have consequences and these consequences may have molecular and biological significance. In synthetic chemistry, concerted reactions involve the simultaneous breaking and forming of bonds. This can result in an inversion of stereochemistry at the carbon atom involved—the Walden inversion (Fig. 2.1).

This is an example of the molecular consequences of a chemical reaction. In biological systems, individual reactions may also have consequences. For example glucosinolates, which occur throughout Brassica species, contain a thioglucose moiety that can be considered to be a form of 'protecting/stabilising group'. The thioglucose moiety is an important aspect of the glucosinolate molecule. When this

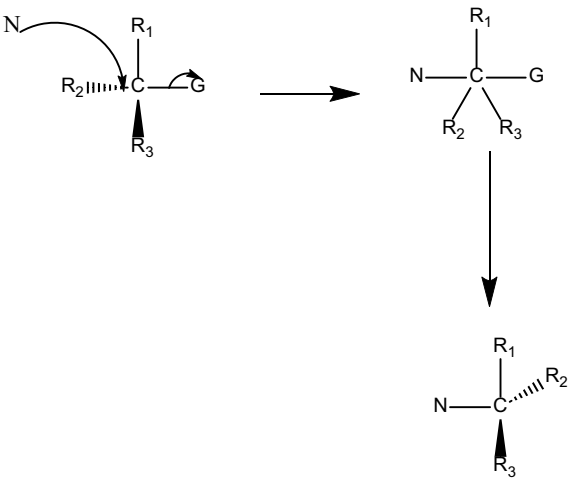


Fig. 2.1 The Walden inversion

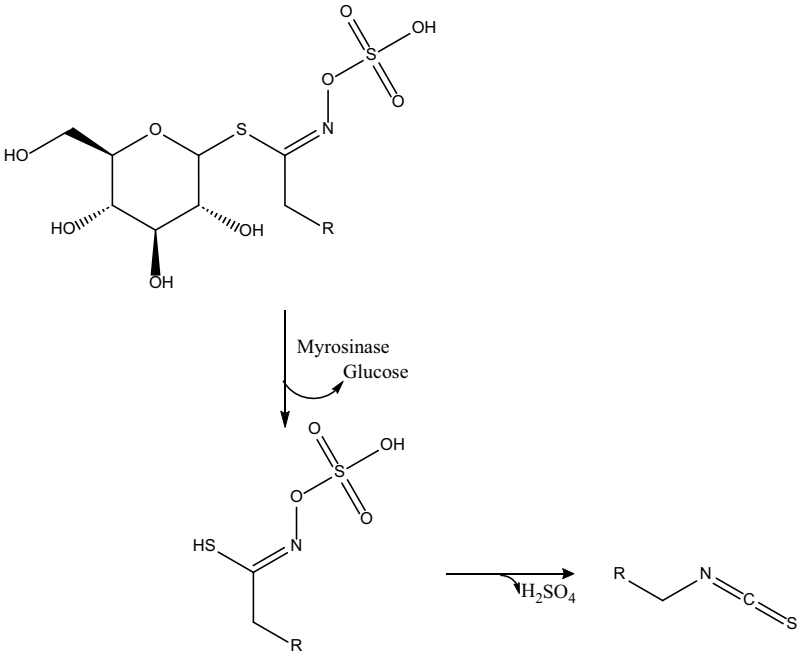


Fig. 2.2 Formation of an isothiocyanate by enzymic Hydrolysis of a glucosinolate

protecting/stabilising group is removed by cleavage of the C–S bond by the thioglucosidase enzyme, myrosinase, the molecule rearranges to give an isothiocyanate which is the bioactive molecule responsible for protecting the plant from attack (Fig. 2.2).

The potentiality of the molecule as a bioactive entity is only released when the protecting/stabilising group is removed. In a similar (but much less elegant) way, in addition to specific molecular consequences such as the Walden inversion, chemists can construct molecules using cascade reactions in organic synthesis [1]. A natural product can be defined as any product that is created by a living system. This means that all synthetic compounds are, in a sense 'natural products' since they are created by a living system (a chemist). This is a rather unusual definition of natural products and must include some very strict confines. This can be summarised in the claim of Protagoras of Abdera who stated simply—"Of all things the measure is Man, of the things that are, that they are, and of the things that are not, that they are not". This means that humans are the ultimate empirical creators and anything they can create is a product of the world. An alternative approach is to consider a different frame of reference. By deciding to adopt a different interpretation of the agency of how things are created, we can decide that it is only things (in chemistry terms, molecules) that are created intrinsically and directly by a living system that are 'natural products' and only defined as such in the context of that living system alone. Everything that lies outside of the living system is an 'unnatural product' from the frame of reference

of that particular living system. By this definition, beta lactams such as penicillin are natural products in the context of the fungi that contain the biosynthetic pathways to create them but are 'unnatural products' from the standpoint of the bacteria which are affected by them. The production of beta lactamases (and, indeed other mechanisms to resist the deleterious effects of beta lactams) are simply a bacterial response to an 'unnatural product'. From the frame of reference of the bacterium, it is irrelevant if the exogenous agent is produced by the biological machinery of a fungus or by the ingenuity of a chemist in the laboratory. If we take the frame of reference definition then, from a human perspective, all molecules that cannot be synthesised, intrinsically and directly by humans are 'unnatural' rather than 'natural' products. In short, our definition of natural and unnatural products is inconsistent and lacks an appreciation of what are appropriate frames of reference. In practical terms, natural products are defined more loosely as small molecules produced naturally by any organism. This excludes any consideration of a frame of reference and purposely excludes molecules synthesised outside an organism.

2.2 Improbable Structures

One of the most celebrated improbable structures is that of the beta lactam ring (Fig. 2.3).

This structure, which was eventually proven by X-ray crystallography [2, 3], was the subject of much debate. Prior to the definitive X-ray analysis, the structure was determined through an elegant series of degradation experiments that were described by Chain in his Nobel prize lecture in 1946. Interestingly, one of the most significant figures in natural product chemistry, Sir Robert Robinson, initially favoured a thiazolidine-oxazolone structure rather than the beta lactam structure (Fig. 2.4) [4].

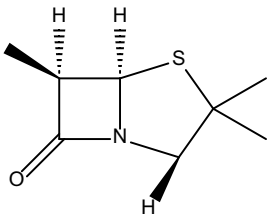
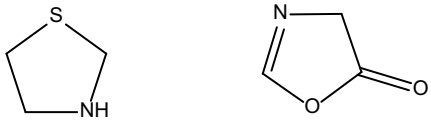


Fig. 2.3 Structure of a beta lactam ring



Thiazolidine ring Oxazolone ring

Fig. 2.4 Thiazolidine and oxazolone rings

This was not a trivial choice since there was some evidence from a group working at Merck in the US that a thiazolidine-oxazolone structure in which the two five-membered rings were joined by a single bond rather than the beta lactam fused ring structure. was plausible and perhaps more likely than the beta lactam structure. If NMR spectroscopy had been available then the structure would have been solved more rapidly and definitively. From a biological perspective, penicillins act by binding to the active site of an enzyme involved in the final step in bacterial cell wall biosynthesis. Bacterial cell walls are composed of a peptidoglycan backbone and penicillin is an analogue of the acyl-D-alanyl-D-alanine terminus that subsequently forms the final step in the cell wall synthesis [5]. The mechanism involves binding to a serine OH moiety in the active site of the enzyme and this is not a reaction that the alternative thiazoline-oxazolone structure would be capable of (Fig. 2.5).

Another example of a misassigned structure is that of the class of compounds called glucosinolates. In this case, the enzymic (natural) and chemical degradation of the original glycoside gives rise to an isothiocyanate. The structure of glucosinolates was initially proposed by Gadamer in 1897 [6] and was based upon the isolated degradation product.

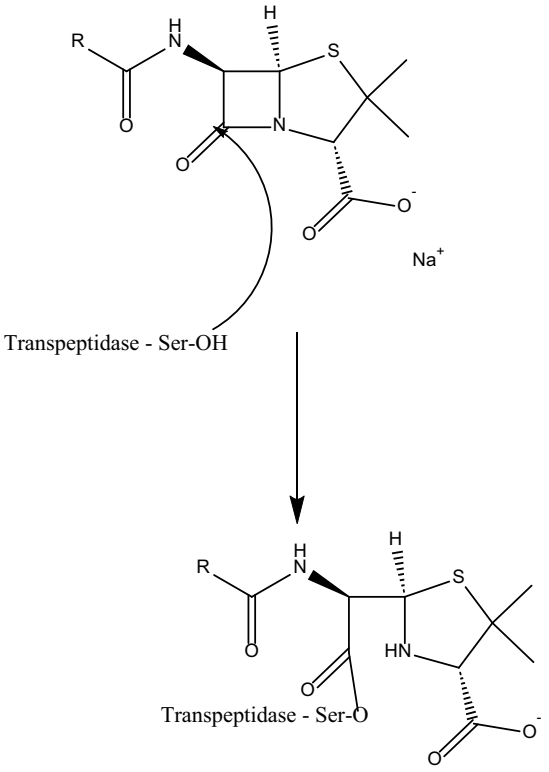


Fig. 2.5 Binding of penicillin to bacterial cell wall transpeptidase

Gadamer's son Hans-Georg has been described as the decisive figure in the development of twentieth century hermeneutics. The essence of hermeneutics is interpretation (traditionally of texts but later in a more general sense) and this can lead us to consider whether natural product chemistry is a form of 'natural hermeneutics'—does it contribute to the interpretation/understanding of living systems? This requires us to extend the principle of hermeneutics beyond textual interpretation to consider all forms of human understanding (in which Hans-Georg Gadamer was a leading figure), and, in particular, the understanding of living systems. By accepting this extension of hermeneutics, it is possible to create the proposition that natural products and the interpretation of their role and purpose is one way that chemists try to understand the natural world. From this it follows that natural product chemistry makes a contribution to the understanding of the natural world and is, therefore, a form of hermeneutics. Such an interpretation places natural product chemistry at the core of one way of inspecting and understanding the world of living systems. It also provides an intellectual link between Hans-Georg Gadamer and his father.

Gadamer's structural assignment of glucosinolates was based both on chemical and enzymic degradation and made some logical sense since it possessed the constituent atoms of the isothiocyanate moiety in the correct relative orientation (R–N=C–S-Glucose). Removal of the glucose moiety would give rise to the isothiocyanate. The correct structure was somewhat different and was proposed and confirmed by Ettlinger and Lundeen in the 1950s both by synthesis and by X-ray crystallography. Indeed since we now know more about the biosynthesis of glucosinolates from amino acids, the Ettlinger and Lundeen structure makes more sense (Fig. 2.6). By considering the whole picture—including the biosynthesis of the natural product and not solely its degradation products—it follows that the 'natural hermeneutics' approach provides more information and yields the correct (although not immediately obvious) structure.

While the assignation of a structure by inference from degradation studies can, as we have seen with penicillin and with glucosinolates, be complex and can lead us to draw an incorrect conclusion about the structure of a given product, this problem has become less pronounced due to the great advances in spectroscopic methods—particularly NMR spectroscopy. Nonetheless, misinterpretation can still occur and a not insignificant number of structural revisions are published each year. This can arise from initial misreading of the spectral data and subsequent reinterpretation, from taking a holistic, natural hermeneutics approach and considering the biosynthesis and biological context of the molecule and, not least, from a confirmatory synthesis of the initially proposed structure bringing a realisation that the synthetic and isolated materials are not identical.

Chemists are generally reluctant to admit to their errors however, in a number of instances, an erratum has been published correcting a previous report [7].

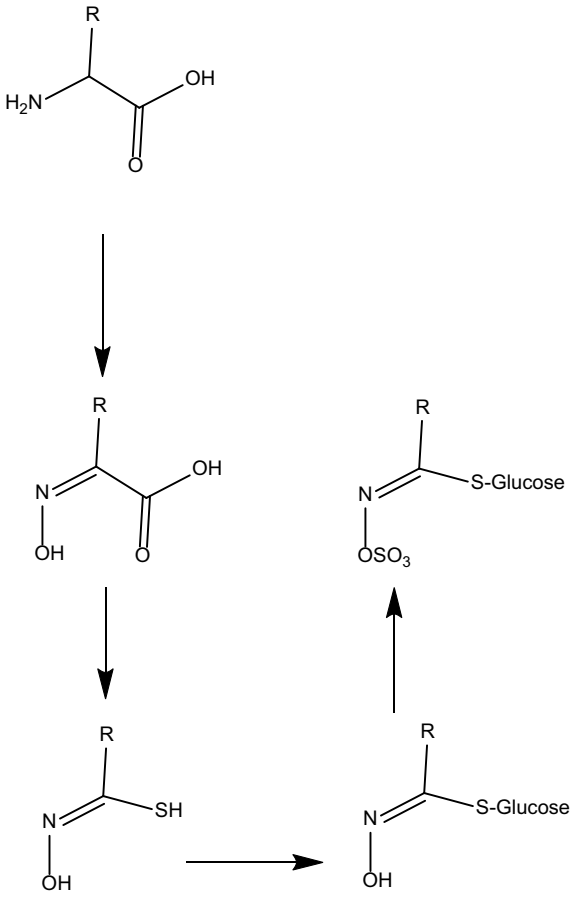


Fig. 2.6 The biosynthesis of glucosinolates

2.3 Proof of Structure Natural Product Isolation and Synthesis

The importance of chemical synthesis, in particular, as a confirmatory technique for proof of a proposed structure has been noted [8]. However complete synthesis of a proposed structure is not always necessary. For example the combination of NMR spectroscopy combined with degradation and synthesis was used to confirm the structure of a ring-substituted indole glucosinolate [9]. The importance of natural product synthesis has been mentioned. In some cases synthetic schemes can be challenging to complete. This is both a consequence of the size and complexity of the target molecule and the amount of functionality present. While the synthesis of relatively small natural products may not be particularly time consuming, the synthesis of complex molecules with multiple chiral centres will be both intellectually challenging and significant in terms of the time and effort expended. In those instances where the natural product

is found to be different to the synthesised molecule then a complete reappraisal may be necessary. Theoretical models can be of great value and can save considerable laboratory time in developing synthetic schemes [10] however there are occasions when they fail. In that case, the simplest method may be to isolate sufficient of the natural product such that a complete suite of spectroscopic and chemical analyses can be carried out and the structure can be confirmed without recourse to a confirmatory synthesis. The most effective path to take depends on the purpose that lies behind the investigation. If, as is often the case, the driver is to find new white space associated with biological activity then the isolation of natural products may not be the most effective strategy since the numbers of products produced may be limited and the purpose for which they are synthesised in vivo is often different to that of the investigator. The discovery, isolation and structural identification of molecules and the subsequent identification of their biological properties can feel somewhat serendipitous. The molecules have frequently been created for a completely different use from that which is of value to the investigator. In contrast, the completely directed, activity-driven approach to medicinal chemistry does not necessarily require input from natural product chemistry. These two approaches are illustrated by the discovery and development of the natural product, penicillin and the synthetic drug Salvarsan. The role of serendipity in drug discovery has been the subject of much commentary [11]. A more precise definition of different types of serendipity has been the subject of a study funded by the European Research Council [12]. One of the key components of serendipitous discovery is the collaborations between networks of researchers. Innovation can occur at the edges where different disciplines meet and this is one of the key aspects of multidisciplinary research. In the case of penicillin, the importance of the properties of the compound produced by the mould *Penicillium notatum* did not attract particular attention until some time after its discovery as an accidental contamination in a petri dish containing a culture of *Staphylococcus.* Fleming was interested in antisepsis and the biological properties of some component of the mould fitted with his scientific view and experience. He interpreted what he saw in the context of his interest in sepsis. Despite his awareness of the potential of a ‘principle’ that could inhibit the growth of pathogens, he experienced great difficulty in isolating the active molecule and might have been aided by the application of natural product chemistry at an earlier stage. Although the initial discovery was made in 1928 by Fleming, it was not until the 1940s that penicillin was purified and even later that its manufacture was put on an industrial footing. Interestingly Fleming also investigated Salvarsan, a purely synthetic organoarsenal compound which was investigated by Ehrlich as a compound to control syphilis. Ehrlich wanted to finding a cure for syphilis—a significant health challenge. He developed the phrase ‘magic bullet’ which based on a German folk story about a hunter who enters into a pact with the devil that enables him to make ‘magic bullets’ that will always hit the intended target. The story was a ghost story collected by Johann Apel and Friedrich Laun in Heidelberg in 1810. The story formed the basis of an opera composed by Carl-Maria von Weber, Der Freischutz (the Freeshooter), which was the first of the great German Romantic operas. Ehrlich adopted the concept of developing agents that could strike specifically against any disease-causing microbe without harming the host [13]. The chemist Albert Bertheim

synthesised a series of compounds and one of these (compound 606—so called because it was the 6th compound of the 6th series to be tested) was tested and found to be effective against the causative microorganism of syphilis, *Treponema pallidum*, by Ehrlich's collaborator, Sahachiro Hata. This organoarsenical compound became the leading treatment for syphilis until it was superseded by penicillin—which had a much lower toxicity profile. Both the biology-driven, observational approach exemplified by Fleming and the organic synthesis route adopted by Ehrlich were effective against the same infectious disease—syphilis, even if the natural product, serendipitous approach produced the better medicinal outcome. An interesting combination approach is the synthesis of putative natural products before they have been discovered. This is a reversal of the more conventional 'discovery then synthesis' paradigm [14].

The development of purely synthetic approaches to the discovery of new bioactive agents led to the development and growth of combinatorial chemistry—a method for preparing and screening large automated libraries of molecules, which will be described in more detail later. This was a contributing factor in the reduction of effort in natural product chemistry—particularly in medicinal chemistry. The period between the isolation and purification of the first biologically useful, characterised natural product (morphine) and the development of combinatorial techniques in the late twentieth century can be described as the 'Golden Age' of natural product chemistry. During this time, the understanding of natural products and their synthesis grew alongside the development of tools to examine their biosynthesis and derivation and the huge advances in understanding how living systems functioned—particularly through the illumination brought about by molecular biology. This has led to a reawakening of natural product chemistry, the development of unnatural product chemistry and the use of biological systems to create complex molecules—a partial reversal of the paradigm that molecules made by nature can invariably be synthesised in useful amounts in the chemistry laboratory.

The role of natural products in the development of isolated bioactive principles was first developed at the beginning of the nineteenth century (morphine) [15] and grew rapidly as isolative and separation techniques grew more sophisticated. The determination of chemical structures, as noted previously, lagged behind isolative techniques and, in effect, active principles were purified to homogeneous bioactivity. The discovery of penicillin led to the development of bioactivity-derived screening of natural products and the isolation of products with defined bioactivity. As noted previously, the isolation of natural products can be a laborious process and rapid throughput screening has served to increase the pressure on isolation as the rate determining step in bioactive compound discovery. This means that there was considerable pressure on pharmaceutical companies to develop more rapid alternatives to natural product isolation. In addition, many complex natural products are only produced in small amounts and are structurally complex (often with multiple chiral centres) so neither the isolative nor the synthetic route are appropriate for producing the large amounts of material that are needed either for commercialisation or even for a full spectrum of confirmatory clinical trials. Two approaches to overcome this significant challenge have been adopted. The first is to identify substructures within which

much of the biological activity of the natural product may reside. This facilitates the simplification of the synthesis of the required part of the natural product where much of the biological activity is found. The second approach (as seen in the case of penicillin and amino-penicillanic acid) is to carry out a partial synthesis in vivo and to then use the semi-developed core structure as a starting point for chemical synthesis of a more complex molecule. There is also in the possibility of using cloning or synthetic biology to transfer the biosynthetic capability into a more suitable host and to thereby simplify the isolation of the bioactive principle or of a modifiable core molecule—generally produced in increased yield.

Given the challenges in isolating and making use of natural products, a different paradigm was required to increase the rate of development of novel bioactive structures. This was provided by the development of high throughput (combinatorial) chemistry. The essence of combinatorial chemistry is the combination of the synthetic chemistry approach as exemplified by Ehrlich's creation of Salvarsan with a rapid biological assay. In this approach, there is no need for natural product chemistry. The essence of combinatorial chemistry has been described in a number of reviews and these can be consulted for a more comprehensive description however, as noted previously, combinatorial chemistry is, of itself, not a technique which requires natural product chemistry input so it will only be dealt with in outline [16]. The essence of combinatorial chemistry is the creation of a large number of structurally diverse motifs which are then screened for activity against biological targets of interest. It is not necessary, in the first instance, to know precisely which molecules provide positive hits and structure:activity information is not required. There are a range of different modes of operation of combinatorial chemistry including the use of biological and synthetic systems however the basis of the approach is that a library is constructed and the members screened. More recently, it has become possible to carry out combinatorial chemistry in silico and the synthesis of an actual library of many thousands of compounds or fragments is not necessary. Biological peptide and DNA libraries can be prepared. One of the most interesting applications is the creation and screening of large libraries of natural product-like macrocyclic peptides [17]. One area that occupies some of the same space as combinatorial chemistry is the de novo synthesis of polyketides. This will be discussed in more detail in Chap. 6. Despite the significant efforts that have gone into combinatorial approaches, the rate of production of small bioactive molecules has not increased dramatically. The increasing use of biologics (including pharmaceutical antibodies) [18] has focussed attention away from traditional small molecule medicinal chemistry however the fraction of plants, microbes and other species that have been mined for natural products remains small and there are a number of areas where further discoveries are likely. One of these, which will be discussed later in this chapter, is the use of microbial modification systems and the investigation of targets based upon multiple biosynthetic pathways rather than traditional, single pathway/mechanism anti-microbial therapy.

The development of computer assisted structure elucidation (CASE) was designed to overcome the errors that may occur as a result of the subjective interpretation of a range of spectroscopic data [19]. Such an empirical approach has a number of advantages including the elimination of many errors however, as the authors of an

important review paper in the area confirm [20], "chemical intuition remains a critical factor in all aspects of the natural product elucidation process." As noted previously in the case of beta lactams, even the most accomplished natural product chemists can be led astray by their knowledge and intuition however the use of automated methodologies (e.g. a reliance on libraries of spectra and properties of molecules to predict the structures of unknown molecules) can be both misleading and can stunt the development of chemical intuition.

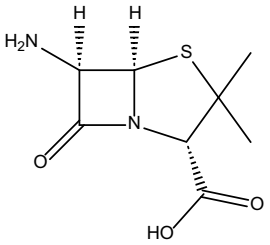
2.4 Beta Lactamases and Anti-microbial Resistance

Natural products evolve and their persistence is a consequence of the advantage they provide for producing species. Although there are a large number of chemical species in nature, natural products only occupy a limited amount of the potential available chemical space and there is a high level of conservation among chemical configurations. This helped chemists to predict the likely structures of natural products from a limited number of core structures and to classify groups of molecules based upon chemical and biosynthetic similarities. This has encouraged the development of pseudo natural products which retain some of the characteristics of known natural products but also have 'unnatural' structural characteristics and different biological properties from the natural parent molecule [21].

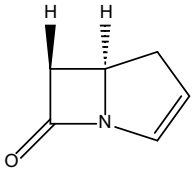
Naturally-occurring lactams comprise a number of different core structures including those that have 4, 5, 6 and 7 membered rings. Of these, the 4-membered rings comprise the class of beta lactams (Fig. 2.7). In simple chemical terms, a lactam is a cyclic amide with the amide N atom in a position beta to the carbonyl moiety—hence the name. Beta lactam antibiotics comprise a number of structures fused onto the lactam ring and these include cephalosporins (with a 6-membered dihydrothiazine ring), carbapenems and monobactams. The latter two are less susceptible to the development of anti-microbial resistance.

The evolution of beta lactamases has been reviewed [22]. The possibility of resistance to penicillins was appreciated by Fleming [23] in his Nobel prize winning lecture and was part of a previously recognised anti-microbial resistance phenomenon. Bacteria had previously developed resistance to sulpha drugs [24]. While beta lactams confer a growth advantage for the fungi that produce them and offer resistance to the outgrowth of pathogens, the bacterial response has been the development of enzymes that are able to hydrolyse the beta lactam ring and render them ineffective or to transport the compounds outside the cell. It is interesting to note that the evolution of beta lactamases predates the evolution of beta lactams. In addition, it has been suggested that beta lactams may have a role in cell–cell communication. Furthermore it has been suggested that a co-regulatory mechanism links beta lactam resistance, alginate production and virulence factor regulation in Pseudomonas aeruginosa [25].

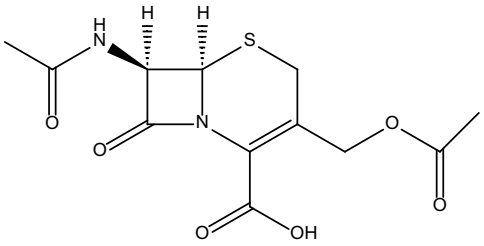
The rate of production of mutations in bacteria is fairly slow and hence the development of beta lactamase or other modes of resistance through the mutation-driven



Beta lactam



Carbapenem



Cephalosporin

Fig. 2.7 Lactam structures

(Darwinian) intrinsic antimicrobial resistance to beta lactams is not the main mode of acquisition of resistance by a bacterial species. This occurs primarily through the transfer of genes for beta lactamases or Penicillin Binding Protein (PBP) between bacteria via plasmids. The widespread use of penicillin only began in about 1945, however, in addition to Fleming's earlier observation alluded to previously, even before the widespread use, in 1940, resistant bacteria were known to employ an enzyme to inactivate penicillin [26]. The resistance mechanism was not purely a result of enzymic inactivation of the beta lactam moiety however—strains with a modified Penicillin Binding Protein (PBP) [27], which has a reduced affinity for the drug were also found (commonly the mode of resistance in Gram positive bacteria). PBPs are transpeptidases which are essential for catalysing the final step in bacterial cell wall synthesis (peptidoglycan) and the binding of penicillin to the active site is the crucial mode of action. The primary target for beta lactam antibiotics is, therefore the PBP. The first plasmid-mediated beta lactamase was discovered in the 1960s [28] and while the introduction of more resistant molecules has helped to increase the efficacy of some beta lactams, the large numbers of evolved beta lactamases mean that any respite from resistance is short lived. Interestingly a number of fungi can produce beta lactamase inhibitors and these can subvert the anti-beta lactam activity of beta lactamase producers. The most well known of these is clavulanic acid from Streptomyces clavuligerus which, like penicillin, is a beta lactam however it does not have intrinsic antibiotic activity but binds to the active site of some beta lactamases, acting as a suicide inhibitor. Given that there are a number of mechanisms

of resistance of bacteria to beta lactams (variation in the PBPs and their affinity for penicillins, reduction in cell permeability to beta lactams, enhanced efflux systems) it appears that there is a degree of biological redundancy however the evolution of beta lactamase inhibitors by fungi suggests that this arsenal of defences against other microbes is both necessary and appropriate. The example of beta lactamases will demonstrate the response of bacteria to what are, to the bacterial species frame of reference, unnatural natural products.

Beta lactamases are able to break down most, if not all beta lactam antibiotics (both naturally occurring (fungally derived) and unnatural (semi-synthetic)) members of the class. In addition, because of the structural diversity of beta lactamases and their different modes of action, enzyme inhibitors are not universally effective. There are 4 classes of beta lactamases. Three of these employ a serine moiety in their active site and are described as PBP-like (A, C and D) while the fourth (B) is a metallo enzyme. The principal mode of resistance to penicillins in Gram positive bacteria is a result of altered affinity of the PBP however in Gram negative bacteria, beta lactamases are a more significant concern. The beta lactamase database [29] contains information regarding known variants of the enzyme. The database is updated weekly and gives information on the biochemical and structural properties of the variants. The evolutionary relationships between beta lactamases has been extensively studied and reviewed [30].

While the three classes of beta lactamase that employ a serine at the active site have different sequence, their mechanism of action is, in essence, the same. They each employ a base to activate the serine which then carries out a nucleophilic attack on the amide producing a covalent acyl enzyme complex which is then hydrolysed by a water molecule which has been activated by a second base. The identity of the bases varies between the different classes of beta lactamase. The mechanistic details of the different beta lactamases have been extensively studied.

The Class B, metallo beta lactamases are related to the metallohydrolase superfamily of enzymes which includes DNA repair and RNA processing enzymes. In their native form, the metal atom on the enzyme is zinc however other metal ions have been found to be suitable [31].

The continued importance of beta lactam antibiotics—and the consequential increase in resistance mechanisms—has fuelled research into the development of new molecular structures and alternative approaches to the control of microbial infections. One of these approaches (the development of strategies around quorum sensing molecules) is discussed below. In summary, the strategies to circumvent anti-microbial resistance (AMR) can be summarised as follows:

a. The discovery of new bioactive molecules with enhanced bioactivity against traditional targets or increased resistance to resistance mechanisms
b. The development of inhibitors of common resistance mechanisms (e.g. the discovery and use of clavulanic acid)
c. The manipulation of microbial ecosystems through the manipulation of factors such as quorum sensing (discussed in Sect. 2.5)
d. Increasing the susceptibility of microbes to existing antibiotics.

This final approach is illustrated in the manipulation of the mevalonate pathway. A study used CRISPR- based gene silencing to investigate genes which, if inhibited or silenced, might increase the susceptibility of bacteria to beta lactam antibiotics. The bioactive molecule chosen was amoxycillin which, in addition to being one of the most widely used antibiotics, is also the first line therapy against *Streptococcus pneumoniae*. Since amoxycillin affects cell wall synthesis (and also cell size and division) the mevalonate pathway was chosen as an ancillary target. Down regulation of the mevalonate pathway genes had been shown to have a significant effect on cell wall biosynthesis [32] and to also upregulate the production of a stress-responsive lactonase in Staphylococcus aureus. This latter effect may have implications for the quorum sensing approach outlined in Sect. 2.5.

It was found that downregulation of the mevalonate pathway led to increased cell size. This was found to be a consequence of a reduced availability of cell wall precursors at the site of cell wall synthesis due to a limitation in the production of an essential lipid carrier, undecaprenyl phosphate. By using a combination of a specific inhibitor of undecaprenyl phosphate synthesis (clomiphene) and amoxycillin, it was found that this combination therapy had the effect of resensitising the amoxycillin resistant *S pneumoniae* to the beta lactam [33]. In addition to providing an effective strategy against other beta lactam resistant bacterial strains, this approach may offer an opportunity for the identification of other primary pathways whose inhibition may increase the susceptibility of pathogenic bacteria to common antibiotics.

The most common bacterial targets for antibiotics are the cell wall, the cell membrane, ribosomal subunits, DNA gyrase, RNA polymerase and folate synthesis. The combination of drugs that are selective for procaryotes together with primary metabolism inhibitors at sub-lethal doses may be of value in selectively inhibiting the growth of bacteria and increasing their susceptibility to antibiotics. The use of CRISPR technology to investigate gene silencing in bacteria may greatly increase the number of available molecular targets [34]. The paradigm is relatively simple. Take a known anti-microbial and investigate genes that are related to the fitness of the gene target (e.g. the mevalonate pathway in the example given previously. Other potential targets could be protein synthesis and folate or the cell membrane). Use CRISPR to silence genes associated with the pathway and determine which is most efficacious in increasing the susceptibility of the bacterium to the antibiotic. Find or design small molecule inhibitors that inhibit the function of the gene product previously identified. Develop combinatorial therapies with enhanced activities. The clinical principle of this approach has already been used in, for example, the use of beta lactams and clavulanic acid and such synergistic therapies have been described [35, 36].

2.5 Approaches to Anti-microbial Resistance (AMR)

Anti-microbial resistance (AMR) has, as we have seen, been an issue from the earliest times of the introduction of anti-microbial agents. Bacteria have had the capacity to overcome the effects of antibiotics even prior to the introduction and widespread use of different antibiotics. The development of strategies to overcome AMR has been likened to an arms race [37] with the inevitable loss of life if we cannot continue to develop effective antimicrobials. The synergistic approach described above may have considerable potential however resistance to any molecular antimicrobial is inevitable, in time. This conventional view of the pharmaceutically-driven progress of AMR solutions has led the World Health Organisation (WHO) and others to think in terms of a conventional drug pipeline which is not being filled sufficiently quickly. The WHO (and others) have lamented the overuse of antibiotics in humans and, in particular, in agriculture where they are used to promote fitness and growth of farm animals. This has led the WHO to a number of actions including the preparation of a report on priority microbes whose control is threatened by AMR [38]. Among Gram negative species, those with the highest scores in terms of their threat were carbapenem-resistant species alongside 3rd generation cephalosporin resistant Enterobacterales. In particular, the problem of finding effective inhibitors of metallo-beta lactamases was noted. Methicillin-resistant *Staphylococcus aureus* (MRSA) and others are already prevalent in some clinical areas. Inhibiting the widespread and inappropriate use of antibiotics together with the development of new, effective therapies were seen as being crucial steps in preventing the spread of multi-drug resistant (MDR) strains. A number of alternative strategies to combat the growth in AMR have been suggested [39]. These include alternatives such as enhanced biosecurity (thereby preventing the spread of infection rather than always fighting an established infection), limiting the widespread use of antibiotics, developing further the use of vaccines against microbes, bacteriophages and phytochemicals. In the case of phytochemicals, the suggested mode of action includes host immune modulation, inhibiting biofilm formation, suppression of virulence factors, effects on bacterial membranes and modulation of antibiotic resistance. There is a lack of mechanistic information for many phytochemicals and, while some that have traditionally been used in diets for many years have a good safety record, the efficacious dose may be much greater than that generally encountered and an assumption of safety should not be applied generally to phytochemicals. Antimicrobial enzymes (for example lysozyme) have been shown to have useful antimicrobial properties however, as with all proteinaceous therapies, stability and delivery of an effective and specific dose are challenges. One interesting approach is the use of enzymes to break down quorum sensing molecules and prevent biofilm formation [40]. This approach will be considered in more detail in Sect. 2.6 where strategies that make use of the ecological properties of quorum sensing in microbial communities. An alternative is to make use of biofilm degrading enzymes directly [41]. By preventing, inhibiting or removing biofilms, pathogens are more susceptible to anti-microbial agents of various types.

2.6 The Microbiome and Quorum Sensing

The microbiome is being more and more appreciated as a key aspect of how microbial communities operate. While the study of individual microbes has been the focus of much of the work of natural product chemists and the reductionist trail from isolated product to biosynthetic pathway to individual reactions and genes has provided considerable understanding of how natural products operate within a single species, it has become increasingly clear that living systems are, on the whole, composed of complex mixtures of species. Indeed the preponderance of a single species (particularly if it is a pathogen) is more frequently deleterious to the host than not therefore using mixed microbial communities (microbiomes) to control pathogens is an attractive option. In addition, modulation rather than destroying microbes is less likely to result in the build up of resistance to specific chemical entities.

The study of microbial communities and interdependencies is replete with unexpected consequences and opportunities for interventions at various junctures in the control systems. In order to develop an understanding of how to modulate the microbiome, it is necessary to have detailed knowledge of which species are present and how they are interrelated. One of the major sites of occupation of the human microbiome is in the gastrointestinal tract and, while the GI microbiome has significant systemic effects, it is difficult to study directly. Examination of the characteristics of faecal material can yield some information however the numbers and properties of bacteria that are shed are different from those which remain in situ. Indeed the human gut microbiome is an important source of AMR microbes whose antimicrobial properties can be transferred from commensal bacteria into acquired pathogens and the increased population of these (together with an acquired resistance to antibiotics) can provoke a number of physiological reactions in the host including inflammation and other deleterious effects [42]. When faced with such complex and interdependent ecosystems, any attempt to break it down into its constituent parts and to the reassemble it to get a picture of the complete system is of only limited value. One way of considering this system is to examine the interactions within and between species in the microbiome since it is the prevalence (or absence) of such signal interactions that will help to describe the totality of the microbiome and its properties. By understanding both the species present and the genomes (and metagenomes) present, the key control points and their evolution can be determined and manipulated. For example, the existence of AMR and antibiotic resistance genes is not a new phenomenon. Bacteria that possess antimicrobial resistance have been found in samples isolated for millions of years [43]. A range of resistance mechanisms were found including likely beta lactamase activity. Many of the bacteria found in the environment that have the capacity for antibiotic resistance are non-pathogenic strains and it has been suggested that transfer of resistance from environmental strains to pathogenic strains is a viable means by which AMR will spread to organisms of clinical significance. Since AMR appears in ancient strains of bacteria which significantly predates the discovery and widespread use of antibiotics by humans, it seems not unreasonable to suppose that the proteins involved in AMR have a significance

beyond AMR itself. By understanding more about the AMR-related genes in the environment, even those for which antibiotics have not been developed, it might be possible to reverse engineer these activities to produce antibiotics which occupy chemical white space that has not been hitherto considered in the search for molecules that may prove to be resilient to AMR and also novel antibiotics, however, as outlined below, this is likely to be of only limited use.

The evidence that AMR is present in the environment even when antibiotics themselves are not extensively used, suggests that the development of novel chemical structures that will have sustained resistance to antibiotics is unlikely, in the long term, to succeed. The potential for the development of AMR exists even before an antimicrobial substance has been discovered. An alternative approach is to make use of microbial communication (quorum sensing) as a way of maintaining a healthy population and preventing the outgrowth of pathogens. This is a multifactorial approach and does not depend on finding a single molecule or a single lethal attribute (the 'silver bullet strategy). Given that the constituents of AMR have been found to greatly predate the widespread use of antibiotics, it would not be unexpected if communities of microbes had also developed ways to modulate microbial growth to prevent pathogen outgrowth, in the competition for resources. This is the basis of quorum sensing and it has the potential to change how antimicrobials and AMR are considered. It requires, however, a paradigm shift away from over 100 years of thought about the selective destruction of pathogenic bacteria which have the potential to cause disease.

Quorum sensing is a cell to cell communication system in which extracellular signalling molecules (autoinducers) facilitate the cellular population's response to changes in their environment—in particular to changes in microbial population dynamics. The concentration of the autoinducers increases in response to increased cell density. The nature and role of quorum sensing has been extensively reviewed both in bacteria [44] and in fungi [45]. One of the most common groups of quorum sensing molecules, which are used throughout a range of bacterial species, are N-acylhomoserine lactones. These molecules can also affect plant growth and mimics of them have been found to be produced by plants and to affect their interactions with microbial pathogens [46].

The impact of quorum sensing molecules extends to the growth of bacteria and other microbial species and also an impact upon specific features such as biofilm formation and filamentous control. In addition, in bacteria, functions such as growth, multiplication, virulence and antibiotic resistance as well as the influence of external factors are all affected by quorum sensing molecules [47]. The use of quorum sensing manipulation is seen as a viable strategy for the control of pathogenic microbes however it is important to recognise that a key aspect of this is the recognition that living systems are not in a state of balanced equilibrium and that retention of homeostasis and resistance to perturbation is neither desirable nor achievable. This is a function of living systems and we shall explore the consequences of this further.

While thermodynamics (and the Second Law, in particular) states that the entropy of closed systems increases with time as systems move from a less likely state to a more likely state, living systems adopt greater and greater levels of complexity and order. This apparent reversal of the Second Law of Thermodynamics has a number

of consequences. The first is that the whole universe (including living systems) will only reach a state of equilibrium when maximum entropy is achieved. Until then, systems will continue to evolve towards increasing entropy. The implication in this is that, if living systems are able to evolve towards more complex and ordered systems (with consequential lowered entropy) then they exist contrary to the Second Law of Thermodynamics. They manage this (and all the processes of life) by absorbing free energy and dissipating heat, thereby increasing the entropy of their environment as they lower their own entropy [48]. While the rate of consumption of energy can be calculated by measuring the energy consumption of each of the components of a living system, in practice this is not possible however both energy consumption and the increase in entropy can be estimated with knowledge of a limited number of variables. A crucial aspect of living systems is that they are in a state of non-equilibrium and this dictates both the properties of the system and its interactions [49].

In the case of quorum sensing, the manipulation of the environment to create optimal states is both an ongoing and a responsive process. If antibiotics are added to the system then the perturbation created will define a new non-equilibrium state and, from that state, consequential states—including the development of AMR will necessarily follow. The addition of antimicrobials to a living system will, therefore inevitably lead to a number of outcomes including the build up of AMR. There will also be unknown consequences and these will be both inevitable and unpredictable. Quorum sensing is one of these adaptations.

Quorum sensing can be considered, therefore, to be an adjustment mechanism in which cells respond to their environment. In the case of bacteria, cells synthesise small molecules (the most common of these, as noted above, are autoinducers and in some cases take the chemical form of lactones). The more bacterial cells that are producing the quorum sensing molecules, the more the molecules accumulate. When the number of these molecules reaches a certain level then a number of genes are transcribed in the whole population of cells. Quorum sensing molecules, therefore, are both an indirect measure of cell density and a mechanism whereby populations respond to cell density. The link with thermodynamics is that quorum sensing creates an increase in order and complexity and a reduction in entropy. Quorum sensing molecules can, therefore, be considered a way that a microbial system responds to a perturbation by increasing the degree of order and making the system more resilient and less disordered. The use of the production of quorum sensing molecules has not been used to measure cell density directly and the linearity of the response has not been established however it would seem to be a proxy both for cell density and the level of entropy of the system. The genes that are expressed are, by implication, those which benefit cell populations rather than individual cells and the system must be sufficiently robust to be able to cope with general perturbations that occur in living systems [50].

There have been a large number of reviews of quorum sensing and how it might influence microbial behaviour. Some of these have been very detailed and comprehensive while others have concentrated upon the mechanistic details of how quorum

sensing works and the effect it has upon the microbial system (including the thermodynamic considerations mentioned above). From the standpoint of natural products, there are two key aspects related to quorum sensing and both of them are a consequence of the reductionist nature of chemistry. These can be considered as follows:

a. To what extent can natural products affect the expression of the genes involved in the biosynthesis of quorum sensing molecules?
b. How might natural products affect the activity of quorum sensing molecules—either through a direct effect on the molecules themselves or affecting their persistence in living systems?

The link between quorum sensing and AMR has been reviewed [51]. In addition to the production of beta lactamases or modifications to the PBP (mentioned previously) the main ways that bacteria exhibit AMR is through the effective removal of drugs (an efflux pump mechanism) or the formation of impenetrable biofilms. Bacteria may employ more than one mechanism to enhance AMR and these can work synergistically to limit the effects of both individual antibiotics and cocktails of drugs. In the case of biofilms, multiple mechanisms operate. The biofilm itself can prevent antibiotic penetration to the pathogen thereby preventing the direct inhibition of a key part of the pathogen's metabolism (e.g. cell wall biosynthesis). In addition, bacteria in different parts of the biofilm will experience different flows of nutrients which can also affect the efficacy of antibiotics which are targeted towards different physiological processes in the bacterial life cycle. Biofilms may also prevent the diffusion of quorum sensing molecules and facilitate the transition of bacteria from an adhered to a planktonic state thereby promoting the spread of bacteria to different parts of the biological system.

One group of quorum sensing molecules are autoinducing oligopeptides (generally used by Gram positive bacteria). These molecules cause upregulation of virulence factors thereby increasing the pathogenicity of bacteria. In the case of Gram-negative bacteria, the main autoinducers are lactones derived from fatty acids. These vary in the nature of the acyl group and different lactones are produced by different species. They are, therefore, a mechanism for intraspecies communication.

The idea that quorum sensing may be a useful strategy to prevent pathogen growth has been suggested [52]. A number of approaches have been suggested including inhibition of quorum sensing signalling pathways, blocking quorum sensing molecule production (particularly lactones in Gram-negative bacteria) and targeting bacterial quorum sensing receptors. There are a number of natural products that are able to inhibit the signalling system of bacteria. Phenolics including flavonoids can inhibit a number of consequences of quorum sensing activation including preventing biofilm formation. One other approach is the secretion of a lactonase enzyme by certain microbes. Lactonase acts by cleaving the ester bond of the homoserine ring thereby rendering the molecule ineffective. In addition to being excreted by microorganisms, lactonase activity has also been found in mammalian tissues. The lactonase activity is found on the class of paraoxonases (PON), of which there are three members. The use of paraoxonases to prevent biofilm formation through inhibition of homoserine

lactones has been suggested [53]. Flavonoids and polyphenols appear to induce PON2 gene expression and this may contribute to the prevention of pathogen growth and thereby limiting pathogen-induced inflammation [54]. The possibility of manipulation of lactonase activity to inhibit bacterial (particularly Gram-negative) growth using natural products could be further investigated. The potential role of different lactonases may lead to the development of optimal enzyme mixtures however, given the stability issues associated with the use of enzymes in a therapeutic context, developing strategies around optimisation of lactonase production by microbes in situ in the microbiome using natural products and thereby optimising the prevention of pathogen outgrowth may be a more promising avenue.

While the escalation of AMR and the difficulty in obtaining antimicrobial agents against which resistance will not easily be attained by pathogens is a clear and present threat, the development of novel strategies based around the alternative activities of natural products has been noted. As previously described, the development of multiple strategies that focus upon numbers of microbial cellular pathways, as was done with mevalonate and ampicillin may also provide a viable alternative strategy. The use of CRISPR-based gene editing will be of value in optimising the inhibition of specific steps in a biosynthetic pathway and therefore determining which of them, in conjunction with existing or novel antimicrobials, is likely to meet with the greatest chance of success. The second approach is to move away from bacterial cell death as an ultimate objective and to focus on maintaining a healthy microbiome and preventing pathogen outgrowth by manipulation of the host response (e.g. using paraoxonase) or microbial production of enzymes such as lactonase or acylase—by genetic manipulation of gene promoter regions, for example. Such approaches will be underpinned by an increasing knowledge of microbial communication, how it might be disrupted or controlled and the relationships between different members of the microbial community.

References

1. Nicolaou KC, Edmonds DJ, Bulger PG (2006) Cascade reactions in total synthesis. Angew Chem Int Ed Engl 45(43):7134–7186. https://doi.org/10.1002/anie.200601872. PMID: 17075967
2. Hodgkin DC (1949) The X-ray analysis of the structure of pencillin. Adv Sci 6:85–89
3. Howard J (2003) Dorothy Hodgkin and her contributions to biochemistry. Nat Rev Mol Cell Biol 4:891–896. https://doi.org/10.1038/nrm1243
4. Curtis R, Jones J (2007) Robert Robinson and penicillin: an unnoticed document in the saga of its structure. J Pept Sci: Official Publ Eur Pept Soc 13(12):769–775
5. Yocum RR, Rasmussen JR, Strominger JL (1980) The mechanism of action of penicillin. Penicillin acylates the active site of Bacillus stearothermophilus D-alanine carboxypeptidase. J Biol Chem 255(9):3977–3986. PMID: 7372662
6. Gadamer J, Über die Bestandteile des schwarzen und des weissen Senfsamens. Archiv der Pharmazie 235

7. Bergstrand LH, Cardenas E, Holert J, Van Hamme JD, Mohn WW (2016) Erratum for Bergstrand et al., Delineation of steroid-degrading microorganisms through comparative genomic analysis. mBio 7(4):e00865–e0086516. https://doi.org/10.1128/mBio.00865-16. Erratum for: MBio. 2016;7(2):e00166. PMID: 27460794; PMCID: PMC4981713
8. Nicolaou KC, Snyder SA (2005) Chasing molecules that were never there: misassigned natural products and the role of chemical synthesis in modern structure elucidation. Angew Chem Int Ed Engl 44(7):1012–1044. https://doi.org/10.1002/anie.200460864. Erratum in: Angew Chem Int Ed Engl. 2005 Mar 29;44(14):2050. PMID: 15688428
9. Hanley AB, Belton PS, Fenwick GR, Janes NF (1985) Ring oxygenated indole glucosinolates of Brassica species. Phytochem 24:598
10. Tantillo DJ (2018) Questions in natural products synthesis research that can (and cannot) be answered using computational chemistry. Chem Soc Rev 47(21):7845–7850. https://doi.org/10.1039/c8cs00298c.PMID:29900461;PMCID:PMC6205925
11. Ban TA (2006) The role of serendipity in drug discovery. Dialogues Clin Neurosci 8(3):335–344. https://doi.org/10.31887/DCNS.2006.8.3/tban.PMID:17117615;PMCID:PMC3181823
12. Yaqub O (2016) Serendipity: towards a taxonomy and a theory. SWPS 2016–17. Available at SSRN: https://ssrn.com/abstract=2841236 or https://doi.org/10.2139/ssrn.2841236
13. Strebhardt K, Ullrich A (2008) Paul Ehrlich's magic bullet concept: 100 years of progress. Nat Rev Cancer 8:473–480. https://doi.org/10.1038/nrc2394
14. Hetzler BE, Trauner D, Lawrence AL (2022) Natural product anticipation through synthesis. Nat Rev Chem 6:170–181. https://doi.org/10.1038/s41570-021-00345-7
15. Goerig M, Schulte am Esch J (1991) Friedrich Wilhelm Adam Sertürner--dem Entdecker des Morphins zum 150. Todestag [Friedrich Wilhelm Adam Sertürner—the discoverer of morphine]. Anasthesiol Intensivmed Notfallmed Schmerzther 26(8):492–498. German. https://doi.org/10.1055/s-2007-1000624. PMID: 1786314
16. Fenniri H (2000) Combinatorial chemistry: a practical approach, Online edn. Oxford Academic, Oxford. https://doi.org/10.1093/oso/9780199637546.001.0001. Accessed 20 May 2024
17. Bashiruddin NK, Suga H (2015) Construction and screening of vast libraries of natural product-like macrocyclic peptides using in vitro display technologies. Curr Opin Chem Biol 24:131–138. https://doi.org/10.1016/j.cbpa.2014.11.011. Epub 2014 Dec 5 PMID: 25483262
18. Walsh G, Walsh E (2022) Biopharmaceutical benchmarks 2022. Nat Biotechnol 40:1722–1760. https://doi.org/10.1038/s41587-022-01582-x
19. Elyashberg ME, Serov VV, Gribov LA (1987) Artificial intelligence systems for molecular spectral analysis. Talanta 34(1):21–30. https://doi.org/10.1016/0039-9140(87)80005-x. PMID: 18964258
20. Elyashberg M, Novitskiy IM, Bates RW, Kutateladze AG, Williams CM (2022) Reassignment of improbable natural products identified through chemical principle screening. Eur J Org Chem, e202200572
21. Grigalunas M, Brakmann S, Waldmann H (2022) Chemical evolution of natural product structure. J Am Chem Soc 144(8):3314–3329. https://doi.org/10.1021/jacs.1c11270. Epub 2022 Feb 21. PMID: 35188375; PMCID: PMC8895405
22. Bush K (2012) Evolution of β-Lactamases: past, present, and future. In: Dougherty T, Pucci M (eds) Antibiotic discovery and development. Springer, Boston, MA. https://doi.org/10.1007/978-1-4614-1400-1_12
23. Fleming A (1999) Penicillin [Nobel Prize Lecture, December 11, 1945]. Nobel lectures in physiology or medicine, 1942–1962. World Scientific Publishing Company, Singapore
24. Podolsky SH (2018) The evolving response to antibiotic resistance (1945–2018). Palgrave Commun 4:124. https://doi.org/10.1057/s41599-018-0181-x
25. Balasubramanian D, Kong KF, Jayawardena SR, Leal SM, Sautter RT, Mathee K (2011) Co-regulation of {beta}-lactam resistance, alginate production and quorum sensing in Pseudomonas aeruginosa. J Med Microbiol 60(Pt 2):147–156. https://doi.org/10.1099/jmm.0.021600-0. Epub 2010 Oct 21. Erratum in: J Med Microbiol. 2011 May;60(Pt 5):696–697. PMID: 20965918; PMCID: PMC3081088

26. Abraham EP, Chain E (1988) An enzyme from bacteria able to destroy penicillin. 1940. Rev Infect Dis 10(4):677–678
27. Zapun A, Contreras-Martel C, Vernet T (2008) Penicillin-binding proteins and β-lactam resistance. FEMS Microbiol Rev 32(2):361–385. https://doi.org/10.1111/j.1574-6976.2007.00095.x
28. Bradford PA (2001) Extended-spectrum beta-lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. Clin Microbiol Rev 14(4):933–951, Table of Contents. https://doi.org/10.1128/CMR.14.4.933-951.2001. PMID: 11585791; PMCID: PMC89009.
29. Naas T, Oueslati S, Bonnin RA, Dabos ML, Zavala A, Dortet L, Retailleau P, Iorga BI (2017) Beta-lactamase database (BLDB)—structure and function. J Enzyme Inhib Med Chem 32(1):917–919. https://doi.org/10.1080/14756366.2017.1344235.PMID:28719998;PMCID:PMC6445328
30. Fröhlich C, Chen JZ, Gholipour S, Erdogan AN, Tokuriki N (2021) Evolution of β-lactamases and enzyme promiscuity. Protein Eng Des Sel 34:gzab013
31. Badarau A, Page MI (2006) The variation of catalytic efficiency of Bacillus cereus metallo-beta-lactamase with different active site metal ions. Biochemistry 45(35):10654–10666. https://doi.org/10.1021/bi0609341. PMID: 16939217
32. Balibar CJ, Shen X, Tao J (2009) The mevalonate pathway of Staphylococcus aureus. J Bacteriol 191(3):851–861. https://doi.org/10.1128/JB.01357-08. Epub 2008 Nov 21. PMID: 19028897; PMCID: PMC2632080
33. Dewachter L, Dénéréaz J, Liu X, de Bakker V, Costa C, Baldry M, Sirard JC, Veening JW (2022) Amoxicillin-resistant Streptococcus pneumoniae can be resensitized by targeting the mevalonate pathway as indicated by sCRilecs-seq. Elife 24(11):e75607. https://doi.org/10.7554/eLife.75607.PMID:35748540;PMCID:PMC9363119
34. Zhang R, Xu W, Shao S, Wang Q (2021) Gene silencing through CRISPR interference in bacteria: current advances and future prospects. Front Microbiol 31(12):635227. https://doi.org/10.3389/fmicb.2021.635227.PMID:33868193;PMCID:PMC8044314
35. Klein JO (2003) Amoxicillin/clavulanate for infections in infants and children: past, present and future. Pediatr Infect Dis J 22(8 Suppl):S139–S148. https://doi.org/10.1097/00006454-200308001-00005. PMID: 14567000
36. Bush K, Bradford PA (2016) β-lactams and β-lactamase inhibitors: an overview. Cold Spring Harb Perspect Med 6(8):a025247. https://doi.org/10.1101/cshperspect.a025247.PMID:27329032;PMCID:PMC4968164
37. Iskandar K, Murugaiyan J, Hammoudi Halat D, Hage SE, Chibabhai V, Adukkadukkam S, Roques C, Molinier L, Salameh P, Van Dongen M (2022) Antibiotic discovery and resistance: the chase and the race. Antibiotics (Basel) 11(2):182. https://doi.org/10.3390/antibiotics11020182.PMID:35203785;PMCID:PMC8868473
38. WHO Bacterial Priority Pathogens List (2024) 2024: bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance. World Health Organization, Geneva. Licence: CC BY-NC-SA 3.0 IGO
39. Gupta R, Sharma S (2022) Role of alternatives to antibiotics in mitigating the antimicrobial resistance crisis. Indian J Med Res 156(3):464–477. https://doi.org/10.4103/ijmr.IJMR_3514_20. PMID: 36751744; PMCID: PMC10101360
40. Thallinger, B et al (2013) Antimicrobial enzymes: an emerging strategy to fight microbes and microbial biofilms. Biotechnol J 8(1):97–109
41. Ramakrishnan R et al (2022) Enzymatic dispersion of biofilms: an emerging biocatalytic avenue to combat biofilm-mediated microbial infections. J Biol Chem 298(9)
42. Brinkac L, Voorhies A, Gomez A, Nelson KE (2017) The threat of antimicrobial resistance on the human microbiome. Microb Ecol 74(4):1001–1008. https://doi.org/10.1007/s00248-017-0985-z. Epub 2017 May 11. PMID: 28492988; PMCID: PMC5654679
43. Bhullar K, Waglechner N, Pawlowski A, Koteva K, Banks ED, Johnston MD, Barton HA, Wright GD (2012) Antibiotic resistance is prevalent in an isolated cave microbiome. PLoS One 7(4):e34953. https://doi.org/10.1371/journal.pone.0034953. Epub 2012 Apr 11. PMID: 22509370; PMCID: PMC3324550

44. Mukherjee S, Bassler BL (2019) Bacterial quorum sensing in complex and dynamically changing environments. Nat Rev Microbiol 17:371–382. https://doi.org/10.1038/s41579-019-0186-5
45. Tian X, Ding H, Ke W, Wang L (2021) Quorum sensing in fungal species. Annu Rev Microbiol 8(75):449–469. https://doi.org/10.1146/annurev-micro-060321-045510. Epub 2021 Aug 4 PMID: 34348030
46. Shrestha A, Grimm M, Ojiro I, Krumwiede J, Schikora A (2020) Impact of quorum sensing molecules on plant growth and immune system. Front Microbiol 16(11):1545. https://doi.org/10.3389/fmicb.2020.01545.PMID:32765447;PMCID:PMC7378388
47. Li Z, Nair SK (2012) Quorum sensing: how bacteria can coordinate activity and synchronize their response to external signals? Protein Sci 21(10):1403–1417. https://doi.org/10.1002/pro.2132. Epub 2012 Aug 21
48. Skinner DJ, Dunkel J (2021) Improved bounds on entropy production in living systems. Proc Natl Acad Sci U S A 118(18):e2024300118. https://doi.org/10.1073/pnas.2024300118.PMID:33906948;PMCID:PMC8106300
49. Aristov VV, Karnaukhov AV, Levchenko VF et al (2022) Entropy and information in the description of biosystems. Biophysics 67:593–599. https://doi.org/10.1134/S0006350922040029
50. Pai A, Tanouchi Y, You L (2012) Optimality and robustness in quorum sensing (QS)-mediated regulation of a costly public good enzyme. Proc Natl Acad Sci 109:19810–19815
51. Zhao X, Yu Z, Ding T (2020) Quorum-sensing regulation of antimicrobial resistance in bacteria. Microorganisms 8(3):425. https://doi.org/10.3390/microorganisms8030425.PMID:32192182;PMCID:PMC7143945
52. Naga NG, El-Badan DE, Ghanem KM et al (2023) It is the time for quorum sensing inhibition as alternative strategy of antimicrobial therapy. Cell Commun Signal 21:133. https://doi.org/10.1186/s12964-023-01154-9
53. Camps J, Pujol I, Ballester F, Joven J, Simó JM (2011) Paraoxonases as potential antibiofilm agents: their relationship with quorum-sensing signals in Gram-negative bacteria. Antimicrob Agents Chemother 55(4):1325–1331. https://doi.org/10.1128/AAC.01502-10. Epub 2011 Jan 3. PMID: 21199929; PMCID: PMC3067127
54. Parween F, Gupta RD (2022) Insights into the role of paraoxonase 2 in human pathophysiology. J Biosci 47(1):4. https://doi.org/10.1007/s12038-021-00234-7.PMID:35092416;PMCID:PMC8721187

Chapter 3
Naturally-Occurring Sulphur Compounds

Abstract The importance of sulphur-containing natural products in primary and secondary metabolism. The role of hydrogen sulphide as a gasotransmitter and its uptake by living systems. The relevance of mustard oil glycosides (glucosinolates) and their enzymic breakdown products. Sulphur compounds of the genus *Allium*. Different approaches to the production of multiple sulphur compounds by *Allium* and *Brassica* species and its implications. The role of sulphur-containing natural products in chronic disease and their impact upon complex biological pathways. Inductive reasoning and natural products. Other potential bioactive compounds in *Brassica* species. Glucosinolates and their breakdown products and plant pathogenic fungal disease. Dithiocarbamates as potential storage forms of hydrogen sulphide. The importance of extra-oral bitter receptors. Natural products, reductionism, biological activity and statins. The weakness of nutritional epidemiology in the investigation of the effects of natural products.

Keywords Glucosinolates · Allium · Brassica · Chronic disease

Studies of the chemistry of naturally-occurring sulphur-containing compounds has made a significant contribution to knowledge and the development of ideas related to other areas of organic and biological chemistry. In the field of natural product chemistry, some atoms (and functional groups) occur more often, in more variations and are therefore subject to greater study than others. Sulphur metabolism in natural products is complex and multivariate. Some of this complexity is a result of the reactivity of the sulphur atom and the different oxidation states it can occupy and it is also in part due to the bonds that sulphur can form. While sulphur occurs in primary metabolism (particularly the amino acids methionine and cysteine) it is also crucial to biological characteristics as diverse as protein structure, enzyme active sites, in the gasotransmitter hydrogen sulphide and even in DNA [1]. In addition, sulphur occurs in important biological molecules such as thiamine, biotin and lipoic acid. In addition sulphur secondary metabolites such as penicillins and cephalosporins (Chap. 2), polysulphides, glucosinolates and the sulphur compounds of plants of the genus Allium [2] have been extensively studied.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_3

Sulphur was recognised in alchemy. Indeed, Paracelsus identified sulphur as one of the three primes (*tria prima*) of alchemy. While this may seem a rather outdated concept and alchemy as a whole is more of a belief system rather than proven empirical facts, it nonetheless recognised a fundamental principle of chemistry—substances can react together to give new products that are not simply the combination of the starting materials. This is an important concept and lies at the core of chemistry. In the alchemical triumvirate of fire, air and earth, sulphur represented fire. In addition, the sulphur-mercury ratio in a substance was thought to define its properties so a metal could be converted into any other metal by the addition or removal of sulphur. Using this principle, it makes sense that lead, for example, can be converted into gold. The idea that materials could be broken apart and then recombined through heating and coagulation and this was a purification cycle was fundamental to alchemical practice.

In this non-comprehensive review of naturally occurring sulphur compounds, the following will be examined:

1. Sulphur-containing primary metabolites
2. Mustard oil glycosides (glucosinolates)—their chemical and biological relevance
3. Sulphur compounds in Allium and Brassica species—important similarities and differences#
4. Sulphur compounds in chronic disease and their effect on quality of life
5. Sulphur compounds in the oral cavity and in the microbiome
6. Sulphur compounds and new approaches to disease and wellbeing. Sulphur and complex biological systems.

By using these as illustrative examples, the systems and mechanisms that underpin the study of natural products will be revealed.

3.1 Sulphur-Containing Primary Metabolites

The sulphur cycle is of great importance in nature. Of particular relevance in living systems is the synthesis of cysteine from serine via O-acetylserine with the addition of hydrogen sulphide. Another important starting point is homoserine which is converted, via O-acetylhomoserine to homocysteine, methionine and S-adenosyl methionine (SAM). Details of these interrelationships, which can be complex and vary between species, can be found in any biochemistry textbook or review [3]. There is a requirement for a cellular store of hydrogen sulphide (or at least the potential to quickly make it), not least for the synthesis of cysteine (mentioned above) and it is also recognised as being of prime importance to living systems [4]. In addition to its role in the sulphur cycle, hydrogen sulphide itself has been hypothesised to be a key component of the redox status of cells and has an important role in the cellular resistance to oxidative stress. Hydrogen sulphide is now generally recognised as the third gasotransmitter (alongside nitric oxide and carbon monoxide). Interestingly all three molecules are toxic and therefore preferably generated in situ when required

rather than stored within the cell. In the case of hydrogen sulfide, in mammalian tissue, the compound is cycled from cysteine. Three primary enzymes are involved in this process—cystathionine beta-synthase (CBS), cystathionine gamma-lyase (CSE), and 3-mercaptopyruvate sulfurtransferase (3-MST). The three are distributed and expressed to different extents in different tissues. While the role of all three of the commonly known gasotransmitters has been extensively studied, other potential (based on structure) gasotransmitters are less well characterised. For example, the possibility that hydrogen cyanide could function as a gasotransmitter has been noted [5], as has sulphur dioxide which is postulated to regulate vascular tone and vascular smooth cell proliferation [6]. The importance of sulphur in various oxidation states is, therefore significant as a component of gasotransmission. In addition, hydrogen sulphide can also be stored as sulphane sulphur in specific cysteine residues of some proteins. While the role of hydrogen sulphide has been extensively studied in mammalian systems, it also has a functional role in plants [7] and is involved in plant growth and development. Interestingly while hydrogen sulphide was first discovered as a product of the bacterial breakdown of living material [8], the role of hydrogen sulphide in prokaryotes is less well studied although it has been suggested to be a generic method of protecting bacteria against antibiotics [9].

In the case of all gasotransmitters, as noted previously, the importance of endogenous production has been recognised. In the specific example of hydrogen sulphide, dithiocarbamates are potential sources of the gas in vivo. Studies have been carried out on the fungicide NABAM (ethylene bis dithiocarbamate) (Fig. 3.1) and it has been shown to release both hydrogen sulphide and carbon disulphide [10].

The endogenous production of hydrogen sulphide in mammalian tissue, as noted above, can also be mediated by the enzymes cystathionine lyase, cystathionine synthase and 3-mercaptopyruvate sulphotransferase [11]. Since the enzymes are produced in different tissues, organ-specific modulation of hydrogen sulphide production may be possible through manipulation of expression or direct enzyme inhibition. The biological effects of hydrogen sulphide have been extensively reviewed and a number of these reviews are freely available. In the case of modulation of inflammation, the importance of natural products in the area of hydrogen sulphide production has been recognised. It has been suggested that the biological properties of a range of natural products is a function of their capacity to release hydrogen sulphide in vivo [12]. In addition researchers have investigated the potential of a range of small molecule hydrogen sulphide donors in order to optimise their bioavailability and pharmacological properties [13]. Included in the list of natural

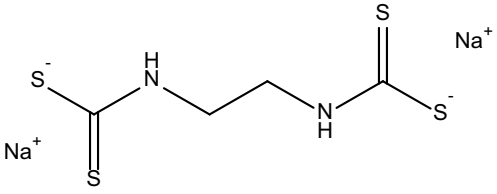


Fig. 3.1 The structure of NABAM (ethylene bis dithiocarbamate)

products that can release hydrogen sulphide are polysulphides produced from Allium species plants (particularly garlic), cyclic polysulphides, epidithiodioxopiperazines which are fungal metabolites, enediyne-containing trisulphides, leinamycin from Streptomyces and a range of other small molecule, sulphur-containing natural products [13]. As further studies are undertaken, the importance of hydrogen sulphide is being increasingly realised and this will lead to significant increases in the application of pro-drugs and other pharmacological sources of the molecule [14]. The formation of dithiocarbamates by the addition of isothiocyanates (commonly found in Brassica species—see Sect. 3.2) to endogenous (bacterially produced) thiols has been suggested as a breath freshening strategy. This reaction also suggests that the reaction between isothiocyanates and thiols can occur under physiological conditions [15] and this may have consequences for the role of diet in health.

3.2 Mustard Oil Glycosides (Glucosinolates)—Chemical Properties

One of the most extensively studied group of sulphur-containing natural products are the mustard oil glycosides (glucosinolates). These compounds first came to prominence through the isolation of the corresponding isothiocyanates which are formed from the glycoside after enzymic hydrolysis of the glucose moiety, rearrangement of the aglycone and expulsion of a sulphate group (Fig. 3.2).

Isothiocyanates (albeit as a component of a mixture) were originally referred to as 'oil of mustard' in a paper in 1892 [16] and it was initially called allyl thiocarbamide. The use of mustard in food and as a medicinal predates their chemical investigation. In antiquity, mustard was used in the treatment of rheumatism, gout, arthritis, colds, and fever. By the eighteenth century, medicinal applications of mustard included use as an emetic, as a digestive, a diuretic, a stimulant, or even as an anti-emetic. By the nineteenth century, it had evolved into an anti-viral and in 1872, advertisements were placed in Household Handbooks for cold remedies such as "Cooper's Sinapine Tissue (the Latin name for white mustard is *Sinapis alba*) [17]. Mustard oil was, therefore known and was prepared by grinding and soaking mustard seeds in water and then distilling the mixture to give an essential oil composed largely of allyl isothiocyanate. The parent glucosinolate, from which allyl isothiocyanate is derived (sinigrin) was first isolated by Bussy [18] with a structure being proposed by Gadamer in 1897 [19, 20] The structure proposed by Gadamer, while apparently logical in that is contained a NH–C–S whose further elaboration into an isothiocyanate was both facile and obvious, was found to be erroneous and it was almost a century later that the correct structure was proposed by Ettlinger and Lundeen [21] and then proven by x-ray crystallography and synthesis [22] (Their structure required that the aglycone undergo a rearrangement reaction (analogous to a Lossen rearrangement—conversion of a hydroxamate ester to an isocyanate) to give the isothiocyanate [23].

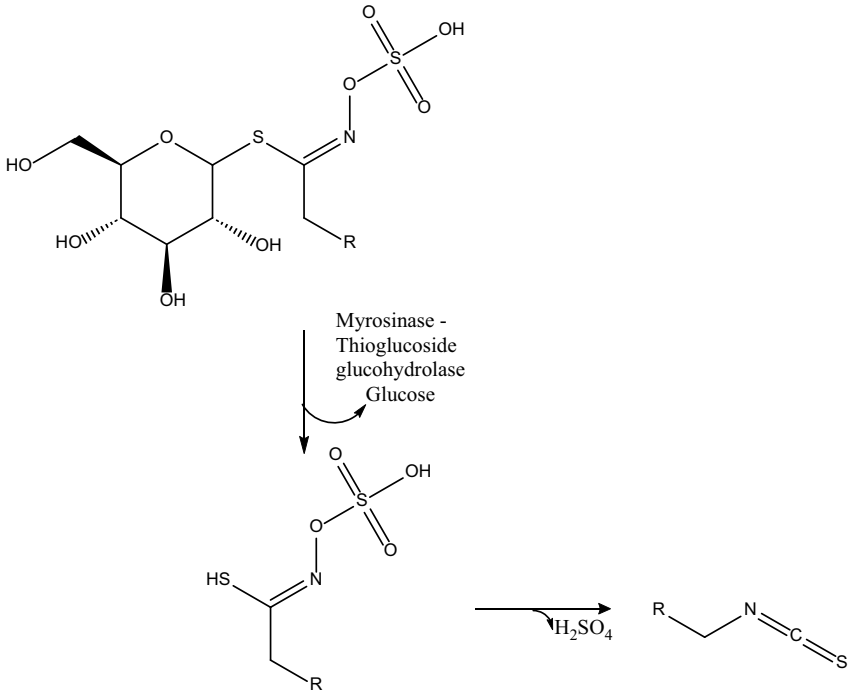


Fig. 3.2 Production of isothiocyanates from glucosinolates

The initial name given to sinigrin was potassium myronate. This is a reflection of the cellular structure of plants that contain glucosinolates. The glucosinolates accumulate in the vacuole of the S cells of plants. In order to ensure that volatile compounds are not formed continually in situ, the thioglucoside glucohydrolase enzyme that hydrolyses the thioglucoside bond of the glucosinolate to give the aglycone which, as noted previously undergoes a spontaneous rearrangement to give the isothiocyanate, is located in a different type of cell (myrosin cells) which are also located in the vacuole but physically separated from the substrate glucosinolate until tissue disruption occurs which leads to release of the enzyme from the myrosin cells. Upon tissue damage, then, the enzyme comes into contact with the glucosinolate and hydrolysis of the thioglucose bond occurs producing the pungent principles noted previously (particularly isothiocyanates) which are potent anti-feedants and part of the stress response of certain plant species [24]. Since the enzyme was located in the myrosin cells, it was given the trivial name of myrosinase. The myrosin cell is a particular type of idioblast and is common in the *Capparales*. Bussy was the first to recognise that, to produce mustard oil, it was necessary to mix a proteinaceous component with "l'acide myronique" (sinigrin). This characteristic also explains why mustard seed must be crushed in water in order to facilitate the hydrolytic reaction which allows for the liberation of the oil which can then be isolated by distillation.

Bussy was the first to demonstrate that the formation of essence of mustard (mustard oil) required water. He also found that production of the oil did not take place if hot water, acid or alcohol were used to mix the essential components and these caused the proteinaceous compound to coagulate (as had been previously observed for albumin). This presumably arises from the denaturation of the myrosinase enzyme as evidenced by its coagulation.

A number of glucosinolates were first identified by steam distillation of aqueous suspensions of plant material. By 1961, a total of 31 mustard oils had been recognised. Many of these were characterised by the group of Anders Kjaer in Copenhagen. The isothiocyanates were isolated by paper chromatography and this facilitated their identification. In a number of cases, Kjaer was able to compare the naturally isolated compound (or the thiourea derivative) with a synthetic version. Significant numbers of glucosinolates have now been isolated, characterised and, in many cases, synthesised. Over 130 glucosinolates have been isolated and all are derived from amino acids after a range of modifications (elongation, hydroxylation etc.). One particular group which has excited a great deal of interest are those derived from the amino acid tryptophan. These compounds, first isolated in 1961 [25] are anomalous in that they do not give rise to a stable isothiocyanate. The initially formed, intermediate isothiocyanate is immediately hydrolysed by water to give indole-3 carbinol with the expulsion of thiocyanate ion (Fig. 3.3).

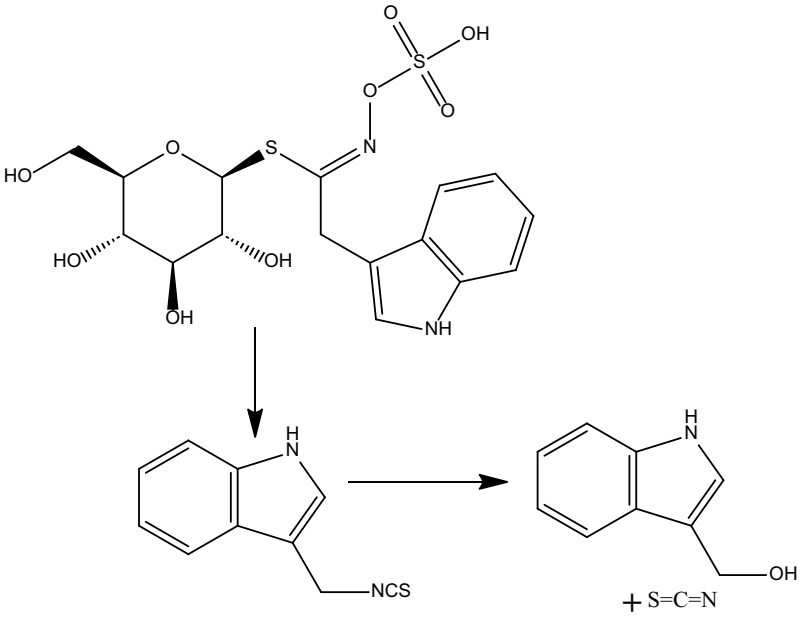


Fig. 3.3 Enzymic Hydrolysis of glucobrassicin

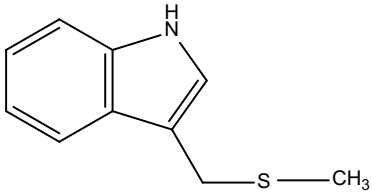


Fig. 3.4 Disulphide from attempted trapping of indolyl-3-methyl isothiocyanate

The reactivity of the 3-position of the indole nucleus and the propensity for substituents at that position to undergo nucleophilic attack has been noted [26]. Attempts to trap the intermediate isothiocyanate by carrying out the enzymic hydrolysis in a low water system or using methyl sulphide to trap the intermediate as a dithiocarbamate were unsuccessful [27]. In the latter case, the only product found to be formed was the corresponding disulphide (Fig. 3.4).

Much of the interest in glucosinolates focuses on the biological properties of the enzymically produced breakdown products. The most significant of these are goitrin, which is produced by intramolecular cyclisation of a hydroxy isothiocyanate to give a vinyl oxazolidine thione, the indole glucosinolate breakdown products (particularly indole-3-carbinol) which have a range of biological activities including anti-inflammation, and effects on proliferation and differentiation [28] and the compound derived from glucoraphanin—methylthiobutyl isothiocyanate (trivially known as sulphoraphane) which has a range of anti-proliferative and anti-inflammatory properties.

The biological activity of sulphoraphane has been the focus of much of the interest in the beneficial properties of brassica vegetables [29]. Many of the effects are observational and are associated with dietary studies [30] although gene expression modulation has been observed and mechanisms have been suggested for some of the biological effects [31]. Sulphoraphane is produced from the hydrolysis of the parent glucosinolate, glucoraphanin, by myrosinase. The mechanism of this reaction, involving a rearrangement of the initially formed aglucone, is the same as that of all other glucosinolates and was described previously. If the vegetables that contain glucoraphanin are treated in a way that inactivates the enzyme (for example cooked for extended periods in boiling water) then either sulphoraphane will not be formed or it may be decomposed by heat [32]. Blanching for a short period of time would appear to be the treatment that optimises sulphoraphane production in fresh vegetables [33], however it should be noted that, in common with other isothiocyanates, sulphoraphane is pungent and this may also restrict intake. The use of blanching to maximise sulphoraphane production correlates with other studies in which heating broccoli florets at 60 °C gave optimal sulphoraphane formation [34].

Glucosinolates and their breakdown products are an example of a group of compounds whose primary purpose (plant protection and aversion to consumption) by virtue of production of a pungent breakdown product has been supplemented by beneficial biological activity for the consuming species. The range of glucosinolates (and hydrolysis products) formed is a consequence of the modification of the core

amino acid-derived molecule through a range of reactions. In the case of plants of the genus Allium, there are only 7 core molecules and the plethora of sulphur compounds found are a consequence of the chemical reactivity of the primary products formed. Once again, the initial step in the reaction pathway occurs after enzymic hydrolysis of precursor molecules as a consequence of damage to the plant tissue. As with glucosinolates and their breakdown products, the biological activities that are provoked in the consuming species are not simply a consequence of the production of pungent volatile compounds with anti-feedant properties.

3.3 Sulphur Compounds of the Genus Allium—Chemical and Biological Properties

As noted previously, the large number of glucosinolates found in nature all derive from amino acids—either directly of after various modifications including chain elongation, oxidation or functionalisation [35]. In the case of the sulphur compounds that occur on plants of the genus *Allium*, they all derive from base units. These are seven members of the class of S-alk(en)yl-L-cysteine sulphoxides. The most common members of the class are the methyl, allyl (alliin), propyl and prop-2-enyl cysteine sulphoxides Once again, as with the glucosinolates, the initiation of the production of a plethora of volatile compounds is the enzymic hydrolysis of an involatile starting material.

The biological pathway from inorganic sulphur compounds to organosulphur molecules has been ascertained. Plants and microorganisms can produce methanethiol and dimethylsulphide. Cysteine is an important component in the incorporation of sulphur in living systems. In the case of allium species, the four primary compounds that are responsible for the large numbers of volatile S compounds are, as noted previously, S-alk-(en)-yl—L-cysteine sulphoxides. These are cleaved by a co-occurring (but as in the case of myrosinase, physically separated) enzyme called alliinase to give initially pyruvate, ammonia, and sulfenic acids. The latter then form a set of thiosulphinates which combine to give a range of sulphides with different constituent groups (Fig. 3.5).

The example of one of the most abundant (and biologically relevant) compounds, allicin (diallyl thiosulfinate) is shown in Fig. 3.5 however each of the S-alk(en)yl-L-cysteine sulphoxides noted above forms similar compounds and these can then combine symmetrically or asymmetrically to give a range di, tri and higher sulphides. There are two molecules that are of particular interest, for different reasons, and which will be discussed separately—the onion-derived lachrymatory factor (thiopropanal-S-oxide) and the compound ajoene.

The lachrymatory factor in onion was first isolated by Virtanen and Spare in 1963 [36]. It was assumed that the substrate was 1-propenyl-L-cysteine sulphoxide which was hydrolysed by the enzyme alliinase to give 1-propenyl sulphenic acid. This unstable intermediate was presumed to undergo a spontaneous reaction to give the

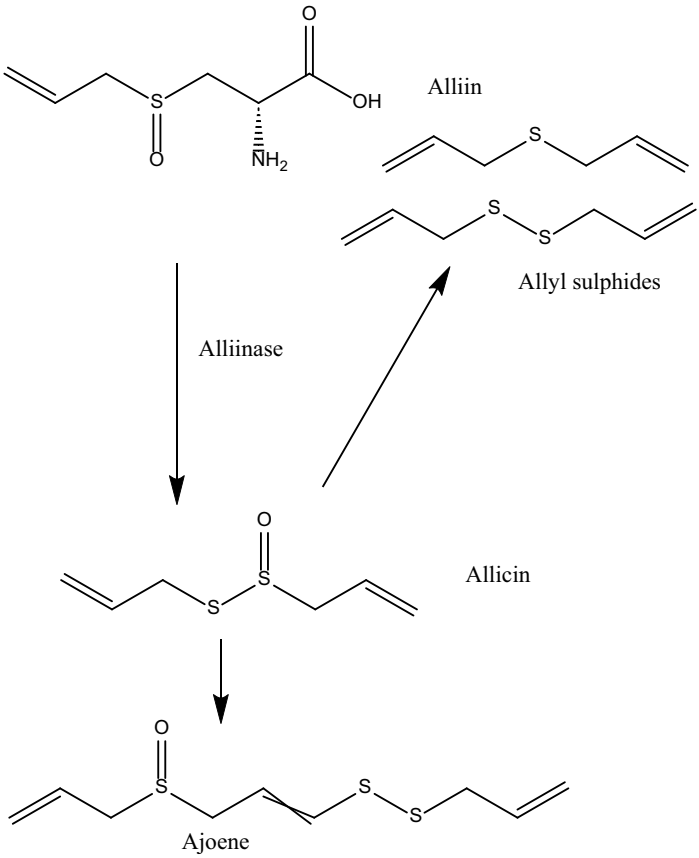


Fig. 3.5 Representative S compounds formed in allium species

lachrymatory factor, thiopropanal-S-oxide (Fig. 3.6). It was later found that a second enzyme—lachrymatory factor synthase—was required [37].

The range of thiosulphinates produced, and their various combination products, are responsible for many of the biological properties ascribed to allium species. In the case of garlic, one of the most studies and investigated products is E/Z ajoene. Ajoene is produced from the post-enzymic product, allicin in garlic (Fig. 3.5). In addition to ajoene (whose name derives from the Spanish word for garlic), garlic contains a number of other bioactive substances [38]. Many of the properties ascribed to garlic (and, indeed to other members of the genus *Allium*) are mechanistically unproven. This has been the driver for synthesis of a range of sulphur compounds with enhanced biological activities—in particular sulphides and related compounds. In addition a small number of derivatives of S-alk(en)yl-l-cysteine sulfoxides have also been synthesised and some of these appear to have promising biological activity [39]. Both physiological activity (for example, suppression of platelet aggregation)

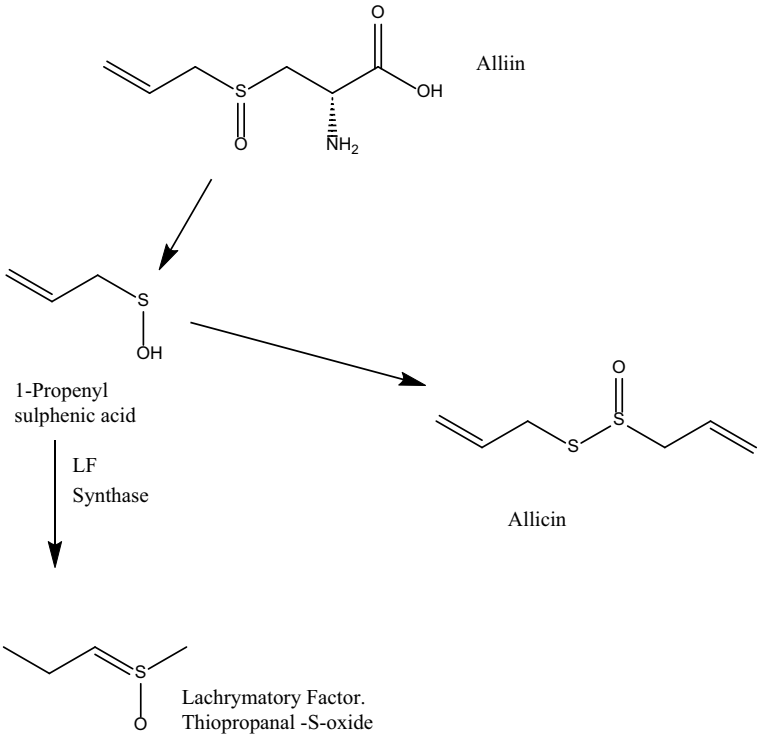


Fig. 3.6 Formation of lachrymatory factor

and antimicrobial activities have been found. In the case of ajoene, it activates the transcription factor Nrf2 [40] and, in addition, is a quorum sensing inhibitor thereby preventing pathogen growth [41]. While definitive studies on the production of hydrogen sulphide by ajoene specifically have not been carried out, the production of hydrogen sulphide by garlic sulphur compounds has been mechanistically shown to be feasible by nucleophilic substitution of allylic sulphur compounds and polysulphides [42]. The same research showed that red blood cells can convert garlic-derived polysulphides into hydrogen sulphide and this correlates with the vasoprotective effect of garlic. The production of hydrogen sulphide might also increase the potency of other antibiotics so, if ajoene (or some other component of garlic that can act as a source of hydrogen sulphide) can be applied in conjunction with an antibiotic, it may be a way of minimising the onset of anti-microbial resistance and increasing the potency of the administered antibiotic [43].

The impact of sulphur compounds on biological systems is both more complex due to the numbers of compounds formed and, mechanistically, simpler that it may appear at first. As we have seen, the structure of mustard oil glycosides was originally misassigned—based on logical extrapolation from the enzymically-produced breakdown products. The challenges of isolating the glycosides encouraged researchers

to focus on characterisation of the isothiocyanates (and simple derivatives such as thioureas) rather than the parent compounds. Even once the true structure of glucosinolates had been established by synthesis and X-ray crystallography, their isolation remained a significant challenge. One significant advance was the development of methods to remove the sulphate group of glucosinolates using sulfatase derived from the snail *Helix pomatia*. Indeed, absorption of glucosinolates onto a suitable resin (e.g. DEAE Sephadex) followed by treatment with sulfatase and elution of the desulpho compounds followed by chromatographic analysis has become a standard method for the analysis and identification of glucosinolates from plants [44]. As chemical techniques grew more powerful and robust, the role of the natural product chemist became increasingly more technically-oriented. The application of more and more sophisticated tools including advanced spectroscopic methods has enabled smaller and smaller amounts of ever more complex compounds to be identified and characterised. Despite this, natural product chemistry remains a highly evidence-based, inductive activity. The key to inductive reasoning is the accumulation of knowledge—filling up the Tabula rasa. In this way, the chemist gradually develops hypotheses that fit with empirical evidence. The problem, as Popper and others have pointed out, is that there is never any definitive proof. Induction implies extrapolation from what can be proven, through the building of evidence, to a likely outcome. While natural product chemists can propose ideas concerning molecules based upon inductive reasoning and on the evidence of analogous examples, these examples serve as a collectively supportive scaffold and if one component of the scaffold is found to be wrong then the whole structure may be undermined and collapse. Inductive reasoning uses specific examples to build a general framework. In the case of natural product chemistry, the structures of individual molecules (or their breakdown products) enables the chemist to build a plausible (but not invariably correct) generalisation. While both glucosinolates and the sulphur compounds that are produced by alliums exist in significant numbers, the means by which such a plethora of compounds is created, as has been discussed, very different. For glucosinolates and their breakdown products, the variety is built into the core structure by variation in the side chain which is biosynthesised into the final product. This suggests that the enzymes implicated in the biosynthesis of glucosinolates have relatively low side chain specificity and will accept a range of substrates. In the case of sulphur compounds of the genus *Allium*, the large number of S-containing breakdown products is a function of the nature and reactivity of the initially formed hydrolysis products and the whole range of compounds are formed from a limited number of precursors. One alternative to inductive reasoning is deductive reasoning. In this case, the logic stream progresses from the general to the specific. While chemistry favours inductive reasoning, the complexity of living systems and their interactions with natural products is better described using deductive reasoning. In the case of the sulphur compounds in both *Allium* and *Brassica* species, biological activity is observed in their interaction with complex living systems. In both cases, using whole animal or complex systems, general effects on oxidative stress or inflammation are found. In both cases, an effect on a specific set of pathways (mediated by Nrf2) is found [45]. By using deductive reasoning, we are able to hypothesise that the general effects (oxidative stress or

inflammation) induced by sulphur compounds from Alliums and Brassicas operate on the same specific pathway. There are, however, a number of conflicting facts that suggest this may not be a complete answer. In particular, if this is a specific effect caused by sulphur compounds, why do some naturally-occurring sulphur compounds fail to exhibit this activity in complex biological systems? Perhaps a part of the solution requires that we revisit the properties of the chemicals found in each type of plant from the standpoint of biological activity rather than chemical motifs. The discovery of hydrogen sulphide as a gasotransmitter has been a significant advance in understanding the nature of the relationship between sulphur and living systems. These investigations have indicated that the half life of hydrogen sulphide in vivo is short and an endogenous source is commonly thought to be required [46]. One interesting aspect of the biological effect of hydrogen sulphide is that it induces the same Nrf2 pathway that has been implicated in the activity of certain isothiocyanates and sulphur compounds from *Allium* species [47]. The sulphur compounds from alliums (particularly allylic sulphur compounds that can undergo nucleophilic attack) are precursors for the production of hydrogen sulphide in vivo. We have previously seen that isothiocyanates can bind to hydrogen sulphide in the oral cavity to form a dithiocarbamate. These compounds are also able to release hydrogen sulphide in vivo. Deductive reasoning suggests that at least some of the general properties of the sulphur compounds of alliums and brassica plants may be mediated through the specific and in situ release of hydrogen sulphide. Deductive reasoning led us to the Nrf2 pathway as one specific potential explanation for the activity of sulphur compounds and an understanding of the chemistry and biology of hydrogen sulphide has suggested a potential common source for the interactions of different classes of compounds with biological systems. From the context of biological activity, using deductive reasoning we can demonstrate a link between the sulphur compounds of *Allium* and *Brassica* species in terms of an effect on the Nrf2 pathway. Furthermore, the same reasoning enables us to deduce that both sets of compounds may be implicated due to their roles as potentiators of hydrogen sulphide release.

3.4 The Biological Consequences of Natural Products of Plants of the Genus Brassica

As is commonly the case with medicinal plants, investigations into the biological properties of natural products like glucosinolates and the corresponding isothiocyanates began with an appreciation of the whole plant. In the case of Brassica vegetables, members of the class including mustard and cabbage were appreciated as medicinal foods long before their chemical constituents were identified. Brassica crops were known in India over 5000 years ago. The spread in brassicas is in part due to their adaptability. Two of the most important species are *Brassica oleracea* and *Brassica rapa.* The natural cross between them (which contains chromosome sets from both parents) is *Brassica napus.* Variants of all three Brassica cultivars (both

naturally and as a result of cultivation) have led to the significant number of species with variable properties. All contain glucosinolates although the precise distribution of the different members of the class vary with plant species and the part of the plant. The presence of glucosinolates has also impacted upon the plant's susceptibility to insect and herbivore predation and to fungal attack. One of the most significant early uses of cabbage juice was as a laxative. This property may be a consequence of the fibre content of the juice rather than any intrinsic bioactive constituent. Cabbage juice has also been found to be effective in the treatment of peptic ulcers [48]. The effect of cabbage juice may be a consequence of its inhibitory activity on *Helicobacter pylori*—a contributory causative agent for peptic ulcers [49]. As the involvement of *H. pylori* in the etiology of peptic ulcers was developed, the positive role of Brassica vegetables may be a consequence of the presence of the anti-microbial glucosinolate breakdown products, isothiocyanates, in macerated cabbage [50]. Cabbage juice was also used as a topical anti-inflammatory. It has been suggested that the mechanism of action is related to the inhibition of a specific human gene (Factor Associated with Neutral Sphingomyelinase activation—FAN) by a cabbage miRNA [51].

In addition to the presence of glucosinolates (and their breakdown products), Brassica vegetables are also a rich source of bioactive phenolic compounds. While reductionism leads to theories regarding *the* active ingredient, the possibility that other components could be involved in any given effect. In addition, the potential for synergistic effects should also be considered by using interaction metabolomics [52]. These other bioactive factors include flavonoids and hydroxy cinnamic acids. The distribution of phytochemicals both between different members of the vegetable family and within individual plants can vary both in terms of the genes that are present, the part of the plant that is analysed and growing conditions (including stress). For example, the sprouting tips of Brassicas contain higher amounts of indole glucosinolates than the roots or stems. Similar profile differences are observed for other phytochemicals. Hydroxy cinnamic acids (particularly ferulic acid) are found in some types of plant (e.g. red cabbage) but not in others in related families (radish). Cyclobutane dimers from ferulic and other hydroxy cinnamic acids are found in many members of the *Gramineae* (grasses) where they seem to have a role in stabilising the physical structure of the cell wall but these compounds have not been reported in Brassicas [53]. Cyclobutane-containing alkaloids have, however been reported in a number of plants [54] and it is possible that the core structure may also occur more widely. Given that cyclobutane natural products have been found to have a number of properties, it is possible that some of the biological properties of crude Brassica extracts could, in part, be due to the presence of these or similar compounds [55], even if they have not yet been reported to be present. The precursors of cyclobutanes (ring substituted cinnamic acids) have been found in plants of the genus *Brassica* [56] and this may suggest a more complete analysis may reveal the corresponding cyclobutanes. Anthocyanins have also been reported in many Brassicas—with once again variation in the amounts and types of compound present in different species and different plant parts [57].

As noted previously, the major bioactive compounds found in plants of the family *Brassicaceae* are those derived from glucosinolates. Glucosinolates are a group of

amino acid-derived compounds that are synthesised and stored by plants. Their primary in situ biological activities (and that of their enzymically produced hydrolysis products) appears to be as anti-microbial (including anti-fungal) agents and as an insect and animal feeding deterrent. The effect of the compounds on human physiology is a secondary feature, from the standpoint of the plants that produce them, but it is this latter activity that has attracted considerable attention. Considerable efforts have been made to understand how different glucosinolates and their breakdown products interact with specific biological systems (see above and Sect. 3.5). There are comparisons between the advances in organic chemistry (particularly isolative and spectroscopic methods) which gradually revealed greater details about the inferred chemical structure and reactions of glucosinolates and the advances in the understanding of the details of biological systems which revealed sites of action and some of the biological consequences of these. In both cases, scientific reductionism has revealed details of the chemical and biological properties of the systems however there are limits to the power of reductionism in science—particularly when attempting to measure chronic effects or the effects on a whole biological system rather than just a part of it.

The starting point for a study of the biological properties of glucosinolates and their breakdown products is the appreciation of their role as plant protection products. It is simplest to think of plant secondary metabolites in the context of their role to the producing species as overtly biologically active agents and, while this may be an important aspect of why they are synthesised, there is also considerable functional diversity in the structures that are created [58]. The traditional separation of plant chemicals into primary metabolites, secondary metabolites and hormones appears to be increasingly over simplistic. In the case of glucosinolates, they are produced in higher levels in response to stress factors such as high salt, drought and physical damage [59]. In addition, glucosinolates have been suggested to be a sulphur storage system in some plants [60]. The major biological and ecological impact of glucosinolates appears to be as a result of, as noted previously, the production of a range of products (particularly isothiocyanates and nitriles) when the tissue of the plant is damaged. This type of damage can occur as a result of microbial infestation, insect predation or attempted herbivore consumption. The effect of fungal attack on the production of isothiocyanates (and other volatile organic compounds) has been reviewed [61] and the potential to produce such compounds in response to fungal attack may provide a plant with an evolutionary survival advantage. A range of antifungal plant-derived isothiocyanates have been assessed including assays against *Aspergillus* and *Penicillium* species as well as other human and plant pathogens. The antifungal activity is associated with the electrophilicity of the C atom of the isothiocyanato moiety which can form a range of adducts with amino acid side chains [61]. Isothiocyanates react rapidly with thiol groups and can be potent succinate dehydrogenase inhibitors. The potential of succinate dehydrogenase inhibitors as anti-fungal agents has been noted [62] and this property may be of particular use to isothiocyanate-producing plants. In addition to anti-fungal activity (including inhibitory activity against plant pests, the anti-microbial activity of isothiocyanates has been extensively studied.

One of the most significant diseases of Brassicas is clubroot. This is caused by the fungal organism, *Plasmodiophora brassicae*. Although some isothiocyanates have, as noted above, some antifungal properties, the presence of indolyl glucosinolates appears to be correlated with increased gall size in clubroot affected plants. This may be a result of a link between indole glucosinolate breakdown products and auxin production [63]. Clubroot remains a major problem and resistance is an accepted method of control however new variants of *Plasmodiophora* can circumvent this. Methods for biocontrol have been reviewed but not of these are commercially viable [64]. Many details of the biology of infection with *Plasmodiophora brassicae* remain unclear however it is known that the co-existing bacterial community can stimulate the germination of spores of *Plasmodiophora brassicae* and if the triggers for this process can be identified then some measure of control of the active fungi and the longevity of its spores in soil may be routes for control [65]. It is possible that CRISPR technology might be a way of identifying potential targets related to spore germination—as has been carried out for restoration of antibiotic efficacy in resistant strains of *Streptococcus pneumoniae* [66].

The impact of glucosinolates and their breakdown products on insect predation of Brassicas has been extensively studied [67]. It is widely accepted that glucosinolates can affect insect predation behaviour. However, far from simply being a deterrent, some insects can accumulate plant defensive toxins such as glucosinolates and to then make use of them [68]. The cabbage aphid can selectively accumulate glucosinolates as a protection against lacewing attack [69]. Sometimes the interaction is more nuanced than simply the accumulation of a toxin. For example, in the case of the horseradish flea beetle, different myrosinases created by the beetle are expressed at different stages in the lifecycle and have different substrate specificities [70]. This enables the insect to produce volatile compounds that are specific for potential predators at different life cycle stages. Glucosinolates and/or their hydrolysis products (particularly isothiocyanates) have been implicated in the behaviour (including feeding and egg laying) of a number of species of insect [71]. Indeed two specific receptors for the glucosinolate, sinigrin, has been found in cabbage butterflies [72]. These are associated with gustatory responses and sinigrin is therefore a feeding stimulant. While glucosinolates are not toxic per se, the isothiocyanates produced as a result of their hydrolysis by myrosinase is harmful to most insect species. This is the case even for those, like cabbage butterflies that use sinigrin as a feeding stimulant. In order to alleviate the toxic effects of isothiocyanates, cabbage butterflies express a nitrile-specifier protein (NSP) that diverts hydrolysis of glucosinolates to less toxic nitriles rather than to isothiocyanates [73].

3.5 Biological Effects of Glucosinolates in Humans

The biological effects of glucosinolates and isothiocyanates in humans have been extensively studied since they have been a virtually ubiquitous part of human diets for a considerable period of time. The deleterious effects of brassicas are primarily

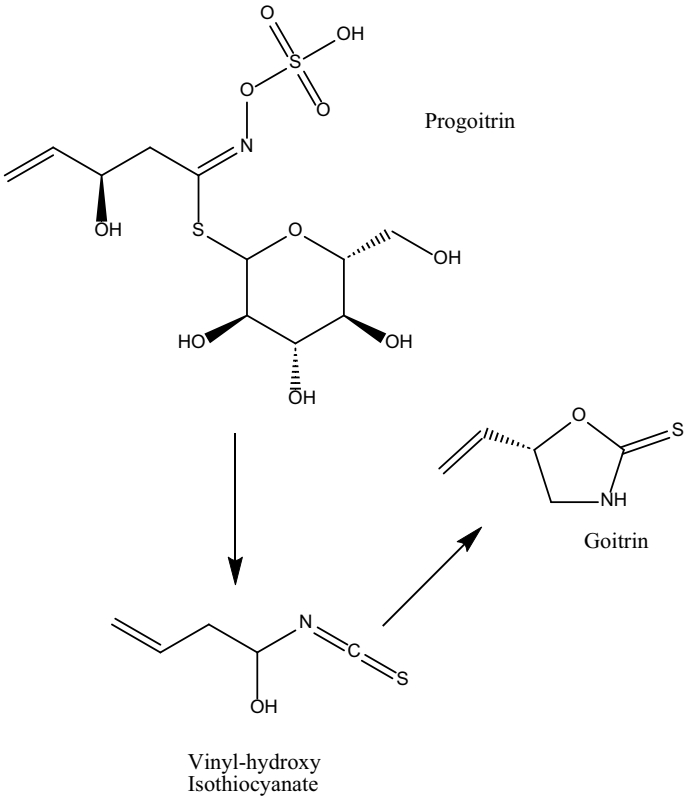


Fig. 3.7 Formation of goitrin

the consequence of the presence of vinyl-oxazolidine thione, trivially called goitrin, (a cyclisation product from a precursor isothiocyanate—Fig. 3.7) and thiocyanate ion which is produced primarily from the hydrolysis of indolyl-3-methyl isothiocyanate.

Goitrin, as the name suggests, has an effect on the thyroid and, while consumption of brassica vegetables is unlikely to make a significant contribution to goitre in humans [74], it is nonetheless an important bioactive metabolite. Thiocyanate ion is a competitive inhibitor of iodine uptake but is not present in brassica vegetables at levels which could impact on the thyroid. There has, however, been a report that cases of thyroid cancer are elevated in iodine-deficient women in New Caledonia whose diet contains high levels of cruciferous vegetables [75], however, as with many associative, observational studies, there are a number of potential confounders. In general, unless iodine deficiency is endemic or the thyroid is compromised, the effects of brassica vegetable consumption on the thyroid is likely to be negligible. As with many natural products, the effects of goitrin are biologically multi-dimensional. In addition to its goitrogenic activity, goitrin was also found to inhibit LPS-induced sepsis through its effect on the caspase-11 non-canonical inflammasome [76]. In addition

the anti-viral activity of goitrin present in extracts of *Isatidis radix* (a traditional Chinese medicine) has been noted [77].

The anti-inflammatory effects of glucosinolate breakdown products has been noted previously with modulation of Nrf2 and the downregulation of NF-kappa B being two of the most studied effects even to the extent of a role being suggested in the prevention of the 'cytokine storm' associated with infection by SARS-COV-2 [78]. The apparent activity of isothiocyanates against SARS-COV-2 my be a consequence of its potential role in the biological storage and release of hydrogen sulphide—which may be the active agent [79]. Indeed, as previously noted, many of the effects of bioactive isothiocyanates appears to be similar to the effects of hydrogen sulphide and the possibility of the formation of dithiocarbamates and subsequent release of hydrogen sulphide is a plausible mechanism.

A further effect of goitrin, which has been comparatively little studied, is its effect on the extra-oral bitter receptor. Goitrin contains an NH–C=S moiety which is also present as an essential factor in the binding of bitter compounds including 6-n-propylthiouracil (PROP) and phenylthiocarbamide (PTC) (Fig. 3.8).

Bitter taste, as noted previously, is a consequence of the binding of compounds to bitter taste receptors—of which 26 members exist. The G-coupled receptors occur in the oral cavity but have been found elsewhere and binding of agonists (or antagonists) to them can have a range of consequences. The receptors are classified as 'type 2 taste receptors' and are abbreviated as TAS2R (or T2R) followed by a number. One member of the class TAS2R38 has a particular affinity for the NH–C=S grouping and has been implicated in a number of extra-oral conditions including upper respiratory infection and chronic rhinosinusitis [80] and in susceptibility to respiratory infections. TAS2R38 is activated in response to acyl-homoserine lactone quorum-sensing molecules secreted by *Pseudomonas aeruginosa* and other gram-negative bacteria [81]. This may have significance for bacterial infection by regulation of the

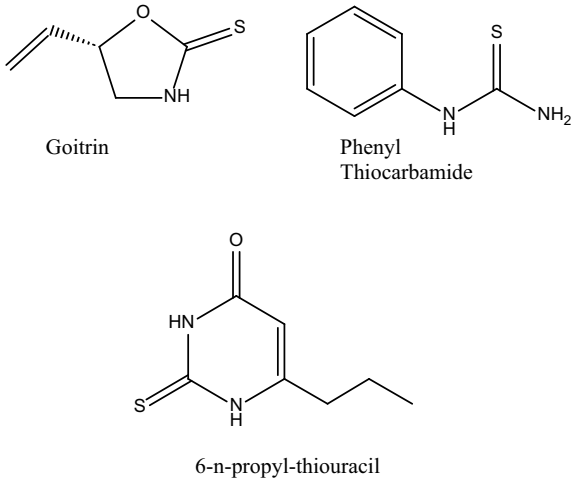


Fig. 3.8 Structures of goitrin, PTC and PROP

innate immune response of the epithelial airway [81]. In pancreatic tumour cells, TAS2R38 can stimulate expression of the multi drug transporter ABCB1 [82] and this may limit the effectiveness of pharmacological treatments. In addition to a potential role in enhancement of chemoresistance mechanisms, other bitter receptors (and ligands) may have anti-cancer or chemotherapy enhancing properties [83] and this may partially explain the apparently contradictory effects of natural products with different bitter taste profiles. The situation, then, is complex and while a reductionist approach to defining the biological properties of natural products may offer some clues in terms of binding to specific receptors, both the isolation of specific compounds and the delineation of specific biological properties may not show the complete picture. This is further exacerbated by the long term (and chronic) properties of many NCDs and the impact of mixtures of natural products that may be obscured by biological test systems using isolated compounds.

In addition to goitrin, the most extensively studied glucosinolate-derived plant component (as noted previously) is sulphoraphane. The effect of sulphoraphane on some specific biochemical pathways including Nrf2 activation, amelioration of oxidative stress, inhibition of histone deacetylase and modulation of an NF-kappaB inflammatory response has been noted [84]. As noted previously, many of the biological effects ascribed to hydrogen sulphide operate on the same pathways and it is of interest to examine whether the effects ascribed to sulphoraphane may be a consequence of the in situ formation of a dithiocarbamate via a reaction between endogenous hydrogen sulphide (which could be produced either from plant or mammalian tissue) and sulphoraphane and then subsequent release of hydrogen sulphide at a site of action in vivo. The formation of the dithiocarbamate would be mediated by the electrophilicity of the C atom in the isothiocyanato moiety and this might be stabilised through proximity to the S=O group in sulphoraphane. The idea that the beneficial properties of brassica vegetables might be linked to the presence of a hydrogen sulphide donor is both logical and able to be tested.

3.6 Natural Product Bioactivity

In most cases, as noted previously, natural product research has favoured an approach based upon the principles of reductionism. Methods to extract and purify natural products have been driven by the gradual exclusion or removal of components of less interest until the pure component is obtained. This is particularly the case for bioactivity-guided exercises when the utility of the source material resides in the existence of some beneficial characteristic but also applies to the isolation of colours and flavours as well as other characteristic natural products. The essence of a bioactivity guided approach is to make an extract of the plant or other biological material and to then fractionate it over several steps until a single bioactive principle has been isolated and can then be chemically and biologically characterised. A key aspect is that an assay must exist (or be developed) for the bioactivity of interest. This means that other aspects of potential benefits of the source material may not be realised. One example

of this was the discovery (and subsequent exploitation) of statins. The story of statins exemplifies many of the standard protocols employed by natural product chemists. The beginning of the story was the discovery of cholesterol—isolated from gall stones in the eighteenth century. In 1910, it was reported that atherosclerotic plaques from aortas of human subjects contained over 20-fold higher concentrations of cholesterol than did normal aortas. The next step was the discovery of a link between arteriosclerosis and plaques and cardiovascular disease. While this was originally an observational observation in the 1930s, the genetic link that suggested that high blood cholesterol levels (familial hypercholesteremia) was linked to premature heart attack was confirmed in the 1960s. Prior to this, two further studies (the Seven Countries study and the Framlingham Heart study) showed that individuals with higher blood cholesterol levels were more likely to develop heart disease and that smoking and high blood pressure were also contributory factors. These various strands suggested a chain of causality. Atherosclerotic plaques were implicated in cardiovascular disease and these plaques contained high levels of cholesterol. Individuals with high blood cholesterol levels were also at higher risk of developing heart disease. The clear implication was that reducing blood cholesterol levels would reduce the development of plaques containing high levels of cholesterol and this would reduce the risk of cardiovascular disease. The question remained, how could excessive cholesterol production be controlled and reduced? Subsequent studies on the biosynthesis of cholesterol showed that a key rate limiting step in the pathway was the conversion of 3-hydroxy-3-methylglutaryl coenzyme A (HMGCoA) to mevalonic acid which is catalysed by HMGCoA reductase (Fig. 3.9).

Knowledge of this pathway, and the rate limiting step, enabled a target and a strategy to be developed. Cholesterol can come from two sources—the diet or de novo synthesis in the liver. If the dietary level is sufficient then synthesis in the liver is suppressed by feedback inhibition to prevent excessive production of cholesterol. In humans, the amount synthesised by the liver exceeds the dietary intake (even if consuming a high cholesterol diet) so inhibition of HMGCoA reductase will be an efficient method of controlling serum cholesterol levels. In addition it was later found that patients with familial hypercholesteremia had little or no feedback control over their HMGCoA reductase activity and this lack of an important control mechanism has important clinical consequences. A number of cholesterol lowering compounds were available in the 1960s including chlofibrate however the cholesterol-lowering capability of chlofibrate was minimal and a WHO study indicated a high (non-cardiovascular) mortality rate associated with chlofibrate use in a clinical trial [85]. Other drugs (e.g. cholestyramine) had a spectrum of side effects and this limits their use. The breakthrough in the discovery of statins was made by the Japanese natural product chemist, Akira Endo. His initial work on the discovery and isolation of HMGCoA reductase inhibitors from fungi mirrored the discovery of penicillin and was designed to find new antibiotics that inhibited the mevalonate to cholesterol pathway in bacteria. While direct measurement of HMGCo reductase inhibitors was generally carried out by measuring the incorporation of (C-14) HMGCoA activity into mevalonate, the radiolabelled substrate was too expensive to use to assay thousands of fungal extracts so he developed an alternative (cheaper) method looking for

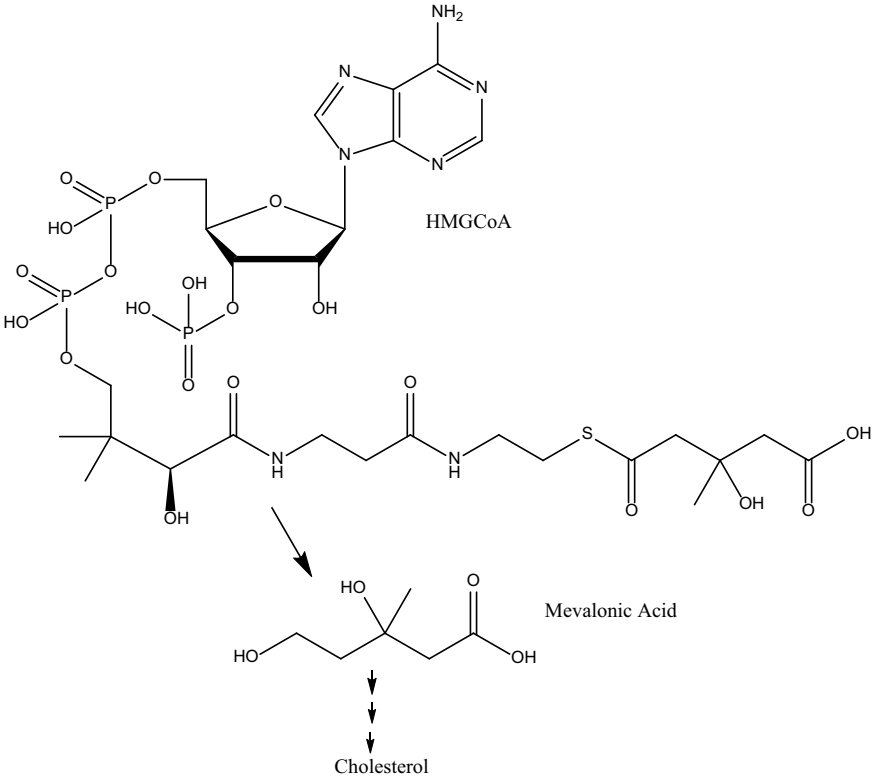


Fig. 3.9 Rate limiting step in cholesterol formation

the inhibition of (C-14) acetate into non-saponifiable lipids. Eventually an inhibitor, the known compound citrinin was isolated and found to be effective in a mouse assay for lowering cholesterol. Unfortunately due to kidney toxicity, the compound was abandoned. After exhaustive extraction and fractionation, a second compound was isolated from *Penicillium citrinum* and the most active compound named compactin. Interestingly another group, in the UK, isolated the same compound but described it as having antifungal properties. It was interesting to note that, with repeated doses of compactin in rats, eventually hepatic HMGCoA reductase activity was increased 8–10fold and cholesterol reduction was no longer observed. This partial failure in rats was one reason why compactin was not developed as a cholesterol-lowering drug by the UK group [86]. Work on compactin was eventually abandoned and other statins were developed including the most prescribed member of the class, Atorvastatin. Clearly the type of activity being assessed is of crucial importance in the development of bioactives from natural products. As we have seen previously, targeting mevalonate as an adjunct therapy to penicillin-based antibiotics could be a successful strategy for antibiotics [66]. The failure to find cholesterol lowering activity by the UK-based researchers may have been a consequence of the model

animal used (rat rather than mouse). In addition, statins also have found utility as anti-inflammatory compounds and this may be a secondary effect in the prevention of cardiovascular disease (improvement in endothelial function and a positive effect on atherosclerosis) [87]. The identification of bioactive principles in complex mixtures of natural products remains a challenge. This is greatly amplified if components in a mixture act synergistically. In some cases the synergic activity will not be detected and one or more fractions may be discarded by mistake. One approach to overcoming this problem is to integrate chemical and biological datasets. This approach (called 'interaction metabolomics') has been used to examine synergy in the antimicrobial activity against Staphylococcus aureus [88]. Other metabolomic approaches have been adopted to better predict disease outcomes. Even if some progress has been made to identifying synergies, two significant challenges remain.

1. Does the reductionist approach allow for the identification of the effect of natural products on overall health—not just specific disease states?
2. How can the long term effects of dietary components on eventual health outcomes be measured?

In the case of the first of these, a complete understanding of the biological system and all its interactions is required. In addition, it is also necessary to measure the effect of perturbations on that system. It is important to realise, however, that biological systems are not in a homeostatic state so it is necessary to measure a constantly changing system. In the case of the second challenge, biomarkers that can be measured at a specific moment but which are predictive of the likelihood of longer term outcomes are required. In clinical terms, a predictive biomarker is one which can be used to identify those who are more likely to respond to a particular treatment. It is, in effect, a susceptibility biomarker (FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet]. Silver Spring (MD): Food and Drug Administration (US); 2016-. Predictive Biomarker. 2016 Dec 22. Available from: https://www.ncbi.nlm.nih.gov/books/NBK402283/ Co-published by National Institutes of Health (US), Bethesda (MD).). Such biomarkers include phenotypes and germline mutational profiles. One of the most frequent use of this type of biomarker is in the area of cancer treatment where the difference between prognostic and predictive biomarkers has been the subject of investigation [89]. One crucial aspect of predictive biomarkers is that they should be based on a mechanistic understanding rather than just being associative.

The development of hypotheses and mechanistic understanding—and the means to test such hypotheses—is crucial to building a better picture of how natural products interact with living systems. The complexity of living systems combined with the large numbers of natural products present (with multifunctional characteristics) encourages the adoption of two approaches—firstly, reductionism, to try to develop an understanding of how an individual compound interacts with one specific (and hopefully important) aspect of a living system and to then apply that to other compounds and other parts of the living system, secondly observation of the impacts of groups of compounds on the complete living system over a period of time. Some

aspects of this latter type of endeavour have been classified as 'nutritional epidemiology' and have been strongly criticised for attempting to provide proof of causality based upon scant evidence and poor techniques [90]. While the authors called for "(1) Stronger Designs, (2) Stronger Measurement, (3) Stronger Analyses, and (4) Stronger Execution and Reporting. Participants from several leading academic institutions explored existing, evolving, and new better practices, tools, and techniques to collaboratively advance specific recommendations for strengthening nutritional epidemiology.", it is unclear how these aims will be achieved within the current experimental and hypothesis-driven paradigm. Simply calling for more rigour within the existing framework will not solve the fundamental problem that lies at the core of developing a better understanding of how two complex systems (natural products and living organisms) interact over time. In addition, the more complex the system, the greater the potential for confounders. Multivariate analysis has been suggested [91], however, some criteria must be adopted to determine which factors are of the greatest significance otherwise a simple analysis of variance (ANOVA) technique may demonstrate which factors vary most but these may not be the most significant. In addition, such approaches may not adequately explain the possibilities for synergy between components. Further studies are likely to arise from the need to develop better mechanistic understandings which can then be synthesised into a more complete picture. The use of CRISPR technology to remove specific biological characteristics of a system combined with the development of more consistent and accurate biological test systems may offer an opportunity to carry out targeted studies in which biological variations are specific and the results are more likely to be mechanistically meaningful and less prone to unanticipated elements or unexpected confounders.

References

1. Chen S, Wang L, Deng Z (2010) Twenty years hunting for sulfur in DNA. Protein Cell 1(1):14–21. https://doi.org/10.1007/s13238-010-0009-y. Epub 2010 Feb 7. PMID: 21203994; PMCID: PMC4875114
2. Hanley AB, Fenwick GR (1989) Naturally occurring sulphur compounds. In: Damani LA (ed) Sulphur-containing drugs and related organic compounds, vol 1, Part A, 41, Ellis Horwood
3. Chatterjee S, Hausinger RP (2022) Sulfur incorporation into biomolecules: recent advances. Crit Rev Biochem Mol Biol 57(5–6):461–476. https://doi.org/10.1080/10409238.2022.2141678. Epub 2022 Nov 20. PMID: 36403141
4. Munteanu C, Turnea MA, Rotariu M (2023) Hydrogen sulfide: an emerging regulator of oxidative stress and cellular homeostasis-a comprehensive one-year review. Antioxidants (Basel) 12(9):1737. https://doi.org/10.3390/antiox12091737.PMID:37760041;PMCID:PMC10526107
5. Pacher P (2021) Cyanide emerges as an endogenous mammalian gasotransmitter. Proc Natl Acad Sci U S A 118(25):e2108040118. https://doi.org/10.1073/pnas.2108040118.PMID:34099579;PMCID:PMC8237670
6. Cai H, Wang X (2021) Effect of sulfur dioxide on vascular biology. Histol Histopathol 36(5):505–514. https://doi.org/10.14670/HH-18-290. Epub 2020 Dec 15 PMID: 33319344

7. Liu H, Wang J, Liu J et al (2021) Hydrogen sulfide (H2S) signaling in plant development and stress responses. aBIOTECH 2:32–63. https://doi.org/10.1007/s42994-021-00035-4
8. Barton LL, Fauque GD (2022) Sulfate-reducing Prokaryotes: changing paradigms. In: Sulfate-reducing bacteria and archaea. Springer, Cham. https://doi.org/10.1007/978-3-030-96703-1_1
9. Shatalin K, Shatalina E, Mironov A, Nudler E (2011) H2S: a universal defense against antibiotics in bacteria. Science 334(6058):986–990. https://doi.org/10.1126/science.1209855. PMID: 22096201
10. Moje W, Munnecke D, Richardson L (1964) Carbonyl Sulphide, a Volatile Fungitoxicant from Nabam in Soil. Nature 202:831–832. https://doi.org/10.1038/202831a0
11. Cirino G, Szabo C, Papapetropoulos A (2023) Physiological roles of hydrogen sulfide in mammalian cells, tissues, and organs. Physiol Rev 103(1):31–276. https://doi.org/10.1152/physrev.00028.2021. Epub 2022 Apr 18 PMID: 35435014
12. Cao X, Cao L, Zhang W, Lu R, Bian JS, Nie X (2020) Therapeutic potential of sulfur-containing natural products in inflammatory diseases. Pharmacol Ther 216:107687. https://doi.org/10.1016/j.pharmthera.2020.107687. Epub 2020 Sep 20 PMID: 32966837
13. Levinn CM, Cerda MM, Pluth MD (2020) Activatable small-molecule hydrogen sulfide donors. Antioxid Redox Signal. 32(2):96–109. https://doi.org/10.1089/ars.2019.7841. Epub 2019 Oct 29. PMID: 31554416; PMCID: PMC6918874
14. Pandey T, Pandey V (2024) Hydrogen sulfide (H2S) metabolism: Unraveling cellular regulation, disease implications, and therapeutic prospects for precision medicine. Nitric Oxide: Biol Chem
15. Tian M, Dodds MJC, Hanley AB (2013) Allyl isothiocyanate is effective in reducing volatile sulphur compounds responsible for intrinsic oral malododor. J Breath Res 7:026001 https://doi.org/10.1088/1752-7155/7/2/026001
16. Dixon AE (1982) Action of bromine on allylthiocarbimide. J Chem Soc Trans
17. Moran FJC (1872) Cassell's family continental handbook. Cassell, Petter, and Galpin, London
18. Bussy A (1840) J Pharmac Chim Sur la formation de l'huile essentielle de moutarde 26:39
19. Gadamer J (1897) Ueber die Sinapinsäure. Arch Pharm 235:44
20. Gadamer J (1897) Ueber Sinapinsäure Ber deutsch chem Ges 30:2322
21. Ettlinger MG, Lundeen AJ (1956) The structure of sinigrin and sinalbin: an enzymatic rearrangement. J Am Chem Soc 78:4172–4173
22. Ettlinger MG, Lundeen AJ (1957) First synthesis of a mustard oil glucoside: the enzymatic Lössen rearrangement. J Amer Chem Soc 79:1764
23. Lossen W (1872) Ueber Benzoylderivate des Hydroxylamins. Justus Liebigs Ann Chem 161(2–3):347–362
24. Del CM-B, Moreno DA, Carvajal M (2013) The physiological importance of glucosinolates on plant response to abiotic stress in Brassica. Int J Mol Sci 14:11607–11625
25. Gmelin R, Virtanen AI, Glucobrassicin, the precursor of SCN∼, 3-indolylacetonitrile and ascorbigen in Brassica oleracea species. Ann Acad Scientiarum Fennicae 107:3–23
26. Kishbaugh, TLS (2010) Reactions of indole with nucleophiles. In: Gribble G (ed) Heterocyclic Scaffolds II: topics in heterocyclic chemistry, vol 26. Springer, Berlin, Heidelberg. https://doi.org/10.1007/7081_2010_35
27. Hanley AB, Parsley KR, Lewis JA, Fenwick GR (1990) Chemistry of indole glucosinolates: intermediacy of indol-3-yl methyl isothiocyanates in the enzymic hydrolysis of indole glucosinolates. J Chem Soc Perkin Trans I 2273
28. Amarakoon D, Lee WJ, Tamia G, Lee SH (2023) Indole-3-carbinol: occurrence, health-beneficial properties, and cellular/molecular mechanisms. Annu Rev Food Sci Technol 27(14):347–366. https://doi.org/10.1146/annurev-food-060721-025531. PMID: 36972159
29. Mangla B, Javed S, Sultan MH, Kumar P, Kohli K, Najmi A, Alhazmi HA, Al Bratty M, Ahsan W (2021) Sulforaphane: a review of its therapeutic potentials, advances in its nanodelivery, recent patents, and clinical trials. Phytother Res 35(10):5440–5458. https://doi.org/10.1002/ptr.7176. Epub 2021 Jun 28 PMID: 34184327

30. Manchali S, Chidambara Murthy KN, Patil BS (2012) Crucial facts about health benefits of popular cruciferous vegetables. J Funct Foods 4:94–106. https://doi.org/10.1016/j.jff.2011.08.004
31. Kim JK, Park SU (2016) Current potential health benefits of sulforaphane. EXCLI J 13(15):571–577. https://doi.org/10.17179/excli2016-485.PMID:28096787;PMCID:PMC5225737
32. Mahn A, Martin C, Reyes A, Saavedra A (2016) Evolution of sulforaphane content in sulforaphane-rich broccoli during tray drying. J Food Eng 186:27–33. https://doi.org/10.1016/j.jfoodeng.2016.04.007
33. Mahn A, Pérez CE, Zambrano V, Barrientos H (2022) Maximization of sulforaphane content in broccoli sprouts by blanching. Foods 11(13):1906. https://doi.org/10.3390/foods11131906.PMID:35804720;PMCID:PMC9266238
34. Matusheski NV, Juvik JA, Jeffery EH (2004) Heating decreases epithiospecifier protein activity and increases sulforaphane formation in broccoli. Phytochemistry 65(9):1273–1281. https://doi.org/10.1016/j.phytochem.2004.04.013. PMID: 15184012
35. Nguyen VPT, Stewart J, Lopez M, Ioannou I, Allais F (2020) Glucosinolates: natural occurrence, biosynthesis, accessibility, isolation, structures, and biological activities. Molecules 25(19):4537. https://doi.org/10.3390/molecules25194537.PMID:33022970;PMCID:PMC7582585
36. Spåre C, Virtanen AI, Munch-Petersen J, Kvande PC, Meisingseth E (1963) On the lachrymatory factor in onion (Allium cepa) vapours and its precursor. Acta Chem Scand 17:641–650
37. Imai S, Tsuge N, Tomotake M et al (2002) An onion enzyme that makes the eyes water. Nature 419:685. https://doi.org/10.1038/419685a
38. El-Saber Batiha G, Magdy Beshbishy A, G Wasef L, Elewa YHA, A Al-Sagan A, Abd El-Hack ME, Taha AE, M Abd-Elhakim Y, Prasad Devkota H (2020) Chemical constituents and pharmacological activities of garlic (Allium sativum L.): a review. Nutrients 12(3):872. https://doi.org/10.3390/nu12030872. PMID: 32213941; PMCID: PMC7146530
39. Liakopoulou-Kyriakides M (1985) Relation between the structure of alliin analogues and their inhibitory effect on platelet aggregation. Phytochemistry 24:1593–1594. https://doi.org/10.1016/S0031-9422(00)81070-5
40. Ngo V, Duennwald ML (2022) Nrf2 and oxidative stress: a general overview of mechanisms and implications in human disease. Antioxidants (Basel) 11(12):2345. https://doi.org/10.3390/antiox11122345.PMID:36552553;PMCID:PMC9774434
41. Jakobsen TH, van Gennip M, Phipps RK, Shanmugham MS, Christensen LD, Alhede M, Skindersoe ME, Rasmussen TB, Friedrich K, Uthe F, Jensen PØ, Moser C, Nielsen KF, Eberl L, Larsen TO, Tanner D, Høiby N, Bjarnsholt T, Givskov M (2012) Ajoene, a sulfur-rich molecule from garlic, inhibits genes controlled by quorum sensing. Antimicrob Agents Chemother 56(5):2314–2325. https://doi.org/10.1128/AAC.05919-11. Epub 2012 Feb 6. PMID: 22314537; PMCID: PMC3346669
42. Benavides GA, Squadrito GL, Mills RW, Patel HD, Isbell TS, Patel RP, Darley-Usmar VM, Doeller JE, Kraus DW (2007) Hydrogen sulfide mediates the vasoactivity of garlic. Proc Natl Acad Sci U S A. 104(46):17977–17982. https://doi.org/10.1073/pnas.0705710104. Epub 2007 Oct 19. PMID: 17951430; PMCID: PMC2084282
43. Ng SY, Ong KX, Surendran ST, Sinha A, Lai JJH, Chen J, Liang J, Tay LKS, Cui L, Loo HL, Ho P, Han J, Moreira W (2020) Hydrogen sulfide sensitizes Acinetobacter baumannii to killing by antibiotics. Front Microbiol 7(11):1875. https://doi.org/10.3389/fmicb.2020.01875.PMID:32849459;PMCID:PMC7427342
44. Hogge RL, Reed DW, Underhill EW (1988) HPLC separation of glucosinolates from leaves and seeds of Arabidopsis thaliana and their identification using thermospray liquid chromatography/mass spectrometry. Chromatog Sci 26(551):556
45. Lin L, Wu Q, Lu F, Lei J, Zhou Y, Liu Y, Zhu N, Yu Y, Ning Z, She T, Hu M (2023) Nrf2 signaling pathway: current status and potential therapeutic targetable role in human cancers. Front Oncol 22(13):1184079. https://doi.org/10.3389/fonc.2023.1184079.PMID:37810967;PMCID:PMC10559910

46. Shen X, Peter EA, Bir S, Wang R, Kevil CG (2012) Analytical measurement of discrete hydrogen sulfide pools in biological specimens. Free Radic Biol Med 52:2276–2283
47. Corsello T, Komaravelli N, Casola A (2018) Role of hydrogen sulfide in NRF2- and Sirtuin-dependent maintenance of cellular redox balance. Antioxidants 7:129. https://doi.org/10.3390/antiox7100129
48. Cheney G, Waxler SH, Miller IJ (1956) Vitamin U therapy of peptic ulcer; experience at San Quentin Prison. Calif Med 84(1):39–42. PMID: 13276831; PMCID: PMC1532869
49. Fahey JW, Stephenson KK, Wallace AJ (2015) Dietary amelioration of Helicobacter infection. Nutr Res 35(6):461–473. https://doi.org/10.1016/j.nutres.2015.03.001. Epub 2015 Mar 6. PMID: 25799054; PMCID: PMC4465045
50. Romeo L, Iori R, Rollin P, Bramanti P, Mazzon E (2018) Isothiocyanates: an overview of their antimicrobial activity against human infections. Molecules 23(3):624. https://doi.org/10.3390/molecules23030624.PMID:29522501;PMCID:PMC6017699
51. Kasarello K, Köhling I, Kosowska A, Pucia K, Lukasik A, Cudnoch-Jedrzejewska A, Paczek L, Zielenkiewicz U, Zielenkiewicz P (2022) The anti-inflammatory effect of cabbage leaves explained by the influence of bol-miRNA172a on FAN expression. Front Pharmacol 24(13):846830. https://doi.org/10.3389/fphar.2022.846830.PMID:35401203;PMCID:PMC8987499
52. Vidar WS, Baumeister TUH, Caesar LK, Kellogg JJ, Todd DA, Linington RG, M Kvalheim O, Cech NB (2023) Interaction metabolomics to discover synergists in natural product mixtures. J Nat Prod 86(4):655–671. https://doi.org/10.1021/acs.jnatprod.2c00518. Epub 2023 Apr 13. PMID: 37052585; PMCID: PMC10152448
53. Hanley AB, Russell WR, Chesson A (1993) Formation of substituted truxillic and truxinic acids in plant cell walls—a rationale. Phytochem 33:957
54. Dembitsky VM (2008) Bioactive cyclobutane-containing alkaloids. J Nat Med 62(1):1–33. https://doi.org/10.1007/s11418-007-0166-3. Epub 2007 Sep 5 PMID: 18404338
55. Yuan X, Liu Y, Zhao H et al (2021) The isolation, structure and fragmentation characteristics of natural truxillic and truxinic acid derivatives in Abrus mollis leaves. Phytochemistry 181:112572. https://doi.org/10.1016/j.phytochem.2020.112572. PMID: 33166750
56. Cartea ME, Francisco M, Soengas P, Velasco P (2010) Phenolic compounds in Brassica vegetables. Molecules 16(1):251–280. https://doi.org/10.3390/molecules16010251.PMID:21193847;PMCID:PMC6259264
57. Guo N, Cheng F, Wu J et al (2014) Anthocyanin biosynthetic genes in Brassica rapa. BMC Genomics 15:426. https://doi.org/10.1186/1471-2164-15-426
58. Erb M, Kliebenstein DJ (2020) Plant secondary metabolites as defenses, regulators, and primary metabolites: the blurred functional trichotomy. Plant Physiol 184(1):39–52. https://doi.org/10.1104/pp.20.00433. Epub 2020 Jul 7. PMID: 32636341; PMCID: PMC7479915
59. Chowdhury P (2022) Glucosinolates and Its role in mitigating abiotic and biotic stress in Brassicaceae. In: Plant stress physiology—perspectives in agriculture. IntechOpen. https://doi.org/10.5772/intechopen.102367
60. Schnug E (1990) Sulfur nutrition and quality of vegetables. Sulphur Agric 16:3–7
61. Plaszkó T, Szűcs Z, Vasas G, Gonda S (2021) Effects of glucosinolate-derived isothiocyanates on fungi: a comprehensive review on direct effects, mechanisms, structure-activity relationship data and possible agricultural applications. J Fungi (Basel) 7(7):539. https://doi.org/10.3390/jof7070539.PMID:34356918;PMCID:PMC8305656
62. Li S, Li X, Zhang H, Wang Z, Xu H (2021) The research progress in and perspective of potential fungicides: succinate dehydrogenase inhibitors. Bioorg Med Chem 50:116476
63. Ludwig-Müller J (2009) Glucosinolates and the clubroot disease: defense compounds or auxin precursors? Phytochem Rev 8:135–148. https://doi.org/10.1007/s11101-008-9096-2
64. Auer S, Ludwig-Müller J (2023) Biocontrol of clubroot disease: how successful are endophytic fungi and bacteria? Eur J Plant Pathol 167:433–451. https://doi.org/10.1007/s10658-023-02701-3
65. Wang Y, Zheng X, Sarenqimuge S, von Tiedemann A (2023) The soil bacterial community regulates germination of Plasmodiophora brassicae resting spores rather than root

exudates. PLoS Pathog 19(3):e1011175. https://doi.org/10.1371/journal.ppat.1011175.PMID:36862655;PMCID:PMC9980788
66. Dewachter L, Dénéréaz J, Liu X, de Bakker V, Costa C, Baldry M, Sirard JC, Veening JW (2022) Amoxicillin-resistant Streptococcus pneumoniae can be resensitized by targeting the mevalonate pathway as indicated by sCRilecs-seq. Elife 24(11):e75607. https://doi.org/10.7554/eLife.75607.PMID:35748540;PMCID:PMC9363119
67. Hopkins RJ, van Dam NM, van Loon JJ (2009) Role of glucosinolates in insect-plant relationships and multitrophic interactions. Annu Rev Entomol 54:57–83. https://doi.org/10.1146/annurev.ento.54.110807.090623. PMID: 18811249
68. Yang ZL, Nour-Eldin HH, Hänniger S et al (2021) Sugar transporters enable a leaf beetle to accumulate plant defense compounds. Nat Commun 12:2658. https://doi.org/10.1038/s41467-021-22982-8
69. Sun R, Jiang X, Reichelt M et al (2021) The selective sequestration of glucosinolates by the cabbage aphid severely impacts a predatory lacewing. J Pest Sci 94:1147–1160. https://doi.org/10.1007/s10340-020-01319-2
70. Körnig J, Ortizo K, Sporer T, Yang ZL, Beran F (2023) Different myrosinases activate sequestered glucosinolates in larvae and adults of the horseradish flea beetle. Insect Biochem Mol Biol 163:104040. https://doi.org/10.1016/j.ibmb.2023.104040. Epub 2023 Nov 22 PMID: 37995833
71. Whiteman NK, Peláez JN (2021) Taste-testing tarsi: gustatory receptors for glucosinolates in cabbage butterflies. PLoS Genet 17(7):e1009616. https://doi.org/10.1371/journal.pgen.1009616
72. Yang J, Guo H, Jiang N-J, Tang R, Li G-C, Huang L-Q et al (2021) Identification of a gustatory receptor tuned to sinigrin in the cabbage white butterfly Pieris rapae. PLoS Genet 17(7):e1009527
73. Wittstock U, Agerbirk N, Stauber EJ, Olsen CE, Hippler M, Mitchell-Olds T et al (2004) Successful herbivore attack due to metabolic diversion of a plant chemical defense. Proc Natl Acad Sci U S A 101:4859–4864. https://doi.org/10.1073/pnas.0308007101
74. Felker P, Bunch R, Leung AM (2016) Concentrations of thiocyanate and goitrin in human plasma, their precursor concentrations in brassica vegetables, and associated potential risk for hypothyroidism. Nutr Rev 74(4):248–258. https://doi.org/10.1093/nutrit/nuv110. Epub 2016 Mar 5. PMID: 26946249; PMCID: PMC4892312
75. Truong T, Baron-Dubourdieu D, Rougier Y, Guénel P (2010) Role of dietary iodine and cruciferous vegetables in thyroid cancer: a countrywide case-control study in New Caledonia. Cancer Causes Control 21(8):1183–1192. https://doi.org/10.1007/s10552-010-9545-2. Epub 2010 Apr 2. PMID: 20361352; PMCID: PMC3496161
76. Ruan D, Yang J, Luo Q, Shi Y, Ding L, Wang Z, Wang R, Yang L (2023) The protective effects of goitrin on LPS-induced septic shock in C57BL/6J mice via caspase-11 non-canonical inflammasome inhibition. Molecules 28(7):2883. https://doi.org/10.3390/molecules28072883.PMID:37049646;PMCID:PMC10096381
77. Nie L, Wu Y, Dai Z, Ma S (2020) Antiviral activity of Isatidis radix derived glucosinolate isomers and their breakdown products against influenza A in vitro/ovo and mechanism of action. J Ethnopharmacol 251:112550
78. Bahoosh SR, Shokoohinia Y, Eftekhari M (2022) Glucosinolates and their hydrolysis products as potential nutraceuticals to combat cytokine storm in SARS-COV-2. DARU J Pharm Sci 30:245–252. https://doi.org/10.1007/s40199-022-00435-x
79. Yang G (2020) H2S as a potential defense against COVID-19? Am J Physiol Cell Physiol 319(2):C244–C249. https://doi.org/10.1152/ajpcell.00187.2020. Epub 2020 Jun 9. PMID: 32515982; PMCID: PMC7381713
80. Lee RJ, Cohen NA (2015) Role of the bitter taste receptor T2R38 in upper respiratory infection and chronic rhinosinusitis. Curr Opin Allergy Clin Immunol 15(1):14–20. https://doi.org/10.1097/ACI.0000000000000120.PMID:25304231;PMCID:PMC5902169
81. Lee RJ, Xiong G, Kofonow JM, Chen B, Lysenko A, Jiang P, Abraham V, Doghramji L, Adappa ND, Palmer JN, Kennedy DW, Beauchamp GK, Doulias PT, Ischiropoulos H, Kreindler JL,

Reed DR, Cohen NA (2012) T2R38 taste receptor polymorphisms underlie susceptibility to upper respiratory infection. J Clin Invest 122(11):4145–59. https://doi.org/10.1172/JCI64240. Epub 2012 Oct 8. PMID: 23041624; PMCID: PMC3484455

82. Gaida MM, Mayer C, Dapunt U, Stegmaier S, Schirmacher P, Wabnitz GH, Hänsch GM (2016) Expression of the bitter receptor T2R38 in pancreatic cancer: localization in lipid droplets and activation by a bacteria-derived quorum-sensing molecule. Oncotarget 7(11):12623–12632. https://doi.org/10.18632/oncotarget.7206.PMID:26862855;PMCID:PMC4914309
83. Stern L, Giese N, Hackert T, Strobel O, Schirmacher P, Felix K, Gaida MM (2018) Overcoming chemoresistance in pancreatic cancer cells: role of the bitter taste receptor T2R10. J Cancer 9(4):711–725. https://doi.org/10.7150/jca.21803.PMID:29556329;PMCID:PMC5858493
84. Ruhee RT, Suzuki K (2020) The integrative role of sulforaphane in preventing inflammation, oxidative stress and fatigue: a review of a potential protective phytochemical. Antioxidants (Basel) 9(6):521. https://doi.org/10.3390/antiox9060521.PMID:32545803;PMCID:PMC7346151
85. WHO cooperative trial on primary prevention of ischaemic heart disease with clofibrate to lower serum cholesterol: final mortality follow-up. Report of the Committee of Principal Investigators. Lancet 2(8403):600–604
86. Fears R, Richards DH, Ferres H (1980) The effect of compactin, a potent inhibitor of 3-hydroxy-3-methylglutaryl coenzyme—a reductase activity, on cholesterogenesis and serum cholesterol levels in rats and chicks. Atherosclerosis 35(4):439–449
87. Proute MC, Kothur N, Georgiou P, Serhiyenia T, Shi W, Kerolos ME, Pradeep R, Akram A, Khan S (2021) The effect of statin therapy on inflammatory biomarkers: a systematic review. Cureus 13(9):e18273. https://doi.org/10.7759/cureus.18273.PMID:34722051;PMCID:PMC8545535
88. Vidar WS, Baumeister TUH, Caesar LK, Kellogg JJ, Todd DA, Linington RG, Kvalheim OM, Nadja B (2023) Interaction Metabolomics to Discover Synergists in Natural Product Mixtures. Cech J Nat Prod 86(4):655–671. https://doi.org/10.1021/acs.jnatprod.2c00518
89. Pritzker KP (2015) Predictive and prognostic cancer biomarkers revisited. Expert Rev Mol Diagn 15(8):971–974. https://doi.org/10.1586/14737159.2015.1063421. Epub 2015 Jul 1 PMID: 26134385
90. Brown AW, Aslibekyan S, Bier D, Ferreira da Silva R, Hoover A, Klurfeld DM, Loken E, Mayo-Wilson E, Menachemi N, Pavela G, Quinn PD, Schoeller D, Tekwe C, Valdez D, Vorland CJ, Whigham LD, Allison DB (2023) Toward more rigorous and informative nutritional epidemiology: The rational space between dismissal and defense of the status quo. Crit Rev Food Sci Nutr 63(18):3150–3167. https://doi.org/10.1080/10408398.2021.1985427. Epub 2021 Oct 22. PMID: 34678079; PMCID: PMC9023609
91. Harris JE, Sheean PM, Gleason PM, Bruemmer B, Boushey C (2012) Publishing nutrition research: a review of multivariate techniques–part 2: analysis of variance. J Acad Nutr Diet 112(1):90–98. https://doi.org/10.1016/j.jada.2011.09.037. Epub 2011 Dec 22 PMID: 22709639

Chapter 4
The Biosynthesis of Natural Products

Abstract The study of biosynthesis is an important aspect of answering two questions that have preoccupied chemists. Where do natural products come from? Why are they produced? The development of ideas from molecules to genes has helped to answer these. Atomic theory from Plato to Democritus was a key pre-requisite. The study of whole, complex systems is challenging. Preformation was developed as a concept of how living systems are formed but it was succeeded by epigenesis. Macrovisualisation of the properties of individual plants is key to hybridisation and crop improvement. The development of the Signaturist approach to plants and disease suggested that remedies were best derived from plants that resembled target organs. The use of 'affinities' was later used to explain how atoms combine. The mechanics of biosynthesis and the involvement of genes can be explained using polyketides as an example of biosynthesis and biosynthetic gene clusters. Manipulation of biosynthesis leads to the development of directed evolution. The growth in synthetic biology for the production of valuable molecules can be studied using basic engineering principles.

Keywords Biosynthesis · Polyketides · Gene cloning · Gene expression

Elucidation of the biosynthesis of natural products has been an important part of chemistry over an extended period of time. Chemists have been intrigued by the study of where molecules come from and how they are assembled. They have also looked for similarities and patterns in how structures are assembled *in vivo* and this has helped them to determine the purpose and workings of living systems. The development of better and more specific techniques, the ability to examine biosynthetic pathways in terms both of the substrates and products and the biological facilitators of these reactions has contributed significantly to the understanding of biological systems and to the greater understanding of chemistry and chemical systems. Most recently, the study of biomimetics has led chemists to look at novel, bio-inspired materials and systems [1]. In addition, speculation concerning biosynthesis has aided biomimetic design and the development of new reaction pathways. Finally, biosynthesis has allowed chemists to manipulate natural products and to understand better

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_4

the interactions between living systems by probing how molecular changes in natural products affect outcomes for living systems.

4.1 The Principles of Natural Product Biosynthesis

It is part of the scientific enquiring mind to wonder where things that you find in the world around you actually come from and how they are made. This applies to the universe as a whole, to humans as individuals (and as a society) and to the things we see around us including plants and other living things. The explanation of where living systems (and their components) come from can be thought of both in terms of classification and historical lineage and of molecular events. In the case of the former, studies have led to theories of evolution and in the case of the latter to an understanding of how molecules are created. The connection between the two is the discovery (and later exploitation) of inheritable genes. It is genes that provide the link between molecular events—which we cannot see but can extrapolate towards—and inheritance that we cannot necessarily follow because of the timescales involved and the complexity of the changes. As we have seen, natural product chemistry is a microcosm of this. In one direction, we cannot see but must find supporting evidence for the molecular interactions that give rise to natural products and on the other we are only able to probe some of the simplest parts of living systems and not the complete features of highly complex and interdependent living systems.

Much of the work on the evolution and development of natural products has concentrated upon the search for bioactive principles [2]. These bioactive natural products are mined for compounds that will be of value. An alternative approach, which is less human-centric, is to consider compounds in the context of those which are essential to the growth of the organism that created them (primary metabolites) and those which have other, specialised purposes (improve the fitness or survival of the species—secondary metabolites). Each group are defined as natural products. Natural products are created as a result of the sequential modification of molecules by enzymes that themselves coded for by genes. In order for a complete sequence of reactions to take place in vivo the enzymes responsible must all be present. This is made possible by the presence of gene clusters so the link between evolutionary, genetic events and molecular structures can at least be envisaged. In short, enzymes create molecules and these molecules confer some advantage to the producing species. The survival of this 'fitter' species means that further modifications to clusters of biosynthetic genes can take place in these species and the selection process is repeated and modified natural products created. The links between genes, natural products and selection are crucial to understanding the role of natural products and to building a picture of chemical (as well as biochemical, ecological and biological) evolution and the understanding that allows chemists to identify approaches to novel natural product systems as well as modified core structures. The example of polyketides will be used to illustrate the evolution of thought and how this has led to a more substantial appreciation of molecules, living systems and their interaction and to the

sequence of natural product discovery, synthetic chemistry and the engineering of biological systems to act as systems to create molecules and value.

4.2 The Development of Natural Products—From Molecules to Genes

The classical world of natural products was dominated by the thinking of the relationship between the external properties of molecules and their molecular structure. This seemed like a logical progression and, from a reductionist standpoint, made sense. However this view was challenged by Plato who proposed that indivisible atoms were all composed of two types of triangles. Other Greek philosophers, including in particular, Leucippus and his pupil Democritus, developed the idea of there being both atoms and voids (the spaces between atoms). Such thinking was crucial to the development of atomic theory. Leucippus, Democritus and Epicurus also developed the idea that atoms reflected the nature of the substances that were composed of them. In this way of thinking, water atoms, for example, were smooth and slippery while iron atoms were solid and strong and (importantly) joined together in a strong matrix. Not surprisingly the idea of atoms and substances as representations of each other also appealed to Descartes. When this was also applied to living systems it appeared axiomatic that whole organisms grew from small but complete versions of the organism. This was called preformation. It became a significant contribution to embryology during the Enlightenment and continued throughout the seventeenth and eighteenth centuries. It could be thought of as the atomic theory of Descartes, interpreted in terms of living systems. The use of microscopes and, with them, the ability to see inside cells at the earliest stages of development was a significant reason for the loss of credibility of Preformatist thinking. The alternative view was that a living organism was formed, part by part. This became known as epigenesis and a major proponent of this way of thinking was William Harvey [3]. Epigenesis proposes that organs are formed gradually from undifferentiated material. This has some echoes in the idea of pluripotent cells and in stem cells. Descartes also belonged to the epigenesis school although many of his followers and adherents were preformationists. The Preformatist way of thinking also had some echoes in the ideas of the signaturists. In essence, signaturists (like Paracelsus) thought that specific plants physically resembled the part of the body they could cure. This link between the macroscopic properties of a whole organism (a plant) and its site of action in the body has some analogy with the idea that an atom should resemble the complete substance that it forms. While the concept of signatures has generally been discredited, the idea of relating physiological processes to the physiology of disease has been suggested [4]. The scientific validity of this concept remains questionable. As noted previously, however, a purely reductionist view of living systems might suggest that the properties of a whole living system might be deduced from the sum of the individual components. This ‘mechanical’ view of life has persisted—particularly as vitalism

has become a less credible alternative. This still leaves the question of the existence of some 'divine agency' (a celestial 'maker') as a plausible explanation but this is one that has been rejected by science (on the whole) in favour of thinking of living systems as active, self-making machines and as being the source of their own agency.

Subsequent deliberations suggested a 'corpuscular' theory for molecules and atoms which was plausibly expressed by Robert Boyle. In addition to developing chemistry as an empirical scientific endeavour, Boyle also rejected both alchemy and the '*tria prima*' of salt, sulphur and mercury. Boyle's theory, which built upon the Cartesian mechanical ideas, held that molecules were composed of clusters of particles and that chemical change came about through rearrangement of these clusters. Etienne Francois Geoffroy took this a stage further and composed tables of 'affinities'. These tables tried to explain the empirical observation that, if you mix two substances together, sometimes you will produce something different from the two starting materials. The tables of 'affinities' were underpinned by experiments in which substances were mixed and their affinity for each other recorded. This is not so distantly removed from alchemy however it provided greater evidence of how substances reacted together. Indeed, such thinking formed the basis for ideas about how molecules reacted together to give substances with different properties. It was against this background that natural product chemistry began to move away from alchemy and towards the ideas of molecules and compounds. This work preceded the development of atomic theory by Dalton and molecules became recognised as containing different atoms in different proportions. While such ideas were capable of being interpreted in quite simple terms for inorganic compounds, most organic compounds (and what came to be known as natural products), contained a limited number of atoms (carbon, hydrogen and oxygen along with some that contain nitrogen and phosphorous) and the proportions of the different atoms present was such that compounds with the same relative proportions of constituent atoms could differ markedly from one another. This encouraged chemists to think in terms of three dimensional structures and different types of bonds. However this advance did not come until the early part of the twentieth century. Indeed, it was only in 1873 that James Clerk Maxwell, in an important paper in Nature, stated that "An atom is a body which cannot be cut in two; a molecule is the smallest possible portion of a particular substance." The following year, a better understanding of the tetrahedral nature of the bonding in the fully saturated carbon nucleus was fully appreciated [5] through the investigation of stereoisomers although this had already been anticipated by Kekule in 1858 [6], Others (for example, Couper and Crum Brown) were also active in the area at this time but the principle of a tetrahedral carbon atom met with significant opposition. Chemists continued to think in terms of balancing the proportions of different types of atoms in order to create molecules of interest (see Perkin, below). This is, in essence, an affinity-driven approach. W H Perkin first synthesised the dyestuff mauve while trying to make quinine from aniline in 1856. At that time, the structural characteristics of many natural products was little understood beyond the calculation (by degradative methods) of the proportions of different types of atoms in molecules and this was the approach that Perkin took. Perkin started with

allyl toluidine (formula $C_{10}H_{13}N$). By taking two equivalents of this readily available compound and adding three equivalents of oxygen, he hoped to create quinine ($C_{20}H_{24}N_2O_2$) and to release one equivalent of water. This approach, lacking knowledge of actual chemical structures, was a failure however Perkin created the first synthetic dyestuff instead (mauveine). One of the most significant advances was the proof, by x-ray crystallography, of the planar structure of benzene (actually hexamethylbenzene) by Kathleen Lonsdale [7]. The development of x-ray crystallography provided the direct proof of the structure of organic compounds rather than proof by inference.

The story of the discovery and subsequent development of quinine has all the components of, in many ways, a typical natural product evolution and development. It was originally, apparently a serendipitous discovery that the bark of what came to be known as the Cinchona tree had fever-relieving properties. This discovery was made by the indigenous population in South America and knowledge of this potent medicinal bark was brought to Europe. The rise of malaria had all the characteristics of a pandemic. It appears to have started as a zoonotic disease in Africa and subsequently spread through southern Europe and eventually as far north as Britain and Denmark. Malaria was brought to the Americas by Europeans and the 'fever tree', later called Cinchona, was an obvious remedy based upon its known fever-relieving properties. Initially the crude bark was powdered and suspended in a suitable liquid (often wine) however quinine itself was extracted from the bark in 1820 and the application of measured amounts of the compound could then be carried out. Interestingly Pelletier and Caventou, who carried out the extraction to isolate the active compound as a hydrochloride, called it quinine rather than cinchonine [8]. Synthetic analogues (particularly chloroquine—Fig. 4.1) became the more commonly used form until resistance built up and, in some cases, reversion to the parent compound, quinine was favoured as a treatment.

Quinine is commonly used in combination with an antibiotic for the management of malaria. The discovery and use of artemesin-based combination therapies has

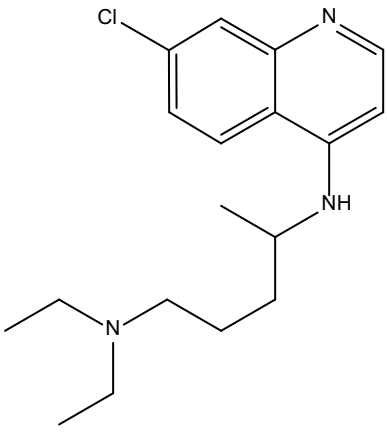


Fig. 4.1 Structure of chloroquine

replaced quinine as the first choice as an anti-malarial strategy. Artemesin, which was isolated in the 1970s, is a sesquiterpene lactone from the plant *Artemesia annua* [9]. While the complete synthesis of quinine has been achieved (most notably and, briefly, controversially by Woodward), the synthetic route has not been adapted for commercial production and extraction from a natural source remains the preferred method of production.

While sesquiterpenes and polyketides have a different biogenesis, at least one group of isolated sesquiterpenes have been suggested to derive from the polyketide biosynthetic pathway [10]. The biosynthesis of polyketides, which will be elaborated upon later, will provide further illumination about the importance of biosynthesis and the development of genetic pathway relationships in developing an understanding of natural product biosynthesis. It should be appreciated, however, that the structural relationships between natural products and suggestions about potential interrelationships between them has often been based upon an imperfect understanding of molecules and how they might be synthesised in nature compared to in the chemistry laboratory. This was a communication, however that ran in both directions. As techniques, reagents and methods were developed, the synthesis of molecules in the laboratory suggested routes that may apply in nature while biomimetic synthesis led to the development of novel synthetic approaches in the laboratory [11]. The synthesis of tropinone by Robinson in 1917 has been widely acknowledged as the first biomimetic synthesis [12]. Robinson reacted succinaldehyde with methylamine and acetone in a one pot synthesis to give tropinone. This was not the first synthesis of tropinone but the earlier synthesis by Willstatter was not biomimetic [13] and started from cyclohexanone in a multiple step series of reaction with an overall yield of < 1%. In the production of tropane alkaloids in plants, the involvement of a polyketide synthase has been proposed [14].

One of the most valuable tools in the determination of biosynthetic pathways is the analysis of gene clusters. Gene clusters are co-expressed genes which are frequently used in the biosynthesis of a range of natural products including polyketides, alkaloids and terpenes. The analysis of gene clusters has led to the elucidation of the biosynthetic routes to a number of secondary metabolites [15]. The evolution of thinking in natural products has, then, developed from an awareness of the structural relations of natural products to an understanding of how molecules are created *in vivo,* the mimicking of such reactions in the laboratory, an understanding of how enzymes act to catalyse biosynthetic reactions [16] and the analysis of groups and clusters of genes that produce the enzymes that constitute a pathway. Each of the steps in the development of knowledge, from molecules to enzymes to mechanisms to pathways and to genes requires the application of techniques whose progression charts the growing understanding of natural products and natural product chemistry in living systems and the increasing role of cross-disciplinary progress and the increased application of biological molecular techniques.

4.3 The Biosynthesis of Polyketides. A Paradigm for Progress of Natural Product Chemistry

Polyketides are natural products that occur extensively throughout nature. While they are produced, most significantly, by bacteria and fungi, they are also found in plants, insects and lichens, amongst other species. They are the largest class of natural products and encompass a range of structural characteristics. Their utility as antibiotics has fuelled investigations into their chemistry and biology. In structural terms polyketides comprise alternating ketone and methylene moieties. Collie first described compounds composed of 'multiple keten groups' (–CH2–CO– units) in 1907 [17]. The biosynthesis of polyketides arises from the decarboxylation and condensation of acyl-CoA subunits catalysed by polyketide synthase enzymes. The biosynthesis of polyketides is similar to that of fatty acids and will be discussed in more detail below. The class of polyketides includes a range of biologically active molecules including antibiotics, statins, insecticides and mycotoxins. The range of biological activities is a consequence of their structural diversity.

The biosynthesis of polyketides follows the route of other natural products—genes produce enzymes that catalyse reactions giving rise to the next molecule in the biosynthesis—and the details of polyketide biosynthesis demonstrates the advances in chemistry that enabled a deeper understanding of natural products to be developed. The process starts with the isolation of members of the class and a recognition of structural characteristics that provide a clue as to their likely biogenesis. This suggests a range of potential precursors whose incorporation into the natural product can be confirmed by adding isotopically labelled putative precursors to the organism at an appropriate time (when secondary product synthesis is taking place) and examining the product for incorporation of the labelled compound. The position of incorporation of the labelled atoms can give further mechanistic information and a biosynthetic scheme may be proposed. Multiply labelled precursors can indicate if specific bonds are formed, retained or broken during the biosynthesis. The use of advanced techniques including metabolomics can map labelling patterns with known biosynthetic gene clusters [18]. The specific case of polyketide biosynthesis is discussed later. The application of reductionist techniques, in which biosynthetic schemes are broken down into individual reactions, enables natural product chemists to study individual steps in the biosynthetic pathway and to isolate enzymes that are responsible for catalysing these reactions. Such detailed mechanistic allows chemists to develop new synthetic chemistry via biomimetic routes to novel polyketides and, in some cases, to introduce novel starting materials which may be acted upon by the enzymes to give 'unnatural' or pseudo natural products [19]. As molecular biology techniques are developed, the genes for enzymes can be cloned and larger amounts expressed in suitable hosts. Finally, the application of synthetic biology allows designer polyketides to be produced either using novel combinations of natural enzymes or modification of natural enzymes themselves to engineer the production of novel compounds. These steps reflect the development of new techniques and their application in natural product chemistry and will be discussed in more detail later.

The biosynthesis of polyketides is a modular process. Polyketide synthases are the enzymes that catalyse the formation of a linear carbon chain. They have ketosynthase activity and this catalyses the extension reaction between a starter unit (or a growing polyketide chain) and extender units. Subsequent functionalisation at the alpha or beta carbon atoms of the growing linear chain leads to the formation of the structurally diverse polyketides. A range of modifications to the linear chain including cyclisation, oxidation, reduction, transesterification and other biotransformations can occur. The growing chain can then be subjected to intramolecular cyclisation to give a range of monocyclic or polycyclic products.

Prior to investigating the potential mechanism of formation of natural products, the molecular biogenesis of the molecules has to be ascertained. In the case of polyketides, initial studies on their biosynthetic derivation were carried out using 14-C labelled precursors. These demonstrated that the polyketide backbone was derived from acetate groups [20]. This provides a link between a primary metabolite (acetate) and secondary metabolism. The discovery of the likely biogenesis of polyketides allowed chemists to speculate upon the likely mechanism of formation. As noted above, the enzymes involved (polyketide synthases) catalyse the condensation of activated primary metabolites (e.g. acetyl CoA and malonyl CoA) to form beta keto acetyl polymers which are linked to an acyl carrier protein (ACP) via thio ester bonds (Fig. 4.2).

It was recognised that polyketide synthases and fatty acid synthases are functionally similar enzymes. The differences arise from processing that occurs after the initial condensation. The genetic basis of the biosynthesis of polyketides has been extensively reviewed and the interested reader is directed towards publications by Hopwood, Staunton, Simpson and others [21]. While a detailed description of the

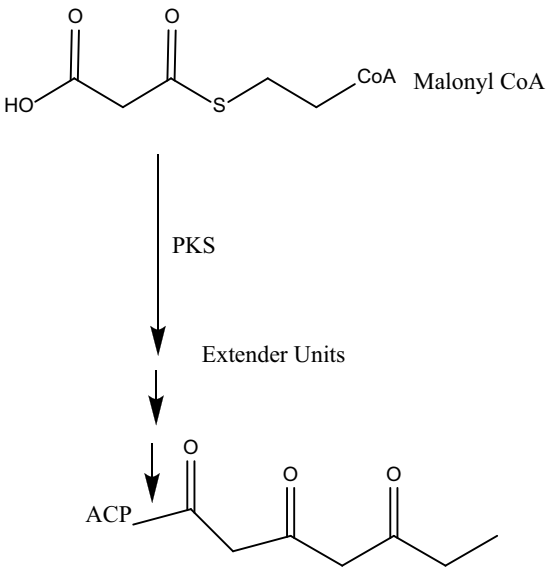


Fig. 4.2 Summary of polyketide biosynthesis

biosynthesis of polyketides and the enzymes that are involved it outside the scope of this chapter, one underlying principle is important. The type of reaction catalysed by the enzyme complex is analogous to reaction pathways employed by synthetic chemists.

While chemistry in general, and organic chemistry in particular, can give rise to very complex structures, these are created by a relatively small number of reactions. Enzyme reactions are also limited in scope however even reactions such as cycloaddition reactions which were formerly thought not to have a biological equivalent, have been found to be catalysed in nature by a group of enzymes called pericyclases [22]. The role of enzymes in biosynthesis has been extensively reviewed and is the subject of much ongoing investigation [16]. Manipulation of the genes involved in polyketide biosynthesis has enabled the planned creation of novel polyketides. This is a consequence of the steps involved in polyketide biosynthesis. Briefly, as noted previously, the process of formation of polyketides comprises three steps—initiation, extension and then functionalisation. Polyketide synthases are modular enzymes and the addition or deletion of some of these three modules can give rise to novel molecules. In essence polyketide synthases operate as an assembly line and, depending on which modules are present (and in what order) different molecules are created. This process has been reviewed and advances are ongoing [23]. The modular assembly characteristic is responsible for the large number of known natural polyketides (> 10,000 members of the class are known) and in the potential for production of unnatural polyketides. The classification of natural products is frequently a consequence of a similarity in structural characteristics or in biosynthesis. Much of the description of natural product chemistry is, therefore predicated upon the idea that molecules can be grouped together based upon molecular structure or biogenesis or a combination of the two approaches. In the case of polyketides, classification is possible based upon structural characteristics (polyethers, lactones, aromatic, polyenes etc.) or on the basis of which of three different types of polyketide synthase are implicated in their biosynthesis. There are three classes of polyketide synthase with different characteristics and while the types of polyketide produced by each is not absolutely definitive, the '3 types' paradigm for polyketide synthases has provided a useful general classification. In addition, while typically natural product chemists will seek to identify new structures and work from them towards an understanding of how they might be formed by rationalising their structural diversity, the identification of novel polyketide synthases is an alternative (biological-based) route to discovering chemical diversity [24]. Such approaches, as noted previously, have also been used for genome mining for biosynthetic gene clusters. This can facilitate the identification of biosynthetic potential that was previously unrealised [25]. While genome mining is a useful tool in the identification of 'hidden' biosynthetic potential and may lead to the optimisation of production of novel products, it relies on the amino acid (and hence gene sequence) of functionally similar enzymes having a high degree of homology. Later in the chapter, the modification of genes (and expression systems) will be considered. Polyketides are modifiable at the enzyme level by manipulation of complete modules rather than changing (mutating) a single base pair in a gene (amino acid modification) level. The genetic modification of polyketide synthases

is challenging. One reason for this is a consequence of the likely evolution of the polyketide synthase clusters. It is possible that the multiple modules of the different polyketide synthase modules arise from gene duplication events [26]. This means that the modules in the clusters of the enzyme complex are highly homologous. Attempts to genetically modify some of the clusters in order to create novel molecules have been challenging since any manipulation is likely to affect multiple sites. The use of specific restriction enzyme recognition points and recombination-based methodology to achieve 'accelerated evolution' of biosynthetic enzymes and hence novel products [27] have been employed with varying degrees of success. Unsurprisingly, site directed mutagenesis (see Sect. 4.5) has been used to manipulate polyketide synthases [28] as an alternative to domain swapping. One route with great potential due to its intrinsically greater level of specificity, is the use of CRISPR technology via a Cas9 system to engineer a specific polyketide synthase gene in order to produce novel analogues of known molecules [29]. The essence of the CRISPR/Cas9 technology is that a small piece of RNA (referred to as a guide RNA or sgRNA) directs the Cas9 nuclease to bind to and cleave at the target site causing a double strand break. Following this a piece of DNA can be specifically introduced or the ends of the DNA can be repaired with concomitant deletion or insertion at the repair site which causes disruption of the function of the target gene. The recognition, cleavage and repair sequence means that a very specific type of genetic modification can take place and only the target gene is affected. In the wild, CRISPR is part of the adaptive immune system of prokaryotes against viruses due to its capability to cleave (and effectively silence) viral DNA. Another application of CRISPR technology is in the activation of silent gene clusters. In essence, this technique involves the specific and strategic 'knock-in' of a promoter leading to the production of novel polyketide metabolites [30]. Such techniques are becoming increasingly sophisticated and will lead to the *de novo* synthesis of novel compounds with a range of activities.

The focus on *Streptomyces* as a producer of secondary metabolites was in large part a consequence of their capacity to produce antibiotics. In addition, each *Streptomyces* species contains about 30 secondary metabolite biosynthetic gene clusters. This number exceeds the number of isolated natural products from many Streptomyces species and this has led to a realisation that there is significant untapped potential within the genome of *Streptomyces*—there are gene clusters whose products have not been isolated or assessed for beneficial properties [31]. In addition, it is also important to recognise that different members of the species are able to make antibiotics with different properties. The first medicinal use of a *Streptomyces*-derived antibiotic was the treatment of tuberculosis by streptomycin in the 1940s [32] and this acted as a spur for other detailed investigations of the whole *Streptomyces* group. As more molecular tools became available, *Streptomyces* became the organism of choice both because of the range of natural products offered and the accumulation of more knowledge about the species. This is a type of self-perpetuating, 'virtuous circle'. While Streptomyces was found to be a very suitable organism for the study of the genetics of secondary metabolite formation and significant advances were made, its versatility and biological importance has also been noted in other areas. For example, a strain of *Streptomyces* has been found to contain compounds

with quorum sensing inhibition activity [33]. Such activity may make an important contribution to the growth advantage for *Streptomyces* species. In *Streptomyces*, quorum sensing is, at least in part, mediated by a series of butyrolactones. Gamma-butyrolactones are also known to influence antibiotic production in *Streptomyces* species and additionally occur in some other Actinobacteria [34]. In addition (and perhaps not surprisingly), gamma-butyrolactone can also bind to the acyl homoserine (AHS) receptor of some Gram negative bacteria so a quorum signalling molecule produced by a Gram positive strain can influence natural product synthesis by Gram negative bacteria [35]. This type of microbial cross talk has become an increasingly apparent feature of microbiomes. *Streptomyces* themselves have, as has been noted been the subject of intense scientific investigation and this has increased the number of unanswered questions regarding the species. While they are classified as Gram positive bacteria, they exhibit mycelial growth, like typical fungi and produce (or have the potential to produce) significant numbers of natural products. While, to a natural product chemist, the plethora of compounds produced and the way they were biosynthesised was of particular interest, *Streptomyces* also provided a significant route into the study of microbial genetics. A significant amount of this work was carried out by Hopwood and collaborators and some of the highlights of this work has been reviewed [36]. The studies of *Streptomyces* and other useful microbes has encouraged the development of a range of techniques to improve the efficiency of utilisation of microbes and their products. In addition, the study of the chemistry and biology of polyketides has increased greatly our understanding of natural products, their structures and functions.

4.4 Polyketides in Nature

As has already been noted, polyketides are biosynthetically similar to fatty acids. Indeed the enzyme fatty acid synthase is similar to a type 1 polyketide synthase. One of the key differences between a fatty acid synthase and a polyketide synthase is reflected in the nature of the final biosynthetic product. In the case of the former, the product is a fully saturated lipid. As noted previously, polyketides have a range of functionalities and structures. It is incorrect to assume that fatty acid synthase is a eukaryotic form of a polyketide synthase since polyketide synthases are found in eukaryotes. Indeed the production of polyunsaturated fatty acids in both prokaryotes and eukaryotes has been suggested to be catalysed by polyketide synthases [37]. This continues the interesting crossover between primary and secondary metabolism referred to previously in the use of acetate. Animal fatty acid synthase-like polyketide synthases (AFPKs) have been discovered and may be an important evolutionary link between fatty acids and polyketides. In addition these AFPKs seem to be implicated in the synthesis of branched hydrocarbons and pheromones [38]. Interestingly AFPKs do not occur throughout the animal kingdom and have only been found in molluscs and arthropods. In apparent contrast, polyketide synthases are found extensively in the animal kingdom including a bird polyketide synthase which is responsible for the

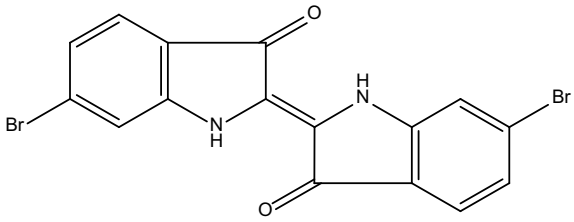


Fig. 4.3 Structure of 6,6′-dibromoindigo

synthesis of polyenes which make budgerigars green [39]. AFPKs are considered to provide a link between type 1 polyketide synthases and animal fatty acid synthases. It should be noted, however, that while identification of these synthases based upon sequence data has been carried out, the function of polyketide synthases and their products in the animal kingdom remains somewhat speculative. Nonetheless, the discovery of synthase enzymes with an apparent function beyond the biosynthesis of lipids opens the possibility of novel chemical structures being discovered from the polyketide-related biogenesis machinery. The existence of novel polyketides (for example, members of the class that includes phenolic lipids and others that use beta amino acids as a starter unit) has already stimulated the study of novel biosynthetic pathways. The occurrence of AFPKs in sea slugs was noted [40] and have been the subject of speculation as to their purpose. Members of this sacoglossan family have retained chloroplasts from algae and these remain functional. Carotenoids, which are commonly found in feathers and are not biosynthesised via a polyketide pathway, have also been found but it is unclear if these are produced *de novo* in the sea slug or simply accumulate as a result of ingestion of algae. Interestingly the colour Tyrian purple is isolated from sea snails and the principal component of the colour is 6,6′-dibromoindigo (Fig. 4.3).

Novel halogenated indole dimers have been isolated from a marine source [41]. While the likeliest source of the indole nucleus in marine and other species is tryptophan, however a tryptophan independent pathway to indole-acetic acid using chorismate being converted to anthranilate and then to indole itself has been proposed. Anthranilate can be incorporated into polyketides as a novel starter unit [42] and the possibility of a novel, polyketide-related route to some marine colour molecules may be possible.

As noted previously, polyketides are the most abundant natural products that exist. While the majority of studies (including in depth biosynthetic investigations) have been carried out in microbes and plants, there remains considerable amounts to be investigated in marine species and in animals. Much of the interest in genetic knock out (or knock in) animal models have concentrated on mimicking human disease [43]. These techniques have also increased the understanding of the consequences of single nucleotide polymorphisms (SNPs) on disease risk and progression. Other model systems have been used to investigate novel biosynthetic products from mixed metabolites. For example, Polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) hybrid systems in *Caenorhabditis elegans*, have been studied and this has shed light on the unusual nememide mixed biosynthesis product [44].

Such novel mixed biosynthetic systems may provide considerable impetus to the use of synthetic biology for the production of novel chemicals.

4.5 Mutations and Expression Systems to Effect Biosynthetic Modifications

The possibility of making use of a site-specific mutation strategy to modify polyketide synthase has been referred to previously. Site directed mutagenesis is a way of altering an amino acid at a specific point in a protein in order to alter its properties. In the case of enzymes, this can result in changes in substrate specificity and/or the nature of the product obtained. Multiple mutations can be carried out—for example using the polymerase chain reaction (PCR) [45]. Another approach which was pioneered by Frances Arnold, is the use of directed evolution. This method produces, first of all, a library of mutated proteins which can then be screened. There are a range of associated techniques including methods to incorporate all 19 natural amino acids at a specific site rather than just those which would result from a single codon change [46]. In addition to changing the properties of a protein, site directed mutagenesis can also be used to alter promoter sequences and hence expression of a given protein. While site directed mutagenesis (particularly when it incorporates creation of a library and rapid screening) has attracted significant attention, it can be a time and technique intensive operation and failures have been reported. Some of these are common and, to an extent predictable [47] while others are a result of unanticipated consequences from changes to the protein active site and can be a consequence of the nature of interactions throughout a target protein (not solely at the site of mutagenesis) [48].

The optimisation of the expression system is an important aspect of the production of a product. In most cases, the expression system is designed to maximise the production of the desired protein although other biological targets are also possible. If the appropriate protein(s) can be produced with the required functionality then the corresponding natural product will follow. There are a number of key attributes of the desired protein which must be considered. The first important consideration is whether the protein derives from a prokaryote or a eukaryote. Secondly (and often related to this) are there any post-translational modifications and if so, what is the nature of them. What is the size of the protein and finally, will it be isolated or used *in situ*. These and other questions have been summarised [49]. In addition, it is often advantageous to have inducible gene systems in order that the production of the protein or enzyme of interest can be controlled. A number of options exist to achieve this objective [50]. Finally cell-free systems are an attractive option in some cases and should also be included in any assessment of potential expression systems [51].

4.6 Synthetic Biology and Biosynthesis

While considerable efforts have led to a greater understanding of the details of the biosynthesis of polyketides, a major driver has remained the large scale production of useful polyketides (e.g. pharmaceuticals) or the investigation of novel chemical structures. A number of such studies have been enabled by the development of techniques which arise from investigations driven by synthetic biology. Some of the techniques developed are of value in the investigation and optimisation of the production of other natural (and unnatural) products. The modular nature of the biosynthesis of polyketides makes this a particularly useful system for investigation and exploitation. The progression in the optimisation of production of natural products can be summarised in terms of the following steps. Some of these have already been mentioned and further details are given below.

- Strain optimisation, development of isolation and extraction protocols.
- Genetic analysis, gene identification, isolation and cloning (including mutation and novel product formation).
- Optimisation of the expression platform.
- Development of a Synthetic Biology platform for scaled production.

One important aspect of the study of natural products is the reductionist approach to the isolation and characterisation of the target compound. There are two key steps in the effective isolation of sufficient amounts of pure samples of natural products—optimised isolation and subsequent purification. While significant advances were made—particularly in the twentieth century—in developing chromatographic and other methods for the isolation of pure compounds, the process is substantially facilitated if a good source of the natural product can be found and the target compound is present as a significant proportion of the mixture of natural products in the source. For example, classes of structurally similar alkaloids are often found to be present in the same plant species and isolation of a single member of the class can be challenging. In some cases an individual compound can be the major component present. This was found to be the case with morphine in the opium poppy which was isolated and crystallised by Serturner in the nineteenth century [52]. In many cases, however, plants and other sources of natural products contain mixtures of compounds that are structurally similar and therefore difficult to separate and therefore a range of techniques have been developed to aid the purification of individual members of the class [53]. The method of choice will depend in part on the chemical characteristics of the compounds but it may also be influenced by the techniques commonly used in the laboratory of choice. If the instrument you have is a hammer, then every challenge looks like a nail. The choice of plant part can also be important. For example, in the case of glucosinolates, there are over 100 members of the class and they occur throughout the order *Brassicales* [54]. It has been suggested that different members of the class may have different biological functions in plants. A possible consequence of this is that different members of the class occur predominately in different plants and plant parts. For example, allyl glucosinolate (sinigrin) can be

isolated from brown mustard seed (*Brassica juncea*) while cress (*Lepidium sativum*) is an excellent source of benzyl glucosinolate (glucotropaeolin). The absence of other glucosinolates from these seeds means that isolation and purification is relatively simple and complex chromatographic techniques are not necessary. In the case of the parent indole glucosinolate (glucobrassicin), relatively high levels are found in the growing apical tips of brussels sprouts and selective harvesting of this part of the plant facilitates the isolation and purification of this compound—although chromatographic steps are required [55]. The selection of a suitable strain is a very important step in the isolation and, in particular, commercialisation of production of bioactive natural products. One of the most significant examples was the isolation, commercialisation and eventual success of penicillin. The original strain that was a fungal contaminant characterised by Fleming was *Penicillium notatum.* Heatley, Chain and Florey developed a method to purify penicillin based on back extraction [56]. It quickly became clear that the amounts of penicillin produced by *Penicillium notatum* were insufficient to enable enough to be produced to be therapeutically useful. By investigating all the *Penicillium* strains available in the US Department of Agriculture collection, a better producer of penicillin (*P. chrysogenum*) was found. While this strain produced about 6 times as much penicillin as the original Fleming strain, the change to using corn steep liquor (readily available in the US) and deep tank fermentation gave a substantial increase in yield. Strain selection was clearly of great importance but in addition, optimisation of growing conditions and extraction/purification led to commercialisation of this very important antibiotic. There are numerous other examples of the optimisation of source and growth of the source of natural products. One final example is the production of indigo. The commercial production of indigo is an important industrial process and over 95% of the chemical being used in textile dyeing. Originally indigo was extracted from plants however in the nineteenth century, commercial production switched to synthetic indigo. One of the driving forces behind this was the complexity of the process but, in addition, the waste plant material after extraction must be disposed of. The synthetic route uses harsh chemicals and is not sustainable therefore a number of researchers have investigated the possibility of using a biotechnological route to indigo production [57]. Identification of the details of a biosynthetic pathway (or of other biocatalytic processes) is essential in the development of biotechnology.

There are a number of examples of the use of biotechnology to produce desired natural products. Optimisation of the process requires significant efforts and understanding. The steps involved include genetic analysis, gene identification, cloning into a suitable host, optimisation of production and isolation and (where necessary) genetic modification to create novel compounds or to optimise reaction conditions. In the case of indigo, the starting point is an understanding of how the molecule is formed. Indigo is a dimer (Fig. 4.4) which is reduced to the corresponding diol in order to enhance its solubility and enable it to be taken up by fabric. The diol, which is pale yellow, is oxidised to reform the diketone and this gives the blue colour.

Indigo itself is not found naturally and precursor glucoside monomers (indican and isatan B) are the most common of these. Cleavage of the sugar moiety of indican gives indoxyl which, in the presence of oxygen, dimerises to give indigo. The plant

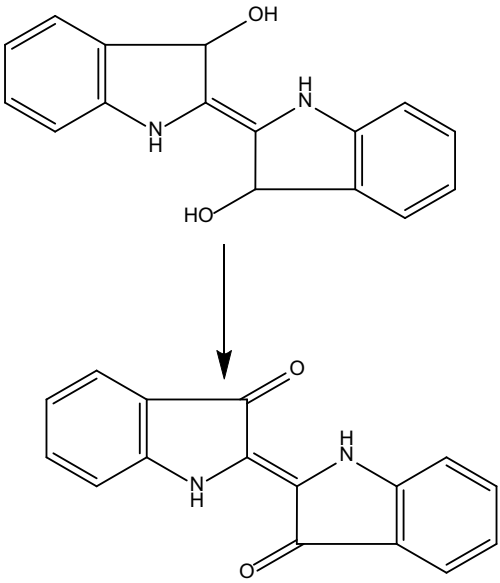


Fig. 4.4 Manipulation of indigo for dying

sources of indigo are woad (*Isatis tinctoria*) and members of the indigofera species (*Indigofera tinctoria*). The former was the most prominent source of indigo in Europe while the latter predominated in Asia. In Germany, in the Middle Ages, woad was cultivated particularly in the region of Thuringia and, in the town of Erfurt, it generated sufficient income for a university to be endowed by the merchants made rich from the proceeds of the dye industry. This was the first university to be created in what is now Germany. Eventually, since *Indigofera* was a better source of indigo, it replaced woad. While the use of plant-based processes to replace synthetic indigo production has been suggested, the amount of land required to meet the global requirement for indigo, the competing need for land to grow crops on and the large amounts of water needed for processing suggest that this route is neither economically nor environmentally viable. The chemical synthesis of indole uses aniline as a starting material. Since aniline is toxic (U.S. Department of Health and Human Services. Hazardous Substances Data Bank (HSDB, online database). (National Toxicology Information Program, National Library of Medicine, Bethesda, MD. 1993) any accidental exposure can be regarded as hazardous and this has led to investigations into potential biotechnological solutions.

The most common route to indigo involves the production of indole (often from the action of an enzyme, tryptophanase on tryptophan) and subsequent reaction catalysed by oxidoreductase enzymes to give indoxyl which will spontaneously dimerise to give indigo (Fig. 4.5).

The assay for microbial species containing the required oxidoreductase systems is a simple procedure since addition of indole to a growing culture will lead to the development of a pink or blue colour if the enzymes are present. One of the

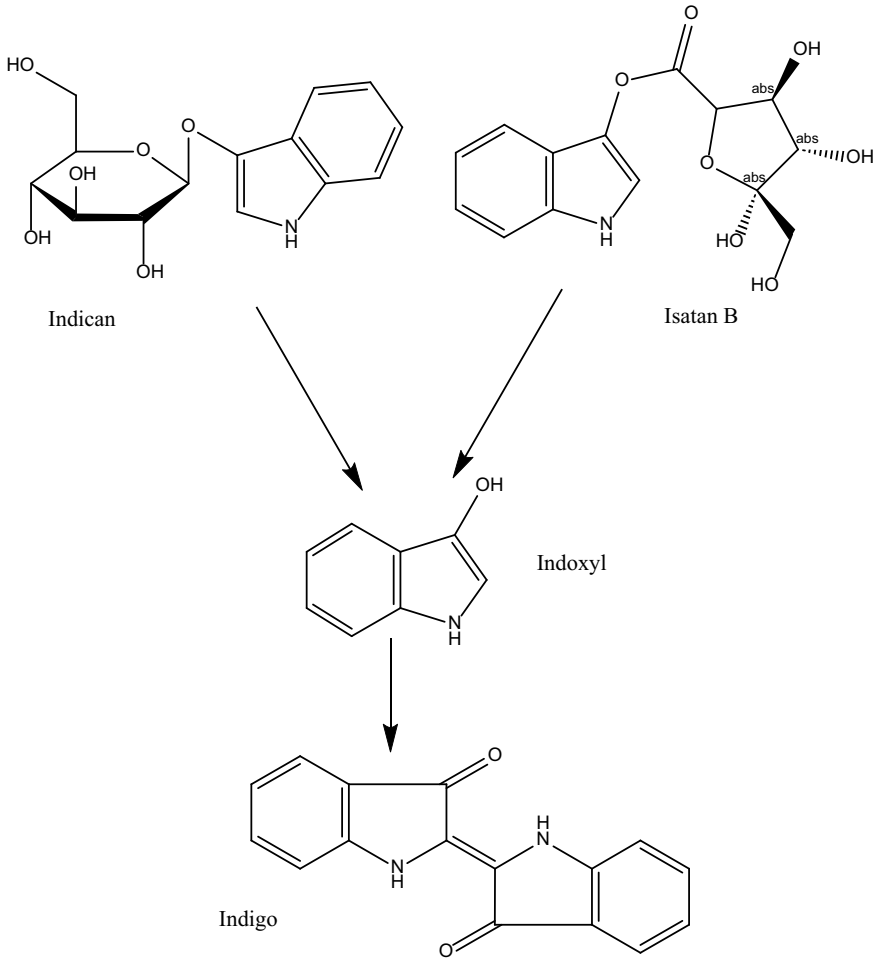


Fig. 4.5 Formation of indigo from precursors

problems with the use of oxidoreductases is the requirement for cofactors (generally NADH or something similar) and, while metabolic engineering has been carried out to try to overcome this and a complete pathway from glucose to indigo was established in E coil [58] production on an industrial scale remains challenging. The use of fermentation in indigo production has been considered to be a feasible challenge [59]. Nonetheless although the various enzyme systems that could be used for indigo production have been assessed [60], no full scale biotechnological solution has been found. An alternative would be the production of colours related to indigo but with better properties or increasing the yield of indigo by plants through strain selection or genetic engineering however such approaches do not completely overcome the drawbacks to plant-based systems mentioned above [61]. Halogenated

indole derivatives are of particular interest since the high value pigment, Tyrian Purple is composed largely of 6,6′-dibromoindigo (Fig. 4.3).

While knowledge of biosynthetic pathways is a major achievement on the route to biotechnological production, even something as simple as indigo production remains a significant challenge. Part of the reason for this is the requirement to find a suitable fermentation system. Cloning and facilitation of maximum expression of the enzymes required to ensure maximum yield of product is only a part (although a significant part) of the solution. Either a suitable host strain must be developed or a purified enzyme reaction system must be used. In the case of the latter, the enzyme is effectively being used as a catalyst and the technological challenges are broadly similar to those in process engineering. If the bioprocess requires the use of cofactors then a whole cell system may be preferable. A further hurdle to the commercialisation of production of natural products by bioprocessing is the relative costs compared to synthetic routes. In the case of high value or complex products such a pharmaceuticals or antibodies, the biological route is either preferred or obligatory. For bulk products such as indigo, bioprocess-based production is often uncompetitive.

The use of the term 'biopharmaceutical' began in the 1980s and was defined broadly in terms of the process by which they were made (products produced using biotechnological processes) rather than the conventional definition of natural products in terms of chemical structure (alkaloids, glucosinolates etc.) or biological targets (antibiotics, beta blockers etc.). They were also considered by some to be a distinct category, separate from 'biologics' which includes conventional (generally small molecule) drugs produced by biological processes. More generally, however, the terms biopharmaceutical and biologic are used interchangeably. While conventional synthetic pharmaceuticals are small molecules, biopharmaceuticals are biological macromolecules. As such, even within the same production system, batch to batch differences may arise and this is of much greater significance than for small bioactive natural products. This extends particularly to the case of preparations with identical biological properties to approved drugs. In the case of small molecules, the term 'generic drug' is commonly used. A generic drug is usually considerably easier and cheaper to bring to market and, because of the reduced cost, can be sold considerably cheaper than the original. In the case of biopharmaceuticals, the term 'generic' cannot be used. Even if the conditions for production (including the expression system, the strain identity etc.) are identical, some variation in characteristics of the biopharmaceutical may occur and so the term 'biosimilar' is used. The regulatory hurdles for biosimilars are more stringent than for generic drugs. The production of biopharmaceuticals is commonly carried out in bacterial (prokaryotic) or various eukaryotic (fungi, mammalian, insect) cell lines. The former are well characterised and are, therefore, well known as expression systems however the latter are preferred for more complex products (including those which are post-translationally modified) and for some specific classes (e.g. antibodies). The use and development of biopharmaceuticals was considerably enhanced by the development of vaccines (including those based on mRNA) for COVID -19 and the development of anti-amyloid antibodies for the treatment of Alzheimer's disease [62]. The use of biotechnology to make pharmaceuticals has grown—particularly as concerns have grown over conventional

synthetic chemistry which relies heavily on petrochemicals and (often) harsh chemical processes. This has led to the development of synthetic biology as a term used for bioengineering of platforms to produce a range of products. Synthetic biology encompasses the application of engineering design principles to biological systems in order to create valuable products and efficient processes. It can include the creation of useful biological devices as well as specific products. Among the products that have been cited as examples of the use of synthetic biology are the cloning and production of hemoglobin from legumes (leghemoglobin) to produce a vegetable-based beef burger that 'bleeds', production of pharmaceuticals by cloning biosynthetic pathways into yeast, synthesis of the diabetes medication sitagliptin using a transaminase subjected to directed evolution, synthesis of a polyimide film and a biological fertiliser based on a proteobacterium [63].

As with many scientific endeavours, the use of biological systems to facilitate natural product chemistry is, at least in part, a question of trial and error. It is important to bear in mind that, although much of biotechnology is defined in terms of the process being used (for example in biopharmaceuticals, the definitions refer to the process almost exclusively), in the case of natural product chemistry, the main driving force is the chemical and its properties rather than the process being used. The core of natural product chemistry lies in the reductionist approaches that first allowed chemistry to develop with the types of precision that were unknown to the world of alchemy. As chemists have developed (or made use of) more advanced techniques, their understanding of the chemicals they encounter as part of nature has increased. This can occur, however, at the expense of an understanding of how natural products impact on whole living systems. Chemists have tended to think of the interaction of natural products with living systems (and, in particular biological macromolecules) in terms of a 'lock and key' relationship. They have, in some ways, been encouraged in this way of thinking by the realisation that, while a large number of natural products exist with a high degree of chemical structural diversity, this vast range of products arises from a very limited set of types of reaction. The development of rapid DNA and protein sequencing methodology has enabled biologists to identify structural motifs that biological molecules have in common. This can result in counter-intuitive ways of thinking. For example, bitter receptors have been identified (and named) on the basis that they can bind bitter compounds. Bitterness (and its detection) has been assumed to be associated with toxicity and therefore bitterness contributes to learned aversion. The use of, for example, an odorant binding protein from the fruit fly linked to an electronic detection system has been used to detect bitter compounds [64]. This implies that certain molecular characteristics of bitter molecules is responsible for the sensation of bitterness and this fits with the 'lock and key' mechanism mentioned previously. There are challenges inherent in this approach. The first of these is the potential multifunctionality of biological macromolecules such as receptors and the second is illustrated by the case of allosteric enzymes. As noted previously, the detection of bitterness depends on the interaction between specific receptors and molecules that contain specific structural characteristics that facilitate their binding to the receptors. The discovery of extra-oral bitter receptors [65] raised an important question—what was their function in other parts of living system? The obvious

answer might be associated with the biological activity of bitter compounds. For example, caffeine is bitter but it also has a number of other activities including

1. Antagonism of adenosine receptors,
2. Inhibition of phosphodiesterase enzyme,
3. Calcium release from intercellular stores and
4. Antagonism of $GABA_A$ receptors.

The question is, do any of these receptor targets have characteristics in common with the bitter receptor and, indeed, with each other? This might imply that the structural aspects of parts of the caffeine molecule might be responsible both for the bitterness of caffeine and its physiological effects. If this proves to be the case and is generally applicable then the spectrum of binding of compounds of known biological activity to bitter receptors may reflect that biological activity and bitter receptors could be used as a surrogate for testing compounds of unknown biological activity. To complete the circle, while caffeine interacts with adenosine receptors, ATP signalling has a role in the detection of taste [66]. In the case of one specific bitter taste receptor, T2R38 has been associated with the biological process of ageing [67] while another bitter taste receptor (T2R14 can detect quorum sensing molecules excreted by the cariogenic bacterium *Streptococcus mutans* [68]. It has been suggested that T2R38 is the physiological receptor for acyl homoserine lactones (AHLs), an important class of quorum sensing molecules, particularly for Gram negative bacteria [69]. In this role they are important in the mediation of the innate immune system and in features such as biofilm formation and mediation of inflammation [70]. The case of allosteric enzymes illustrates another challenge to the simple 'lock and key' model for the interaction of natural products with biological systems.

Allostery, in a biological sense, is typified by action at a distance. In essence this means that an event at one place has an effect on a distant place. One of the first examples of allostery that was described was the effect of the binding of carbon dioxide on the subsequent binding of oxygen to haemoglobin [71]. Monod was the first to define allosteric as an effect that occurs when an inhibitor is not a steric analogue of a given substrate [72]. In the case of enzymes, allostery is an effect at one site in an enzyme which affects the activity at the active site. In the case of myrosinase, the enzyme responsible for the hydrolysis of the thioglucose moiety of glucosinolates, addition of ascorbic acid to the reaction substantially increases the reaction rate. It has been suggested that the binding of ascorbic acid to a position other than the active site causes a conformational change at the active site thereby exhibiting an allosteric effect [73]. Allosteric effects, whether through interactions with enzymes or with other biological macromolecules, is an important consideration for natural products. When carrying out activity-guided fractionation, it is important to be aware that some fractions may contain the active principle while others may contain an allosteric activator. This might account for the apparent 'loss' of activity during the fractionation procedure and recombination of fractions may lead to the rediscovery of activity.

Biosynthesis is an important aspect of natural product chemistry. The isolation of products and the delineation of putative biosynthetic pathways is both an intellectual

and practical challenge. Determination of pathways by using radioactive or stable isotope tracers and the synthesis of putative intermediates is a key step which may yield information both about pathways and mechanism of formation of natural products. Continuing the reductive route can lead to the isolation of individual enzymes and a more detailed study of their mechanism of action. Isolation of individual genes in a biosynthetic pathway can lead to the development of bioreactors and biotechnological approaches to the production of large amounts of valuable products. Modification of enzymes and/or the introduction of novel substrates can enable the chemist to create novel products. In the case of polyketides, the modular nature of the assembly of products facilitates the reorganisation of gene pathways to give novel products. Ultimately natural product chemistry starts with the isolation of the chemicals produced by nature, their synthesis by chemical means and ultimately, through an understanding of biological synthetic processes, to the use and manipulation of biology in order to create both valuable and novel products. By starting and finishing with natural (or unnatural) products, greater understanding and better ways of manipulating complex living systems are achieved.

References

1. Sanchez C, Arribart H, Giraud Guille M (2005) Biomimetism and bioinspiration as tools for the design of innovative materials and systems. Nat Mater 4:277–288. https://doi.org/10.1038/nmat1339
2. Barona-Gómez F, Chevrette MG, Hoskisson PA (2023) On the evolution of natural product biosynthesis. Adv Microb Physiol 83:309–349. https://doi.org/10.1016/bs.ampbs.2023.05.001. Epub 2023 May 18 PMID: 37507161
3. Goldberg B (2017) Epigenesis and the rationality of nature in William Harvey and Margaret Cavendish. HPLS 39:8. https://doi.org/10.1007/s40656-017-0134-5
4. Richardson-Boedler C (1999) The doctrine of signatures: a historical, philosophical and scientific view (I). Br Homeopath J 88(4):172–177. https://doi.org/10.1054/homp.1999.0322.PMID:10582648
5. Cintas P (2002) On the origin of tetrahedral carbon: a case for philosophy of chemistry? Found Chem 4:149–161. https://doi.org/10.1023/A:1016063004499
6. Kekulé A (1858) Ueber die Constitution und die Metamorphosen der chemischen Verbindungen und über die chemische Natur des Kohlenstoffs. Ann Chem Pharm 106(2):129–159
7. Lonsdale K (1929) The structure of the benzene ring in $C_6(CH_3)_6$. Proc R Soc A: Math, Phys Eng Sci 123(792):494–515. Bibcode:1929 RSPSA. 123..494L
8. Yip A, Sivarajah S, Natkunarajah J (2023) H22 Barking up the right tree: history of quinine. Brit J Dermatol 188(Supplement_4):ljad113.304
9. Tu Y (2016) Artemisinin—a gift from traditional Chinese medicine to the world (Nobel lecture). Angew Chem Int Ed 55(35)
10. Bansal R, Sethy SK, Khan Z, Shaikh N, Banerjee K, Mukherjee PK (2022) Genetic evidence in Favor of a polyketide origin of acremeremophilanes, the fungal "sesquiterpene" metabolites. Microbiol Spectr 10(4):e0179322. https://doi.org/10.1128/spectrum.01793-22. Epub 2022 Aug 8. PMID: 35938791; PMCID: PMC9430172
11. Bao R, Zhang H, Tang Y (2021) Biomimetic synthesis of natural products: a journey to learn, to mimic, and to be better. Acc Chem Res 54(19):3720–3733. https://doi.org/10.1021/acs.accounts.1c00459. Epub 2021 Sep 22 PMID: 34549936
12. Robinson R (1917) A synthesis of tropinone. J Chem Soc 111:762

13. Willstätter R (1901) Synthesis in the tropin groups. Liebigs Ann Chem 317:204–265
14. Huang JP, Fang C, Ma X, Wang L, Yang J, Luo J, Yan Y, Zhang Y, Huang SX (2019) Tropane alkaloids biosynthesis involves an unusual type III polyketide synthase and non-enzymatic condensation. Nat Commun 10(1):4036. https://doi.org/10.1038/s41467-019-11987-z.PMID:31492848;PMCID:PMC6731253
15. Chen R, Wong HL, Burns BP (2019) New approaches to detect biosynthetic gene clusters in the environment. Medicines (Basel) 6(1):32. https://doi.org/10.3390/medicines6010032.PMID:30823559;PMCID:PMC6473659
16. Dickschat JS (2022) Enzymes in biosynthesis. Beilstein J Org Chem 30(18):1131–1132. https://doi.org/10.3762/bjoc.18.116.PMID:36105723;PMCID:PMC9443342
17. Collie JN (1907) CLXXI—derivatives of the multiple keten group. J Chem Soc Trans 91:1806–1813. https://doi.org/10.1039/CT9079101806.ISSN0368-1645
18. McCaughey CS, van Santen JA, van der Hooft JJJ et al (2022) An isotopic labeling approach linking natural products with biosynthetic gene clusters. Nat Chem Biol 18:295–304. https://doi.org/10.1038/s41589-021-00949-6
19. Asai T, Tsukada K, Ise S et al (2015) Use of a biosynthetic intermediate to explore the chemical diversity of pseudo-natural fungal polyketides. Nature Chem 7:737–743. https://doi.org/10.1038/nchem.2308
20. Birch AJ, Donovan FW (1953) Studies in relation to biosynthesis. I. Some possible routes to derivatives of orcinol and phloroglucinol. Aust J Chem 6:360–368
21. Hopwood DA (1997) Genetic contributions to understanding polyketide synthases. Chem Rev 97(7):2465–2498. https://doi.org/10.1021/cr960034i. PMID: 11851466
22. Jamieson CS, Ohashi M, Liu F, Tang Y, Houk KN (2019) The expanding world of biosynthetic pericyclases: cooperation of experiment and theory for discovery. Nat Prod Rep 36(5):698–713. https://doi.org/10.1039/c8np00075a.PMID:30311924;PMCID:PMC6461539
23. Nivina A, Yuet KP, Hsu J, Khosla C (2019) Evolution and diversity of assembly-line polyketide synthases. Chem Rev 119(24):12524–12547
24. Shen B (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. Curr Opin Chem Biol 7(2):285–295. https://doi.org/10.1016/s1367-5931(03)00020-6. PMID: 12714063
25. Scherlach K, Hertweck C (2021) Mining and unearthing hidden biosynthetic potential. Nat Commun 12:3864. https://doi.org/10.1038/s41467-021-24133-5
26. Jenke-Kodama H, Dittmann E (2009) Evolution of metabolic diversity: insights from microbial polyketide synthases. Phytochemistry 70:1858–1866
27. Wlodek A, Kendrew SG, Coates NJ, Hold A, Pogwizd J, Rudder S, Sheehan LS, Higginbotham SJ, Stanley-Smith AE, Warneck T, Nur-E-Alam M, Radzom M, Martin CJ, Overvoorde L, Samborskyy M, Alt S, Heine D, Carter GT, Graziani EI, Koehn FE, McDonald L, Alanine A, Rodríguez Sarmiento RM, Chao SK, Ratni H, Steward L, Norville IH, Sarkar-Tyson M, Moss SJ, Leadlay PF, Wilkinson B, Gregory MA (2017) Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. Nat Commun 8(1):1206. https://doi.org/10.1038/s41467-017-01344-3.PMID:29089518;PMCID:PMC5663706
28. Drufva EE, Hix EG, Bailey CB (2020) Site directed mutagenesis as a precision tool to enable synthetic biology with engineered modular polyketide synthases. Synth Syst Biotechnol 5(2):62–80. https://doi.org/10.1016/j.synbio.2020.04.001.Erratum.In:SynthSystBiotechnol.2020Aug06;5(4):268.doi:10.1016/j.synbio.2020.07.007.Erratumin:SynthSystBiotechnol.2020Oct14;5(4):332.doi:10.1016/j.synbio.2020.10.004.PMID:32637664;PMCID:PMC7327777
29. Kudo K, Hashimoto T, Hashimoto J et al (2020) In vitro Cas9-assisted editing of modular polyketide synthase genes to produce desired natural product derivatives. Nat Commun 11:4022. https://doi.org/10.1038/s41467-020-17769-2
30. Zhang MM, Wong FT, Wang Y, Luo S, Lim YH, Heng E, Yeo WL, Cobb RE, Enghiad B, Ang EL, Zhao H (2017) CRISPR-Cas9 strategy for activation of silent Streptomyces biosynthetic gene clusters. Nat Chem Biol. https://doi.org/10.1038/nchembio.2341. https://doi.org/10.1038/nchembio.2341. Epub ahead of print. PMID: 28398287; PMCID: PMC5634907

31. Lee Y, Hwang S, Kim W, Kim JH, Palsson BO, Cho B-K (2024) CRISPR-aided genome engineering for secondary metabolite biosynthesis in Streptomyces. J Ind Microbiol Biotechnol 51:kuae009. https://doi.org/10.1093/jimb/kuae009
32. Albert S, Waksman SA (1944) Effect of streptomycin and other antibiotic substances upon Mycobacterium tuberculosis and related organisms. Proc Soc Exper Biol Med 57:244–248
33. Zhang Z, Sun Y, Yi Y, Bai X, Zhu L, Zhu J, Gu M, Zhu Y, Jiang L (2023) Screening and identification of a Streptomyces strain with quorum-sensing inhibitory activity and effect of the crude extracts on virulence factors of Pseudomonas aeruginosa. Microorganisms 11(8):2079. https://doi.org/10.3390/microorganisms11082079.PMID:37630639;PMCID:PMC10458028
34. Takano E (2006) Gamma-butyrolactones: Streptomyces signalling molecules regulating antibiotic production and differentiation. Curr Opin Microbiol 9(3):287–294. https://doi.org/10.1016/j.mib.2006.04.003. PMID: 16675291
35. Liu X, Wang W, Li J, Li Y, Zhang J, Tan H (2021) A widespread response of Gram-negative bacterial acyl-homoserine lactone receptors to Gram-positive Streptomyces γ-butyrolactone signaling molecules. Sci China Life Sci 64(10):1575–1589. https://doi.org/10.1007/s11427-021-1956-8. Epub 2021 Jul 26 PMID: 34319534
36. Hopwood DA (2019) Highlights of Streptomyces genetics. Heredity 123:23–32. https://doi.org/10.1038/s41437-019-0196-0
37. Metz JG, Roessler P, Facciotti D, Levering C, Dittrich F, Lassner M, Valentine R, Lardizabal K, Domergue F, Yamada A, Yazawa K, Knauf V, Browse J (2001) Production of polyunsaturated fatty acids by polyketide synthases in both prokaryotes and eukaryotes. Science 293(5528):290–293. https://doi.org/10.1126/science.1059593. PMID: 11452122
38. Torres JP, Lin Z, Winter JM, Krug PJ, Schmidt EW (2020) Animal biosynthesis of complex polyketides in a photosynthetic partnership. Nat Commun 11(1):2882. https://doi.org/10.1038/s41467-020-16376-5.PMID:32513940;PMCID:PMC7280274
39. Cooke TF, Fischer CR, Wu P, Jiang TX, Xie KT, Kuo J, Doctorov E, Zehnder A, Khosla C, Chuong CM, Bustamante CD (2017) Genetic mapping and biochemical basis of yellow feather pigmentation in budgerigars. Cell 171(2):427-439.e21. https://doi.org/10.1016/j.cell.2017.08.016.PMID:28985565;PMCID:PMC5951300
40. Lin Z, Li F, Krug PJ, Schmidt EW (2024) The polyketide to fatty acid transition in the evolution of animal lipid metabolism. Nat Commun 15(1):236. https://doi.org/10.1038/s41467-023-44497-0.PMID:38172109;PMCID:PMC10764717
41. Hu Y, Chen S, Yang F, Dong S (2021) Marine indole alkaloids-isolation, structure and bioactivities. Mar Drugs 19(12):658. https://doi.org/10.3390/md19120658.PMID:34940657;PMCID:PMC8708922
42. Jackson DR, Tu SS, Nguyen M, Barajas JF, Schaub AJ, Krug D, Pistorius D, Luo R, Müller R, Tsai SC (2016) Structural insights into anthranilate priming during Type II polyketide biosynthesis. ACS Chem Biol 11(1):95–103. https://doi.org/10.1021/acschembio.5b00500. Epub 2015 Nov 3. PMID: 26473393; PMCID: PMC7886370
43. Doyle A, McGarry MP, Lee NA, Lee JJ (2012) The construction of transgenic and gene knockout/knockin mouse models of human disease. Transgenic Res 21(2):327–349. https://doi.org/10.1007/s11248-011-9537-3. Epub 2011 Jul 29. PMID: 21800101; PMCID: PMC3516403
44. Feng L, Gordon MT, Liu Y et al (2021) Mapping the biosynthetic pathway of a hybrid polyketide-nonribosomal peptide in a metazoan. Nat Commun 12:4912
45. Zeng F, Zhang S, Hao Z et al (2018) Efficient strategy for introducing large and multiple changes in plasmid DNA. Sci Rep 8:1714. https://doi.org/10.1038/s41598-018-20169-8
46. Nirantar SR (2021) Directed evolution methods for enzyme engineering. Molecules 26(18):5599. https://doi.org/10.3390/molecules26185599.PMID:34577070;PMCID:PMC8470892
47. Wassman CD, Tam PY, Lathrop RH, Weiss GA (2004) Predicting oligonucleotide-directed mutagenesis failures in protein engineering. Nucleic Acids Res 32(21):6407–6413
48. Sunden F, Peck A, Salzman J, Ressl S, Herschlag D (2015) Extensive site-directed mutagenesis reveals interconnected functional units in the alkaline phosphatase active site. Elife 22(4):e06181. https://doi.org/10.7554/eLife.06181.PMID:25902402;PMCID:PMC4438272

49. Schütz A, Bernhard F, Berrow N, Buyel JF, Ferreira-da-Silva F, Haustraete J, van den Heuvel J, Hoffmann JE, de Marco A, Peleg Y, Suppmann S, Unger T, Vanhoucke M, Witt S, Remans K (2023) A concise guide to choosing suitable gene expression systems for recombinant protein production. STAR Protoc 4(4):102572. https://doi.org/10.1016/j.xpro.2023.102572. Epub 2023 Oct 31. PMID: 37917580; PMCID: PMC10643540
50. Kallunki T, Barisic M, Jäättelä M, Liu B (2019) How to choose the right inducible gene expression system for mammalian studies? Cells 8(8):796. https://doi.org/10.3390/cells8080796.PMID:31366153;PMCID:PMC6721553
51. Brookwell A, Oza JP, Caschera F (2021) Biotechnology applications of cell-free expression systems. Life (Basel) 11(12):1367. https://doi.org/10.3390/life11121367.PMID:34947898;PMCID:PMC8705439
52. Krishnamurti C, Rao SC (2016) The isolation of morphine by Serturner. Indian J Anaesth 60(11):861–862. https://doi.org/10.4103/0019-5049.193696.PMID:27942064;PMCID:PMC5125194
53. Leitão GG, Leal CM, Mendonça SC et al (2021) Purification of alkaloids by countercurrent chromatography. Rev Bras Farmacogn 31:625–647. https://doi.org/10.1007/s43450-021-00163-4
54. Nguyen VPT, Stewart J, Lopez M, Ioannou I, Allais F (2020) Glucosinolates: natural occurrence, biosynthesis, accessibility, isolation, structures, and biological activities. Molecules 25(19):4537. https://doi.org/10.3390/molecules25194537.PMID:33022970;PMCID:PMC7582585
55. Hanley AB, Heaney RK, Fenwick GR (1983) Improved isolation of glucobrassicin and other glucosinolates. J Sci Food Agric 34:869
56. Chain E, Florey HW, Gardner NG, Heatley NG, Jennings MA, Orr-Ewing J et al (1940) Penicillin as a chemotherapeutic agent. Lancet 236:226–268
57. Linke JA, Rayat A, Ward JM (2023) Production of indigo by recombinant bacteria. Bioresour Bioprocess 10:20. https://doi.org/10.1186/s40643-023-00626-7
58. Murdock D, Ensley BD, Serdar C, Thalen M (1993) Construction of metabolic operons catalyzing the de novo biosynthesis of indigo in Escherichia coli. Biotechnology (N Y) 11(3):381–386. https://doi.org/10.1038/nbt0393-381. PMID: 7763440
59. Chandel N, Singh BB, Dureja C et al (2024) Indigo production goes green: a review on opportunities and challenges of fermentative production. World J Microbiol Biotechnol 40:62. https://doi.org/10.1007/s11274-023-03871-2
60. Fabara AN, Fraaije MW (2020) An overview of microbial indigo-forming enzymes. Appl Microbiol Biotechnol 104(3):925–933. https://doi.org/10.1007/s00253-019-10292-5. Epub 2019 Dec 13. PMID: 31834440; PMCID: PMC6962290
61. Fräbel S, Wagner B, Krischke M, Schmidts V, Thiele CM, Staniek A, Warzecha H (2018) Engineering of new-to-nature halogenated indigo precursors in plants. Metab Eng 46:20–27. https://doi.org/10.1016/j.ymben.2018.02.003. Epub 2018 Feb 18 PMID: 29466700
62. Wu W, Ji Y, Wang Z, Wu X, Li J, Gu F, Chen Z, Wang Z (2023) The FDA-approved anti-amyloid-β monoclonal antibodies for the treatment of Alzheimer's disease: a systematic review and meta-analysis of randomized controlled trials. Eur J Med Res 28(1):544. https://doi.org/10.1186/s40001-023-01512-w.PMID:38017568;PMCID:PMC10683264
63. Voigt CA (2020) Synthetic biology 2020–2030: six commercially-available products that are changing our world. Nat Commun 11:6379. https://doi.org/10.1038/s41467-020-20122-2
64. Chen Z, Zhang Q, Shan J, Lu Y, Liu Q (2020) Detection of bitter taste molecules based on odorant-binding protein-modified screen-printed electrodes. ACS Omega 5(42):27536–27545. https://doi.org/10.1021/acsomega.0c04089.PMID:33134717;PMCID:PMC7594143
65. Behrens M, Lang T (2022) Extra-oral taste receptors-function, disease, and perspectives. Front Nutr 4(9):881177. https://doi.org/10.3389/fnut.2022.881177.PMID:35445064;PMCID:PMC9014832
66. Kinnamon S, Finger T (2022) The role of ATP and purinergic receptors in taste signaling. Handb Exp Pharmacol 275:91–107. https://doi.org/10.1007/164_2021_518.PMID:34435233;PMCID:PMC10078839

67. Melis M, Errigo A, Crnjar R et al (2019) TAS2R38 bitter taste receptor and attainment of exceptional longevity. Sci Rep 9:18047. https://doi.org/10.1038/s41598-019-54604-1
68. Medapati MR, Singh N, Bhagirath AY, Duan K, Triggs-Raine B, Batista EL Jr, Chelikani P (2021) Bitter taste receptor T2R14 detects quorum sensing molecules from cariogenic Streptococcus mutans and mediates innate immune responses in gingival epithelial cells. FASEB J 35(3):e21375. https://doi.org/10.1096/fj.202000208R. PMID: 33559200
69. Coquant G, Grill JP, Seksik P (2020) Impact of N-acyl-homoserine lactones, quorum sensing molecules, on gut immunity. Front Immunol 28(11):1827. https://doi.org/10.3389/fimmu.2020.01827.PMID:32983093;PMCID:PMC7484616
70. Coquant G, Aguanno D, Brot L et al (2022) 3-oxo-C12:2-HSL, quorum sensing molecule from human intestinal microbiota, inhibits pro-inflammatory pathways in immune cells via bitter taste receptors. Sci Rep 12:9440. https://doi.org/10.1038/s41598-022-13451-3
71. Bohr KCH, Krogh A (1904) Ueber einen in biologischer beziehung wichtigen Einfluss, den die kohlensäurespannung des blutes auf dessen sauerstoffbindung übt. Skandinavisches Archiv Für Physiologie 16:402–412
72. Monod JF (1961) General conclusions: telenomic mechanisms in cellular metabolism, growth, and differentiation. Cold Spring Harbor Symp Quant Biol 26:289–401
73. Botti MG, Taylor MG, Botting NP (1995) Studies on the mechanism of myrosinase. Investigation of the effect of glycosyl acceptors on enzyme activity. J Biol Chem 270(35):20530–20535. https://doi.org/10.1074/jbc.270.35.20530. PMID: 7657629

Chapter 5
Genetic Manipulation in Agriculture

Abstract Classification of plants (Linnaeus) is a necessary precursor to effective cross breeding. Breeding by phenotype and adoption of hybridisation techniques were early stage advances. The Green Revolution led to enhanced yields of major crops. The Malthusian theory of population growth suggests it will outstrip increased food production unless the paradigm can be altered. The development of marker assisted selection of useful traits and homologous gene expression can be used to improve crops using existing genetic material. The development of heterologous gene expression in suitable hosts—for example the production of long chain polyunsaturated fatty acids in oilseed crops—offers a way of conserving resources. Gene editing and CRISPR technology will become increasingly impactful. The parable of the Ship of Theseus and replacement of functional parts of the genome is an important thought experiment. The production of natural and unnatural products. The regulatory landscape and the Precautionary Principle (Europe) compared to the GRAS system (US) will influence the pace of technological change. The potential for the enhanced production of bioactive secondary natural products to improve heath and wellness is comparatively little studied compared to methods to induce phenotypic changes.

Keywords Genetic manipulation · CRISPR · Regulation · Technology

As we have seen, the knowledge and understanding of the details of biosynthetic pathways opens the door for making use of biological systems to produce specific valuable (or novel) natural products. Understanding the triggers for natural products in biological systems enables scientists to exert a degree of control far in excess of the level of their understanding of the details of the complexity of biological systems. In short, by adopting a reductionist philosophy and finding what the most significant lever are that determined how plants and other living species can be manipulated, some of the most useful and significant attributes can be controlled in a way that produces benefits. Just as genes make proteins (enzymes) which make other components of living systems, a small level of knowledge enables the manipulation of genes themselves and this in turn has more far-reaching consequences than trying to

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_5

manipulate enzymes or other components. The challenge is to be able to alter genes in such a way that the modification leads to a defined benefit and to rapidly measure the change so that successful changes (or those with the potential to develop a success pathway) can be quickly identified. Advances in molecular biology have enabled chemists and biologists to alter the agronomic properties of plants and to change their characteristics to improve yield, reduce waste and optimise their usefulness. The progress in the manipulation of crops has occurred in stages. Domestication of crops and the creation of new varieties by natural breeding was the first step and led to a revolution in agriculture. This was followed by the development and application of molecular biology techniques and the introduction of, firstly homologous and then heterologous genes which introduced novel traits not seen in nature. This also brought considerable regulatory constraints. More advanced gene editing techniques (particularly CRISPR technology) has shifted the paradigm for genetic engineering and introduced further challenges to the acceptance of genetically manipulated food. Finally, the potential for science to accelerate the optimisation of agriculture and food production has been recognised however considerable scientific, ethical and societal challenges remain.

5.1 Traditional Plant Breeding—Opportunities and Impacts

Traditional plant breeding has been both a planned and a serendipitous progression. In order to carry out planned breeding of plants, two concepts were crucial. The first of these was the proposition that plants reproduced sexually. This was proposed by the French botanist Sebastian Vallant in 1717 in a short treatise (Discours sur la structure des fleurs, leurs differences et l'usage de leurs parties). The idea that plants had male and female parts was a groundbreaking concept that greatly influenced the author of the second important concept—the systematic naming of plants by Linnaeus. The important breakthrough that Linnaeus authored was not simply the two-word name for plants and other species but the fact that these words were based in part upon the characteristics of the species (plants and animals) and naming was based on specific rules. The first part of the name is the genus to which the plant belongs while the second part differentiates it from the other members of the genus. Furthermore the rules also advocated the use of Latin nomenclature and so were not constrained by language since Latin was still commonly used by scientists. In these two insights, scientists recognised that there were significant differences between species that allowed them to be identified according to phenotype and purpose and that there was a potential to change the characteristics of species by combining certain of them and harvesting the seeds. This is analogous to the suggestion of 'similars' discussed in Chapter Four and the alchemical idea that some substances could react together to produce something different. The Linnaean ordering of species is a Periodic Table of living entities with a classification underpinned by the properties of each member.

The earliest examples of plant breeding predated knowledge of how plants reproduced. Farmers recognised that some individual plants were more useful than others (e.g. produced more seed, were more resistant to disease etc.) and so they saved seed from these and used it for the next year's cultivation. Over a period of time, plants were selected based on physical characteristics such as size of the seed head, resistance to disease, degree of vigour of growth and so there was a gradual increase in 'fitness' for specific purposes. In addition, hybrid forms of some crops arose through cross-fertilisation and these were also favoured although this was a very slow process over many generations (and harvest seasons). The progress from primitive domestication to planned breeding to create varieties with specific characteristics has been reviewed [1]. All modern major food crops were created from progenitors initially deriving from primitive selection made on the basis of desired characteristics rather than planned breeding. There is a consistent trend in this process. Empirical evidence has generally accumulated in advance of theoretical understanding. Primitive selection and even domestication of plants pre-dated an understanding of how plants reproduced. A knowledge of the characteristics of heredity and how it worked was developed by Mendel before the concept of genes and generational stability of characteristics was understood. One of the most important principles that Mendel elucidated was that of inheritance. While the idea that a particular trait can be passed from one generation to the next seems obvious, this is perhaps not the feature of inheritance that is of greatest significance. Mendel's work disproved the concept of 'blending'. Briefly, if you take two species with particular traits, their offspring will not be a blend of these two traits but will have the traits of one or other of the parents. This is not a trivial point. Although such considerations lie outside the scope of this book, the principles of 'good breeding' and the dilution of particular desirable characteristics is a significant part of the underpinning of eugenic thought [2]. It is important to state unequivocally that, not only is eugenics and its progression, morally reprehensible, it is also thoroughly bad science.

The idea of primary and secondary metabolites and how the latter might be modified was known before genetic manipulation and the planned transfer of genes, operons and biosynthetic gene clusters (BGC) between species was either understood or was possible. From a philosophical standpoint, natural products have been used before they were understood and this has affected the way that chemists have developed and used tools to study them in detail. On the one hand, scientific reductionism seeks to develop tools that improves the understanding of more and more finely divided and detailed representations of the components of a living system while on the other, the effects of the interactions of living systems on a macro scale are easier to observe—even if these are not so readily understood.

As centuries passed, plant breeding moved from the field into the laboratory. This was at least in part a response to the slow pace of change of agriculture. In order for the effects of the crossing of two variants to be known, a whole growing season had to take place. In addition, seed would have to be collected and it would be necessary to show that any change in the progenitor species was stably inherited. Even without a knowledge of genetics, rules for hybridisation were being developed. Hybridisation is defined as the process of combining different species or varieties to create a

hybrid successor organism. In chemistry, hybridisation refers to the combination of orbitals. For example, the electronic configuration of a carbon atom is 1s(2), 2s(2), 2p(1),2p(1),2p(0). Rather than having differently filled orbitals, the 2s and the three 2p orbitals combine to give four equal energy sp^3 hybrid orbitals. This also explains the inferred tetrahedral shape of the bonds of a fully saturated carbon atom. For living species, the classification system developed by Linnaeus helped in the identification of related species which were much more likely to undergo successful hybridisation. Similarly, knowledge of how pollination worked and of the male and female nature of plants in practice was an essential prerequisite for the controlled and purposeful production of hybrids with predicted properties. The work by Mendel (and later the insights developed by Darwin) enabled plant scientists to develop rules for cross fertilisation and the exchange of genetic material—even if the genetic material itself had not been identified. A key appreciation was the idea that genetic information was, in a sense, quantised rather than blended (as mentioned previously). What this means is that, for example, if you cross a tall variety with a short variety then you will produce, as offspring, a mixture of tall and short plants. You will not produce plants that are midway in height between the two parental plants. The extent to which a particular characteristic (phenotype) is derived from genetic material is defined as heritability. Working backwards from phenotype, it is possible to determine the extent to which a given trait is genetically derived. Observation over multiple years is essential in order to determine heritability and the stability of a given hybrid. One of the consequences of the rules developed by Mendel is that he created a model within which it is possible to predict the success of selective breeding—even if the mechanism of the process is unknown—and the importance of using statistics to determine likely outcomes.

An important consequence of the development of predictive plant breeding techniques was the ‘Green Revolution’ of the 1950s and 1960s. The American plant scientist, Norman Borlaug began developing disease resistant wheat varieties in Mexico in the 1940s. In essence the Green Revolution was developed around the concept of picking varieties of grain crops (including wheat and rice) which responded best to fertiliser and other agronomic practices thereby producing increased crop yields and increased amounts of food. [4] The Green Revolution was necessary to deal with one particular aspect of human life—the need to grow more food over a wider part of the available surface of the earth. The need for other revolutions to address the interface between agriculture, the living world and society has been noted [5]. While the Green Revolution can perhaps be described as the pinnacle of the response to the Malthusian principle that, since populations grow geometrically and the food supply can only grow arithmetically it is therefore inevitable that population size will outstrip the capacity to grow enough food, further revolutions will be necessary too if human populations continue to grow and resources are limited. The Green Revolution was a way of responding to the Malthusian dilemma. If Malthus was correct then eventually the population would outstrip the available resources leading to starvation and a population readjustment. These ideas were outlined by Malthus in his treatise ‘An Essay on the Principle of Population’ in 1798. While it has been argued that the continued growth in the global population suggests either that the Malthusian idea is

false (or has been subverted by technological progress) or has not yet come to pass, the continued existence of inequality and extreme poverty suggests that, in some places, the Malthusian trap has been sprung. One of the major obstacles to the rapid development of crops with new desirable properties is the selection of offspring. The desirable traits of one plant may be transferred together with undesirable traits (poor agronomy or yield, lack of fitness, susceptibility to disease or pests). This means that traditional plant breeding can take 10–15 years to produce a new viable variety with a mixture of desirable traits and an acceptable level of undesirable traits. An alternative approach is to induce changes in plant DNA through causing mutations that alter the phenotype of the plant. In essence, however, mutation (whether induced by radiation or by chemicals) is a random event and the development of a new variety can still be a lengthy process. Some ways to speed up this selection process will be considered later.

In terms of significant advances in plant breeding, the twentieth century saw the introduction of significant beneficial changes. The principal targets have been improved nutrition, enhanced yields and other commercially-beneficial trait optimisations. In terms of natural products, the relative absence of verifiable health benefits that can be attributed to natural products means that breeding has focused more on the removal of undesirable traits (toxicity, undesirable taste characteristics, properties that limit processing) rather than optimisation of natural product formation.

One example of the use of traditional plant breeding to eliminate undesirable properties is the development of double low rapeseed (*Brassica napus*). Rapeseed is an important oilseed crop however its use as a high quality vegetable oil was limited by two features of the oil derived from the wild type plant. The first of these is the presence of high levels of erucic acid (up to 40% of the total oil content). Erucic acid is a monounsaturated fatty acid produced by plants of the *Brassicaceae* family—in particular those which accumulate oil such as rape and mustard. It has been implicated in cardiotoxic effects in animal models and therefore a programme was established in Canada in the 1960s to find varieties of rape with naturally low levels of erucic acid. This led to the development and commercialisation of a low erucic acid rapeseed which was called ‘Canola’ to distinguish it from wild type cultivars with high levels of erucic acid. The second limiting factor in the use of rapeseed oil (and post-processed rapeseed meal as animal feed) was the presence of high levels of glucosinolates. The breakdown products from these natural products include anti-thyroid components (goitrin and thiocyanate ion) and compounds that cause palatability issues for farm animals (isothiocyanates, goitrin and nitriles). A potential solution was found in a naturally low glucosinolate variety of *B. napus* from Poland (Bronowski). By crossing different varieties, ‘double low’ (low erucic acid and low glucosinolate) varieties of *B. napus* and *B. rapa* were developed in the 1970s. Subsequent breeding programmes led to increased oil and seed yield and early maturity and reduced plant height. Some high erucic acid, low glucosinolate varieties were also developed to produce oil for non-food use (particularly industrial processing to produce biodiesel and polymers). Subsequently a low glucosinolate rapeseed plant (*B. napus*) was created by knocking out two glucosinolate biosynthetic genes [5]. The biosynthesis of the three broad classes of glucosinolates (aromatic,

aliphatic and indolic) is carried out by three distinct gene clusters however over 90% of the glucosinolates in *B. napus* are aliphatic [6], so disruption of this specific biosynthetic pathway will lead to a significant reduction in overall glucosinolate content.

While glucosinolate and erucic acid reduction are perhaps two of the most significant successes of plant breeding to reduce anti-nutritive factors in commercial crops, plants used as food or feed contain a large number of other deleterious natural products. These include anti-nutritional factors such as saponins, lectins, gossypol, phytic acid, tannins and also overtly toxic components such as alkaloids, oxalic acid and cyanogenic glycosides. Among food plants, legumes are perhaps the most significant in terms of their anti-nutrient/toxin content [7] and have therefore been the focus of considerable efforts to optimise their nutritional profile [8]. In addition legumes (and other crops) can also be a source of fungal toxins (mycotoxins) and strategies to eliminate these have also been developed [9]. A number of strategies have been developed to reduce unwanted compounds in foods including autolysis, fermentation, milling, washing and cooking. While breeding and strain selection can be successful, this has frequently been carried out with little detailed molecular genetic knowledge and therefore the impact on other agronomic characteristics such as yield, resistance to pests and macronutrient profile may also be affected. The development of molecular techniques to carry out planned manipulation of pathways to manipulate the production of specific natural products has been the focus of considerable efforts over a significant period of time however this has led to considerable opposition by pressure groups on a number of levels.

5.2 Molecular Biology of Plants

The importance of Mendelian genetics and Darwin's theories about selection and survival of the fittest provide the key intellectual underpinning for plant molecular biology. While the use of mutagenesis—using either chemical agents or radiation—has already been mentioned as a way of inducing changes in the genotype of plants that can have a subsequent effect on phenotype, it is a random process and, although considerably more rapid than the time taken for natural genetic changes and selection to occur, it remains a lengthy and essentially random process. There are a number of factors that can influence the effectiveness of phenotypic breeding programmes. In general the likelihood of genetic gain (defined as the change in the mean value of the trait of interest, in other words, the anticipated genetic improvement) is affected by a number of variables including the probability of passing on a trait from parent strain to offspring, the length of time needed for a cycle of selection to be completed and the phenotypic variation in the starting population. One key development which has helped in the rapid development of desired traits is the development of whole genome sequencing and the use of molecular markers. Molecular marker-based trait monitoring can take a number of forms including direct and indirect methods. Traditionally molecular markers have included tracking morphological characteristics.

However there are some issues and limitations associated with the use of morphological markers. The number and specificity of morphological markers is generally insufficient and there may not be any suitable markers linked specifically to desired trait characteristics (e.g. yield or quality). In some cases the markers may actually be negatively correlated with desired traits or with robustness and growth. A more direct correlation is of greater value and applicability. Some characteristics of plants are associated with particular biochemical markers. These may include the presence of certain functional proteins including enzymes. Monitoring biochemical markers during strain crossing and development may be a direct way of assessing the success in optimisation of strain development. Another direct measure is the use of DNA markers and this can be of value. Whichever method is used, detection of changes (polymorphisms) associated with specific traits can be a rapid way of assessing strains. The use of DNA markers, in particular has provided significant advances in the speed of development of new crop varieties [10]. Many of these rely on established associations between specific molecular markers and genes (quantitative trait loci—QTLs) of agronomic interest. This enables marker assisted selection (MAS) to be utilised. Molecular markers and other genomic-based technologies have been a mainstay in the development and modification of crops within individual species. In some cases desirable traits are either not present or are lost during strain development. In this case, heterologous genes from other species can be inserted.

5.3 Manipulation and Expression of Non-homologous Genes

The development of techniques to transfer and then express foreign genes in a particular species relied upon an understanding of what the function was of products from individual genes in the donor species and how these might be expressed in the recipient species. The normal process of gene transfer within species is the basis of heritable (vertical) gene transfer—from parents to progeny. The transfer of genetic material between, rather than within species is known as horizontal gene transfer. It most commonly occurs between bacteria but can also occur from bacteria to eukaryotes [11]. Studies have demonstrated that horizontal gene transfer makes an essential contribution to the genomes of prokaryotes and eukaryotes. A range of mechanisms are known to be used for horizontal gene transfer and these have been reviewed [12]. Horizontal gene transfer—either engineered or naturally occurring – is, as noted previously, distinct from vertical (inherited) gene transfer and is a major source of phenotypic variation and adaptation. For example, the prevalence of antibiotic resistance and its spread between bacteria is substantially due to horizontal gene transfer.

Genetic modification is a laboratory-based process whereby the DNA of a species is changed. This definition clearly includes gene editing although, as we will see, some scientists and regulators have tried to draw distinctions between the insertion of

foreign genes and the modification of 'native' genes in an organism. While genetic modification can arise from the modification of a single base pair (and therefore a coding change that can affect the amino acid sequence of a protein), it can also include the insertion of a whole gene or multiple genes (biosynthetic gene clusters [13]) to code for the biosynthesis of a natural product. After isolation of the gene or genes of interest, the next stage in the process of genetic modification is transferring the DNA into the host cell. One simple method is to coat small metal particles with DNA and to fire these into a cell. This is commonly referred to as a 'gene gun' or 'biolistic' approach [14]. There is variability in the effectiveness of the method and measuring the degree of transfection can also be challenging.

An alternative is to make use of the capability of some biological agents to transfer DNA into a host cell. There are also a number of bacteria and viruses which, as part of their life cycle, are able to transfer DNA into the genome of other species (including plants). In essence, in engineered gene transfer, the bacterium or virus is manipulated then used as a vector to carry the DNA of interest into the host genome where it becomes functionally viable. One example of genetic modification of plants is the acquisition of resistance to the weedkiller, glyphosate. In one example, a synthetic gene coding for glyphosate acetyl transferase activity was transferred into maize enabling the transgenic maize to grow in the presence of glyphosate [15]. In the case of natural products, there has been less interest in enhancing their production in plants since the commercial (health) benefits are less clear cut and the use of conventional plant breeding has been of greater significance in the transfer and optimisation of desirable traits among natural products [16]. Nonetheless, tomatoes have been genetically altered to contain high amounts of anthocyanins [17]. Such an approach is certainly superficially attractive however the health promoting properties of different anthocyanins remains a matter of debate and, in any case, natural sources of the compounds occur in other plants. It is unclear that the significant time and effort (and the regulatory hurdles) required to produce transgenic plants containing mixtures of natural products with putative, but unproven benefits is a worthwhile activity.

A second example, which will demonstrate the potential of heterologous gene expression with additional advantages, is that of a particular type of fatty acid. The omega-3 long chain polyunsaturated fatty acids (LC-PUFAs) have been implicated in a number of benefits to human health [18]. The two most studied members of the class, from this standpoint, are docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA) (Fig. 5.1).

They occur naturally in oily fish and have been implicated in cardiovascular health as well as in the prevention of diseases associated with inflammation including cognitive decline. The supplementation of infant formula with LC-PUFAs is postulated to improve visual acuity and cognitive development. In the case of the former, a meta analysis suggests a positive correlation [19] however results supporting a link with cognitive development are less convincing [20, 21] An important dietary precursor of DHA and EPA is alpha linolenic acid (ALA) however the conversion of ALA to EPA in mammals varies from 8 to 12% while conversion to DHA is as low as 1% [22]. Dietary intakes of DHA and EPA in particular are highly dependent upon

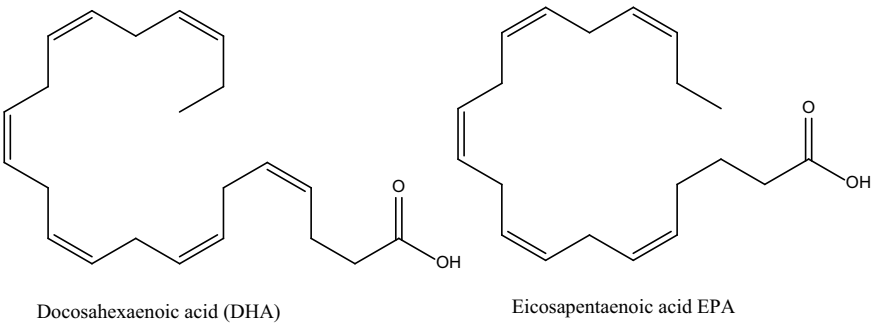


Fig. 5.1 Structures of docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA)

fish consumption and, in most regions, this is low so intakes are sub-optimal for the prevention of chronic diseases. The US Food and Drug Administration has approved a qualified health claim for foods that contain EPA and DHA "Supportive but not conclusive research shows that consumption of EPA and DHA omega-3 fatty acids may reduce the risk of coronary heart disease." Given the decline in global fish stocks and the lack of availability (and consumption) of oily fish, the production of DHA and EPA by the transfer of genes into suitable species has been investigated. The production of EPA and DHA by microalgae has formed the basis of commercial exploitation [23]. While many plants can be a source of ALA, the low conversion rate of ALA to DHA and EPA noted previously, means that the human dietary requirement is primarily met by fish consumption rather than by ingestion of ALA-rich plants. Manipulation of the polyketide synthase pathway from a marine organism (*Colwellia psychrerythraea* 34H) enabled the production of a strain of *E. coli* that was able to synthesise DHA [24]. These and other studies suggest that heterologous expression of biosynthetic genes for DHA and EPA production is possible and such studies were noted previously in the case of *Camelina sativa.* The production of LC-PUFAs by plants is a credible and sustainable alternative. Unlike the case of anthocyanins noted previously, an alternative and convenient dietary source (apart from oily fish) for LC-PUFAs is not available. The use of an alternative food crop will, therefore, help to preserve a diminishing and increasingly expensive source (fish) and so for both economic and sustainability reasons heterologous gene expression in a suitable host is a viable option. An obvious starting point is an oil seed crop. For this reason, *Camelina sativa* has been chosen as a suitable host. *Camelina sativa* is a member of the Brassicaceae family. It is an oil seed crop related to mustard and has a promising nutritional profile [25]. The success both in the laboratory and in field trials suggests that this alternative source of 'fish oil' may be both viable and sustainable [26]. Further work is necessary to deal with the practical and regulatory hurdles that are likely to arise and the existence of a natural algal source of DHA and EPA [27] may limit the commercial opportunities for a transgenic source—particularly if infants are the main target population [28].

5.4 Gene Editing

One of the problems associated with the insertion of foreign genes into a host species (as noted above) is public and regulatory perception. The application of the Precautionary Principle in Europe [29] combined with strong and effective lobbying by environmental groups in different countries in Europe [30] has increased the size of the challenge faced by scientists and innovating commercial concerns. At the core of many of the arguments that are presented on environmental and safety grounds is the rule of unintended consequences. As genetic modification becomes technically more advanced an increasing number of characteristics of a host species may be altered. This can result in unanticipated consequences and also raises the question of when (outside of primary metabolism) do the changes in a host species change it so much that the nature of the modified species is different from the parent. One analogy which might illustrate this is the story of the Ship of Theseus. The story of the Ship of Theseus is also known as Theseus' paradox [31]. It poses the question whether an object that has had all its components replaced can still be considered to be the same object. This has some analogies with the concept of living things containing some unique property (and is that property unique to a given species) that is not shared by non-living objects—is single species vitalism a viable concept? The ship in which Theseus and the Athenians returned to Athens was considered to be so important that each plank of the ship was, over time, replaced with fresh timber. The question of whether the ship remains the same (is it still Theseus' ship) or is it in some way a different ship. Hobbes later added a refinement to the paradox by suggesting that, if each plank that was replaced was then used to build another ship, would either or both ships be considered to be the ship of Theseus? Various solutions to the paradox have been suggested. It was proposed that an appropriate analogy would be a river which is always different because of the flow of water. In that case, it is impossible to step into the same river twice since it is an ever-changing entity. Aristotle suggested that the important aspect of an object is the use to which it is put. In that case the ship of Theseus remains the same because the design and the use to which it was put remains constant even if the material from which it is constructed is different. To extend the analogy, a plant remains a plant even if it produces a range of unnatural (to it) products. A further interpretation suggests that time is an important criterium. By that consideration, an object is identical to itself only at a specific time. At all other times, it is different even if it appears to be invariant. In terms of biology and gene transfer, the question is whether an organism remains the same before and after functionally active genes are transferred into it. In one sense, the host species remains the same—and this is perhaps reflected in the retention of its original name with the addition of the prefix 'transgenic' to denote that it contains altered and/or foreign DNA. A plant which has been transfected with genes that code for the enzymes involved in synthesis of algal LC-PUFAs and therefore has the capacity to produce DHA and EPA does not 'become' an algae. It remains a plant that shares some specific characteristics in terms of natural product profile with the particular algae from which the genes are derived. In the case of conventional plant breeding,

while the crossing of two varieties to give a third variety also involves transfer of functional DNA, the derived variety may, in some cases, be considered a separate species but this is not invariably the case. Crossing yellow and black seeded varieties of *Brassica napus* can give progeny that have a different natural product (in this case flavonoid) profile but they are still classified as *B napus* [32].

Brassica napus itself is an important crop and illustrates an important aspect of classification. It arose from hybridisation between *Brassica rapa* and *Brassica oleracea* some 7500 years ago. Plants are, therefore, defined in terms of phenotypic similarities and differences according to Linnaeus and other classification botanists at a particular point in time when they are classified. Being aware of the lineage of species does not invalidate their classification. Genome sequencing has helped to confirm phenotypic relationships and to identify previously unknown taxonomic data as well as reveal gaps in species' lineage. In that sense genomic data is similar to the periodic table with the genetic components being analogous to the nuclear and electronic configuration of the atoms of an element and the phenotype being reflected in the bulk properties of the elements. The analogy with the ship of Theseus while being incomplete, can help us to consider how species are linked and that, while vitalism is not an underpinning concept, living systems are more than just the sum of their genetic information. Species are generally classified according to physiological and genetic similarities however the concept of functional similarity has been suggested for microbes and the proposers suggest that one of the attributes of this method of classification is that it "can use organism meta-data to highlight the specific environmental factors that drive microbial diversification" [33]. The extension of this approach to include the case when the complexity and interlinked nature of natural products, outside of primary metabolism, are concerned and the consequences of environmental factors in the promotion of gene expression could be considered.

The recent discovery and exploitation of gene editing could be considered to add a further level of complexity to the genetic manipulation of living systems. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) forms the basis of a gene editing technique which has become increasingly important. The technology involves the recognition of a sequence, cleavage and repair. A guide RNA fragment (gRNA) recognises a complementary sequence in a gene of interest. A Cas-9 nuclease then makes a double stranded cleavage which is subsequently repaired by non-homologous end joining or homology-directed repair. The origin of CRISPR-based technology is considered to derive from the discovery of short palindromic sequences of unknown function in *Escherichia coli* [34]. As rapid genome sequencing became a common tool for biologists, more instances of these repeat sequences were found in bacteria and archaea. At about the same time, co-occurring genes which encoded DNA repair proteins were also found. These were called Cas (CRISPR associated) genes. Further analysis suggested that CRISPR and Cas worked together as an acquired immune system to protect prokaryotes against invading phages, plasmids and viruses [35]. In essence, a bacterium which is attacked by a virus will acquire viral DNA (spacers) which are inserted into the bacterial DNA at a position called

the CRISPR array. These spacer sequences enable the bacterium to recognise subsequent attack by the same virus. The bacterium synthesises CRISPR-RNA which can bind to the viral RNA and a co-expressed Cas nuclease then destroys the viral nucleic acid. The next important step which converted this defence mechanism into a method for genetic modification, was the realisation that, by manipulation of the recognition sequence, any DNA sequence could be targeted by a modified CRISPR/Cas-9 system. This means that any target sequence can be cleaved in a very specific fashion and the breakage then repaired by the normal cellular machinery. It is possible to repair DNA using non-homologous end joining or homology-directed repair. The former is more common but is prone to error and to the insertion of a frameshift mutation or a stop codon. In this way a target gene is inactivated. In the case of homology-directed repair, if an excised piece of DNA has a fault in it then a piece of exogenous donor DNA with sequence homology to the cleaved DNA but without the error is inserted. In this way repair of a faulty gene can be carried out. The CRISPR/Cas-9 system has then the potential to repair (and thereby cure) human genetic diseases. The use of CRISPR/Cas-9 (and other modification) has great potential—particularly in the context of human disease however a number of practical and technical challenges remain [36].

In the case of agriculture, as has been noted, the use of selective breeding was followed, by the application of radiation or chemicals to introduce random genetic mutations and the organism containing a desirable trait identified by screening and regeneration. Some of the drawbacks to this approach have been mentioned including its lack of specificity, the need to screen large numbers of mutated organisms and the likelihood that multiple mutation events with unknown (or undesirable) consequences will take place. The use of foreign (heterologous) DNA to make transgenic plants with genes added to provide specific traits has been discussed. The advantage of using CRISPR technology is twofold. First of all it is precise and there will be no unintentional consequences such as may occur if foreign DNA is inserted or random mutagenesis is carried out. Secondly, CRISPR/Cas-9 enables the editing of the genome of the target organism without insertion of any foreign DNA. These advantages mean that gene editing may come to be the preferred route for genetic manipulation [37]. In addition, as we will see, the regulatory hurdles may be less in the case of gene editing. While much of the interest in using CRISPR/Cas-9 technology has concentrated on standard crop characteristics (e.g. grain length of rice), the manipulation of natural products has also been carried out and this has led to the production of fruits with varying levels of carotenoids, anthocyanins and polyphenols [38]. While some of this work was carried out in response to consumer demand for a preferred colour, the manipulation of health properties is also possible. The same caveats apply to CRISPR derived products as noted previously for anthocyanin-rich tomatoes—the health benefit has to be clear and unequivocal to justify the time and effort expended to create the modified crop. It should be emphasised, in addition, that the use of CRISPR/Cas-9 still requires a clear knowledge of biosynthetic pathways and the consequences of manipulation of specific components both on the target pathway and on any other characteristics. This also applies in the case of adjustments

that may affect the quality of plants (e.g. reduction of autolysis of fruit and vegetables and improvement in flavour characteristics). For example, as noted previously many of the negative health properties of Brassica vegetables are due to the production of the glucosinolate progoitrin (and its hydrolysis product, goitrin—which is also responsible for much of the perceived bitterness of Brassica vegetables). Two key genes involved in the biosynthesis of progoitrin have been identified in Chinese kale [39] and downregulation of these by gene editing may be a suitable strategy to improve the health and qualities of Brassica vegetables more generally. This may make the manipulated plants more prone to predation by insects [40]. Further work is needed on the precise function of the target genes and the ecological and other consequences of disruption of their expression.

5.5 The Nature of Biosynthesis—Natural and Unnatural Products

The use of genetic manipulation has incited regulatory challenges that do not occur with conventionally bred crops. This arises in part from the perception that genetic modification produces species that contain 'unnatural products'. The idea that a product can be deemed to be natural or unnatural depending on the environment in which it is produced is somewhat counter-intuitive. In the early modern period, a distinction began to be drawn between the Principles of Nature which are innate to substances and the Laws of Nature which are external and imposed from outside the substance itself. Kant believed that nature was 'ordered and lawful' and these laws govern the properties of the natural world. In contrast the Principles of Nature define the nature of substances that are part of the natural world and define their attributes. This might imply that an approach based upon Principles of Nature would define all substances based upon their attributes no matter where they originate. The Laws of Nature are more all-encompassing and suggest that a substance must also reflect its origin and this would imply that a valid distinction can be made between natural products in their natural setting (or place) and those which occupy a space where they do not occur naturally according to Nature and are, therefore, unnatural.

If the nature and origin of natural products are considered in the context of the Laws of Nature then there are various interpretations that attempt to explain why organisms create natural products. One such theory which has been widely adopted is the Chemical Co-evolution model [41]. This model states that a natural product confers some fitness advantage to the producer (this could be deterring herbivores or attracting insects, for example) and so both the organism itself and the context of the production of the natural product are important. The problem with this interpretation is that the chances of a single natural product conferring a significant advantage to the producer is relatively low and so retention of the ability (which is energetically costly) is an unsustainable concept since the fitness advantage will be limited. The situation is, in part, ameliorated by the biosynthetic machinery for the natural product

only being activated when required however this will affect speed of response and hence effectiveness. An alternative view (the Screening Hypothesis) suggests that metabolic traits which favour the production of a diverse set of natural products will be retained and this will be advantageous because multiple products have the potential to act by several different mechanisms. At one level then, the production of multiple products from the same (or closely related) starting materials or the capability of biosynthetic enzymes to accept a number of substrates will increase the range of natural products formed. This concept should not be confused with the exogenous process of screening bioactive principles for desired bioactivities in the context of other organisms. It is, rather, a host-specific screening process. One consequence of the Screening Hypothesis is that the introduction of biosynthetic pathways through the addition of foreign DNA will increase the complexity of natural products being synthesised by an organism. This may occur as a result of production of the intended natural product or from the biosynthetic enzymes being exposed to other novel potential substrates that are already present in the host plant but not in the donating species. No matter which route operates, the Screening Hypothesis suggests that the new, manipulated species will, over time, possess enhanced fitness compared to the wild type since it will have a larger pool of natural products that can be utilised for host protection and this could be an unanticipated consequence of the creation of a transgenic organism. The converse situation could also arise. If a gene editing procedure leads to a reduction in the number and variety of natural products formed then, over time, the manipulated species may prove to be less hardy (and successful) than the wild type. Experimental (or empirical) philosophy rather than natural philosophy is required to change the nature of organisms. This may be analogous to microbiomes. It has been noted that in both the oral and the gut microbiome, an increased diversity of microbial populations is associated with better health outcomes [42], possibly by nutrient blocking and other competitive mechanisms including quorum sensing [43].

Both the Principles of Nature and the Laws of Nature are valid concepts when considering natural products and their role in nature. Substances do not simply impart their properties to the place they are situated and therefore the Principles of Nature are complicit in their existence. Their impact is more broadly seen as intrinsic to the host and the response of the host is an important aspect of how significant that impact is. At the same time, the Laws of Nature are important since a natural product in one place may have different effects from its impact in another (original host) place. Products can, therefore, be considered to be both natural and unnatural and this dichotomy is reflective of how they are considered.

5.6 Regulatory Aspects of Genomic Modification

The underlying principles of how transgenic organisms are treated is governed by a sense that a risk:benefit analysis is an appropriate way of establishing a regulatory framework. The development of regulations governing the use of genetic modification has tended to focus on the final products and the exposure of individuals to

them rather than a determination of any risk associated with their production or use. Genetically modified foods (the most likely route of exposure for most of the population) can be defined as organisms in which the genetic material has been altered in a way that does not occur naturally. There have been a number of technically viable developments both with regard to agronomy (plant protection) and improvements in crop characteristics however, in the case of the latter, in particular, there have been legislative and societal challenges that have proven to be insurmountable.

The first reported use of genetic modification can be ascribed to Cohen and fellow researchers who, in 1973, claimed to have constructed biologically viable bacterial plasmids in vitro [44]. Cohen linked together fragments of plasmids created by the action of restriction endonucleases. These new plasmids were inserted into E. coli and found to be biologically viable. The procedure that they followed (fragmentation and production of DNA fragments with overlapping compatible ends, separation of these by gel electrophoresis, ligation using DNA ligase and insertion of the novel plasmid into E. coli) became a standard protocol and the first example of the planned insertion of foreign DNA into a species. Subsequent studies included the cloning of the gene to synthesise insulin. From an agronomic standpoint, the production of plants containing the gene to produce a toxin from *Bacillus thuringiensis* (Bt toxin) was a significant step [45]. The toxins was specific for certain insects and therefore transgenic plants were both safe for humans (and animals) to consume and had an induced resistance to insect predation. While the application of Bt toxin to crops using sprays is an alternative mode of plant protection, the advantage of developing the capability to produce it in crops is both more cost effective and specific.

While protection against predation was a significant and cost effective measure for agronomy, one of the most significant advances that had importance beyond crop production was the development of a tomato with improved quality characteristics. The Flavr savr tomato was genetically engineered to last longer and was the first GM product approved for sale in the US in 1994 [46]. The technology was a significant breakthrough and used antisense oligonucleotides to prevent ripening. Antisense technology is a method of gene silencing. It uses short pieces of RNA or DNA (called antisense oligonucleotides) into cells which then bind specifically to RNA produced from a specific gene, preventing the RNA from being translated into a functional protein. In this way the gene is switched off (silenced). Despite the potential to reduce food waste and spoilage, reduce usage of chemicals for ripening and a competitive pricing policy, a vigorous campaign by anti-GM interests led to the eventual withdrawal of Flavr savr tomatoes [47].

One of the most significant nutritional benefits that was developed using genetic modification was the development of ‘golden rice’. The rice contained a precursor to vitamin A and has attracted extensive publicity however it has been a commercial failure [48].

While some of the advances that have been made possible by genetic modification have been welcomed [49], there has also been significant opposition by some NGOs and environmental groups which has also seen some overspill into consumer attitudes. On the whole, therefore, the impact of genetic techniques on agriculture has been less than anticipated by many scientists. This is, in part, a function of different global

attitudes to new technology. In Europe the approach to any challenge where the evidence base is limited is governed by the 'Precautionary Principle'. This approach, mentioned previously, was developed and grew in significance during the 1970s. It was first enacted in German law and was later adopted by the EU in the Treaty of Maastricht. While different definitions have been adopted by different organisations, the general approach has an impact upon the governance of risks and on the approach to innovation. The approach adopted in Europe is, in part, influenced by events where scientific evidence was insufficient and decisions made were subsequently shown to have deleterious consequences. One of these was the addition of organo-lead compounds to petrol and a second was the human-derived release of mercury into the environment where it accumulated in fish and other aquatic species. There are other examples where complex situations cannot easily be evaluated at a specific time point. The precautionary principle is adopted by those who believe that this approach can help to protect human health and/or the environment from complex situations and/or situations where the risk to health is not immediately apparent. Others, however, regards the application of the precautionary principle as unscientific and as something that hinders progress and leads, inevitably, to over-regulation. The application of the precautionary principle is designed to control potential hazards before a causal link has unequivocally been established by absolute and clear scientific evidence. It is clear that, after the decisions made at Maastricht, in EU law, the potential risk to humans and the environment take precedence over economic interests. The three interpretations of the precautionary principle can be expressed as follows:

- Lack of a firm and unequivocal evidence base should not prevent legislation being enacted.
- The uncertainty about cause and effect is such that legislation should be enacted as a matter of course.
- Legislation is mandatory until the absolute absence of hazard has been proven.

These three approaches can be seen as escalations in the interpretation of the precautionary principle and have clear consequences in terms of steps to be taken. In the specific case of genetically-modified plants, the EU has adopted the precautionary principle and most member countries have restricted or banned their cultivation [50]. While it has been suggested that the US takes a less stringent approach to the precautionary principle (with, for example, the allowance of GRAS - Generally Regarded As Safe – as a principle), much of the divergence is over specific instances rather than the general principles. Nonetheless, the US tends to allow commercial concerns more freedom to operate than the EU legislative position. The overall approach in the US is that a hazard must be proven before measures to protect the public are taken. This is summed up in the concept of 'prudent vigilance'. It has been argued that adoption of the precautionary principle may actually limit the breadth of research into potential harm to humans or the environment since it encourages researchers to focus on known risks and hazards and to accumulate further evidence on these rather than on unidentified issues. In the case of genetic modification (GMOs), three agencies (the Food and Drug Administration, the Environmental Protection Agency and the US Department of Agriculture) are responsible. This tripartite approach is

designed to ensure that GMOs are safe for humans, animals and plants. This is carried out by the relevant agencies who work with developers and commercial interests. In the interests of public disclosure, food for human consumption that contains GMOs must be labelled as being bioengineered.

The development of gene editing has been suggested to be a special case which should not be subject to the regulations that govern GMOs. The EU has defined gene editing as a New Genomic Technique (NGT) and plants produced using technology that does not introduce new DNA (i.e. no heterologous genes are inserted) have been assessed by the European Food Standards Authority (EFSA). One of the potential problem areas concerns material produced using NGTs which does not contain any extra genetic material and therefore cannot be detected using genetic analysis. This is largely a question of transparency rather than an issue of food safety issue per se and further studies, driven by the interests of EU member States, is likely. Decisions regarding NGTs will impact upon the development of new products and techniques and the sustainability of agricultural systems. Organisations such as Greenpeace have already lobbied against any relaxation of rules on GMOs in the context of NGTs. It is unclear if the development of more advanced NGTs will lead to any change in the position of environmental organisations such as Greenpeace since they seem to have adopted the position that no research should be carried out in this area and that any form of genetic manipulation—even if it is environmentally benign (or even beneficial) should be outlawed. This is, in some ways, the precautionary principle *in extremis.*

5.7 Long-Term Options for Genetic Technologies in Natural Products

The importance of standardisation has been a driving force for natural product research. Indeed, one of the reasons that the isolation of morphine crystals from the juice of poppy seeds was considered such a significant breakthrough was that, for the first time, studies could demonstrate the optimal dose for a beneficial effect [51]. The isolation of pure pharmacological agents has been a driving force for natural product chemistry and was a key factor in the emergence of chemistry from the shadow of alchemy.

Investigation of natural products (and, in particular, their biosynthesis) has been greatly aided by the development of molecular genetic techniques and related technologies. The elucidation of pathways for the production of bioactive principles has become increasingly important and, in addition to the optimisation of the production of many naturally occurring compounds from their normal hosts (e.g. penicillin from *Penicillium* species) the use of bioengineering has been used to optimise production systems for heterologous products in a biotechnologically adaptable and suitable host species. [52]. And the principle of ‘Design-Build-Test-Learn has enabled the

significant growth in synthetic biology and the use of living systems as 'factories' for the production of natural and unnatural products [53].

In the case of high value products, there is considerable impetus to develop known and novel products. Targeted genetic engineering has been used for the production of new entities [54]. The most recently discovered gene editing techniques have found applications in modifying *Streptomyces* species and, in particular, their production of natural products [55]. Genetic analysis can also aid in the discovery of unknown natural products and their exploitation [65]. In summary, genetic techniques (including pathway and individual enzyme manipulation) can be of great benefit in the discovery, exploitation and manipulation of valuable natural products by making use of living systems both as a source of natural products and as a suitable host for their production.

Many natural products have lower value than compounds which are pharmacologically active because their physiological relevance and use is unclear. For example, the bioactive compound ajoene (Fig. 5.2) occurs in garlic and has a range of biological activities however most of these are chronic in nature and the long-term effects of ingestion over an extended period of time are unknown.

In some cases this is because the measurable effects are chronic—for example minimisation of reactive oxygen species [65]—or they impact upon a number of pathways either directly or indirectly. In addition, in the case of ajoene and many other bioactive natural products it is unknown if all of the biological activity resides in the target molecule or whether some aspect of its beneficial properties are linked with other compounds present in garlic and if the effects are additive or synergistic. Another extensively studied molecule is sulphoraphane (Fig. 5.3) which occurs in many Brassica species.

Its biological profile resembles that of ajoene in some ways (e.g. an effect on pathways associated with Nrf2) [57]. This would suggest a classification based upon functional properties rather than chemical structure or biological origin might be of value. In many cases, bioactive natural products appear to have a more significant protective effect against accumulated damage rather than curative properties in the context of specific diseases. In addition the effects are often confounded by inter-individual variation which may be a result of human genetic polymorphisms. In order to develop a more comprehensive view of the value of chronically bioactive natural

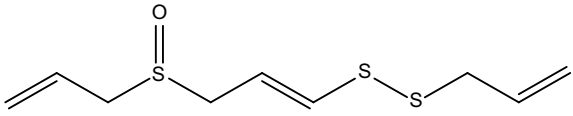


Fig. 5.2 Structure of ajoene

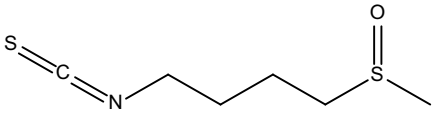


Fig. 5.3 Structure of sulphoraphane

products, it is necessary to carry out trials over lengthy periods of time with relevant (and predictive) endpoints. The same principles apply to assessment of chronic toxicity and to developing a better understanding of synergies between compounds of related (and different) classes. Lifetime observational studies are expensive, lengthy and, if not mechanistically based, of only limited value.

Many of these chronic conditions are defined as non-communicable diseases (NCDs) and, as life expectancy increases, they are assuming greater importance as a cause of morbidity and making a significant contribution to mortality. Attempts have been made to stratify NCDs on a societal and country level [58] however prediction of individual risk and therefore mitigation measures remain elusive. The development of medical condition-specific biomarkers collected through the use of wearable devices has the potential to greatly assist in the management of NCDs [59]. In most cases such biomarker-based devices are used to detect abnormalities associated with disease states such as cancer. Management of NCDs such as diabetes and cardiovascular disease are available including, for example, continuous blood pressure monitoring and wearable devices to provide a continuous eletrcardiogram (ECG). For diabetes, continuous blood glucose monitoring, stress measurement and dietary intake of glucose can be measured. Most devices are disease indicators and the more subtle, long term effects of natural products are not amenable to such data gathering. Until a predictive paradigm (mechanistically underpinned) is developed that can provide a verifiable link between natural products and health endpoints, opportunities to optimise natural product levels will be of largely academic interest rather than being of practical use. Unfortunately, although considerable efforts are made to develop such optimisation using both conventional breeding/strain selection and the use of various types of genetic manipulation it is unclear what the most efficacious amounts (or proportions) of a set of natural products will be. The value of such research efforts is scientifically valid but, in the context of human health and disease, questionable.

References

1. Fang J (2023) Brief history of plant breeding (III): evolution from primitive selection to domestication. Mol Plant Breed 14(17):1–5. https://doi.org/10.5376/mpb.2023.14.0017
2. Lombardo PA (2018) The power of heredity and the relevance of eugenic history. Genet Med 20:1305–1311. https://doi.org/10.1038/s41436-018-0123-4
3. Briney A (2023) History and overview of the green revolution. ThoughtCo, 5 April 2023, thoughtco.com/green-revolution-overview-1434948
4. Pingali PL (2012) Green revolution: impacts, limits, and the path ahead. Proc Natl Acad Sci U S A 109(31):12302–12308. https://doi.org/10.1073/pnas.0912953109. Epub 2012 Jul 23. PMID: 22826253; PMCID: PMC3411969
5. Jhingan S, Harloff HJ, Abbadi A et al (2023) Reduced glucosinolate content in oilseed rape (*Brassica napus* L.) by random mutagenesis of *BnMYB28* and *BnCYP79F1* genes. Sci Rep 13:2344. https://doi.org/10.1038/s41598-023-28661-6

6. Velasco P, Soengas P, Vilar M, Cartea ME, del Rio M (2008) Comparison of glucosinolate profiles in leaf and seed tissues of different Brassica napus crops. J Am Soc Hortic Sci 133(4):551–558
7. Samtiya M, Aluko RE, Dhewa T (2020) Plant food anti-nutritional factors and their reduction strategies: an overview. Food Prod Process and Nutr 2:6
8. Enneking D, Wink M (2000) Towards the elimination of anti-nutritional factors in grain legumes. In: Knight R (ed) Linking research and marketing opportunities for pulses in the 21st century. Current plant science and biotechnology in agriculture, vol 34. Springer, Dordrecht. https://doi.org/10.1007/978-94-011-4385-1_65
9. Pavicich MA, Roose L, Meerpoel C et al (2024) Unraveling the fate of mycotoxins during the production of legume protein and other derived products. npj Sci Food 8:59. https://doi.org/10.1038/s41538-024-00303-9
10. Hasan N, Choudhary S, Naaz N et al (2021) Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. J Genet Eng Biotechnol 19:128. https://doi.org/10.1186/s43141-021-00231-1
11. Husnik F, McCutcheon JP (2018) Functional horizontal gene transfer from bacteria to eukaryotes. Nat Rev Microbiol 16(2):67–79. https://doi.org/10.1038/nrmicro.2017.137. Epub 2017 Nov 27. PMID: 29176581
12. Emamalipour M, Seidi K, Zununi Vahed S, Jahanban-Esfahlan A, Jaymand M, Majdi H, Amoozgar Z, Chitkushev LT, Javaheri T, Jahanban-Esfahlan R, Zare P (2020) Horizontal gene transfer: from evolutionary flexibility to disease progression. Front Cell Dev Biol 8:229. https://doi.org/10.3389/fcell.2020.00229. PMID: 32509768; PMCID: PMC7248198
13. Liu Z, Cheema J, Vigouroux M et al (2020) Formation and diversification of a paradigm biosynthetic gene cluster in plants. Nat Commun 11:5354. https://doi.org/10.1038/s41467-020-19153-6
14. Miller K, Eggenberger AL, Lee K et al (2021) An improved biolistic delivery and analysis method for evaluation of DNA and CRISPR-Cas delivery efficacy in plant tissue. Sci Rep 11:7695. https://doi.org/10.1038/s41598-021-86549-9
15. Castle LA, Siehl DL, Gorton R, Patten PA, Chen YH, Bertain S, Cho HJ, Duck N, Wong J, Liu D, Lassner MW (2004) Discovery and directed evolution of a glyphosate tolerance gene. Science 304(5674):1151–1154. https://doi.org/10.1126/science.1096770. PMID: 15155947
16. Li Z, Liu Y, Yuan S et al (2021) Fine mapping of the major QTLs for biochemical variation of sulforaphane in broccoli florets using a DH population. Sci Rep 11:9004. https://doi.org/10.1038/s41598-021-88652-3
17. Butelli E, Titta L, Giorgio M, Mock HP, Matros A, Peterek S, Schijlen EG, Hall RD, Bovy AG, Luo J, Martin C (2008) Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nat Biotechnol 26(11):1301–1308. https://doi.org/10.1038/nbt.1506. Epub 2008 Oct 26. PMID: 18953354
18. Van Dael P (2021) Role of n-3 long-chain polyunsaturated fatty acids in human nutrition and health: review of recent studies and recommendations. Nutr Res Pract 15(2):137–159. https://doi.org/10.4162/nrp.2021.15.2.137. Epub 2021 Jan 4. PMID: 33841720; PMCID: PMC8007408
19. Qawasmi A, Landeros-Weisenberger A, Bloch MH (2013) Meta-analysis of LCPUFA supplementation of infant formula and visual acuity. Pediatrics. 131(1):e262–e272. https://doi.org/10.1542/peds.2012-0517. Epub 2012 Dec 17. PMID: 23248232; PMCID: PMC3529943
20. Verfuerden ML, Dib S, Jerrim J, Fewtrell M, Gilbert RE (2020) Effect of long-chain polyunsaturated fatty acids in infant formula on long-term cognitive function in childhood: a systematic review and meta-analysis of randomised controlled trials. PLoS ONE 15(11):e0241800. https://doi.org/10.1371/journal.pone.0241800.PMID:33152012;PMCID:PMC7644261
21. Willatts P et al Effect of long-chain polyunsaturated fatty acids in infant formula on problem solving at 10 months of age. The Lancet 352(9129):688–691
22. Gerster H (1998) Can adults adequately convert alpha-linolenic acid (18:3n–3) to eicosapentaenoic acid (20:5n–3) and docosahexaenoic acid (22:6n–3)? Int J Vitam Nutr Res 68(3):159–173 PMID: 9637947

23. Santigosa E, Brambilla F, Milanese L (2021) Microalgae oil as an effective alternative source of EPA and DHA for Gilthead Seabream (Sparus aurata) aquaculture. Animals (Basel) 11(4):971. https://doi.org/10.3390/ani11040971.PMID:33807244;PMCID:PMC8065835
24. Peng Y-F, Chen W-C, Xiao K, Xu L, Wang L, Wan X (2016) DHA production in Escherichia coli by expressing reconstituted key genes of polyketide synthase pathway from marine bacteria. PLoS ONE 11(9):e0162861. https://doi.org/10.1371/journal.pone.0162861
25. Zanetti F, Alberghini B, Marjanović Jeromela A et al (2021) Camelina, an ancient oilseed crop actively contributing to the rural renaissance in Europe. A review. Agron Sustain Dev 41:2. https://doi.org/10.1007/s13593-020-00663-y
26. Han L, Usher S, Sandgrind S, Hassall K, Sayanova O, Michaelson LV, Haslam RP, Napier JA (2020) High level accumulation of EPA and DHA in field-grown transgenic Camelina—a multi-territory evaluation of TAG accumulation and heterogeneity. Plant Biotechnol J 18(11):2280–2291. https://doi.org/10.1111/pbi.13385. Epub 2020 May 8. PMID: 32304615; PMCID: PMC7589388
27. Harwood JL (2019) Algae: critical sources of very long-chain polyunsaturated fatty acids. Biomolecules 9(11):708. https://doi.org/10.3390/biom9110708.PMID:31698772; PMCID:PMC6920940
28. Fernbach PM, Light N, Scott SE et al (2019) Extreme opponents of genetically modified foods know the least but think they know the most. Nat Hum Behav 3:251–256. https://doi.org/10.1038/s41562-018-0520-3
29. Goldner Lang I (2023) "Laws of Fear" in the EU: The precautionary principle and public health restrictions to free movement of persons in the time of COVID-19. Eur J Risk Regul 14(1):141–164. https://doi.org/10.1017/err.2020.120
30. Seifert F (2020) National specificity and convergence in the European anti-GM movement: the cases of Austria, Germany, France, Spain and the UK. Innov: Eur J Soc Sci Res 34(4):511–532. https://doi.org/10.1080/13511610.2020.1766950
31. Scaltsas T (1980) The Ship of Theseus. Analysis 40(3):152–157. http://www.jstor.org/stable/3327668
32. Qu C, Zhu M, Hu R et al (2023) Comparative genomic analyses reveal the genetic basis of the yellow-seed trait in Brassica napus. Nat Commun 14:5194. https://doi.org/10.1038/s41467-023-40838-1
33. Zhu C, Delmont TO, Vogel TM, Bromberg Y (2015) Functional basis of microorganism classification. PLoS Comput Biol 11(8):e1004472. https://doi.org/10.1371/journal.pcbi.1004472.PMID:26317871;PMCID:PMC4552647
34. Shino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol 169:5429–5433. https://doi.org/10.1128/jb.169.12.5429-5433.1987
35. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol Direct 1:7. https://doi.org/10.1186/1745-6150-1-7
36. Asmamaw M, Zawdie B (2021) Mechanism and applications of CRISPR/Cas-9-mediated genome editing. Biologics 21(15):353–361. https://doi.org/10.2147/BTT.S326422.PMID:34456559;PMCID:PMC8388126
37. Liu Q, Yang F, Zhang J, Liu H, Rahman S, Islam S, Ma W, She M (2021) Application of CRISPR/Cas9 in crop quality improvement. Int J Mol Sci 22(8):4206. https://doi.org/10.3390/ijms22084206.PMID:33921600;PMCID:PMC8073294
38. Göttl VL, Schmitt I, Braun K, Peters-Wendisch P, Wendisch VF, Henke NA (2021) CRISPRi-library-guided target identification for engineering carotenoid production by Corynebacterium glutamicum. Microorganisms 9(4):670. https://doi.org/10.3390/microorganisms9040670.PMID:33805131;PMCID:PMC8064071
39. Wu S, Zhang T, Wang Y, Chen M, Yang J, Li F, Deng Y, Zhu Z, Lei J, Chen G, Cao B, Chen C (2022) BocODD1 and BocODD2 regulate the biosynthesis of progoitrin glucosinolate in

Chinese kale. Int J Mol Sci 23(23):14781. https://doi.org/10.3390/ijms232314781.PMID:36499110;PMCID:PMC9739482
40. Singh A (2017) Glucosinolates and plant defense. In: Mérillon JM, Ramawat K (eds) Glucosinolates. Reference Series in Phytochemistry. Springer, Cham. https://doi.org/10.1007/978-3-319-25462-3_8
41. Firn R (2010) 'Why do organisms make NPs?' Nature's chemicals: the natural products that shaped our world (Oxford, 2009; online edn, Oxford Academic, 2010), https://doi.org/10.1093/acprof:oso/9780199566839.003.0005
42. Manor O, Dai CL, Kornilov SA et al (2020) Health and disease markers correlate with gut microbiome composition across thousands of people. Nat Commun 11:5206. https://doi.org/10.1038/s41467-020-18871-1
43. Hartmann A, Binder T, Rothballer M (2024) Quorum sensing-related activities of beneficial and pathogenic bacteria have important implications for plant and human health. FEMS Microbiol Ecol 100(6):fiae076. https://doi.org/10.1093/femsec/fiae076
44. Cohen SN, Chang AC, Boyer HW, Helling RB (1973) Construction of biologically functional bacterial plasmids in vitro. Proc Natl Acad Sci U S A 70(11):3240–3244. https://doi.org/10.1073/pnas.70.11.3240.PMID:4594039;PMCID:PMC427208
45. Van Vaeck M et al (1987) Transgenic plants protected from insect attack. Nature 328:33–37
46. Lobato-Gómez M, Hewitt S, Capell T, Christou P, Dhingra A, Girón-Calva PS (2021) Transgenic and genome-edited fruits: background, constraints, benefits, and commercial opportunities. Hortic Res 8(1):166. https://doi.org/10.1038/s41438-021-00601-3.PMID:34274949;PMCID:PMC8286259
47. Kramer MG, Redenbaugh K (1994) Commercialization of a tomato with an antisense polygalacturonase gene: the FLAVR SAVR™ tomato story. Euphytica 79:293–297. https://doi.org/10.1007/BF00022530
48. Potrykus I (2001) Golden rice and beyond. Plant Physiol 125(3):1157–1161. https://doi.org/10.1104/pp.125.3.1157
49. Brookes G, Barfoot P (2020) Environmental impacts of genetically modified (GM) crop use 1996–2018: impacts on pesticide use and carbon emissions. GM Crops Food 11(4):215–241. https://doi.org/10.1080/21645698.2020.1773198.PMID:32706316;PMCID:PMC7518756
50. Directive (EU) 2015/412 of the European Parliament and of the Council of 11 March 2015 amending Directive 2001/18/EC as regards the possibility for the Member States to restrict or prohibit the cultivation of genetically modified organisms (GMOs) in their territory text with EEA relevance OJ L 68, 13.3.2015, pp 1–8
51. Krishnamurti C, Rao SC (2016) The isolation of morphine by Serturner. Indian J Anaesth 60(11):861–862. https://doi.org/10.4103/0019-5049.193696.PMID:27942064;PMCID:PMC5125194
52. Carsanba E, Pintado M, Oliveira C (2021) Fermentation strategies for production of pharmaceutical terpenoids in engineered yeast. Pharmaceuticals (Basel) 14(4):295. https://doi.org/10.3390/ph14040295.PMID:33810302;PMCID:PMC8066412
53. Garner KL (2021) Principles of synthetic biology. Essays Biochem 65(5):791–811. https://doi.org/10.1042/EBC20200059.PMID:34693448;PMCID:PMC8578974
54. Hug JJ, Krug D, Müller R (2020) Bacteria as genetically programmable producers of bioactive natural products. Nat Rev Chem 4:172–193. https://doi.org/10.1038/s41570-020-0176-1
55. Tong Y, Whitford CM, Blin K, Jørgensen TS, Weber T, Lee SY (2020) CRISPR-Cas9, CRISPRi and CRISPR-BEST-mediated genetic manipulation in streptomycetes. Nat Protoc 15(8):2470–2502. https://doi.org/10.1038/s41596-020-0339-z. Epub 2020 Jul 10. PMID: 32651565
56. Chen J, Wu Q, Hawas UW et al (2016) Genetic regulation and manipulation for natural product discovery. Appl Microbiol Biotechnol 100:2953–2965. https://doi.org/10.1007/s00253-016-7357-3
56. Cho SJ, Ryu JH, Surh YJ (2019) Ajoene, a major organosulfide found in crushed garlic, induces NAD(P)H:quinone oxidoreductase expression through nuclear factor E2-related factor-2 activation in human breast epithelial cells. J Cancer Prev 24(2):112–122. https://doi.org/10.15430/JCP.2019.24.2.112. Epub 2019 Jun 30. PMID: 31360690; PMCID: PMC6619855

57. Houghton CA, Fassett RG, Coombes JS (2016) Sulforaphane and other nutrigenomic Nrf2 activators: can the clinician's expectation be matched by the reality? Oxid Med Cell Longev 2016:7857186. https://doi.org/10.1155/2016/7857186. Epub 2016 Jan 6. PMID: 26881038; PMCID: PMC4736808
58. Wang Y, Wang J (2020) Modelling and prediction of global non-communicable diseases. BMC Public Health 20(1):822. https://doi.org/10.1186/s12889-020-08890-4.PMID:32487173; PMCID:PMC7268487
59. Smokovski I, Steinle N, Behnke A et al (2024) Digital biomarkers: 3PM approach revolutionizing chronic disease management—EPMA 2024 position. EPMA J 15:149–162. https://doi.org/10.1007/s13167-024-00364-6

Chapter 6
Bioengineering

Abstract The principles of bioengineering underpins the use of living systems as biological machines to effectively make desirable products. The concepts of vitalism and agency are important when considering biological systems. The quantum mechanical understanding of consciousness suggests that natural products may act at a quantum level. Collections of organisms allow for the concept of ultra-reductionism (Descartes) to be explored. The example of polyketides can be used to explain bioengineering and the manipulation of systems to create new products. The concept of silenced genes, metagenomics and their incorporation and expression in suitable hosts, increases the numbers of possible natural products. The BioInternet of Things (BioIoT) has significant bioengineering potential. Communication within and between biological systems is important if this potential is to be realised. Homologous and heterologous gene expression are possible and are a valuable source of specific natural products. The development of systems biology is essential to support the concept of bioengineering.

Keywords Polyketides · Heterologous gene expression · Combinatorial PKS

Bioengineering can be defined as the application of the principles of engineering (including the use of engineering design principles and the modification or replacement of biological systems with engineered substituents) in biological systems. It is, by its nature, an interfacial discipline and, as with many interfaces, a potent source of new ideas and synergies. It is worth noting, however, that while many applications seem to see the flow of ideas, as implied in the above definition, as being from engineering to biology, the application of biological principles to engineering is equally valid and, indeed, underpins areas such as biocomputing, self-replication, self-repair and aspects of biomimicry. The basic, traditional principles of bioengineering have been referred to previously (Chapter Five). In the context of natural products, bioengineering can be directed towards the products themselves (particularly when they have a high intrinsic value), to the system used to produce the natural product, to modification of the natural product to produce an unnatural product and to the development of sustainable sources of products. The breadth of products considered includes,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_6

therefore, bioactive constituents of various types, primary products that can be used as feedstocks and speciality chemicals. The importance of bioengineering was identified, along with other areas of biological relevance and importance in a report by McKinsey in 2020 (The Bio Revolution Innovations transforming economies, societies, and our lives McKinsey, May 2020). In this report, it was suggested that the 4 areas that were likely to be at the forefront of advances were:

(1) biomolecules—the mapping, measuring and engineering of molecules;
(2) biosystems—the engineering of cells, tissues, and organs;
(3) biomachines—the interface between biology and machines (note this refers to the way that biological systems are linked to machines and is different to the concept of living systems as machines—see Sect. 6.1); and
(4) biocomputing—the use of cells or molecules such as DNA for computation.

In addition it was estimated that the biorevolution would have an annual value of up to US$4T in the period 2030–2040.

6.1 Living Systems as Machines

The concept that living systems are basically mechanisms is one that has been the subject of study over many centuries and is a consequence of the application of reductionism to the study of how things operate. Trying to emulate how things work has provided the intellectual basis for the design and construction of automata and, in addition to signifying the application of engineering to living systems, is also a form of biomimicry. The development of viable models for what happens in nature extends from the molecular (atomic) scale to complete systems. Discussions about how things work has been underpinned by scientific endeavours and philosophical imaginings. Bioengineering itself, as noted previously, brings the principles of engineering design to the study of how biological systems work but, more than simply being an observational exercise, it also deals with how to manipulate systems using engineering principles.

Although it is not explicitly stated, the principle of bioengineering is underpinned by particular assumptions. The first of these is that the concept of vitalism—the idea that living systems have some unique characteristic that contributes to their being alive and differentiates them from non-living objects—is false. The refutation of vitalism is important since it implies that living systems can be created without external intervention. During this process, there is no point at which a collection of inanimate fragments becomes ‘alive’. The second important fact underpinning bioengineering is that living systems can be interpreted in terms of mechanistic science and therefore living systems do not possess ‘agency’. There is no such thing as a living system having some unique ‘will’ that dictates its behaviour. The mechanical nature of life was developed in the seventeenth century and considerable effort was expended demonstrating that creation of mechanical equivalents to living systems was possible. Such automata were popular but there was also a serious point to them.

They can be thought of as an early example of what later became known as the Turing test. While the Turing test (and its consequences) has been the subject of considerable interpretation and thought, in its simplest incarnation it asks whether a machine can demonstrate intelligent behaviour [1]. The Turing test poses the question of whether a machine can respond in a way that a human cannot distinguish it from another human. There are profound philosophical consequences to this question and, while some have argued that there is nothing unique about humans (or other living systems) and therefore it is simply a question of creating the right machinery, others have argued that machines cannot possess a mind or any form of understanding or consciousness. Indeed, Penrose and Hameroff have argued that consciousness is only possible if quantum effects are included as part of the explanation [2, 3]. Without going into the theory in detail, it suggests that structures within neurones (microtubules) are capable of quantum superposition and it is the collapse of these quantum states that gives rise to self awareness and hence consciousness. One of the proposed proofs of the effect is the claim that drugs affecting microtubules within neurons delay the onset of unconsciousness caused by anaesthetic gases [4]. It should be noted, however, that the attribution of the effect to a quantum explanation is only made in the absence of any other hypothesis. In that context, it is an absence of alternative explanations rather than a validation of the quantum theory of consciousness. Nonetheless, if this concept can be substantiated then natural products which bind to and affect microtubules may have an impact on consciousness. Natural microtubule disruptors are commonly used in cancer chemotherapy. One class of these compounds (Vinca alkaloids) are known to be neurotoxic through disruption of microtubule dynamics. Poisoning with another microtubule-interacting natural product, colchicine, produced symptoms including dizziness (the patient later died following severe organ failure) [5].

Others have also developed ideas about consciousness and mind in the context of quantum events. This is not as outlandish as it may at first appear since the interaction of quantum physics with biology is an established area of research [6]. It has been postulated that nature makes use of quantum coherence in energy transfer. One of the problems in the rationalisation of quantum effects in biology is assembling proof that biological events take place on a timescale that is congruent with quantum mechanics [7].

One of the problems with the banishment of the concept of an external agency for living systems is that there are some aspects of living systems that cannot be explained as the sum of all the molecular parts. The lack of an internal agency in living systems suggests also that living systems are passive and are acted upon by mutation and selection. This is somewhat akin to the idea that the proof that the world exists and that the existing life forms are the best adapted to the world is a consequence of the somewhat circular argument that current living systems must, by reason of their existence, be the best adapted option. This concept has been challenged and the existence of ‘internal agency’ in a living system has been postulated [8]. This theory suggests that living systems can purposely act upon and control interactions with their environment and this enables living systems to control and manage their own existence. This suggests a relationship between agency and evolution and that the latter is not purely random. Taking this a stage further, it

is possible to conceive of multicellular systems (such as microbiomes) as single, multi-faceted entities with a degree of shared agency. In this model, communication between the different components could arise through the production of quorum sensing molecules and these, then, would assume the role of neurotransmitters like serotonin. It would also imply that the individual cellular components (microbes) exhibit a degree of integrated behaviour. The question that is raised is twofold—can living systems possess an internal agency and what are the limits of that agency if it exists? One way to examine the question is to take, as an initial premise, a complete lack of agency and to examine the rational consequences of this.

One of the premises that underpins the argument for the absence of agency is the notion that all living systems can be fully explained as the sum of their constituent parts. The lack of complete knowledge of all the components of all living systems which was noted previously, is then either a consequence of a lack of knowledge of all the constituents or a gap in awareness of the nature of their interdependencies. In the absence of an acceptance of incomplete knowledge, there is a tendency to ascribe these parts of living systems which are not completely understood to some external agent (this is the basis of much theocratic thought) or internal agency. In the case of the latter, if we assume, a priori, that nature cannot, of itself, possess agency in the form of will, consciousness or perception then the fact that these exist at all must be due to some external guiding force. Furthermore, the existence of some inexplicable biological features implies that there must be some divine 'maker' who has responsibility for their creation. Much of this thinking—at least from the context of natural products—has been discredited by the development of the concept of evolution through natural selection by Darwin and others. Incremental change on a molecular level (mutation) will change the nature of a protein and this in turn will have consequences for the pathway, cell and (eventually) organism. Such random events and their consequences do not require that we postulate the existence of an agency (internal or external). In addition, the acceptance of some external agency is even more problematic since it is an ever-changing concept. As scientific knowledge increases, things which were unknown and previously ascribed to some form of external agency become explicable in terms of the revised knowledge base and so the nature and requirement for agency will alter. In short, scientific and technical progress provides further insights and these can help in the formulation of new hypotheses which can then be tested and may reveal further insights—and so the process continues. For example, nuclear magnetic resonance (nmr) spectroscopy can provide, by inference, information about the structure of a molecule. Being able to propose the molecular structure, a chemist can predict how the molecule will respond to a specific set of circumstances (reaction conditions) and what the likely outcome will be. This hypothesis can be tested and the resultant structure examined again using nmr spectroscopy to confirm not only that the reaction pathway has been followed but also that the original structural assignment was correct. This accumulation and expansion of knowledge is an ongoing process which means any external agency cannot be a fixed point.

While Descartes (and others including Leibniz) in the seventeenth century continued to fixate upon the physical manifestations of nature as a mechanical

entity, this led them to suggest that mechanical systems like clocks and automata could be considered as models whose basis for operating could also explain how animals, plants, and the human body operate. By explaining complex living systems in this way, Descartes wanted to prove the universality of mechanistic science as a way of explaining all life. This is, in some ways, the ultimate in scientific reductionism. Progress in science, in reality, came more through observation than mechanical hypotheses and the development of instruments to enable magnification and later molecular (particularly spectroscopic) techniques enabled chemists and biologists to examine the world on a microscopic scale and to increase their powers of classification of living systems based on closer examination. This would eventually, in the twentieth century, lead to the understanding of molecular interactions on an individual molecule basis.

The ultra-reductionist approach to living systems promoted by Descartes (reducing them to their component parts) was mechanistically very useful as an underpinning for the observational studies however an understanding of individual reactions and the mechanism by which they occur is only of limited and fractional value. Living systems are, as Descartes suggested, composed of multiple interacting components—even if we do not know what all the constituent parts are or how precisely they interact. By increasing our understanding of how these components work together, it is possible not only to develop a greater understanding of systems but also to manipulate both the individual components and the way they interact. This is the basis of the principal approaches to bioengineering. Not all systems can currently be bioengineered but the basic principles remain the same for all of them. The reverse process, which is somewhat less frequently examined is the use of biological knowledge to advise our engineering progress and this is also an observational and knowledge-based process.

6.2 Principles of Bioengineering

Bioengineering is the development of methods which combine the breadth and flexibility of biological systems with the control exerted by the application of engineering. While it can be a mutually advantageous process (of use to developing better biological and engineering solutions), there is a tendency to look at systems purely from the perspective of the addition of engineering principles to biological systems. While many of the current applications of bioengineering apply to health and medicine, its potential stretches beyond this area. In addition, there are also links to the concept of living systems as machines which was discussed earlier. Unsurprisingly much of the interest in bioengineering involves communication at the interface between mechanical and living systems. Rather than being two systems running side by side, bioengineering ideally involves the development of synergies between living and mechanical systems and not simply their additive interaction. Examples of this can come from several directions.

In order to carry out bioengineering, a prerequisite is an understanding of the language used by biological systems in order that effective communication can be enabled. In addition, it is also necessary to have sufficient knowledge of the key principles of engineering. The development of engineering principles has a long and distinguished history. Newton codified engineering by developing equations that correctly described the mechanics of how things work. It is not true, however, to suggest that Newton explained how natural phenomena like gravity function. His was more of an observational approach which explained the consequences of the existence of gravity rather than the nature of gravity itself. While the equations he developed enable mechanics to be understood they also created the framework that allowed empirical scientists and engineers to develop practical tools such as steam engines. The idea of the existence of 'caloric' (a fluid that travels from hotter to colder bodies) had to be discarded and heat had to reimagined in terms of atomic motion. This enabled engineers to manipulate systems in a predictable fashion and to create machines that would function according to anticipated criteria. The language of engineering describes the equations and these reflect the properties of the atoms and molecules that make up a substance. In addition, by using equations that described the nature of mechanical systems, scientists and engineers could develop new mechanical systems that, because they obeyed the laws implicit in the equations, would have properties that not only were able to be predicted but also could be manipulated to give a desired and useful endpoint.

In a similar fashion, the description of biological systems required an understanding both of the alphabet that biological systems used and how this could be manipulated to create predictable and novel biological systems. Some of the underpinning ideas concerning how living systems worked and could be manipulated was derived, by inference, from the work of Mendel and Darwin. Some concepts, however (just like the caloric theory) were wrong. For example, the theory of the homunculus had a significant impact on how scientists thought of heredity. According to this theory, a germ cell contained a tiny, fully formed individual which then grew into a complete human. It was originally thought that the origin of this cell was as a sperm cell. Much of the theory of homunculus died with the demise of alchemy and the advent of microscopy—which allowed for the examination of single cells in great detail. Paracelsus even developed a recipe for the creation of a homunculus—from which humans of various types could be derived and the creation of artificial life was tied to the concept of spontaneous generation. There may even be echoes of this in later work such as Mary Shelley's Frankenstein which was published in 1818—some 40 years before the publication of 'On the Origin of Species' by Darwin but 10 years after the publication by Dalton of his atomic theory on the nature of matter. Reductionism and the theory of atoms provided an intellectual underpinning for engineering before the development of theories to explain how living systems evolved.

The importance of genetics in establishing the rules of heredity—which was an essential prerequisite to understanding how living systems functioned—required other discoveries and observations. The reductionist nature of chemistry, which had revealed the nature of matter as being composed of atoms, suggests that determining

how living systems function would require the study of individual components first and then the investigation into how they interact to create a complete system. This approach was not readily available since there were no tools that could be used to examine genetic material in detail until the second half of the twentieth century. Models could only be created by inference and by studying complete living systems and their interactions. Similarly, determining how living systems were replicated was dependent upon observation of whole systems and extrapolation to the likely characteristics of models.

While Darwin developed the theory of evolution by means of natural selection of variants with beneficial traits and Mendel formulated laws governing inheritance, the mechanism by which changes in species were created on an organismic or molecular level remained unclear. The relationships between how living systems resembled each other formed the basis of classification. These phenotypic relationships were established primarily by observation with the naked eye and later with magnifying systems. Science looked outwards to the planets and stars using telescopes and inwards to a whole new 'invisible' world using microscopes. Before the beginning of the twentieth century, the discovery of the cell nucleus and that genetic material in a cell was located in the nucleus were key findings. The latter discovery was by the German biologist Ernst Haekel. In addition to being a Professor of Zoology at Jena, Haekel was also an artist and an ardent supporter of Darwin. Indeed he was referred to as the 'German Darwin' [9]. While some aspects of heritability were known and the cellular location of genetic material was appreciated, there was still considerable work to be done to identify how biological machinery worked and, perhaps more importantly, how it could be manipulated. As with many other areas of science, the advances were driven both by ingenuity (the development of theories) and by the application of new and better techniques.

On a macro level, the understanding of how heredity worked was empirically understood by Mendel including the principles that lay behind dominant and recessive traits. The next stage was the discovery of chromosomes and this was illustrated by Morgan on the sex-linked inheritance of the white eye trait in fruit flies. In addition Morgan discovered the role of the organisation of genes into chromosomes and the linkage between certain genes which are found on the same chromosome and are inherited together. Overall, Morgan (with others) re-popularised Mendelian theories about the nature and organisation of genes and the role of chromosomes [10].

Despite the huge amount of observational work carried out by Morgan and coworkers, the nature and the role of genes remained controversial. The dilemma was solved through the work of Avery and others. In 1940, while working on the 'transforming factor' that converted a live but harmless form of a bacterium into a lethal form, he proposed that this was composed of nucleic acids [11]. While the structure of DNA was elucidated by Watson and Crick in 1953 [12], the nature of the mechanism of transfer of genetic material and the codon theory of inheritance was also developed. The structure of DNA could not have been confirmed without the development of x-ray crystallography and it was this and subsequent molecular techniques that facilitated the manipulation of genetic material to give desired and predictable traits. These concepts formed the basis of bioengineering. It can be seen, therefore,

that the lineage of bioengineering involves classification of living systems, knowledge of the important aspects of phenotype and physiological characteristics, the concept of heredity and reproduction, the mechanics of how heredity works through a genetic framework and the molecular events that provide an entryway for genetic manipulation which can then influence, in a predictable way. Phenotype and physiological characteristics. In the context of natural products, manipulation at the whole gene and within-gene levels can be carried out using the molecular biology tools and understanding built up from a cascade of knowledge from Linnaeus to modern biological pathway engineering.

6.3 Bioengineering of Natural Products—The Example of Polyketides

One way of considering biological manipulation is to classify changes according to size and function. The simplest classification is at a molecular level. Compounds such as polyketides can be isolated from living entities and then used or further manipulated. The next stage, in terms of function or scale, is the biological system that produces the molecule. The system (and the machinery that provides it with functionality) can be manipulated. The third level is the interface between biological machinery and physical machines. This forms the basis of many of the medical applications of bioengineering—most commonly using a read out from a biological system and converting it into something that can be interpreted (and manipulated). This is the basis of much of the thinking behind the Internet of Things (IoT) [13] and the BioInternet of Things (Bio-IOT). The essence of a Bio-IoT action sequence is that a biological entity in one part of a system will store data then transmit it to another part of the system where an action occurs that affects the system as a whole. Just as a conventional IoT can be used to connect individual objects, manufacturing systems and networks of systems, Bio-IoT may operate on individual cells, complex systems and whole organisms. The level of complexity will increase with size, as will the value and flexibility of the data being transmitted. The concept of the Bio-IoT provides an interface with the fourth aspect of bioengineering—the idea of using biological systems to store data that can then be accessed both by biological systems and by physical entities. The three aspects of a Bio-IoT (information (data) collection and storage, data transmission and action) are scalable on a molecular, cellular and organism level. The idea of classifying biological manipulation in terms of size and function gives rise to the concept of manipulation of molecules, systems, the interface between the biological and physical worlds and the use of the Bio-IoT can be useful when considering how natural products can be useful entities in each. Natural products can be the biological molecules themselves, the indicator of the manipulation of a biological system through production of molecules with novel chemical characteristics, the mode of communication with the physical world (for example as indicators of redox status) and the mode of transmission (signalling)

between different biological systems and between the systems and their environment (for example quorum sensing). In addition, the development of bioengineering has opened up the possibility of using biomolecules for data storage and preservation.

Polyketides (and the cellular machinery that makes them) are among the most widespread examples of natural products. Polyketides are an important class of natural products whose biosynthesis is similar to that of fatty acids. They are formed from the successive Claisen-type condensation reaction of acetyl and malonyl groups with concomitant decarboxylation. The reactions are catalysed by polyketide synthases. Polyketide synthases form modular biosynthetic units which act like an assembly line. The sequence of biosynthesis of a polyketide is in essence a modular process. The key initial step is the transfer of an acyl starter subunit onto an acyl carrier protein (ACP) catalysed by an acyl transferase enzyme. The polyketide synthase family of enzymes then catalyses a range of reactions. Different starter or extender units are used to create structural diversity. The modular (type I) KTS enzymes contain different catalytic domains including ketoreductase, dehydratase and enoyl reductase activities and each set of modules is responsible for a single iteration of polyketide chain elongation as well as other modification reactions (for example, reduction, dehydration). Each domain is linked and polyketides are synthesized in a collinear fashion therefore addition, removal and/or substitution of these domains or modules will theoretically result in defined alterations of the end product. Post-translational modification (glycosylation, methylation, hydroxylation) adds a further level of complexity. In addition, if the sequence of the constituent enzymes is known then the structure of the final product can be predicted. Two other types of PKS enzymes (type II and type III) are known and have slightly different characteristics but these will not be described in detail. In essence polyketide biosynthesis is an assembly line process. As such it has characteristics which are particularly suitable for engineering and it is this aspect of polyketide biosynthesis that will form the focus of this discussion. The simplest way to modify the biosynthetic process is to rearrange the modules involved in the various stages of the process. In addition to their modular nature, modification of the enzymic sub-units has also been a goal of natural product chemists since this also offers the possibility of creating completely novel molecules. These products have been suggested to be ‘unnatural products’ and their exploitation has been anticipated [14]. Polyketides are complex structures and therefore the post-biosynthetic modification of the structures in a precise way is challenging. For this reason, the biological engineering of the molecule *in situ* is the preferred option. This approach, which makes use of a number of genes for different enzymes from different sources which can be modified in vitro prior to being introduced into a suitable host, has been termed ‘combinatorial biosynthesis’ [15]. It should be noted that a combinatorial biosynthesis strategy has been used for the production of other ‘unnatural products’ but the case of polyketides is particularly illustrative.

Perhaps the simplest way of finding alternative polyketides makes use of existing pathways that are poorly expressed or not expressed at all. Genome mining can be used to find new PKS-related genes based on sequence similarity with known members of the class (or genes for other natural products that can be employed in

mixed biosynthetic strategies) and this can allow for the development of novel products. One way of changing the metabolic profile of secondary metabolite production is through gene silencing of some competing pathways thereby changing the metabolic flux through these related pathways by changing substrate availability. This approach can, in principle, work with any secondary metabolite for which details of the biosynthetic pathway are known. This type of approach can be used in the case of any natural product. For example, in the case of glucosinolates, silencing of the methylthioalkylmalate (MAM) synthase gene using RNAi can change the profile of, particularly aliphatic, glucosinolate production in *Brassica napus* [16]. Modification of the precursor pathway in the case of polyketides can alter the polyketide being produced (see Lechner A, Wilson MC, Ban YH et al. (2013) Designed biosynthesis of 36-methyl-FK506 by polyketide precursor pathway engineering. ACS Synth Biol 2(7):379–383 for an example of this). The third option is activation of silenced genes. There are a number of approaches to the activation of silenced biosynthetic gene clusters [17] and such approaches are of value.

An alternative approach is to make use of hybrid gene clusters incorporating genes from different species. The modular structure of type I PKSs mean that they are the most popular choice of biosynthetic gene clusters to incorporate other genes. This can be carried out by adding or deleting modules from a biosynthetic gene cluster. In addition, directed mutations in the acyl transferase domain of PKS clusters can change the substrate specificity and hence the nature of the starter or extender units that can be used. This can result in the formation of a suite of unnatural products. Changes to other units in the gene cluster (ketoreductase and enoylreductase) can affect the chirality of the polyketide being produced. Other strategies based on modular rearrangement of PKS clusters have been suggested [18].

Finally, enzymes responsible for post-translational modification of polyketides have also been modified (or unnatural substrates have been used) to modify the attributes of the produced polyketide structure. These modifications can result in the formation of a range of structures with different chemical and biological properties.

The modular nature of PKS gene clusters and the understanding of how the different modules work has then made polyketides an attractive option for modification of biosynthesis either by module rearrangement or by mutation of individual genes. This can be part of a strategy to find novel molecules (natural and unnatural) or to optimise production of a particular product in a suitable host. Further work will extend the methodology for manipulation of polyketides (and other natural products), and this will greatly enhance the options available for the further exploitation of natural and unnatural products through bioengineering.

6.4 Gene Manipulation and Heterologous Gene Expression Systems

The use of gene manipulation (rearrangement or mutation) was referred to in the context of polyketide biosynthesis. In addition to this specific example, gene manipulation can be carried out as part of a number of optimisation processes. These include the following:

- Novel product formation and discovery of new or optimised bioactive compound formation
- Formation of molecules with enhanced stability, more convenient isolative properties, better physicochemical properties.
- Choice of a suitable host with optimal fermentative properties, enhanced thermal or chemical stability, enhanced production, rapid growth.

The development of methods to genetically engineer living species was dependent upon two key advances—the knowledge of how traits were passed on and the link between those traits and the genetic material (genes) in each cell. Understanding the link between genotype and phenotype and how altering the former could change the latter—and, indeed, provide the mechanism and means for the planned alteration of traits—was perhaps the single most significant enabling discovery that allowed bioengineering to be developed.

While enzyme reactions are often elegant and precise, this is not invariably the case. Some biosynthetic pathways (and the enzymes that constitute them) are tolerant of a range of starting materials. For example the class of enzymes responsible for the synthesis of terpenes (terpene synthases) can also catalyse a reaction to produce non-terpenes [19]. Many other examples exist. Previously the potential for manipulation of polyketide pathways has been noted and this has formed the basis of chemoenzymatic synthesis [20]. The next stage is to manipulate biosynthetic enzymes such that they will accept a range of substrates and this has been applied to industrial processes [21]. This enables optimisation of the producer system such that production is optimised.

While in many cases the natural host is the preferred and simplest option as producer species, this can prove to be a significant barrier if, for example, yields are low or the host is slow growing. The following is an example of this challenge. The anti-malarial drug, quinine, is produced by the bark of a tree (*Cinchona pubescens*) and this remains the most commercially viable route to the alkaloid. Quinine has a complex structure (Fig. 6.1) and, while various routes have been developed, synthesis on an industrial scale is not a viable option.

The biosynthesis of quinine has been extensively studied however attempts to produce the alkaloid in plant cell culture rather than harvesting mature tree bark were unsuccessful [22]. Transcriptomic and metabolomic studies have revealed details about the biosynthesis of quinine [23]. Detailed knowledge of the biosynthetic pathway for a natural product is a crucial step in optimising (or manipulating)

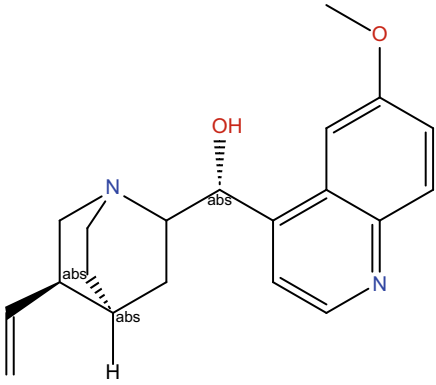


Fig. 6.1 Structure of quinine

its production. Since quinine is only produced by mature trees, a route for its production using biotechnology and a faster growing (and more amenable) host would be of value. Two of the most frequently used organisms for production of natural products are the bacterium *E. coli* and the yeast *S. cerevisiae* and either may be viable alternatives.

The use of *E. coli* as a heterologous host for the production of natural products is well known [24]. It has a number of advantages over other systems and, by exploring the use of *E. coli* and *S. cerevisiae*, it will be possible to create a more general overview of the use of heterologous gene expression and natural product synthesis in biological systems.

One of the major advantages in the use of *E. coli* is its high growth rate. This means not only can compounds be produced quickly but the success or failure of particular strategies rapidly becomes apparent and changes can be made. It should be noted that the production of foreign compounds by heterologous gene transfer (and therefore, foreign proteins) is a non-trivial task. In the case of *E. coli*, it is a widely used host and, for the expression and production of simple proteins, it is the host of choice. There are problems if the protein is subjected to post-translational modification. Prokaryotes such as *E. coli* are unable to accomplish this and a eukaryotic host (for example the yeast *Picia*) is often preferred. Details of protocols that are known to work are collected by the Cold Spring Harbor in their journal, Cold Spring Harbor protocols and the interested researcher is referred to this excellent source of information. There are a number of factors that can lead to sub-optimal outcomes. In addition to the production and isolation of natural products themselves, it is important that the appropriate precursors are available in the new producer species and this may require further genetic manipulation. The transplanted genes must be stable in the new host and must be expressed at an optimal level. There are a number of steps and precautions that must be included. The first of these is to ensure that sufficient copies of the plasmid containing the genes that code for the biosynthetic pathway of interest are present and are expressed. In addition, it is important to distinguish between bacterial cells which contain the plasmid of interest from those which have not taken up (or are not able to express the proteins of interest). The simplest way to achieve this is to

incorporate a gene that codes for a protein that confers antibiotic resistance. If a mixed population of bacterial cells are grown on agar plates that contains the antibiotic, only those which have the antibiotic resistance gene will form colonies. Various other manipulations can be carried out and the progress of these followed by amplification of specific DNA fragments using polymerase chain reaction (PCR) followed by gel electrophoresis. Manipulation of the *E. coli* strains can be followed by monitoring for antibiotic resistance. This is a phenotypic change. The phenotypic change can be augmented by monitoring changes in the DNA (genotypic change) by using PCR and gel electrophoresis, as noted above. The ultimate phenotypic change is production of the required secondary metabolite. The limitations noted above concerning post-translational modification means that yeast and other eukaryotic microorganisms are increasingly used to produce complex biological molecules. Such processes are at the core of industrial biotechnology and, just as factories are designed with specific outputs in mind but are also able to be used flexibly and have features in common no matter what the endpoint of production happens to be, so industrial microorganisms have both the specificity of purpose and the capacity to be used as a more generic production platform. The inherent difficulties in using plants and simple harvesting as a source material for biological molecules has already been noted. In addition to the complex nature of the biological machinery of a plant and the large number of variants of natural products that many plants contain, there is only a limited degree of optimisation of yield that can be carried out. The complex metabolism, genetics and physiology of plants has been a major limitation in their exploitation as 'biological factories'. For these and other reasons, microbial factories have been the preferred route for production of valuable, plant-derived secondary metabolites [25]. Plant cell or tissue culture has developed however this is more often used for the sterile and controlled propagation of plants rather than as a source of natural products [26]. Nonetheless, with the pressure on agricultural resources and difficulty in ensuring adequate supply of plant material in some cases, plant cell culture as a source of high value natural products for industrial applications has been promoted [27]. Further progression in this area is likely however the microbial route is likely to be favoured as a production platform since there remain a number of obstacles to the agricultural production of plant natural products. One of these is the amount of waste plant material that must be discarded once the natural products of interest have been removed. Valorisation of waste is a significant issue in agriculture. Building a business based on the use of waste material has a number of dangers—not least ensuring a consistent supply of material. In some cases (for example indigo extraction from *Indigofera* species), the residue can be used as a ruminant feed however this use is limited and further studies are likely to be required to optimise the circular economy aspects of dye production from plants [28]. A similar challenge occurs with rapeseed. The seedcake produced after the extraction of the oil can be used as a high protein feedstock for several agricultural animals (particularly ruminants) however use of the seed meal is limited due to the presence of glucosinolates and palatability issues associated with their breakdown products—particularly bitterness and the presence of isothiocyanates—and the effect on the rumen microbiome [29]. It

is also possible to find other uses for the waste material from the cultivation of plants. For example, vegetable waste has been promoted as a source of natural colours [30].

As noted previously, the use of microbes as a platform for the commercial production of natural products has been a widely studied area. Bacterial systems (such as *E. coli*) are frequently used because of their rapid growth rate and the ease with which they can be genetically manipulated. It has generally been considered that post-translational modification of proteins was a relatively rare event in bacterial species. While it is perhaps more widespread than previously thought [31], it remains a barrier to the use of prokaryotic organisms and eukaryotic yeast-based systems remain the option of choice [32]. Protein engineers, like many other manipulators of living systems, have an understandable tendency to remain wedded to systems and operations that have a demonstrable track record rather than developing more esoteric methods. For example, more than more than 30,000 recombinant proteins have been successfully expressed and purified using *E. coli* and this has added to the evidence base to support this system [33]. However, many post-translational modifications are different in prokaryotes compared to eukaryotes and are often associated with virulence/pathogenicity rather than natural products [34], and this also limits their effective use for specific natural product production.

The problem of post-translational modification is not immediately solved by using yeast rather than a bacterium as a host however a greater range of products are possible. The use of *S. cerevisiae* is, in part, a historical artefact. It is one of the most studies strains and has a long history of use both industrially and domestically. The history of the use of yeast and details of the wild type organism have been reviewed [35]. One of the key discoveries, which resulted from the availability of molecular genetic methods, was the appreciation that the most studied reference strain (S288c) is somewhat anomalous in many of its characteristics. This does not affect its suitability as a host for the development of bioengineering but may limit the broader applicability of techniques developed in this strain to other strains of *S. cerevisiae* (or, indeed, to other yeasts). A second aspect that has not, so far, been mentioned is the requirement for proteins to be correctly folded if they are to be biologically active. One of the obstacles to the correct folding of proteins is the occurrence of inclusion bodies [36]. These can occur in most species but are a particular problem in *E. coli* and result in the formation of protein aggregates rather than correctly folded, soluble proteins.

While a range of eukaryotic hosts have been used including insect and mammalian cell lines, the advantages of yeast lie in the ease of growth, the need for a relatively simple growth medium and the robustness of the organism which means that growth in a bioreactor or other fermentation system is more certain to be successful. Yeasts other than *S. cerevisiae* have also been used and the relative advantages and disadvantages of these have been reviewed [37]. As noted previously, gene editing techniques such as CRISPR/Cas9 have been used in bacteria and in mammalian cells. They have also found application in yeast and this has permitted genome editing. This means that genes for post-translational modification of recombinant proteins can be inserted—as, indeed can many other genes of interest [38].

In the context of post-translational modification, a range of alterations to proteins can be carried out including the addition of specific functionality and, perhaps most importantly, glycosylation. Failure to properly glycosylate a protein can result in misfolding of the protein, greatly reduced functionality and/or a significantly reduced half life. The pattern and type of glycosylation can vary in different yeasts therefore this may impact upon the choice of host strain. As mentioned previously, proteins may also be misfolded and this can result in the accumulation of inactive protein in the endoplasmic reticulum. The correct folding of proteins is often a consequence of the presence of 'chaperones' [39]—of which perhaps the most significant are heat shock proteins (HSPs). Organisms have a significant investment in preventing protein mis-folding or aggregation and the mechanisms they use have been studied. In addition to the production of individual recombinant proteins, yeasts are also suitable hosts for the production of natural products through the incorporation of complete biosynthetic pathways. Yeasts have the capacity to produce the essential building blocks for natural product synthesis in vivo (amino acids, pyruvate, acetate etc.). By addition of specific plant genes, desired natural products can be synthesised. For example the expression of relevant plant genes can give rise to the production of alkaloids, phenolic compounds, flavonoids, anthocyanins and even complex polyketides [40]. While these manipulations may appear to be relatively simple, in practice there are a plethora of factors including feedback inhibition, deletion of competing gene pathways and the consequences of over-expression of certain products that must be taken into account. It is not simply a case of finding the requisite gene(s) and cloning them into a convenient host.

6.5 Polyketide Synthases. Systems Biology and Bioengineering

Some of the key aspects of the biosynthesis of polyketides including the modular nature of polyketide synthase (PKS) enzymes has already been referred to. In addition to the manipulation of PKS-mediated pathways PKS enzymes are a useful example of how systems biology and bioengineering can lead to the commercialisation/industrial application of the exploitation of natural products. The idea that PKS enzymes might be used in a modular fashion to create an 'assembly line' of molecules has been mentioned ant this is attractive both from the standpoint of creating novel, bioactive structures and to 'bio-synthesise' polyketides with novel properties in other areas such as fuels and chemicals [41]. Reverse engineering of the biosynthetic pathway of a given polyketide has been suggested as a route to creating considerable numbers of novel structures [42]. PKS is, therefore, not only of interest of itself in that it offers the possibility to create new structures in a purposeful fashion, it is also a paradigm for protein modification and, subsequently, an impact on natural product synthesis in vivo. However, protein structural engineering may provide a powerful approach that goes beyond the modular gene shuffling that has, until now, been

one of the major approachable, modifiable facets of PKS enzymes. This approach has the potential to take bioengineering to a new level. While the potential for the commercial production of valuable polyketide pathway-derived products has been noted [43], the translation of this into a viable process remains challenging and process-specific. It is likely that routes to valuable lipids will be at the forefront of efforts—indeed the nutritional lipids EPA and DHA are produced using fermentation methods and further opportunities are likely [44] including, in particular, the commercial production of statins, pullulan production [45] and others. Application of the principles of engineering to 'commodify' any process such that a number of valuable products can be created rapidly and at scale from a given platform, will remain a significant challenge. Considerable obstacles remain including those which are common to large scale fermentation of any biomolecule—the isolation of the final product using 'green' technology, optimising the growth characteristics of the producing organism, eliminating unwanted side reactions, flow rather than batch production and strain stability. Meeting and overcoming these and other challenges will enable the realisation of the potential of bioengineering for the production of any required polyketide and related useful molecules as well as other valuable natural products using the same basic paradigm. Alongside this commercialisation of known opportunities, the discovery of new molecules with novel functionality and properties will be important [46]. The acquisition of skills and knowledge will mean that the pathway from initial discovery (or proposition in the case of the development of novel 'unnatural' products) to commercial production will be quicker and more profitable.

6.6 Natural Products and the World of Utility

Natural products are an integral part of the living world however, like many things that occupy our shared space, they are complex and challenging to study. This is in part because the normal scientific approach is to simplify complex situations in order that they make sense in terms of known frames of reference. Natural products can be both structurally and biologically complex and may not exist in terms of a simple understandable context. The approach of isolation and verification of the utility of natural products is, as we have seen, an important step in defining what they are and how we might use them. It is the combination of structural and functional properties that underpins our understanding of natural and unnatural products. In order to develop a more substantial understanding of natural products, it is important to simultaneously contextualise them in terms of acquired knowledge and embrace the complexity of living systems and their role in the functioning of natural products.

The application of reductionism, as has been noted previously, allows us to determine the structure of individual natural products, how they are synthesised in vivo by living systems, how they are structurally (and biologically) related and what their natural function might be. Chemists follow structural identification by developing synthetic methods which may follow the route of biological systems (biomimetic

schemes) or may be completely independent of them and may rely on purely chemical logic, intuition and experience. Until the development of the use of enzymes in synthetic chemistry (a relatively recent trend) chemists would make use of reagents and reaction conditions that do not occur naturally. While this approach has facilitated the exploitation of natural products and, in many cases production of the products themselves or a more efficacious analogue demand for some complex natural products outstrips the synthetic or biological resources available to meet requirements. In addition many natural products are too complex to be synthesised in the required bulk properties or are only present from natural sources in small amounts or as part of a complex mixture and therefore not easily isolable in pure form. Considerable efforts have been made, therefore, to apply reductionism at a biological level to determine the essential biological pre-requisites for natural products—genes which code for the synthesis of proteins which can then catalyse the reactions that produce natural products of interest from easily obtainable precursors. Deconvoluting the precise machinery required for the synthesis of specific natural products requires two things—a knowledge of which parts of the cellular machinery are essential for their production (i.e. which genes are required for the complete biosynthetic pathway) and an organism (generally a microbe) that can act as a host to the pathway and produce significant amounts of the natural product—and little else. This is important for two reasons—first of all to enable the desired product to be isolated as the major product and secondly to ensure that the system is as efficient as possible so the energy of the host is directed towards synthesis of the desired product. This was in part touched upon previously (Sect. 6.4) in the discussion of heterologous gene expression.

In some cases, individual reactions are sufficient to produce the required product. Such enzymic-based bioreactors are popular and relatively simple to operate since conditions for only one reaction needs to be optimised [47]. In addition, such bioreactors can operate both under batch and flow conditions [48] and this flexibility is an important aspect of the use of bioreactors. The alternative is the use of complete biosynthetic pathways, as was touched upon (Sect. 6.5) in the discussion about polyketides. In this case, reactor design is considerably more complex and features such as post-translational modification of the product may be important. Significant advances have been made in this area through the use of synthetic biology in which biology is coupled to engineering principles and some of the advantages (and challenges) of this approach have been mentioned.

Natural products are a crucial part of living systems. They are not, however, an essential component of every living system in which they occur. This is perhaps why their production may not be optimised. For example, in some cases, the advantage they confer is resistance to predator attack (as is the case, for example, in the production of the chemicals that make up the ‘mustard oil bomb’ which can be used both by plants themselves [49] and other species that have come to rely on the plant’s mixture of potent, bioactive compounds [50]. For humans, much of the interest in the components of the ‘mustard bomb’ are a consequence of the biological activity in controlling inflammation by impacting on the Nrf2 pathway, cancer preventative properties and even their capacity to remove thiols from breath. This raises the question of what the purpose is of natural products to the producing species and this is often

overshadowed by the use to which they are put. A significant number of medicinally active agents are derived (directly or indirectly) from natural products. Significant efforts have been made to ensure that compounds used as pharmaceuticals are 'pure' and consist of a single active species since this is crucial to the development of a therapeutic window and to understanding the mechanism of action. There are a few exceptions—for example the drug co-amoxiclav which contains both a beta-lactam antibiotic (amoxicillin) and an inhibitor of the microbially-produced beta lactamase (clavulanic acid). This increases the effectiveness of the antibiotic and inhibits one of the principal resistance mechanisms.

Apart from the use of significantly biologically active natural products—which in turn has spurred work on bioengineering to produce large amounts of specific bioactive compounds—the effect of less acutely bioactive principles and their effect on human long-term health has grown. In some cases (for example phytoestrogens) the primary mode of action appears to be known however secondary effects are also evident in many cases. For others (for example, bitter compounds) the primary mode of action may be assumed (a feeding deterrent) however other activities may also become apparent when the compound translocates to other parts of the body. This may be because the function of the principal biological target of the natural product has a location-specific activity (as is the case with some types of bitter receptor) or the biological target may be different. For example bitter compounds can interact with bitter receptors in the GI tract and this influences GI hormone release [51]. Such activity might suggest that bitter taste receptors confer a survival advantage for mammals that consume natural products by inducing greater consumption of energy dense materials which have a higher sugar content and lowering consumption of species that taste bitter and whose consumption is less advantageous when food availability is a limiting factor. In another example from bitter taste receptors, it has been found that one receptor (T2R14) can use a compound produced by an oral bacterium as a ligand and this stimulates an innate immune response thereby limiting the bacterial infection [52]. A further example is the mediation of a response by another of the class of bitter taste receptors (T2R38). The bacterial quorum sensing molecule, N-(3-oxododecanoyl)-L-homoserine lactone (AHL12) is known to bind to this particular bitter receptor and this induces phagocytosis and prevents biofilm formation. In addition, the effect noted previously on the innate immune response might suggest a novel activity for bitter receptor antagonists in terms of minimising inflammation and mitigating inflammatory diseases. Prevention of bacterial biofilm formation and induction of phagocytosis may help to minimise bacterial-infection induced disease and this may be a previously unidentified property of bitter compounds which might confer a survival advantage for consumers. This suggests that taste receptors in general (and bitter receptors in particular) may be polyfunctional. And this might imply that, where natural products are ligands (agonists or antagonists) their ingestion may have a range of biological consequences.

While significant advances have been made in the development of methods to exploit natural products, these have followed a traditional path. Compounds have been isolated then synthesised. Their biogenesis has been determined—often applying experience and logic with proof being supplied by using radioactive and

stable isotopes. Further reductive reasoning has led to the identification of individual enzymes that comprise the biosynthetic pathway. Advances in molecular biology and protein chemistry has enabled individual enzymes and complete pathways to be cloned and expressed in suitable (generally microbial) hosts. These transgenic species have been used, in some cases after modification, to produce desired molecules. Bioengineering and using the tools developed by synthetic biology has culminated in the production of complex molecules by making use of nature rather than the ingenuity of chemists. While these achievements have been significant a full understanding of the totality of importance of natural products to both the producing and the consuming species remains elusive. Multiple targets and consequences seem to be possible. In addition chronic (long term) effects are still laden with complexity. It is always tempting for chemists to try to solve every problem by extending the scope of reductive reasoning however such an approach is likely to produce ever diminishing returns of clarity and new approaches to determining the holistic effects of natural products and their significance will be required.

References

1. Turing, Alan (1950) Computing machinery and intelligence. Mind, LIX(236):433–460. https://doi.org/10.1093/mind/LIX.236.433, ISSN 0026-4423
2. Hameroff SR, Watt RC (1982) Information processing in microtubules. J Theor Biol 98:549–561. https://doi.org/10.1016/0022-5193(82)90137-0
3. Penrose R (1991) The Emperor's new mind. Neuroscience 37:285–286
4. Khan S, Huang Y, Timuçin D, Bailey S, Lee S, Lopes J, Gaunce E, Mosberger J, Zhan M, Abdelrahman B, Zeng X, Wiest MC (2024) Microtubule-stabilizer Epothilone B delays anesthetic-induced unconsciousness in rats. eNeuro 11(8) ENEURO.0291-24.2024; https://doi.org/10.1523/ENEURO.0291-24.2024
5. Seixas R, Lopes D, Couto M, Pereira J, Paiva J (2021) Severe colchicine intoxication: a case report and review of literature in adults. Cureus 13(10):e19151. https://doi.org/10.7759/cureus.19151. PMID: 34868785; PMCID: PMC8629332
6. McFadden J, Al-Khalili J (2018) The origins of quantum biology. Proc R Soc A.47420180674 https://doi.org/10.1098/rspa.2018.0674
7. Lambert N, Chen YN, Cheng YC et al (2013) Quantum biology. Nature Phys 9:10–18. https://doi.org/10.1038/nphys2474
8. Virenque L, Mossio M (2024) What is agency? A view from autonomy theory. Biol Theory 19:11–15. https://doi.org/10.1007/s13752-023-00441-5
9. Aveling E (1886) Die Darwinsche theorie. Verlag von J. H. W, Dietz, Stuttgart
10. Morgan TH, Sturtevant AH, Muller HJ, Bridges CB (1915) The mechanism of Mendelian heredity. Henry Holt, New York
11. Avery OT et al (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types. J Exp Med 79:137–157
12. Watson JD, Crick FH (1953) The structure of DNA. Cold Spring Harb Symp Quant Biol 18:123–131. https://doi.org/10.1101/sqb.1953.018.01.020
13. Gotovtsev P (2020) How IoT can integrate biotechnological approaches for city applications—review of recent advancements, issues, and perspectives. Appl Sci 10:3990
14. Zhang C, Ke D, Duan Y et al (2018) The combinatorial biosynthesis of "unnatural" products with polyketides. Trans Tianjin Univ 24:501–512. https://doi.org/10.1007/s12209-018-0151-9

15. Wong FT, Khosla C (2012) Combinatorial biosynthesis of polyketides--a perspective. Curr Opin Chem Biol 16(1–2):117–23. https://doi.org/10.1016/j.cbpa.2012.01.018. Epub 2012 Feb 16. PMID: 22342766; PMCID: PMC3328608
16. Liu Z, Hammerlindl J, Keller W et al (2011) MAM gene silencing leads to the induction of C3 and reduction of C4 and C5 side-chain aliphatic glucosinolates in Brassica napus. Mol Breeding 27:467–478. https://doi.org/10.1007/s11032-010-9444-y
17. Rutledge P, Challis G (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. Nat Rev Microbiol 13:509–523. https://doi.org/10.1038/nrmicro3496
18. Xu Y, Zhou T, Zhang S et al (2014) Diversity-oriented combinatorial biosynthesis of benzenediol lactone scaffolds by subunit shuffling of fungal polyketide synthases. Proc Natl Acad Sci U. S.A 111(34):12354–12359
19. Harms V, Kirschning A, Dickschat JS (2020) Nature-driven approaches to non-natural terpene analogues. Nat Prod Rep 37(8):1080–1097. https://doi.org/10.1039/c9np00055k. Epub 2020 Feb 18. PMID: 32068211
20. Hahn F, Guth FM (2020) The ambruticins and jerangolids—chemistry, biology and chemoenzymatic synthesis of potent antifungal drug candidates. Nat Prod Rep 37(10):1300–1315. https://doi.org/10.1039/d0np00012d. Epub 2020 May 18. PMID: 32420573
21. Galanie S, Entwistle D, Lalonde J (2020) Engineering biosynthetic enzymes for industrial natural product synthesis. Nat Prod Rep 37(8):1122-1143. https://doi.org/10.1039/c9np00071b. PMID: 32364202
22. Robins RJ, Hanley AB, Richards SR, Fenwick GR, Rhodes MJ (1987) Uncharacteristic alkaloid production by suspension cultures of Cinchona pubescens fed with L-trytophan. Plant Cell, Tissue Organ Cult 9:49
23. Trenti F, Yamamoto K, Hong B, Paetz C, Nakamura Y, O'Connor SE (2021) Early and late steps of quinine biosynthesis. Org Lett 23(5):1793–1797. https://doi.org/10.1021/acs.orglett.1c00206. Epub 2021 Feb 24. PMID: 33625237; PMCID: PMC7944568
24. Zhang H, Fang L, Osburne MS, Pfeifer BA (2016) The continuing development of E. coli as a heterologous host for complex natural product biosynthesis. Methods Mol Biol 1401:121–134. https://doi.org/10.1007/978-1-4939-3375-4_8. PMID: 26831705; PMCID: PMC9904415
25. Gao J, Gou Y, Huang L, Lian J (2024) Reconstitution and optimization of complex plant natural product biosynthetic pathways in microbial expression systems. Curr Opin Biotechnol 87:103136. https://doi.org/10.1016/j.copbio.2024.103136. Epub 2024 May 4. PMID: 38705090
26. Loyola-Vargas, V.M., Ochoa-Alejo, N. (2018). An introduction to plant tissue culture: advances and perspectives. In: Loyola-Vargas V, Ochoa-Alejo N (eds) Plant cell culture protocols. Methods in molecular biology, vol 1815. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-8594-4_1
27. Chandran H, Meena M, Barupal T, Sharma K (2020) Plant tissue culture as a perpetual source for production of industrially important bioactive compounds. Biotechnol Rep (Amst) 26:e00450. https://doi.org/10.1016/j.btre.2020.e00450. PMID: 32373483; PMCID: PMC7193120
28. Gunun N, Kaewpila C, Khota W, Polyorach S, Kimprasit T, Phlaetita W, Cherdthong A, Wanapat M, Gunun P (2022) The effect of indigo (Indigofera tinctoria L.) waste on growth performance, digestibility, rumen fermentation, hematology and immune response in growing beef cattle. Animals (Basel) 13(1):84. https://doi.org/10.3390/ani13010084. PMID: 36611693; PMCID: PMC9818010
29. Gao J, Sun Y, Bao Y, Zhou K, Kong D, Zhao G (2021) Effects of different levels of rapeseed cake containing high glucosinolates in steer ration on rumen fermentation, nutrient digestibility and the rumen microbial community. Br J Nutr 125(3):266–274. https://doi.org/10.1017/S0007114520002767
30. Benucci I, Lombardelli C, Mazzocchi C, Esti M (2022) Natural colorants from vegetable food waste: recovery, regulatory aspects, and stability—a review. Compr Rev Food Sci Food Saf 21(3):2715-2737. https://doi.org/10.1111/1541-4337.12951. Epub 2022 Apr 2. PMID: 35368140

31. Macek B, Forchhammer K, Hardouin J, Weber-Ban E, Grangeasse C, Mijakovic I (2019) Protein post-translational modifications in bacteria. Nat Rev Microbiol 17(11):651–664. https://doi.org/10.1038/s41579-019-0243-0. Epub 2019 Sep 4. PMID: 31485032
32. Baghban R, Farajnia S, Rajabibazl M et al (2019) Yeast expression systems: overview and recent advances. Mol Biotechnol 61:365–384. https://doi.org/10.1007/s12033-019-00164-8
33. Graslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, Knapp S, Oppermann U, Arrowsmith C, Hui R et al (2008) Protein production and purification. Nat Methods 5:135–146
34. Duan H, Zhang X, Figeys D (2023) An emerging field: post-translational modification in microbiome. Proteomics 23(3–4):e2100389. https://doi.org/10.1002/pmic.202100389. Epub 2022 Oct 26. PMID: 36239139
35. Gianni L (2015) The natural history of model organisms: the fascinating and secret wild life of the budding yeast S. cerevisiae eLife 4:e05835
36. Bhatwa A, Wang W, Hassan YI, Abraham N, Li XZ, Zhou T (2021) Challenges associated with the formation of recombinant protein inclusion bodies in Escherichia coli and strategies to address them for industrial applications. Front Bioeng Biotechnol 9:630551. https://doi.org/10.3389/fbioe.2021.630551. PMID: 33644021; PMCID: PMC7902521
37. Vieira Gomes AM, Souza Carmo T, Silva Carvalho L, Mendonça Bahia F, Parachin NS (2018) Comparison of yeasts as hosts for recombinant protein production. Microorganisms 6(2):38. https://doi.org/10.3390/microorganisms6020038. PMID: 29710826; PMCID: PMC6027275
38. Akhmetov A, Laurent JM, Gollihar J, Gardner EC, Garge RK, Ellington AD, Kachroo AH, Marcotte EM (2018) Single-step precision genome editing in yeast using CRISPR-Cas9. Bio Protoc 8(6):e2765. https://doi.org/10.21769/BioProtoc.2765. PMID: 29770349; PMCID: PMC5951413
39. Hartl F, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. Nature 475:324–332. https://doi.org/10.1038/nature10317
40. Zhao M, Zhao Y, Yao M et al (2020) Pathway engineering in yeast for synthesizing the complex polyketide bikaverin. Nat Commun 11:6197. https://doi.org/10.1038/s41467-020-19984-3
41. Yuzawa S, Keasling J, Katz L (2016) Insights into polyketide biosynthesis gained from repurposing antibiotic-producing polyketide synthases to produce fuels and chemicals. J Antibiot 69:494–499. https://doi.org/10.1038/ja.2016.64
42. Nava AA, Roberts J, Haushalter RW, Wang Z, Keasling JD (2023) Module-based polyketide synthase engineering for de Novo polyketide biosynthesis. ACS Synth Biol 12(11):3148–3155. https://doi.org/10.1021/acssynbio.3c00282. Epub 2023 Oct 23. PMID: 37871264; PMCID: PMC10661043
43. Ma W, Liu M, Zhang Z et al (22022) Efficient co-production of EPA and DHA by Schizochytrium sp. via regulation of the polyketide synthase pathway. Commun Biol 5:1356. https://doi.org/10.1038/s42003-022-04334-4
44. Neves RPP, Ferreira P, Medina FE et al (2022) Engineering of PKS megaenzymes—a promising way to biosynthesize high-value active molecules. Top Catal 65:544–562. https://doi.org/10.1007/s11244-021-01490-5
45. Chen S, Zheng H, Gao J, Song H, Bai W (2023) High-level production of pullulan and its biosynthesis regulation in Aureobasidium pullulans BL06. Front Bioeng Biotechnol 11:1131875. https://doi.org/10.3389/fbioe.2023.1131875. PMID: 36777253; PMCID: PMC9909216
46. Katz L, Baltz RH (2016) Natural product discovery: past, present, and future. J Ind Microbiol Biotechnol 43(2–3):155–176. https://doi.org/10.1007/s10295-015-1723-5
47. Jankowska K, Bachosz K, Zdarta J, Jesionowski T (2020) Application of enzymatic-based bioreactors. In: Ochowiak M, Woziwodzki S, Mitkowski P, Doligalski M (eds) Practical aspects of chemical engineering. PAIC 2019. Springer, Cham. https://doi.org/10.1007/978-3-030-39867-5_12
48. Zhang Z, Gao L, Boes A et al (2024) An enzymatic continuous-flow reactor based on a pore-size matching nano- and isoporous block copolymer membrane. Nat Commun 15:3308. https://doi.org/10.1038/s41467-024-47007-y
49. Kissen R, Rossiter JT, Bones AM (2009) The 'mustard oil bomb': not so easy to assemble? Localization, expression and distribution of the components of the myrosinase enzyme system. Phytochem Rev 8:69–86. https://doi.org/10.1007/s11101-008-9109-1

50. Sporer T, Körnig J, Wielsch N, Gebauer-Jung S, Reichelt M, Hupfer Y, Beran F (2021) Hijacking the mustard-oil bomb: how a glucosinolate-sequestering flea beetle copes with plant myrosinases. Front Plant Sci 12:645030. https://doi.org/10.3389/fpls.2021.645030. PMID: 34093609; PMCID: PMC8173161
51. Rezaie P, Bitarafan V, Horowitz M, Feinle-Bisset C (2021) Effects of bitter substances on GI function, energy intake and glycaemia-Do preclinical findings translate to outcomes in humans? Nutrients 13(4):1317. https://doi.org/10.3390/nu13041317. PMID: 33923589; PMCID: PMC8072924
52. Medapati MR, Singh N, Bhagirath AY, Duan K, Triggs-Raine B, Batista EL Jr, Chelikani P (2021) Bitter taste receptor T2R14 detects quorum sensing molecules from cariogenic Streptococcus mutans and mediates innate immune responses in gingival epithelial cells. FASEB J 35(3):e21375. https://doi.org/10.1096/fj.202000208R. PMID: 33559200

Chapter 7
Enzymes and Natural Products

Abstract The classification of natural products can be by chemical structure (e.g. alkaloids), by biological properties (e.g. anti-inflammatory properties) or by the types of enzymic reactions they have undergone (e.g. glycosides). Biosynthesis of natural products by enzymes will vary between primary metabolites (very specific since structural integrity is essential) and secondary metabolites (less specific and a range of substrates can be acted upon to give several products with the same core). Enzymes can operate in 'novel' environments—non-optimal reaction conditions. Extremophiles can produce more stable enzymes to give novel applications (e.g. polymerase chain reaction to amplify DNA). Properties of enzymes (e.g. thermal, pH or ionic stability) can be altered or the reaction being catalysed may change if the gene for the enzyme is mutated in a controlled fashion. Can lose enzyme activity due to unforeseen consequences of amino acid changes (e.g. misfolding, altered substrate specificity). Enzymes operate through a transition state so if an antibody is raised to a stable transition state analogue then this antibody may operate as an enzyme (abzyme). A similar effect may be found with molecular imprinting of a transition state analogue using a globular, cross-linked protein.

Keywords Enzymes · Natural products · Enzyme engineering · Environmental effects · PCR · Abzymes · Molecular imprinting

As noted previously, throughout the various chapters in this volume, the reductionist approaches of chemists as applied to natural products and the living systems that they interact with (both the producers and the consumers), has, almost as a tangential activity, led to extensive discoveries in the field of enzymology both in terms of biosynthetic pathways and the interaction of living systems and the enzymes that they contain with natural products. As has been noted earlier, living systems are rarely in a state of homeostasis. This is in part because they must react and evolve both to external circumstances and by internal motivations—enzymes lie at the core of this feature. It is insufficient, however, to simply understand each individual enzyme reaction in terms of its substrate requirement, optimal rate constant or even mechanism of action. The framework for understanding enzymes lies with the enzymes themselves

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_7

specifically when they are in the environment in which they operate. This is the crucial component of understanding the role of enzymes in living systems.

7.1 Enzymes and Natural Products

The classification of natural products has traditionally been driven by chemistry and, at first, the measurable properties of individual compounds. This means that nomenclature and classification was dominated by those groups of compounds that were most abundant and/or easiest to isolate. Later, classes of compounds were grouped by implied structural similarities based on physico-chemical properties. Perhaps one of the best examples of this is illustrated by the delineation of alkaloids. These were initially defined as alkali-like (basic) natural products that contained at least one nitrogen atom. Further classification was carried out on the basis of structural similarity, based primarily on decomposition studies, solubility properties and the reactions of specific functional groups with reagents which in turn reflected their biogenesis. For example, the Lassaigne test melts a substance with sodium and, depending on the constituents, sodium halides, sodium cyanide, sodium sulphide or sodium phosphate are formed. This provides a limited elemental analysis for the most frequently found constituents of natural products. Other reagents can be used to identify specific molecular characteristics. For example, Ehrlich's reagent (p-dimethyl amino cinnamaldehyde) is used to identify compounds that contain indole and Tollen's reagent distinguishes between aldehydes and ketones. By piecing together the core chemical units and the attached substituents, a range of (sometimes interlinked) pathways can be identified and putative relationships between them hypothesised. In short, by breaking molecules into identifiable parts, wider associations and biosynthetic relationships can be suggested.

An alternative approach is to group natural products based upon biological function. Since structure:activity relationships are a known paradigm, classification based on biological activity will reflect, in many cases, structural similarity due to molecular interactions at a biological target site. For example, if two molecules bind to a particular receptor then they will most likely share some molecular characteristics. This is not, however, invariably the case and similarity of biological effect may be misleading. In part this is because of different modes of action. For example, analgesics (pain killers) may operate in a number of different ways and the structural variation is, therefore, notable. Even within classes of bioactives (for example, non-steroidal anti-inflammatory drugs—NSAIDs) there can be considerable structural variation (Fig. 7.1), and this can be reflected in the mechanism of action [1] and/or their binding differently to different receptors.

From a non-pharmacological perspective, the activity of very structurally diverse bitter compounds (Fig. 7.2) is consistent with each binding to a different spectrum of bitter receptors.

Classification by biosynthetic origin or structural similarity rather than by biological function is a less ambiguous approach. Having derived such a classification,

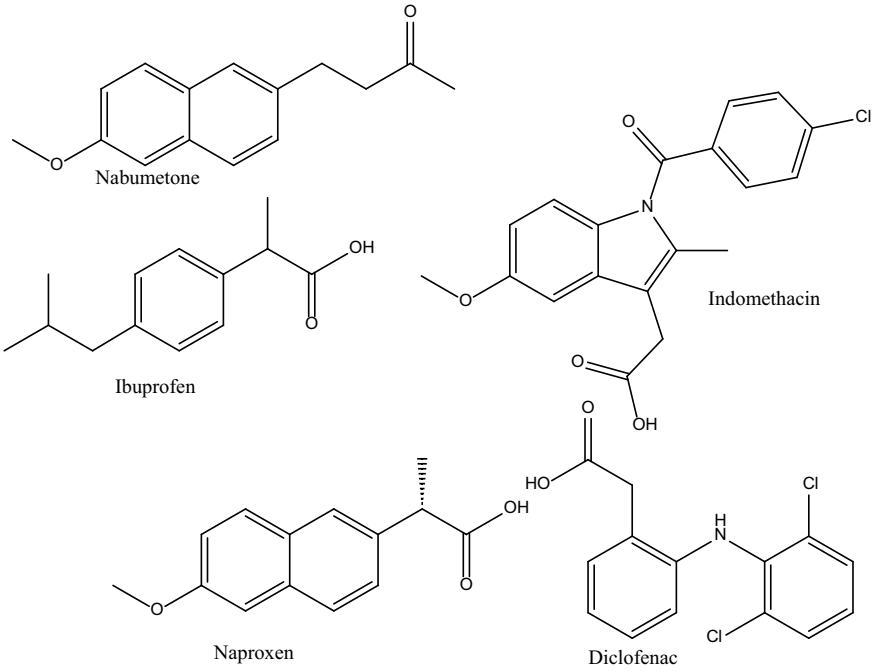


Fig. 7.1 Structures of NSAIDs

the reductionist chemistry approach then suggests that greater understanding can be gained from a better understanding of the individual steps in a pathway. In essence, this is analogous to the structural approach of decomposition described above but applied to a biological system. By breaking a biosynthetic sequence into individual steps, it may be possible to retrace the gradual building of a molecule. This will then allow for the modification of some of the specific steps in a planned way in order to create different natural (or even unnatural) products. In most cases, the individual reactions are catalysed by enzymes. While it may seem logical to then group natural products based upon the types of reaction that their precursors undergo thereby providing a different type of functional ordering from simple biological functionality, this degree of granularity is only of limited use. It can result in the grouping together, for example, of all glycosides whose only common feature is the presence of a glycoside moiety (Fig. 7.3).

The development of more advanced technologies including DNA and protein sequencing, X-ray crystallography of proteins themselves and of complexes formed from the binding of natural products to proteins and precise spectroscopic methods has increased the transparency with which natural products can be classified and defined thereby increasing the development of the links between individual and groups of natural products. In addition, these structural comparisons have been underpinned by mechanistic studies developed from the generation of plausible hypotheses

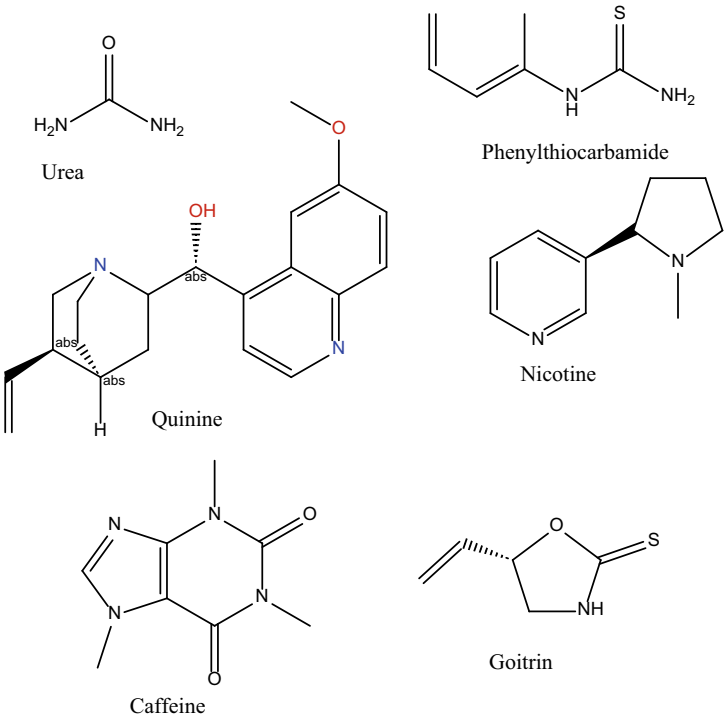


Fig. 7.2 Representative bitter compounds

about the generation of natural products and the confirmation of those hypotheses using stable or radioisotopes to track the pathway of natural product formation and therefore elucidation of the likely sequence of biosynthetic events. The use of inhibitors and genetic manipulation to confirm details of the active site of an enzyme, (most recently using CRISPR-Cas 9) have also contributed to a greater understanding of how natural products are formed and how the process can be optimised or manipulated. Classification of natural products simply on the basis of plausible structural links and the biosynthetic reaction sequence that is required to form them does not always provide a complete picture. Metagenomics has resulted in the discovery of some enzymes and gene clusters that may suggest the existence of undiscovered natural products. These help to fill in gaps where natural products can be hypothesised but have not been discovered. This has also resulted in the proposal of a link between amino acid sequence and the structure of some natural products [2]. Certain structural motifs in a protein are associated with particular types of reactions (for example, the catalytic triad of some proteases). It is, therefore, reasonable to extend this concept (by analogy) to suggest that certain elements of the sequence of a protein cluster may be associated with the structural motifs of particular types of natural products. This would enable the potential for specific products to be formed to be assessed without having to switch the genes on and create an active enzyme. Refinement of

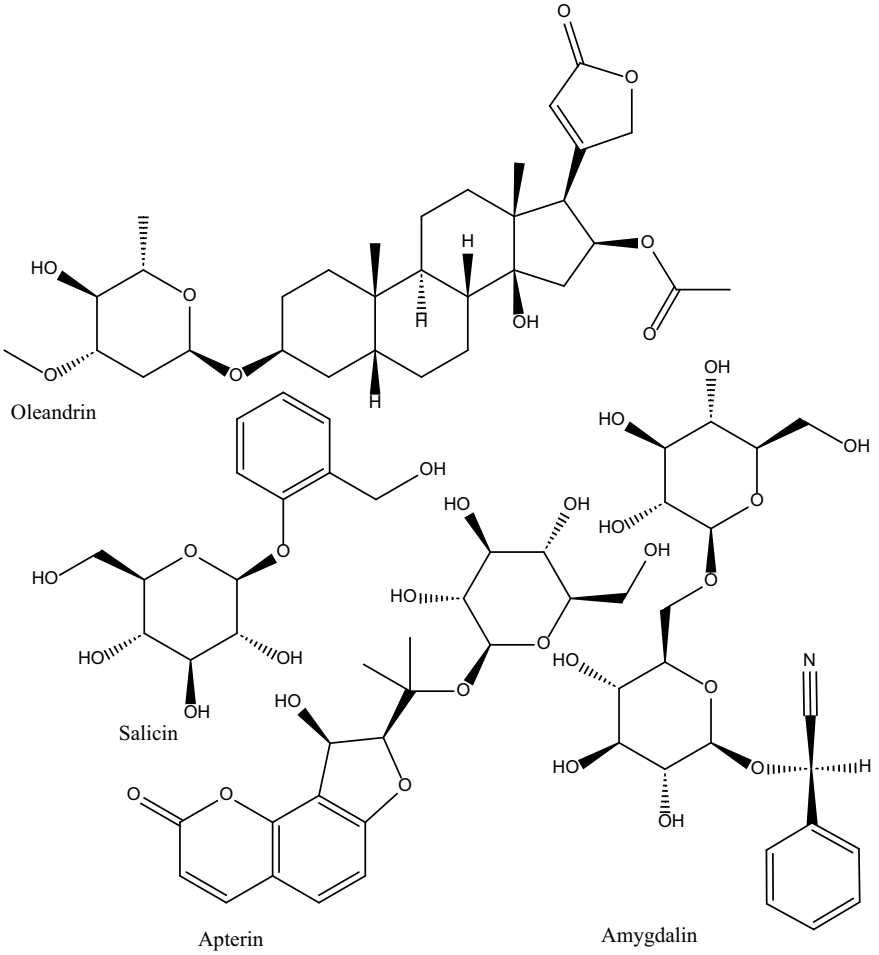


Fig. 7.3 Examples of glycosides

this model is still required however a biological rather than a purely chemistry-driven approach still requires the incorporation of structural as well as biological reaction information. For example, an extremely structurally diverse group of natural products can be formed from reactions catalysed by cytochrome-P450 enzymes even within a single species [3]. Such diversity of structures argues against the establishment of a specific direct gene (or protein) sequence—natural product link.

Enzymes involved in natural product formation have a core function—the synthesis of a specific product. In addition, however many of them are also multi-functional in that they can act on a number of substrates. In this way a diverse range of natural products can be formed as a consequence of the operation of a single biosynthetic pathway. This introduces a dilemma—the enzymes involved in

the biosynthesis of natural products must combine reaction specificity (they must catalyse a single reaction) with a degree of substrate flexibility. This is in contrast with enzymes involved in primary metabolism whose role and function is fixed. If an enzyme involved in primary metabolism undergoes any alteration or loses substrate specificity then any such random change is likely to be deleterious to metabolic flux and balance. There is no physiological pressure for change and no advantage to be gained. In addition the pathways of primary metabolism are closely linked to one another and must be expressed and function at a high level. They must continue to carry out the same reaction without any deviation or alteration or reduction in efficiency. The integrity of the organism is dependent upon this consistency. Secondary natural products, on the other hand, are generally produced in smaller amounts and are more diverse than primary metabolites. As a result, the consequences of their lowered catalytic efficiency or reduced substrate specificity are less traumatic for the living system. Structural diversity can also confer an advantage. Under these conditions the consequences of changes in specificity are an increase in numbers of secondary natural products that can be formed and this be consequential for enhanced biological function. The result of the changes in the biosynthesis of secondary natural products compared to structurally confined primary natural products (and the extension of the spectrum of products formed) may create an advantage for the organism as changes in biological activity of the natural products being produced may, for example, deter predation by a wider range of species and improve survival and fitness for the environment. In short, the loosening of the specificity of enzymes involved in natural product biosynthesis will result in the increase in the diversity of secondary natural products being produced and this may provide the producing organism with an adaptational advantage [4]. In terms of artificial manipulation of secondary natural products, the development of novel natural products (unnatural products) outside of the gradual evolution and subsequent enhanced fitness process will necessitate the combination of chemical knowledge and reaction engineering together with a deeper understanding and appreciation of how enzymes function and can be manipulated. One aspect of this is the modification of the reaction conditions that are not simply suited to maximisation of the rate and specificity of an isolated, single enzyme-catalysed reaction but also the production of particular pathway products in a timely fashion that ensures the whole biosynthetic pathway is able to operate using, if appropriate, novel substrates or the accumulation of unnatural intermediates. It is also important to recognise that manipulation of biosynthetic pathways by scientists and engineers is often objectively different to changes that are advantageous to the producing organism. The purpose of manipulation by humans is optimisation of the ‘useful’ properties of the secondary unnatural product. These will often not confer a ‘fitness’ advantage for the producing organism but will be directed towards some other property which is of physiological importance to the manipulating species rather than to the producing organism.

7.2 Enzymes in Novel Environments

Many studies on enzymes started with an essentially reductionist approach—isolation of the enzyme purification and then the investigation of the reaction that is catalysed. Enzymes are complex proteins and (with the exception of those isolated from extremophiles) generally operate under non-denaturing conditions. Denaturing can take a number of forms including loss of solubility, aggregation and, ultimately, irreversible loss of structural integrity. The denaturation process generally follows a series of reversible steps before the process becomes irreversible [5]. Extremes of temperature, pH or ionic strength (for example the presence of detergents) can render many enzymes inactive. Some processes including immobilisation and operation in a low water activity environment can enhance the stability of some enzymes [6]. The use of bioengineered enzymes to carry out reactions has a number of advantages in addition to improvement in enzyme stability. Enzyme bioreactors are commonly used in industrial processes [7] and considerable efforts have been made to optimise the reactions themselves, the reactor design and the isolation of the product.

In order to carry out comparative studies on different enzymes, reactions catalysed by enzymes are frequently measured in buffered solutions at a pH close to 7 and at a 'physiological' temperature of 37° C. While this may not represent accurately the conditions under which most enzymes operate *in vivo*, it provides baseline measurements for the purposes of comparing different enzymes and can be important in finding out if specific inhibitors can affect the reaction catalysed by a given enzyme. The rate of a reaction catalysed by an enzyme is defined in terms of the Michaelis–Menten equation. This expression states simply that an enzyme binds reversibly to a substrate to form a dual entity. The substrate then undergoes an irreversible transformation to form a product and the enzyme is released and can join with another substrate. This way of considering an enzyme catalysed reaction leads to the development of certain definitions which are commonly used throughout studies of enzymes and the reactions they catalyse. Vmax is defined as the maximum rate of the reaction –and this occurs when all the enzyme active sites are occupied by substrate molecules. The other term that is defined by the Michaelis–Menten kinetic model is Km which is the substrate concentration at which the reaction rate is 50% of Vmax. By plotting the rate of the reaction against the substrate concentration, it is possible to derive Km. Another way of measuring an enzyme-catalysed reaction is a Lineweaver–Burk Plot in which the inverse of the reaction rate is plotted against the inverse of the substrate concentration (it is, therefore, a double reciprocal plot) [8]. This gives a straight line and the form of this line provides information about the enzyme reaction—the type of enzyme inhibition (particularly competitive or non-competitive or a mixture) that is occurring and the value of Km and Vmax. In physiological terms, measuring the activity of some enzymes can be used in diagnosis. High activities in the bloodstream can be indicative of organ damage and leakage of cell content while an abnormally low level of activity can be a sign of a poorly functioning enzyme due to the presence of inhibitors or a genetic polymorphism in the enzyme itself. The development of

diagnostic tools that make use of enzyme activity makes use of computer models, bioengineering, biotechnology, chemical biology and molecular biology [9].

While the operator has considerable control over a bioreactor, in the case of plants, there has been some interest in the relevance of variables such as temperature on complex, multifactorial biological processes including photosynthesis [10]. While this makes a useful contribution to understanding the mechanism of biological processes *in situ,* losses in the yield of some major crops have been ascribed to elevated growing temperatures and the consequences of non-optimal conditions has agronomic significance. The investigation of complex biological systems suggests that reductionism has limitations and that trying to reduce biological systems to an empirical set of additive reactions is not valid. This is illustrated in the case of the effect of temperature on reactions. In principle reaction rates increase exponentially with increased temperature. This is, in simple terms, a consequence of two factors—colliding molecules having more energy therefore more collisions occurring and the colliding molecules being more likely to have sufficient energy to overcome the activation energy barrier for a reaction to occur. This effect is described by the Arrhenius equation. However, in biological systems, this increase occurs up to a maximum value beyond which the reaction rate can drop markedly because of the effect of increased vibration on the integrity of the enzyme active site and (eventually) denaturation of the enzyme itself. In addition to the impact on primary metabolism of temperature (or other stress factors), secondary metabolite profiles (as a result of biosynthesis modulation) are also influenced by external factors beyond a simple acceleration of the reaction rates of enzymes in a biosynthetic pathway as a result of the modulation of kinetic factors mentioned above. Plants (and other biological systems) may respond to a range of stress factors by producing increased levels of secondary metabolites [11]. This can occur as a result of a switch in production from primary to (defensive) secondary metabolites. Abiotic stress response by plants has been suggested to be a strategy for the production of increased levels of secondary natural products [12]. This effect is not limited to plants. It has been demonstrated that microbial production of compounds can switch from primary to secondary metabolite pathways in response to environmental stress [13]. As with plants, the extent of knowledge of the metabolic events that trigger production of elevated levels of secondary metabolites remains unclear. Optimisation of production of natural products through (as we have seen) strain selection or even promoter engineering remains a long term goal of natural product chemists [14]. A range of tools including Flux Balance Analysis (FBA) have been used to probe directly the effect of different factors on secondary metabolite production. FBA is a mathematical model that can be used to determine the metabolite flow through metabolic pathways [15]. The focus of many of these efforts have been towards optimisation of production of specific compounds such as antibiotics. Key steps include the identification of natural products and the genes which code for the enzymes of the corresponding biosynthetic pathway, construction of the pathway and then modelling of the flux through the pathway under different sets of conditions. The use of FBA enables optimal conditions for the production of specific natural product profiles to be achieved.

In addition to the development of biosynthetic pathway manipulation, (and indeed before tools such as FBA were available) plants were routinely chosen on the basis of strain selection for optimisation of natural product profile. One of the main drawbacks to the use of whole plants is the time taken for them to come to maturity. Cell culture techniques are able to overcome these problems but are not efficacious in all cases [16]. While proof of principle has been demonstrated, the use of plants as hosts for bioreactor-based production of natural products is not widespread. The advantages of plant cell cultures includes fewer regulatory hurdles than the use of microbes and the capability of the plant cellular machinery to carry out post-translational modification—a key requirement for many natural product biosynthetic enzymes and something that can be a significant failing in microbial systems used for heterologous gene expression. As gene editing and expression systems improve then it seems likely that plant culture technologies will find greater applicability as a source of valuable natural products [17]. Nonetheless there remain significant challenges to the rapid commercialisation of plant bioreactors including yield and efficiency of the process.

One approach which has been suggested is the use of enzymes in 'unusual' environments since, as mentioned previously, this may affect the stability of the enzyme. In addition, enzymes may demonstrate different substrate specificities and even catalyse different reactions if operating under a range of different conditions. In this context an unusual environment is one in which the standard conditions of an aqueous systems operating at a neutral pH and a temperature of about 37° C are deviated from. Such an approach may be of value when considering how enzymes might function *in vivo* (for example at interfaces rather than in wholly aqueous systems). Another example is enzymes which catalyse reactions in extremophiles and which are, therefore, more accustomed to diverse reaction conditions [18]. One of the most frequently cited examples of an enzyme from an extremophile is that of thermostable DNA polymerases. These enzymes (Taq polymerases) are derived from *Thermus aquaticus* [19]. They are crucial components of the development and use of the polymerase chain reaction (PCR) to selectively amplify fragments of DNA. The PCR is a way of selectively amplifying the number of copies of a selected piece of DNA exponentially. In addition to making use of an extremophile enzyme, PCR is also able to be easily automated and this has helped it to become one of the most useful and powerful of biologically-based analytical tools. The essence of the technique relies on the thermostability of Taq polymerase. The enzyme is stable at the relatively high temperature that denatures double stranded DNA and is optimally functional at 70-80C. The process of PCR comprises the following steps.

1. Identification of the fragment of DNA of interest.
2. Creation of single stranded DNA by heating the DNA containing the fragment of interest.
3. Cool slightly. Bind short primer pieces of DNA to the target DNA. These acts as a starting point for the polymerase enzyme which requires double stranded DNA and this is a requirement for the DNA copying reaction to proceed.

4. Add the polymerase which copies the fragment of interest so from a single, undifferentiated piece of double stranded DNA, the reaction produces two double stranded specific pieces of DNA.
5. Repeat the steps 2–4 a number of times (25–45) to get millions of copies of the specific piece of DNA of interest.

The heating and cooling cycles are carried out automatically and the reaction requires no operator intervention.

One of the reasons that enzymes are susceptible to extreme environments (high temperature, salinity, the presence of denaturing agents etc.) is because water is structured around proteins generally and enzymes, in particular. Extreme conditions often limits the availability of normally structured water and means that many extremophile enzymes have structural features that improve their hydration capability and resilience to the absence of structured water. One result of these features is that the enzymes maintain a stable hydration shell around the protein to prevent deleterious events such as protein aggregation or denaturation. A number of the characteristics of such enzymes (for example the ordering of particular amino acid residues on the surface) have been investigated [20].

Non-extremophile enzymes have also been examined under unusual conditions—particularly in organic solvents and under low water conditions. In many cases, the major effect of suspension of a normo-philic enzyme in an organic solvent is a reduction in the reaction rate. This was observed in the cleavage of lambda DNA by a restriction enzyme in a reverse micellar system [21] when, instead of the expected fragmentation pattern, limited hydrolysis was observed and certain sites on the DNA fragment were found to be more susceptible than others. Since the reaction being carried out was a hydrolysis in which water plays a crucial role, the loss in reactivity is perhaps not surprising.

The use of enzymes in organic solvents in multiphase systems has been attempted with some success. If polar, water miscible solvents are used, these may remove the layer of ordered water from around the enzyme surface resulting in denaturation and/or loss of enzyme activity. The use of enzymes in water immiscible organic solvents can have a number of advantages including the maintenance of a ‘shell’ of tightly bound, ordered water molecules on the surface of the enzyme. Enzymes in organic solvents can be more thermostable and rigid than those in aqueous media and inefficient side reactions (e.g. hydrolytic reactions) are less common. Strategies using low water systems to limit post-reaction hydrolysis are not invariably successful, for example attempts to isolate or trap a putative isothiocyanate intermediate by carrying out the enzymic glycosidic cleavage of 1-methoxy indolyl glucosinolate in a low water system met with only limited success (Fig. 7.4) and 1-methoxy indolyl-3-methyl isothiocyanate was only detected transiently by mass spectrometry [22].

One other advantage of enzyme reactions in organic solvents is that it increases the range of substrates that may be reacted in the system and it can reduce enzyme inhibition by the formation of the product in the active site.

One additional capability of enzymes in restricted water environments is the capacity to carry out reverse hydrolytic reactions. A reverse hydrolytic reaction is one

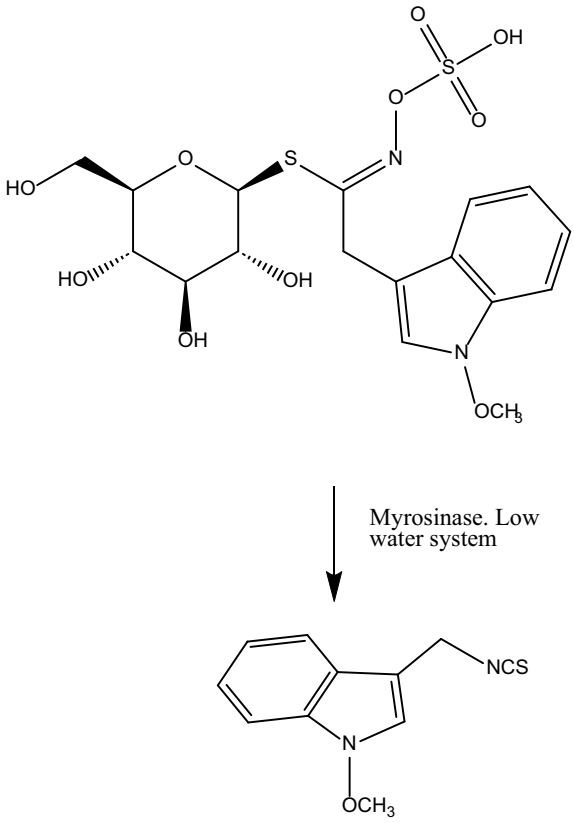


Fig. 7.4 Hydrolysis of 1-methoxy indole glucosinolate in a low water system

in which a bond is formed with the concomitant elimination of water [23]. In nature, some reactions that appear to be hydrolytic in nature are actually reversible. Hydrolases include proteases, peptidases, glycohyrolases and phosphatases. For example, glycoside hydrolases can perform both hydrolytic reactions and reverse hydrolysis/trans glycosylation. There are a number of industrial applications to the use of enzymes in novel environments. These are particularly important with polyfunctional molecules which require the level of specificity exhibited by enzymes and a rapid reaction rate [24]. A second advantage of using biological catalysts is that they are often enantiospecific. Such techniques are often driven by the need (and value associated with) enantiopure compounds. Various strategies have been adopted and those which use enzymic approaches have been reviewed [25]. The concept of enantioconvergent synthesis is important and is widely investigated. While the enhanced stability of some enzymes in low water systems has been mentioned, perhaps the simplest was of reducing water activity (which also has the effect of reducing protein molecular motion) is by freezing reactions containing enzymes. Freezing a reaction mix can prevent competing reactions and has the same effect as carrying out a reaction in a low water environment—the amount of free water is limited. This has been used

in the formation of peptide bonds [26]. The use of enzymes is therefore not limited to reactions typically carried out under Michaelis–Menten kinetic conditions. While enzymes lower the activation energy of a reaction they do not, a priori, alter the position of the equilibrium. In many biological processes (for example gluconeogenesis) the forward and reverse reactions are in equilibrium and the direction is dependent on the concentrations of 'substrates' and 'products'. In the case of natural products, the desired molecule is often only produced in small amounts and is non-essential. While this can encourage molecular and evolutionary flexibility, it is less useful if large amounts of a single product are required. The use of enzyme cascades in the synthesis of natural products is a powerful tool which can optimise the production of complex natural products [27] and is likely to find further applications.

7.3 Enzyme Engineering and Directed Evolution—Successes and Failures

While natural product chemists have worked extensively with complex biological systems, the need to create novel (unnatural) products and to optimise the production of desired, valuable molecules has led them towards, as we have seen, manipulation of enzymes by altering the environment they operate in. In addition, natural product chemists have also used the tools they have available to manipulate enzymes and biosynthetic pathways at a molecular (amino acid) level using techniques such as site directed mutagenesis, directed evolution and a combination of chemo-enzymic and biological approaches. These approaches have met with varying degrees of success.

Initial studies of site-directed mutagenesis sought to modify the activity of an enzyme by making 'conservative' changes that would not disrupt the folding of the protein and therefore would have only a minimal and site-specific effect on the enzyme. The use of 'conservative' mutagenesis changes the amino acid in a protein to one of a similar size and with similar physicochemical characteristics to the one being replaced [28]. However a more intrusive approach will often be required. Changes to key residues in the active site of an enzyme, for example, can have far reaching effects on substrate specificity and reaction rate. For example, PKS enzymes have been the focus of much interest in this area since such experiments can supplement domain swapping investigations and illuminate aspects of the mechanism of action of the enzyme [29]. Site directed mutagenesis has also been used to alter the cofactor affinity, enhance the stability and improve the activity of a monooxygenase [30]. There are numerous examples of the use of mutagenesis to improve the characteristics of an enzyme. In many cases, however, there is a necessary trade off between activity and stability. In part this is a consequence of the competing requirements for flexibility near the enzyme active site and the need for rigidity to prevent denaturation. In order to develop strategies which are able to solve this contradiction, it is necessary to develop robust models to link genotype with phenotype and to use these to engineer changes resulting in predictable properties [31]. A range of models

have been developed to predict phenotype from genotype and these are able to help in the design of proteins and to predict phenotype from a given natural protein DNA sequence [32, 33] as well as protein structure from amino acid sequence data [34] or even from the DNA sequence [35]. While prediction of protein structure from sequence data may provide useful insights into the structure (and perhaps function) of 'non-crystallisable' proteins, it is important to remember that the putative structure predictions are only as valid as the data on which they are built and, ultimately, chemistry is an empirical exercise and proof of structure and function will come from actual data rather than computer predictions.

A simpler (in principle) approach to the modification of the activity of proteins is to carry out random mutagenesis and to assay the derived protein for a desired activity or property. This is the basis of directed evolution which mimics the evolutionary process but makes use of much shorter timescales and more specific changes. The essence of directed evolution is the creation of a set of genetically diverse proteins through mutagenesis followed by identification of those with the desired set of properties [36]. This process of diversification, creation of a library and screening for desirable properties has been used to engineer biomolecules such as proteins all the way to complete living systems. It has the potential to derive improved molecules and systems or to create completely novel entities. On an individual enzyme level, random mutagenesis coupled to directed evolution has been used to create libraries of artificial metalloenzymes with enhanced properties [37]. In addition to conventional mutagenesis, the errors that occur in PCR can also be used to generate mutations. One of the advantages to an activity-driven approach is that the protein sequence does not have to be known. The essence of the approach has some similarity to combinatorial chemistry in that a library of species is created and then inspected for activity. The choice of which to take forward is made on the basis of the screen rather than any structure:activity relationship. The problem with such an approach is that it provides a significant number of examples which can be screened rapidly for hits but it contributes little to the development of theorems or proofs. This is analogous to the development of vast algorithms to prove individual cases rapidly without considering if a more generic proof might be achieved. One example of how to derive specific data from theorem development, taken from mathematics, is the concept of magic squares.

A magic square is a square with a set number of cells. Each cell is filled with an integer such that each vertical, horizontal and main diagonal always adds up to the same value—known as the magic number or magic constant. The first of these is the 3×3 square and it was known to the ancient Chinese and called the Lo Shu. It was used by them (using your birth date as a starting point) to determine your character and destiny [38]. It is possible, if you have enough computing power to calculate the magic square components and the magic constant for a large number of grids very quickly. Generic methods have been developed for the construction of magic squares and, while some of them at least are somewhat tedious, they have wider applicability than constructing a separate solution for each magic square—no matter how great the computational power you have at hand. The value of the Magic Constant (M), for example, is given by the formula:

$$M = n(n^2 + 1)/2$$

where n is the order (number of cells along the side) of the magic square.

Magic squares were also known to ancient Arabic and Indian mathematicians and were thought to have special properties. They were also used in astrological calculations. This fusion of science which was driven by astronomical observation and the adoption of ancient Greek mathematics (in translation), was an important aspect of the spread of ideas and concepts. In biological chemistry the drivers were, in addition to the development of concepts and techniques, the development of knowledge about the relationship between biologically active compounds in plants and fungi, in particular. By using the principles of the 'magic square' and knowing how diverse data is related, it is possible to predict the identity of each component of a magic square even if they have not yet been discovered. The conceptual basis of the magic square is that, no matter which variables you choose, they remain consistent and predictable. The tradition has continued—particularly in China but, also as we have seen in the development of statins, by Japanese and other scientists [39]. The properties of natural products allows them to be fitted into a biochemical magic square from which the constants that link them can be derived.

It is even in principle possible to construct a chemistry 'magic square' using the atomic numbers of the elements however, unlike the Periodic Table, the components in the rows, columns or even diagonals have no common theme. The driver for the creation of the Periodic Table was chemical similarity (albeit largely driven by outer shell electron configuration) rather than the extension of an esoteric mathematical principle. For that reason a chemistry 'magic square' has no underlying logic and is not mathematically provable even if it could be developed using a different (non-mathematical) set of underlying principles.

The idea that living things are, in essence, a form of machinery (like a clock) was expressed by Descartes and others [40]. While this was seen as being overly simplistic, the more modern expression of this is that all life can be explained as being analogous to a computer. The Turing test has already been discussed as a way of distinguishing humans from machines however the concept of directed evolution may also suggest that a purely computational approach is an unlikely option. Calculation and quantification of properties will not create the objectively best solution however it may allow for comparative assessments (i.e. X is better than Y).

The essential aim of directed evolution is the development of components with improved properties through fast tracking the evolutionary process by coupling it with rapid screening methods to select the best solution after each round of gene diversification. It can have a number of different targets however, from the standpoint of natural products, the aim is to create new molecular white space that can be exploited. The 'best' solution can be thought of as the expression of the Magic Constant for the particular system being investigated and the solution that is being sought. The magic square and the magic constant become the representation of the attributes of the perfect solution and therefore become crucial to finding the closest match to that perfect solution. Perhaps the most direct (and simplest) example of an endpoint is the production of proteins with enhanced properties [41]. This approach

can make use of advanced techniques in machine learning to close in on a solution more rapidly. This has been used extensively in pharmaceutical science and for the discovery of novel entities [42]. It has even been used 'beyond sequence' to predict the effect of post-translational modifications on the properties, stability and therapeutic potential of proteins [43]. In the context of natural products, directed evolution is a complex aim. The simplest way of looking at the potential of directed evolution is somewhat similar to the use of a bioreactor containing an enzyme that carries out a single reaction. If an enzyme is picked that is able to carry out the reaction but is limited in some way (it produces a related, but not identical compound to the one required or the rate of the reaction is lower than optimal) then there is scope for improvement. If the gene coding for the enzyme is subjected to random mutations then the multiplicity of slightly different enzymes that are produced will have slightly different properties from the starting point enzyme and from each other. Some will be slightly better and some will be worse. If a selection of those that are better are taken and subjected to a second (or more) rounds of mutation, eventually this process of 'directed evolution' will give a much more optimised enzyme that can then be used commercially to produce the required product. While this sounds like a very random process with mutations spread throughout the enzyme, in fact it is possible to direct the process a little more by focusing the mutation site on a particular part of the enzyme thereby increasing the chances of obtaining a functionally improved enzyme. Directed evolution of a single enzyme is just one of a number of techniques that can be used to create novel natural products. As well as directed evolution of specific reactions and individual pathways, the directed evolution of complete biosynthetic pathways has been achieved. For example, diverse carotenoid biosynthetic pathways have been created using directed evolution [44]. This approach has been used with specific enzyme targets to alter specific aspects of natural product formation (e.g. formation of an unnatural long chain carotenoids through directed evolution of a specific desaturase) [45]. There are additionally two important ideas that underpin the concept of directed evolution. The first of these is that enzymes are potentially multi-functional. Changes to an enzyme may change the type of reaction it is capable of catalysing. In order to exploit this potential, it is necessary to have an understanding of the mechanisms of the biological reactions that take place through enzyme catalysis and the conditions under which the reaction takes place. While random mutation may provide part of the mechanistic answer and enable changes in function to be planned, replacement of one region of a protein with a region from another protein (homologous recombination) may produce a protein with additive or very different properties from the progenitor proteins. The nature of these 'chimeric' proteins can be, at least in part, predicted by understanding the structure:function relationships of the contributing proteins. Apart from random mutation, a range of other techniques can be used to create libraries of protein structures and to screen these for functional benefits [46]. The second point is that functionally neutral mutations may facilitate future, functionally important mutations. These may be revealed by a further series of mutations and/or by a deeper understanding of the nature of the link between structure and function.

While directed evolution has had a number of successes, some of these have been built upon a number of less successful excursions into the area of protein engineering. One of the challenges in carrying out the directed alteration of proteins is to conclusively determine the link between the primary sequence of a protein and the precise structure of the folded molecule. A number of complex approaches have been advanced however this remains, as noted previously, a non-trivial problem [47]. In any case, it may be simpler to continue with a strategy of carrying out random mutations and assaying for desired activities and properties. It is important to bear in mind the ultimate objective—whether that is increasing knowledge of protein structures or improving the properties of a particular protein. One other area which has developed in response to the latter, is the creation of synthetic enzymes. These may contain unnatural units and be derived from synthetic genetic polymers [48] or may contain a mixture of biotic and abiotic active sites [49]. By using an existing enzyme with a single active site, it is possible to induce novel activity by inserting a coordination metal atom into a functionally convenient site thereby allowing dual functionality within a single enzyme to be created [50]. Such an approach can be used to create new to nature enzymes with novel functionalities and to carry out reactions that do not occur normally. It is important that the advantages of this approach be carefully assessed to ensure that the enzymes being developed contain a useful activity rather than just showing proof of principle [51]. The idea of using non-protein material to design and create enzymes from first principles has been advocated [52]. Such an approach lies within the broad tradition of chemistry since it is oriented towards designing a solution to a specific problem rather than a biological approach which relies more on finding a use for properties that already exist within a given biological system. Although there have been a number of successes, there have also been (less frequently advertised or published) failures. In some cases, as noted previously this may be due to the failure of an engineered protein to fold as anticipated or to the development of inclusion bodies. Even trying to carry out relatively conservative changes to the engineered protein (e.g. not adding, removing or changing charged residues, only changing amino acids for others of a smaller, not larger, size), can lead to an unanticipated loss in activity or functionality. Another source of error is failure at the nucleotide mutation introduction level [53]. Nonetheless, engineering new catalytic function in enzymes, whether through random mutagenesis and screening or directed evolution, remains an area of considerable interest [54]. In addition to developing novel catalytic activity, such approaches also support the trend towards more sustainable systems and away from the development of limited or environmentally harmful solutions.

7.4 Abzymes and Molecular Imprinting

The twin concepts of antibody enzymes (abzymes) and molecular imprinting to create catalytically active entities are related. Both make use of the concept that enzymes work by stabilising the transition state between reactant and product (see

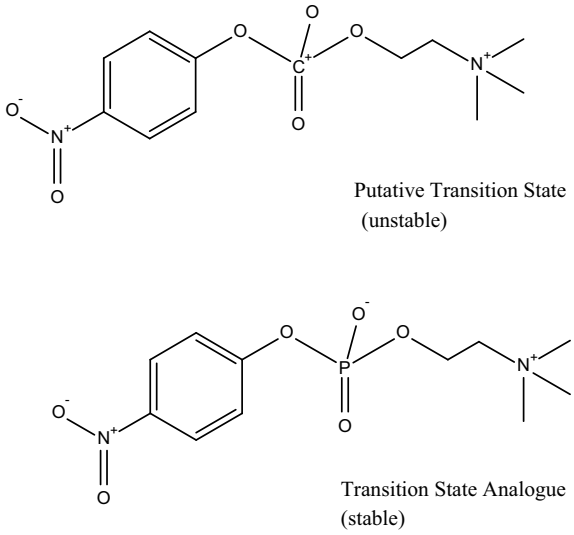


Fig. 7.5 Putative transition state and stable transition state analogue

below for a fuller explanation). This short-lived transition state does not have a sufficiently long biological lifetime to be able to be the target for raising antibodies against. It was suggested, however, that if a chemically stable transition state analogue could be created, then it may be possible to generate monoclonal antibodies against this type of molecule and these might have catalytic competence since they would stabilise a putative reaction transition state. This was achieved by using a pentavalent phosphorous atom as a stable transition state analogue for a reaction in which a transient pentavalent carbon atom formed in the transition state. This approach was taken by two groups who used p-nitrophenyl phosphoryl choline [55] and monoaryl phosphonate esters [56] as transition state analogues for esterase reactions (Fig. 7.5) and to raise monoclonal antibodies against them.

The biological molecules created were catalytically active antibodies and were abbreviated to the term 'abzymes'. Naturally occurring catalytic antibody enzymes have also been found to exist and to play a role in some human diseases [57]. Not only have a substantial number of these transition state analogues been created, they have also been designed to catalyse reactions that do not occur naturally [58] although in this specific case it should be noted that a natural Diels Alderase was subsequently discovered [59], nonetheless, the principle remains. Abzymes have been found to have clinical relevance in, for example, autoimmune diseases [59].

While a detailed description of abzymes is out of scope, the basic principles derive from molecular imprinting (which can be considered the basis of the generation of all antibodies) and of the transition state theory of enzyme-catalysed reactions developed by Linus Pauling in 1940. Pauling suggested that enzymes bind to substrates and that they then distort them towards an activated transition state conformation. The corollary of this is that compounds that mimic the transition state will bind more closely to the enzyme active site than the substrate does [60]. This seemed to be

confirmed by the tight binding of transition state analogues to enzyme active sites [61]. Dynamic models have also been developed however the idea that a transition state exists on the reaction pathway between reactant and product remains a key concept [62]. As noted above, if a macromolecule (such as an antibody) can be constructed such that it recognises and binds to the transition state (or a transition state analogue) in the reaction pathway, it will have catalytic capability because it will stabilise (and lower the energy of) the molecule in the transition state. The principle of molecular imprinting moves this concept out of a living (animal) system and into the realm of chemistry. The central idea of molecular imprinting is that a polymer can be formed around a template molecule and, when the template is removed, the remaining polymer will be able to be used to bind to the template (or to close analogues of the template). Depending on the nature of the template molecule used for the imprinting, such imprinted polymers (commonly called 'plastic antibodies') can operate as specific separation materials to selectively bind molecules, as antibody and receptor binding site mimics in recognition and assay systems, as enzyme mimics for catalytic applications (as non-biological abzymes, in effect) and as recognition elements in bio-sensors [63]. Proof of principle that cinnamic acid diesters could be used as transition state analogues for the 2 + 2 cycloaddition of phenyl propenoid units in the plant cell wall has been demonstrated [64]. The same unit was used in a variation of molecular imprinting in which a globular protein (bovine serum albumin) was imprinted with a set of cinnamic acids and diesters containing various ring functionalities. The protein was then cross linked with glutaraldehyde and the imprinting molecules washed out. The principle being established was that globular proteins contain a number of amino acids that were also present in enzyme active sites but, in the native state, the amino acids were in the wrong relative orientation to catalyse a reaction. By creating a molecular imprint of a transition state analogue, it was hypothesised that this would create an abzyme-like structure in vitro and this could be used to catalyse reactions without the need to raise antibodies. In preliminary studies, a number of the features of the imprinted protein were identified. For example, it was found that charged functional groups (for example carboxylic acid moieties) were more significant drivers for molecular recognition than groups that occupied uncharged space).

7.5 Conclusion

The development and utilisation of enzymes and their use in natural product chemistry has been a significant driver of biotechnology and the manipulation of natural products. Modification of enzymes by genetic engineering, directed evolution, random mutation or manipulation of reaction conditions has led to the development and creation of unnatural products as well as the optimisation of production of natural products. Many of the advances have not led to commercially viable systems although the impetus for increased sustainability is a driver for the adoption of green methods which make use of enzymes (generally operating under mild conditions) rather than

metal catalysts. Further work is still needed on the optimisation of biological systems for the production of both natural and unnatural products. Some challenges remain including cofactor recycling, optimisation of reaction conditions and the elimination of side reactions. Some aspects of the use of enzymes such as immobilisation and how to optimise aspects of the use of enzymes such as turnover number have not been dealt with in detail since they are very generic issues and not specific to natural products [65]. In addition, there are a number of specialist reviews and texts that deal appropriately with these (and other) areas. The specificity and precision of enzyme reactions mean that they will continue to be a core aspect of the development of new and valuable biotechnological processes. The use of enzymes, if it is to reach its full potential, will require the continued use of a range of multidisciplinary techniques including genetic level manipulation, systems engineering, synthetic biology, molecular microbiology and natural product chemistry [67]. Significant challenges remain however there are a number of avenues that remain to be explored and exploited.

References

1. Jacobson AE (1978) Analgesics and their Antagonists: structure-activity relationships. In: Iversen LL, Iversen SD, Snyder SH (eds) Drugs of abuse. Handbook of psychopharmacology. Springer, Boston, MA. https://doi.org/10.1007/978-1-4684-3186-5_2
2. Minami A (2023) A new trend in biosynthetic studies of natural products: the bridge between the amino acid sequence data and the chemical structure. In: Ishikawa H, Takayama H (eds) New tide of natural product chemistry. Springer, Singapore. https://doi.org/10.1007/978-981-99-1714-3_7
3. Rudolf JD, Chang CY, Ma M, Shen B (2017) Cytochromes P450 for natural product biosynthesis in Streptomyces: sequence, structure, and function. Nat Prod Rep 34(9):1141–1172. https://doi.org/10.1039/c7np00034k.PMID:28758170;PMCID:PMC5585785
4. Noda-Garcia L, Tawfik DS (2020) Enzyme evolution in natural products biosynthesis: target- or diversity-oriented? Curr Opin Chem Biol 59:147–154. https://doi.org/10.1016/j.cbpa.2020.05.011. Epub 2020 Aug 6. PMID: 32771972
5. Masson P, Lushchekina S (2022) Conformational stability and denaturation processes of proteins investigated by electrophoresis under extreme conditions. Molecules 27(20):6861. https://doi.org/10.3390/molecules27206861. PMID: 36296453; PMCID: PMC9610776
6. Khan MR (2021) Immobilized enzymes: a comprehensive review. Bull Natl Res Cent 45:207. https://doi.org/10.1186/s42269-021-00649-0
7. Jankowska K, Bachosz K, Zdarta J, Jesionowski T (2020) Application of enzymatic-based bioreactors. In: Ochowiak M, Woziwodzki S, Mitkowski P, Doligalski M (eds) Practical aspects of chemical engineering. PAIC 2019. Springer, Cham. https://doi.org/10.1007/978-3-030-39867-5_12
8. Lineweaver H, Burk D (1934) The determination of enzyme dissociation constants. J Am Chem Soc 56(3):658–666
9. Soleimany AP, Bhatia SN (2020) Activity-based diagnostics: an emerging paradigm for disease detection and monitoring. Trends Mol Med 26(5):450–468. https://doi.org/10.1016/j.molmed.2020.01.013. Epub 2020 Apr 5. PMID: 32359477; PMCID: PMC8290463
10. Moore CE, Meacham-Hensold K, Lemonnier P, Slattery RA, Benjamin C, Bernacchi CJ, Lawson T, Cavanagh AP (2021) The effect of increasing temperature on crop photosynthesis: from enzymes to ecosystems. J Exp Bot 72(8):2822–2844. https://doi.org/10.1093/jxb/erab090. PMID: 33619527; PMCID: PMC8023210

11. Isah T (2019) Stress and defense responses in plant secondary metabolites production. Biol Res 52:39. https://doi.org/10.1186/s40659-019-0246-3
12. Yeshi K, Crayn D, Ritmejerytė E, Wangchuk P (2022) Plant secondary metabolites produced in response to abiotic stresses has potential application in pharmaceutical product development. Molecules 27(1):313. https://doi.org/10.3390/molecules27010313. PMID: 35011546; PMCID: PMC8746929
13. Qiu S, Yang A, Zeng H (2023) Flux balance analysis-based metabolic modeling of microbial secondary metabolism: current status and outlook. PLoS Comput Biol 19(8):e1011391. https://doi.org/10.1371/journal.pcbi.1011391. PMID: 37619239; PMCID: PMC10449171
14. Del Carratore F, Hanko EK, Breitling R, Takano E (2022) Biotechnological application of Streptomyces for the production of clinical drugs and other bioactive molecules. Curr Opin Biotechnol 77:102762. PMID:35908316
15. Orth J, Thiele I, Palsson B (2010) What is flux balance analysis? Nat Biotechnol 28:245–248. https://doi.org/10.1038/nbt.1614
16. Wu T, Kerbler SM, Fernie AR, Zhang Y (2021) Plant cell cultures as heterologous bio-factories for secondary metabolite production. Plant Commun 2(5):100235. https://doi.org/10.1016/j.xplc.2021.100235. PMID: 34746764; PMCID: PMC8554037
17. Kowalczyk T, Merecz-Sadowska A, Picot L, Brčić Karačonji I, Wieczfinska J, Śliwiński T, Sitarek P (2022) Genetic manipulation and bioreactor culture of plants as a tool for industry and its applications. Molecules 27(3):795. https://doi.org/10.3390/molecules27030795. PMID: 35164060; PMCID: PMC8840042
18. Karan R, Capes MD, DasSarma S (2012) Function and biotechnology of extremophilic enzymes in low water activity. Aquat Biosyst 8:4
19. Eom S, Wang J, Steitz T (1996) Structure of Taq polymerase with DNA at the polymerase active site. Nature 382:278–281. https://doi.org/10.1038/382278a0
20. Kohli I, Joshi NC, Mohapatra S, Varma A (2020) Extremophile—an adaptive strategy for extreme conditions and applications. Curr Genomics 21(2):96–110. https://doi.org/10.2174/1389202921666200401105908. PMID: 32655304; PMCID: PMC7324872
21. Furniss CSM, Hanley AB, Mackie AR, Mingins J (1991) DNA manipulation in low water systems. Enzyme Microb Technol 13:525
22. Hanley AB, Parsley KR, Lewis JA, Fenwick GR (1990) Chemistry of indole glucosinolates: intermediacy of indol-3-yl methyl isothiocyanates in the enzymic hydrolysis of indole glucosinolates. J Chem Soc Perkin Trans I:2273
23. Lai J, Huang H, Lin M, Xu Y, Li X, Sun B (2023) Enzyme catalyzes ester bond synthesis and hydrolysis: the key step for sustainable usage of plastics. Front Microbiol 13:1113705. https://doi.org/10.3389/fmicb.2022.1113705. PMID: 36713200; PMCID: PMC9878459
24. Shrivastava S (2020) Introduction to glycoside hydrolases: classification, identification and occurrence. In: Shrivastava S (ed) Industrial applications of glycoside hydrolases. Springer, Singapore. https://doi.org/10.1007/978-981-15-4767-6_1
25. Schober M, Faber K (2013) Inverting hydrolases and their use in enantioconvergent biotransformations. Trends Biotechnol 31(8):468–478. https://doi.org/10.1016/j.tibtech.2013.05.005. Epub 2013 Jun 25. PMID: 23809848; PMCID: PMC3725421
26. Hänsler M, Jakubke HD (1996) Reverse action of hydrolases in frozen aqueous solutions. Amino Acids 11:379–395. https://doi.org/10.1007/BF00807943
27. Wu D, Lei X (2022) Enzymatic cascade reactions for the efficient synthesis of natural products. Tetrahedron 127:133099
28. Wong TS, Tee KL (2020) Gene mutagenesis. In: A practical guide to protein engineering. Learning materials in biosciences. Springer, Cham. https://doi.org/10.1007/978-3-030-56898-6_8
29. Drufva EE, Hix EG, Bailey CB (2020) Site directed mutagenesis as a precision tool to enable synthetic biology with engineered modular polyketide synthases. Synth Syst Biotechnol 5(2):62–80. https://doi.org/10.1016/j.synbio.2020.04.001. Erratum in: Synth Syst Biotechnol 5(4):268. https://doi.org/10.1016/j.synbio.2020.07.007. Erratum in: Synth Syst Biotechnol 5(4):332. https://doi.org/10.1016/j.synbio.2020.10.004. PMID: 32637664; PMCID: PMC7327777

30. Mansouri HR, Gracia Carmona O, Jodlbauer J, Schweiger L, Fink MJ, Breslmayr E, Laurent C, Feroz S, Goncalves LCP, Rial DV, Mihovilovic MD, Bommarius AS, Ludwig R, Oostenbrink C, Rudroff F (2022) Mutations increasing cofactor affinity, improve stability and activity of a Baeyer-Villiger monooxygenase. ACS Catal 12(19):11761–11766. https://doi.org/10.1021/acscatal.2c03225. Epub 2022 Sep 13. PMID: 36249873; PMCID: PMC9552169
31. Vanella R, Küng C, Schoepfer AA et al (2024) Understanding activity-stability tradeoffs in biocatalysts by enzyme proximity sequencing. Nat Commun 15:1807. https://doi.org/10.1038/s41467-024-45630-3
32. Orgogozo V, Morizot B, Martin A (2015) The differential view of genotype-phenotype relationships. Front Genet 6:179. https://doi.org/10.3389/fgene.2015.00179. PMID: 26042146; PMCID: PMC4437230
33. Medvedev A, Mishra Sharma S, Tsatsorin E, Nabieva E, Yarotsky D (2022) Human genotype-to-phenotype predictions: Boosting accuracy with nonlinear models. PLoS One 17(8):e0273293. https://doi.org/10.1371/journal.pone.0273293. PMID: 36044406; PMCID: PMC9432766
34. Jumper J, Evans R, Pritzel A et al (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2
35. Braberg H, Echeverria I, Kaake RM et al (2022) From systems to structure—using genetic data to model protein structures. Nat Rev Genet 23:342–354. https://doi.org/10.1038/s41576-021-00441-w
36. Packer M, Liu D (2015) Methods for the directed evolution of proteins. Nat Rev Genet 16:379–394. https://doi.org/10.1038/nrg3927
37. Yang H, Swartz A, Park H et al (2018) Evolving artificial metalloenzymes via random mutagenesis. Nature Chem 10:318–324. https://doi.org/10.1038/nchem.2927
38. Cammann, S (1960) The evolution of magic squares in China. J Am Orient Soc 80:116–124
39. Shen Y, Hao X (2020) Natural product sciences: an integrative approach to the innovations of plant natural products. Sci China Life Sci 63:1634–1650. https://doi.org/10.1007/s11427-020-1799-y
40. Ben-Yami H (2022) Automata in the early modern period. In: Jalobeanu D, Wolfe CT (eds) Encyclopedia of early modern philosophy and the sciences. Springer, Cham. https://doi.org/10.1007/978-3-319-31069-5_239
41. Yang KK, Wu Z, Arnold FH (2019) Machine-learning-guided directed evolution for protein engineering. Nat Methods 16:687–694. https://doi.org/10.1038/s41592-019-0496-6
42. Xu Y, Verma D, Sheridan RP, Liaw A, Ma J, Marshall NM, McIntosh J, Sherer EC, Svetnik V, Johnston JM (2020) Deep dive into machine learning models for protein engineering. J Chem Inf Model 60(6):2773–2790. https://doi.org/10.1021/acs.jcim.0c00073. Epub 2020 May 5. PMID: 32250622
43. Ertelt M, Mulligan VK, Maguire JB, Lyskov S, Moretti R, Schiffner T, Meiler J, Schoeder CT (2024) Combining machine learning with structure-based protein design to predict and engineer post-translational modifications of proteins. PLoS Comput Biol 20(3):e1011939. https://doi.org/10.1371/journal.pcbi.1011939. PMID: 38484014; PMCID: PMC10965067
44. Umeno D, Tobias AV, Arnold FH (2005) Diversifying carotenoid biosynthetic pathways by directed evolution. Microbiol Mol Biol Rev 69(1):51–78. https://doi.org/10.1128/MMBR.69.1.51-78.2005. PMID: 15755953; PMCID: PMC1082795
45. Furubayashi M, Umeno D (2022) Use of directed enzyme evolution to create novel biosynthetic pathways for production of rare or non-natural carotenoids. Methods Enzymol 671:351–382. https://doi.org/10.1016/bs.mie.2022.03.008. Epub 2022 Apr 18. PMID: 35878986
46. Sellés Vidal L, Isalan M, Heap JT, Ledesma-Amaro R (2023) A primer to directed evolution: current methodologies and future directions. RSC Chem Biol 4(4):271–291. https://doi.org/10.1039/d2cb00231k. PMID: 37034405; PMCID: PMC10074555
47. Chen Y, Xu Y, Liu D et al (2024) An end-to-end framework for the prediction of protein structure and fitness from single sequence. Nat Commun 15:7400. https://doi.org/10.1038/s41467-024-51776-x

48. Taylor A, Pinheiro V, Smola M et al (2015) Catalysts from synthetic genetic polymers. Nature 518:427–430. https://doi.org/10.1038/nature13982
49. Palomo JM (2021) Artificial enzymes with multiple active sites. Green Sustain Chem 29:100452
50. Alonso S, Santiago G, Cea-Rama I, Fernandez-Lopez L, Coscolín C, Modregger J, Ferrer M (2020) Genetically engineered proteins with two active sites for enhanced biocatalysis and synergistic chemo-and biocatalysis. Nat Catal 3(3):319–328
51. Hanreich S, Bonandi E, Drienovská I (2023) Design of artificial enzymes: insights into protein Scaffolds. Chembiochem 24(6):e202200566. https://doi.org/10.1002/cbic.202200566. Epub 2023 Jan 3. PMID: 36418221
52. Lovelock SL, Crawshaw R, Basler S et al (2022) The road to fully programmable protein catalysis. Nature 606:49–58. https://doi.org/10.1038/s41586-022-04456-z
53. Wassman CD, Tam PY, Lathrop RH, Weiss GA (2004) Predicting oligonucleotide-directed mutagenesis failures in protein engineering. Nucleic Acids Res 32(21):6407–6413. https://doi.org/10.1093/nar/gkh977
54. Chen K, Arnold FH (2020) Engineering new catalytic activities in enzymes. Nat Catal 3:203–213. https://doi.org/10.1038/s41929-019-0385-5
55. Pollack SJ, Jacobs JW, Schultz PG (1986) Selective chemical catalysis by an antibody. Science 234:1570–1573. https://doi.org/10.1126/science.3787262
56. Tramontano A, Janda KD, Lerner RA (1986) Catalytic antibodies. Science 234:1566–1570. https://doi.org/10.1126/science.3787261
57. Tolmacheva AS, Nevinsky GA (2022) Essential protective role of catalytically active antibodies (abzymes) with redox antioxidant functions in animals and humans. Int J Mol Sci 23(7):3898. https://doi.org/10.3390/ijms23073898. PMID: 35409256; PMCID: PMC8999700
58. Hilvert D, Hill KW, Nared KD, Auditor MT (1989) Antibody catalysis of the Diels-Alder reaction. J Am Chem Soc 111:9261–9262. https://doi.org/10.1021/ja00208a037
59. Byrne MJ, Lees NR, Han LC, van der Kamp MW, Mulholland AJ, Stach JE, Willis CL, Race PR (2016) The catalytic mechanism of a natural Diels-Alderase revealed in molecular detail. J Am Chem Soc 138(19):6095–6098. https://doi.org/10.1021/jacs.6b00232. Epub 2016 May 6. PMID: 27140661
60. Gabibov AG, Ponomarenko NA, Tretyak EB, Paltsev MA, Suchkov SV (2006) Catalytic autoantibodies in clinical autoimmunity and modern medicine. Autoimmun Rev 5:324–330
61. Wolfenden R (1969) Transition state analogues for enzyme catalysis. Nature 223(5207):704–705. https://doi.org/10.1038/223704a0. PMID: 4979456
62. Lienhardt GE (1973) Enzymatic catalysis and transition-state theory. Science 180(4082):149–154. https://doi.org/10.1126/science.180.4082.149 PMID: 4632837
63. Schramm VL (2007) Enzymatic transition state theory and transition state analogue design. J Biol Chem 282(39):28297–28300. https://doi.org/10.1074/jbc.R700018200. Epub 2007 Aug 9. PMID: 17690091
64. Mosbach K (1994) Molecular imprinting. Trends Biochem Sci 19(1):9–14
65. Hanley AB, Russell WR, Chesson A (1993) Formation of substituted truxillic and truxinic acids in plant cell walls—a rationale. Phytochem 33:957
66. Smejkal GB, Kakumanu S (2019) Enzymes and their turnover numbers. Expert Rev Proteomics 16(7):543–544. https://doi.org/10.1080/14789450.2019.1630275
67. Jaeger KE, Liese A, Syldatk C (2024) Introduction to enzyme technology. In: Jaeger KE, Liese A, Syldatk C (eds) Introduction to enzyme technology. Learning materials in biosciences. Springer, Cham. https://doi.org/10.1007/978-3-031-42999-6_1

Chapter 8
Receptors and Natural Products

Abstract One of the key concepts that underpin the interaction between natural products and receptors is the principle of causality that was explained by Kant. In the case of receptors, they can control both a local (granular) and a physiological (systemic) effect. It is not always necessary to understand the details of all the steps that lie between these since knowledge of molecular binding and physiological endpoint and the substantive nature of the relationship between them may enable manipulation and control. One of the most common classes of receptors are G Protein Coupled Receptors (GPCRs). Taste, odour and vision receptors are examples of GPCRs. Taste receptors seem to have a significant extra-oral function and are involved in more than simply taste transduction. In terms of physiological effect, enzyme inhibitors are an important class of medicinal compounds. Studies of antimicrobial resistance (AMR) suggests that targeting multiple enzymes associated with a particular microbial physiological function (for example cell wall construction or energy metabolism) may facilitate overcoming AMR. This may be important as increasing numbers of disease states (including, for example, Alzheimer's Disease) may be exacerbated by inflammation which is provoked by infection.

Keywords Causality · Signalling · GPCR · Taste receptors · Agonists · Antagonists

The concept that natural products interact with living systems and this interaction is mediated through receptors is a key concept in understanding the biological activity of such natural products. It has been explained, in global terms, by Kant's ideas of causality. The basis of the interaction between natural products and living systems is a two way process with, on the one hand, natural products engaging with aspects of a range of biological targets within a single biological system and, on the other, biological targets interacting with several different natural products at the same time. This can create several levels of complexity. It can therefore be challenging to deconvolute the complexity of the totality of the biological response and to determine which particular aspect of that response is most important. Such studies form the basis of much of the progress in research into the interactions of natural products

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_8

with biological systems. It is not necessary to understand the complexity of living systems in all of their detail in order to study aspects of the interactions and their significance. Biological systems (even extremophiles) operate under specific sets of limiting conditions and this implies that they operate within homeostatic boundaries to ensure that the system as a whole is biologically stable. Kant proposed that an organism is both cause and effect within itself and these must be balanced. The different aspects of living systems, indeed, are purposely designed to maintain the system in a state of balance. This does not mean, however, that systems are insulated from external influences or that homeostasis is a wholly static condition. However, maintaining a non-equilibrium state can come at a cost. For example, while the cells of living systems can operate a gradient (for example in the concentration of calcium), this is energetically unfavourable and requires energy in order to be maintained. The maintenance of a calcium gradient comes with an associated energetic cost. Receptors are both interlinked (as part of a whole organism and its systemic response) and granular (they respond locally to a specific ligand). At one end of the scale, chemists examine properties of specific ligands and how they interact with individual receptors on a molecular level. At the other end of the scale, biologists study how a complete organism responds systemically to a perturbation or challenge. Between these two extremes, the dense interlinked cascade of effects that start with a molecular interaction is understood to a lesser or greater extent in the context of the organism as a complete entity.

The Kantian doctrine of the 'law of final causality' suggests that causality is a function of human experience rather than being an additive summation of the mechanical steps that make up a specific result. This suggests that causality cannot be understood solely in terms of empirical experience or discovery (for example, the additive effects of the binding of compounds to multiple receptors or of multiple compounds interacting with one or several receptors). Turning this argument around, defining living systems in terms of overall causality—even if all of the steps in the chain are not known, is a viable option since the overall effect is non-additive so knowing the impact of some of the steps can provide sufficient evidence for a likely chain of causality even if some components are not visible. Since causality is non-additive, the total effect will not be able to be divided into a set number of equal steps. Some will be more impactful than others. Ultimately this leads to the concept of the rate determining step. This is the step upon which the rate of the biological response depends. The binding of a ligand to a receptor can be considered to be the first step (and, in some cases will be the rate determining step) in the chain of causality that leads to the maintenance of the stability of a living system. Knowing the consequences of the loss of the laws of causality (the catastrophic disintegration of the living system thereby bringing the causality network to an end) allows us to contextualise the knowable first step (natural product ligand binding to a receptor) in terms of final outcome even if we do not have all the details of the chain of causality. An analogy of this might be a car journey. We know that, in order to take the journey, we need to begin by starting the car. Subsequent actions that take us to the final destination may not be known but we can tell if we arrive at the correct or the wrong place even if we do not know the precise details of how that journey was undertaken

and which route was chosen. In that sense, we have causal knowledge of the eventual consequences of the interaction of a natural product with a living system but not the details that would allow us to make an additive (and hence mechanical summation) of the individual steps.

8.1 Receptors and Cell Signalling—A Cascade of Events

While a single molecule may react with a specific biological macromolecule to produce a singular and measurable event, biological systems have a range of complex and interlinked components and receptor binding can also act as a trigger for a systemic response. In simple terms, a cell interacts with components in its environment through the binding between a natural product (generally a small, functionalised molecule) and a receptor. The binding of the ligand induces a change in the receptor which then (usually) induces a cascade of biological consequences. These biological changes are thought of in terms of the transfer of information through a range of chemical messengers to other components of the cell. Such a process is referred to as 'signal transduction' and it is the basis of much of our understanding of pharmacology, toxicity and hormonal responses. While physiological responses have been extensively studied and signal transduction pathways are a useful means of visualising the cellular basis of the interaction of living systems with natural products, there are also instances where the interaction of ligands with biological systems can be thought of in terms of quantum effects. Some of these have been discussed in the context of quantum events such as tunnelling and quantum mechanical vibration [1]. While this may appear to be a somewhat esoteric field of study, quantum biology seems to be significant in photosynthesis and in the detection of odorant molecules. Since many natural products affect the perception of taste and smell, the possibility that quantum effects might govern either is both attractive and challenging and may suggest alternative ways of considering structure:activity relationships [2]. In a more directly relevant and practical application, the importance of quantum effects in the discovery and design of pharmaceuticals has been noted [3]. Both molecular modelling and quantum mechanics have particular attributes and hybrid methods that combine aspects of both have been suggested, for example, in drug design and discovery [4]. It is unclear if rational design will have enough advantages in terms of numbers of useful pharmacological entities discovered per dollar of research spend compared to rapid screening methods which can be designed to identify biologically useful molecules from a large number of automatically generated leads.

The more traditional approach to the investigation of natural products in terms of their interaction with living systems makes use of the type of receptor-based paradigms mentioned previously. The chemistry approach adopts a reductionist and empirical approach by defining the properties of receptors (classification) and then building up an understanding of their interactions with ligands and with the rest of the biological system.

The simplest classification of receptors is based upon where they operate since that also defines their properties—internally or on a cell surface. Cell surface receptors generally comprise a cell surface domain which binds the ligand, a transmembrane domain that spans the cell membrane and an internal domain. The implication of the three domain structure is that a transmembrane receptor is closely involved in signal transduction—converting the signal received by the binding of the ligand extracellularly into an intracellular event (or series of events). Possibly the simplest of the transmembrane receptors are those that operate by opening an ion channel. These receptors, upon binding to a ligand, undergo a conformational change that allows ions (for example, calcium) to flow through to the interior of the cell. While the operation of the ion channel receptors is relatively simple, the impact can be profound. For example, changes in the intracellular calcium ion concentration can affect phosphate concentration and high levels of calcium will disrupt cellular homeostasis. In addition, a number of proteins contain calcium binding sites and binding to these can induce conformational changes that affect the protein function. More generally ion channels have an important function in a range of responses (particularly neural and muscle responses) and diseases [5]. The simple nature of the primary function of ion channel receptors, therefore, belies their importance to living systems. A second major class of transmembrane receptors are G-protein coupled receptors (GPCRs). These comprise a multifunctional group of receptors that are involved in a plethora of physiological responses and as such represent a major target for pharmaceutical investigation. GPCRs are activated by a range of extracellular ligands and, through signal transduction, translate these external binding events into a series of internal consequence through interaction with G proteins (rather than opening an ion channel). The binding of GPCRs to G proteins results in modulation of the prime function of G proteins (guanine nucleotide binding). The G protein acts as a switch whose function depends on the bound nucleotide. The family of G proteins include a number of different functional sets which govern diverse physiological outcomes including both stimulation and inhibition [6]. The physiological functions that are controlled by G proteins include taste, olfaction and vision. While G proteins are activated by GPCRs, they can also be regulated by non-GPCRs. Much of the work done to delineate the physiological role of GPCRs and G proteins has been carried out using knock out gene mutations in a rodent model. With the advent of CRISPR/Cas9 technology, it is likely that knock out studies will become less technically challenging and more precise. Studies of pathways combined with physiological investigations suggest that different types of G proteins are involved in functions as diverse as cardio function, immune regulation, taste and olfaction, muscle control and even in cell division [6]. A full appreciation of the role of any biological molecule can only arise from knowledge both of how it operates and the pathway that it acts upon. In this way, the consequences of the introduction of natural products can be determined.

In addition to GPCRs, there is a third type of cell surface receptor that interacts with natural products—the enzyme-linked receptor. This type of receptor binds a ligand on the cell surface which then leads to activation of a linked enzyme which occurs on the portion of the receptor in the cytoplasm. This type of allostery has been found in a number of cases including, most notably, that of tyrosine kinase [7].

While GPCRs and ion channels respond to the extracellular environment (including the presence of natural products), some receptors (intracellular receptors) are only available to compounds that can migrate (or are actively transported) to the interior of cells. Such molecules must generally be hydrophobic in order to pass through the cell membrane. Cells contain a number of intracellular receptors and these are involved in a range of cellular responses. One of the major functions of intracellular receptors is the modulation of mRNA synthesis and hence protein production. For example a range of hydrophobic steroids bind to receptors and, upon translocation to the nucleus, the hormone-receptor complex then binds to DNA at a gene regulatory site and induces gene expression.

Receptors are one of the most significant mediators of the interaction of natural products with living systems. While molecules did not develop in response to physiological benefits for the non-host organism, the interaction of natural products specifically with GPCRs involved in taste perception can provide a significant advantage in discouraging predation. These will be considered in more detail later.

8.2 Natural Products and Enzyme Inhibition

The interaction of natural products with receptors makes a significant contribution to their biological activity however some natural products are also able to interact directly with another class of biologically active, proteinaceous macromolecules—enzymes. We have previously considered the role of enzymes in the biotransformation of natural products (Chap. 7). Many important aspects of biological systems are mediated by enzyme-catalysed reactions and some of these can be disrupted by natural products. The importance of enzymes as molecular pharmacology targets has been noted. Indeed it has been suggested that almost half of the small molecule drugs that are available have enzyme inhibition as their core activity [8].

There has been substantial interest in the discovery of pharmacologically important natural enzyme inhibitors [9]. The combination of screening and in silico modelling has enabled new potentially useful drugs to be discovered. In addition to their potential as enzyme inhibitors, some natural products can also act as enzyme activators. For example AMP-activated protein kinase (AMPK) activators have been found. The link between AMPK activation and prevention of cardiovascular disease has been mechanistically probed and signalling pathways have been identified [10]. A number of natural products including berberine, resveratrol, quercetin and other phenolics have been found to have a positive effect on AMPK. Further studies may help to elucidate more details of the link between cardiovascular and other metabolic diseases and the AMPK pathway. As studies on AMPK-related pathways has illustrated, metabolic diseases are an important target for enzyme-related interventions. In particular, attention has focussed on amylases, glucosidases and lipases and their link to metabolic diseases such as diabetes and obesity. The major requirements when developing assays to try to focus on potential active agents, no matter which biological target is the focus of investigation, is that assays should be specific, rapid,

cost-effective and robust. This has been noted and reviewed [11]. Such requirements are also obligatory for the development of activity guided isolation and fractionation protocols.

Another aspect of the use of enzymes and the potential role of natural products can be seen in the example of the use of plant-derived isothiocyanates in the prevention of cancer. While a number of pathways have been proposed to be involved including the inhibition of cytochrome P450 enzymes that are involved in the activation of carcinogens in vivo [12] a number of other pathways [13] have been proposed to be important. This illustrates the potential of following a particular mechanism because a suitable enzyme assay exists rather than being guided by the totality of the biological information that is available. It is unclear to what extent prevention of activation of specific procarcinogens by cytochrome P450 inhibitors has a significant role to play in the apparent anti-cancer effects of isothiocyanates. Other pathways (some involving other natural products) may also be significant and the Kantian doctrine of final causality should be remembered when trying to assign biological significance to a particular step in the development of a complex disease such as cancer. One comparatively little investigated possibility concerns the potential role of glutathione. It is known that glutathione is involved in the regulation of reactive oxygen species and thereby preventing DNA damage. As a conjugation system, glutathione is essential in the conjugation and removal of activated carcinogens. In addition, however, elevated levels of glutathione are associated with some types of cancer [14] and some of the possible mechanisms have been suggested [15]. The cellular metabolism of isothiocyanates includes the reversible reaction with thiol groups to form dithiocarbamates. These thiol groups include glutathione [16]. It is possible, therefore, that isothiocyanates may act by modulating the cellular levels of glutathione. Such a mechanism would help to explain some of the effects of isothiocyanates as impactful agents of carcinogenicity. The role of isothiocyanates as targets of biological thiol groups (including those present in enzymes) has been comparatively little studied. The discovery that isothiocyanates would react readily with hydrogen sulphide to form the corresponding dithiocarbamates under physiological conditions [17] and that this reaction is reversible in vivo, suggests that dithiocarbamates could be a physiological storage system (pro-drug) for hydrogen sulphide. The importance of hydrogen sulphide as a gasotransmitter has been noted [18] and the complex relationship between the class of natural products that have an isothiocyanato functionality and living systems will benefit from further study.

The possible role of the fungal natural product 3-nitropropanoic acid (Fig. 8.1) as an enzyme inhibitor may be of relevance to the extension of the activity of some antibiotics. This hypothesis can be postulated as follows.

Antibiotic resistance in common human pathogens such as *Streptococcus pneumoniae* is increasing. The first line therapy is the beta lactam antibiotic, amoxycillin however resistance is growing and strategies to ‘re-sensitise’ the pathogen to beta lactams have been suggested. In one example, disruption of the mevalonate pathway using CRISPR gene silencing technology led to the identification of a genetic variant in which the cell is elongated and weakened due to disruption of the supply of the important cell wall precursor, undecaprenyl phosphate. The activity of amoxycillin

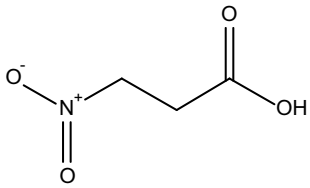


Fig. 8.1 Structure of 3-nitropropanoic acid

involves disruption of the peptidoglycan cell wall formation so it was reasoned that a strategy that involved two different ways of attacking the same bacterial target (cell wall synthesis) would be more likely to synergise and the effect of putting bacterial cell wall synthesis under pressure from two sources may 'resensitise' a resistant strain of *Streptococcus pneumoniae*. With this in mind, a known inhibitor of undecaprenyl phosphate synthesis, clomiphene was used and found to re-sensitise cells of *Streptococcus pneumoniae* to amoxycillin [19]. It was therefore demonstrated that an inhibitor of part of the mevalonate pathway involved in bacterial cell wall synthesis can work synergistically with a beta lactam antibiotic to resensitise a resistant bacterial strain. The fungal toxin 3-nitropropanoic acid (Fig. 8.1) is an inhibitor of succinate dehydrogenase—a key enzyme in the tricarboxylic acid (TCA) cycle [20]. A study of chronic application of 3-nitropropanoic acid to human cells suggested that it affected energy metabolism and the ribosome increased in size to overcome the energy depletion caused by the fungal toxin [21]. Importantly, the cells remained viable. The activity of aminoglycoside antibiotics and the development of resistance appears to involve the upregulation of proteins associated with the TCA cycle. The mechanism of action of aminoglycoside antibiotics involves inhibition of ribosomal protein synthesis and resistant strains compensate for the disturbance in their energy metabolism by increasing their TCA cycle activity [22]. If such a mechanism is operating then the combination of two approaches (chronic inhibition of succinate dehydrogenase with 3-nitropropanoic acid combined with treatment with an amino glycoside) might mitigate against the development of resistant strains of bacteria by reducing the capacity of bacterial cells to modulate their ribosomal protein synthesis. Application of 3-nitropropanoic acid will increase ribosomal protein expression and thereby prevent the inhibition of ribosomal protein synthesis that is a feature of resistance to amino glycoside antibiotics.

Alzheimer's disease is one of the most severe threats to an ageing population and will affect a significant and growing proportion of the population. Neuropathologic indicators of Alzheimer's disease include neurofibrillary tangles derived from a protein called tau which can become hyperphosphorylated, and amyloid plaques. Cholinergic deficiency (lack of acetylcholine) in the brain leading to loss of neurotransmission is a key feature of Alzheimer's disease. Acetylcholinesterase is the primary enzyme responsible for cleaving acetyl choline to give acetate and choline therefore considerable efforts have been made to mitigate the effects of cognitive decline using acetylcholinesterase inhibitors [23], including a number of natural products [24]. A number of these are derived from natural products including galantamine (Fig. 8.2).

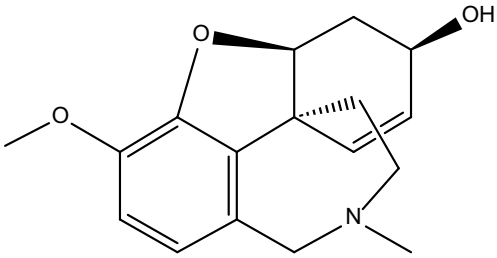


Fig. 8.2 Structure of galantamine

While cholinesterase inhibitors are designed to reduce the natural cycle of degradation of acetylcholine, two other important targets which appear to be implicated in the progression of the disease are the accumulation of beta amyloid plaques and neurofibrillary tangles of tau protein. Beta amyloid occurs naturally in the brain and may be a response to infection [25]. In addition, evidence for the substantiation of the link between inflammation and Alzheimer's disease appears to be accumulating. If neuroinflammation is induced by infective species leading to an enhanced level of expression of beta amyloid then there may be a causative link than encompasses both the clinical features of Alzheimer's disease and a putative causative agent. This may be mediated through the increased expression of IL 17. The role of IL 17 as a response to microbial infection [26] and the use of IL 17 inhibitors to treat a range of inflammatory diseases suggests there may be a link between microbial infection, IL 17 expression and Alzheimer's disease. This is not a simple causative relationship since IL 17 is also protective against pathogens [27] (as is beta amyloid) so one of the challenges is to determine the nature of the effect of IL 17 specifically in neuroinflammation. One of the diseases that IL 17 inhibitors are used to treat (and to prevent IL 17-mediated inflammation) is psoriasis. It has been noted that patients suffering from psoriasis are at an increased risk of developing Alzheimer's disease [28] so some link between different diseases that provoke IL 17 release and Alzheimer's disease may be plausible. One of the key aspects of the beta amyloid hypothesis is that it is a fragment (1–42) of the complete amyloid precursor protein (APP) that forms plaques. The enzymic cleavage of the APP by secretase enzymes to give the aggregable form [29] is therefore a target for therapeutic intervention. The possibility of using natural products directly as secretase inhibitors has been noted [30]. In addition, inhibition of phosphorylation of the tau protein [31] may also be a target. Indeed, in the case of tau phosphorylation, inhibition of the key enzyme GSK3-beta has been suggested as a strategy for preventing tau hyperphosphorylation and the subsequent formation of tau tangles [32]. This may suggest (as noted previously) that dithiocarbamates may be a useful 'pro-drug' form of hydrogen sulphide and this may have pharmacological significance in preventing the progression of Alzheimer's disease (Fig. 8.3).

While the onset of cognitive diseases like Alzheimer's disease are associated with a loss in taste sensitivity [33], it has also been suggested, as noted previously, that Alzheimer's disease has some aspects of an inflammatory disease [34].

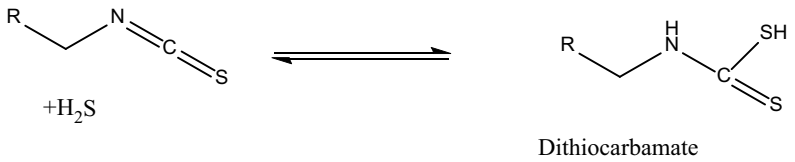


Fig. 8.3 A Dithiocarbamate as a 'pro-drug' for hydrogen sulphide

When combined with the observation that inflammation as a result of a viral infection can cause greater sensitivity to bitter compounds [35], this may suggest that people suffering from the early stages of Alzheimer's disease may avoid bitter foods (including Brassica vegetables) and this may limit their intake of isothiocyanates and hence lower their potential to produce dithiocarbamates as 'pro-drugs' for hydrogen sulphide. In such cases, they may have an increased tendency to produce hyperphosphorylated tau protein and this may increase the progression of the disease with an increased risk of production of tau tangles.

A range of other enzyme targets have been suggested to have potential in the prevention of Alzheimer's disease [35] however the establishment of a causal link between specific enzymes and aspects of the pathology of Alzheimer's disease remains challenging.

8.3 Taste Receptors and Natural Products

Taste receptors can be considered to be an example both of varying types and also the different functionalities of receptors. They are much more complex than their functional classification may imply and their role outside that simply of biosensors of taste is being increasingly realised. This also impacts upon how they interact with natural products and, as alluded to in the discussion about Alzheimer's disease, their physiological functions.

Taste receptors occur (amongst other places) in the oral epithelium and it is their presence here that has been largely responsible for their classification. The most common type of taste receptor are G-protein-coupled Receptors (GPCRs) and members of this class are responsible for detecting sweet and umami (Type I receptors) and bitter (Type II receptors). The receptors are named according to their type and then a sequential number. Sweet receptors are classified as T1R2 and T1R3. In addition to binding to (and thereby initiating a signal transduction pathway) sugars, receptors of the T1R2 and T1R3 type can also bind certain other molecules including intensely sweet peptides (e.g. aspartame). Other mechanisms for detection of sugars involving the glucose transporter type 4 (GLUT4) and sodium/glucose cotransporter 1 (SGLT1) have been suggested as an alternative to the T1R2 and T1R3 routes. Umami taste is mediated by heterodimers of T1R1-T1R3 types T1R1 receptor has no functional properties except when part of the umami detecting heterodimer. The

receptor for salt is an epithelial sodium channel. Protons can also access this channel (but more slowly than sodium ions) and this has been suggested as a mechanism whereby the addition of acid can enhance the salt sensation. The sour taste receptor (also an ion channel for protons) is found in one particular type of taste cell and responds to low pH. The sour taste receptor (OTOP1) has been identified and its mechanism of action elucidated [37].

From the standpoint of natural products the bitter receptor has been subjected to the greatest amount of study. This is due to three main factors. First of all, larger numbers of variants of bitter taste receptors occur than for any other taste receptor. Secondly a large number of natural products elicit a range of bitter taste responses and finally, the existence of significant numbers of extra-oral bitter taste receptors suggests that signal transduction to initiate aversive behaviours towards bitter compounds is not their sole physiological function.

Bitter receptors are Type 2 GPCRs. They are named, in the same way as sweet receptors, sequentially using the format T2RNN. The range of T2Rs can be specific to a particular chemical motif or may be more broadly responsive to bitter compounds. There are 25 functionally responsive T2R genes in humans. The T2Rs trigger signal transduction pathways that create the sensation of taste. There are a number of reviews and articles that explain the molecular basis of taste sensitivity and taste transduction [38]. In addition to the range of T2Rs with different specificities for different bitter compounds, there are also polymorphic differences between taste sensitivity, for example between sensitivity to phenylthiocarbamide (PTC) and propylthiouracil (PROP) [39]. Both contain the NH–C=S motif (as does a bitter component from Brassica vegetables, goitrin—Fig. 8.4) [40] however sensitivity can vary significantly.

All these bitter compounds containing the NH–C=S moiety have a feature in common, they bind to one particular bitter taste receptor—T2R38. It has been found that approximately 25% of the population are 'super bitter tasters' and have polymorphic forms of the T2R38 gene that predispose them to having significant bitter taste sensitivity while another 25% are non-tasters. The remainder (50%) have what is referred to as a normal level of taste sensitivity. The nature of this polymorphism and how it affects bitter taste sensitivity is discussed in more detail below. In addition to T2R38, there are different allelic forms of some other bitter taste receptors. For example T2R16 occurs in different forms with a variable sensitivity to several glucopyranosides [41]. In addition, the extra-oral location of several bitter taste receptors suggests that polymorphisms may be associated with a range of disease susceptibilities [42]. The possibility that taste preference as dictated by polymorphisms in T2R38 could influence dietary preference has led to suggestions that some types

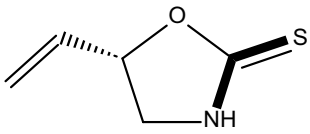


Fig. 8.4 Goitrin with NH–C=S feature highlighted in bold

of cancer that are hypothesised to be affected by the intake of Brassica vegetables (e.g. colorectal cancer) could be indirectly linked through bitter taste responses. The results from associations between taste receptor polymorphisms and cancer have been ambiguous but this may reflect the large number of confounding factors and the limited evidence for a link between diet and the prevention of cancer [43]. Polymorphisms in other taste receptors associated with sweet taste perception (T1Rs) have also been found. Perhaps one of the most interesting attributes of both T1Rs and T2Rs (as noted above) is that they occur in extra-oral sites in the human body and have been associated with a number of physiological effects [44].

T1R (sweet) taste receptors have been found at a range of sites including the nasal epithelium, respiratory system and pancreatic islet cells. In addition T1Rs are found in the gut on enteroendocrine cells where they are thought to be implicated in nutrient sensing and they may be important in the regulation of glucose balance. It has been suggested that T1Rs are also implicated in the control of GLP-1 and other gut hormone production processes [45]. In addition sweet taste is correlated with plasma leptin levels in humans [46]. The significant occurrence of sweet taste pathways and T1Rs in various extra-oral organs and their link to gut and other hormone production suggests that modulation of a range of health-related outcomes through modulation of T1R agonists and antagonists might be achievable. Nutrient homeostasis is essential for all living systems and nutrient sensing is an important component of the whole nutrient balance system [47] as well as being important in many metabolic diseases. Understanding how to manipulate nutrient homeostasis could be important in the development of treatments for metabolic diseases and in their control.

There has been significant progress in the study of other extra-oral bitter receptors [48]. These have been found at a number of sites including, as noted above, the nasal cavity, the GI tract and elsewhere. Not only does this suggest that classification of bitter receptors by primary perceived function (taste) may not be sufficient, it also suggests alternative routes to elicit physiological effects and might imply that some natural products which have been thought of as having simply a chemosensory effect, may have a range of potential physiological purposes. A number of natural products which are physiologically active were previously found to be bitter including many alkaloids. It is unclear if some of the physiological effects ascribed to these compounds is a result of their binding to extra-oral bitter receptors which have an organ-specific function. Polymorphisms in the T2R38 gene were previously mentioned and these have been studied on a molecular level. The wild type (normal) receptor has three important functional amino acids that are changed in the polymorphic forms. These are Pro49, Ala262 and Val 296—as a result this polymorphism is referred to by the abbreviation PAV. The non-functional allele has Ala49, Val262 and Isoleucine 296 and is, therefore referred to as AVI. If an individual is homozygous (PAV/PAV) then they are very sensitive to bitter compounds with the NH–C=S group and are referred to as ‘super bitter tasters’. Heterozygous individuals (AVI/PAV or PAV/AVI) have a ‘normal ‘sensory experience while those who are homozygous with the inactive allele (AVI/AVI) are non-tasters. Other polymorphisms have been noted but in many cases the physiological consequences in terms of taste perception are unknown. The sites of location of bitter taste receptors may provide some

clue as to their alternative physiological functions. For example, it has been found that T2R38 receptors are a primary target for acyl homoserine lactone (AHL) type quorum sensing molecules which are produced by Gram negative bacteria. Binding of AHLs to T2R38 receptors in the human sinonasal epithelium cells leads to NO production and enhanced clearance of inhaled pathogens [49] and to inhibition of biofilm formation [50]. In addition to possibly having a role in controlling airway infections, bitter taste receptors have also been implicated in the control of inflammation. As noted previously, a range of neurological disorders (including Alzheimer's and Parkinson's diseases) are considered to be at least partially triggered by inflammation. The presence of bitter taste receptors in the brain and their potential role in mediating inflammation and in minimising microbial infection [51] suggests that the assay of natural products as bitter receptor agonists at extra-oral sites might provide new leads for pharmacological investigation. As noted previously, isothiocyanates from Brassica vegetables may, through the formation of dithiocarbamates, act as precursors for the physiological storage and presentation of hydrogen sulphide. By binding to T2R38, it is possible that another compound from Brassica vegetables, goitrin, may potentiate the production of NO. This would implicate Brassica vegetables in the production of two of the most significant gasotransmitters. While the extra-oral importance of several bitter receptors has been noted, some members of the class appear to be more significant than others. Implicit in this is the concept that the role of bitter receptors extends beyond the sensory area and the determination of food preference. For example, it is becoming apparent that the T2R14 receptor has a significant extra-oral presence. It is expressed in human muscle smooth airway cells and has been considered to be a potential target for bronchodilators [52]. In addition, the T2R14 receptor is subject to agonist mediated down regulation and mutant studies also revealed some details of the binding site. The receptor has been linked to a number of other physiological outcomes including the innate immune response and T2R14 is accepted as a drug target [53]. While associations between the occurrence of bitter (and other) taste receptors (and polymorphisms) and physiological states have been noted, the mechanistic details in many cases, remain to be elucidated. The binding of structurally diverse compounds (including natural products) to the T2R14 receptor has been studied [54] and these studies suggest that a single flexible binding pocket (and not multiple binding pockets) is implicated. In addition to specific physiological effects, bitter taste receptors have also been associated with generic longevity. In particular the PAV/PAV variant of T2R38 has been implicated in this effect [55]. While these results have been questioned in a later meta-analysis, an association between a particular genotype of T2R16 and the longevity of populations in southern Italy was observed [56]. While observational studies (even if associated with genetic data) should be treated with caution it has been suggested that the observed effects may be a result of more than simple modulation of taste preferences and phytochemical intake but may be associated with the influence of extra-oral bitter taste receptors on phytochemical absorption, processing and metabolism [57]. The realisation that a number of chronic, non-communicable diseases have an inflammatory aspect [58] and the link between taste receptors and anti-inflammatory effects [59] suggests that reduction in inflammatory response may

be a generic explanation for the role of bitter receptors and the extension of longevity. This may be further evidenced by the role of inflammation in a number of chronic disease conditions.

While the primary anticipated role of taste receptors appears to be associated with the sensory detection of chemical stimuli and associated food preference (and avoidance), the existence of multiple polymorphisms and the widespread occurrence of taste receptors in extra-oral tissues suggest that they may have a range of other properties including resistance to microbial infection and prevention of chronic inflammation. Taste receptors are implicated in a number of responses and the taste-receptor-mediated response to natural products may be of significantly greater importance than realised hitherto. This is reflected both in the extra-oral occurrence of taste receptors and also in the number of different receptors (with different specificities) that are found. For example, if the sole purpose of bitter receptors was as a toxin warning and avoidance system, then the most effective solution would be a single receptor with a flexible (or even multiple) binding pocket that can accommodate the maximum structural diversity. The existence of multiple bitter receptors suggests that there are multiple biological processes that are affected by different agonists and antagonists and are dependent upon the site of action of the receptor. This illustrates the importance of the action of natural products as both agonists and antagonists.

8.4 Natural Products as Agonists and Antagonists

One of the key properties of bioactive molecules in interactions with receptors is their property of being agonists or antagonists. The concept of agonists and antagonists is one which was known to philosophers before the terms were adopted by chemists to describe the interactions of natural products (or any other ligand) with a receptor. The concept of agonism was developed to describe how some types of conflict are beneficial. The physiological (pharmaceutical) description of an agonist is an exogenous compound that binds to a receptor in place of the natural ligand. In so doing, an agonist can modulate, enhance or balance the physiological response and will, therefore, have a positive effect. This fits with the philosophical definition—something that competes (conflicts) with the natural ligand and produces an overall beneficial effect. Agonists can be complete, partial or inverse and each has slightly different properties and different physiological effects. Compounds which are agonists (or partial agonists) for one receptor may also be antagonists for other members of the same class of receptor [60].

Similarly an antagonist is something that lies between a situation and a preferred outcome. It has connotations of social negativity. In literature an antagonist is someone (or thing) that lies between (and therefore blocks) a protagonist and their goal. In pharmacology and medicinal chemistry an antagonist is a compound that binds to a receptor and thereby subsequently blocks any other molecule from binding but, of itself, does not enable the functionality of the receptor. There are different types of antagonist—those which compete directly for the receptor active site and,

like agonists, bind reversibly to the receptor, those which bind at an alternative (allosteric) site and induce a change in the receptor conformation such that it can no longer bind an agonist and is thereby rendered inactive and those which bind irreversibly to the receptor active site and block it permanently so that it is only by creating another receptor molecule that functionality can be restored. For competitive antagonists, increasing the levels of the agonist present may restore at least partial activity while for non-competitive (allosteric) or irreversible antagonists, increasing the concentration of the agonist will not restore the physiological activity of the receptor. Within the language of physiology and pharmacology there is a further definition that is sometimes used to describe another type of component—an inhibitor. The distinction between antagonist and inhibitor is sometimes difficult to envision. Specifically, an antagonist prevents the binding of an agonist to a receptor. In this way it prevents the receptor from responding to an agonist. This implies that an antagonist can only operate in the presence of an agonist. In the absence of an agonist (natural or synthetic), the binding of an antagonist will have no direct physiological effect because, in the absence of an agonist, the receptor does not function anyway. In pharmacological terms, an inhibitor is something that prevents (or reduces) the functioning of a receptor (often an enzyme) and does not completely block a pathway but rather reduces its effectiveness.

Agonism and antagonism is often the basis for the biological activity of natural products and this can provide useful tools for the isolation and separation of bioactive natural products. The development of natural product medications from their roots in folk medicine are sometimes entangled with belief systems as much as with scientific proof. The trend—perhaps emphasised by Ehrlich's concept of the 'magic bullet'—is towards reductionist chemistry. Compounds are isolated on the basis of their biological activity. While whole organism assays (particularly using rodents) are still commonplace these can be both misleading (giving rise to false positive and false negative results) and, for the screening of large numbers of candidate compounds or extracts, expensive and both labour and time expensive as well as being costly. In addition animal model systems are precisely that—models based upon the physiological characteristics of animals. As such, their applicability to humans (or, indeed, to other animals remains questionable and should not be taken as definitive. For example, an extensive review of three different animal models (mice, rats and rabbits) in the assay of statins [61] found that there were significant differences in the effect level between the different models with rats having the lowest effect size. One approach to overcoming the inter-species problems associated with animal models is the use of 'humanised' animal models. One of the key differences between humans and animal models is in the metabolism of drugs. This affects the amount of the bioactive compound that reaches the active site and the rate of clearance of the compound from the body. The differences in metabolism between humans and mice is, in part, a function of differences in the cytochrome P450 enzymes present in the two species [62]. More generally, transgenic humanised rodent models have therefore been developed to provide a better and more accurate picture of how drugs (including natural products) might be active in humans before committing to expensive clinical trials [63] using volunteers.

An alternative approach to the use of whole animal models is to apply reductionist thinking to complex biological systems and to define the most significant biological interaction in terms of the binding of a natural product to a human receptor which has specific properties and is the cause of some specific effects. Receptor binding (and either measuring the binding per se or the physiological consequences of binding (signal transduction) can be considered a form of biomarker for physiological activity. Receptor-based assays have a long history and have been at the forefront of pharmacological chemistry for a number of years. While the information on absorption, distribution, metabolism and excretion (ADME) is limited in such systems, coincubation with cytochrome P450 enzymes (microsomal preparations) have been used in conjunction with receptor binding assays to further refine the basic assay [64]. Even the simplest of binding assays can be complicated by confounders and these must be taken into account in the design of a receptor-based assay.

The purpose of any bioassay can be twofold. From a biological standpoint, the assay should be capable of discriminating between active and non-active compounds and additionally between, for example, agonists and antagonists. In order to be scientifically and pharmacologically useful. An assay should also identify chemical uniqueness. Spending time and effort 'rediscovering' a known compound is of limited value. This led to the concept of 'dereplication' which can be carried out using a combination of high performance liquid chromatography (hplc) and a receptor binding assay [65]. Another widely adopted approach is 'affinity-based screening'. In essence this approach, rather than using functionality as a defining characteristic, simply measures the binding of a natural product (or mixture of natural products) to a biological macromolecular target such as a receptor, an antibody or an enzyme. A range of these techniques are available including surface plasmon resonance, biological chromatography, affinity chromatography and ligand fishing [66]. Such approaches rely on considerable technical advances and on knowledge of how different biological macromolecules function. An example of a target-oriented discovery programme is the group of histamine receptors whose biological activities are multi-faceted and can therefore be used to differentiate different biological activities of ligands. Histamine was the first inflammatory biogenic amine to be discovered. It is defined by a capability to bind to histamine receptors. There are 4 different types of histamine receptor (chronologically named H1-H4) which occur in different organs and have different biological activities. The H1 receptor mediates allergic inflammation while the H2 receptor, which occurs in the GI tract, affects acid secretion. Both of these are described as 'low-affinity' receptors. The H3 receptor is implicated in neuronal function while H4 is mediates inflammation. Both H3 and H4 are high affinity receptors for histamine so the biological response to histamine is a function of the type of receptor and the amount of histamine present. Much of the work carried out on histamine receptors has been focussed on synthesising and screening for antagonists that will limit their inflammatory-related responses. Interestingly many of the studies of histamine in the early twentieth century were carried out using whole animal models and apparently conflicting results were obtained. While the H1 receptor was not characterised until the 1960s, the development of first generation histamine antagonists (anti-histamines) was carried out with a focus on

prevention of conditions such as allergic rhinitis. While Mast cells had been characterised by Ehrlich in 1879, they were not linked to anaphylactic shock until the 1940s [67]. A significant breakthrough arose from the linkage between acid secretion in the gut and histamine mediated (presumably, although not proven until later) by a gut histamine receptor. By carrying out a series of bioassays using known H1 receptor antagonists and some novel thiourea compounds, it was recognised that the histamine receptor in the gut was of a different type to that implicated in the H1 receptor response. A novel imidazole derivative called cimetidine was created and this became the first 'blockbuster' drug. While H2 antagonists have been overtaken by proton pump antagonists for the treatment of acid secretion in the gut, this remains an interesting area of research for the development of H3 and H4 antagonists. Perhaps the most interesting aspect of the studies on the histamine receptors was that, even before the receptors themselves had been fully characterised and cloned, imaginative chemistry allowed some features of the receptors related to antagonist activity to be ascertained. The discovery of the detailed structure of histamine receptors (particularly H3 and H4) may enable new structural characteristics of potential agonists and antagonists to be determined and mapped onto libraries of natural products [68]. A range of natural histamine antagonists have been proposed to be effective including stinging nettle (*Urtica dioica*) however detailed pharmacological investigations are rare and a range of physiological effects are associated with allergic responses including down regulation of expression of the histamine receptor itself and it is unclear if specific pathways are being affected by natural products [69]. Nonetheless, histamine receptors have been extensively studied and characterised and there may be new chemical 'white space' worth investigating.

The use of computational methods to investigate novel compounds that may bind to histamine (and other) receptors has shown a significant amount of potential however there are some important caveats. Molecular docking predicts the binding affinity of chemical ligands to receptors [70] however this may not translate directly into biological activity. Having acknowledged this however, techniques are becoming increasingly sophisticated [71] and this, combined with the enhanced cost-effectiveness of in silico approaches, will increase the applicability. In addition to the functional binding of a compound to a receptor, biological activity requires the migration of an active compound to the receptor. A number of guidance rules are used to determine the probable bioavailability of a given molecule. A key signifier is the logarithm of the partition coefficient (logP) which measures the extent to which a molecule partitions between octanol and water. Other measures that can help in the calculation of the hydrophobicity/lipophilicity of a molecule includes the number of hydrogen donors, the dynamic or static total polar surface area and the ionisation constant. A number of methods exist to enable the measurement of key physicochemical properties [71]. These techniques can be of value in pre-screening a molecule and can help to eliminate some candidates while also indicating molecular adjustments that may improve overall bioavailability (for example, the addition or removal of hydrophobic residues).

8.5 Summary and Conclusions

Receptors are a key aspect of how living systems interact with natural products. They are also a reductionist window into complex living systems. While much of the research effort has focussed on proteinaceous receptors, they can also be composed of other biological macromolecules such as nucleic acids. The original idea expressed by Ehrlich and others was of a 'receptive side chain' which was different in toxic species compared to the host. Ehrlich was among the first to develop a link between molecules and the cells upon which they had an effect. This was not a refutation of reductionism but rather a recognition that what happened on a molecular scale (and which could not be directly observed) could be deduced as a result of the visible consequences of that action. Different molecules had different effects on the biological system as a whole.

Although the work initially carried out by Ehrlich and later developed by him and others had profound implications for the understanding of how biological systems worked, his approach was underpinned by an understanding of chemistry. His ideas about the responsiveness of cells to molecules, combined with the 'Magic Bullet' idea of a drug that would affect only a pathogen without impairing the host cell was key to the development of the first effective treatment for syphilis (Salvarsan, arsphenamine) and to the progression of early medicinal chemistry. The beginnings of theories about receptors was derived from the idea that certain types of cells had side chains attached to them that allowed them to bind to specific toxic agents such as diphtheria or other microorganisms. These 'antitoxins' could stimulate the production of more antitoxins which would bind specifically to any other 'toxins' of that type and prevent them from exerting a toxic effect on the cell. This is, in essence (although without a deep understanding of the cellular machinery involved) the concept of production of antibodies by the immune system. The key aspects of the theory were that the response was specific and occurred subsequent to a challenge. The next step in receptor theory was to take the concept of specificity and to extend it to the surface of cells. This led Ehrlich, together with John Newport Langley, to propose the idea of cell surface 'chemoreceptors' that could bind small bioactive compounds and these could then exhibit an effect on the cell as a whole. These ideas underpinned the concept of a 'Magic Bullet'—a compound that demonstrated specific toxicity against a disease-causative agent but which had little toxic effect on the host. The concept of specific ligand-receptor interactions was an essential aspect of drug screening and drug synthesis developed and carried out by Ehrlich.

While a correlation between the binding of a compound and a biological endpoint is conceptually simple—even if the cellular machinery is not understood—the link between specific compounds and sensations such as taste or effects on mood and behaviour may not be as intuitively obvious. The concept of a receptor and of compounds that activate (agonist) or block (antagonist) the effect on the cell to which it is attached was described by Newport Langley at the beginning of the twentieth century. While conceptual evidence for receptors continued to increase, these were based on observing effects of treatments on whole organisms or cells and not

binding of certain molecular entities to individual receptors. The standardisation of the effect of morphine required its isolation by Serturner in the early part of the nineteenth century and the administration of known amounts to 'willing' volunteers, in the twentieth century, however finding a way of predicting the biological effect of a given dose of a substance without a trial and error process remained an elusive goal. The concept of different ligands having different avidities for a specific receptor was developed and alongside this, it was realised that binding and efficacy were linked but not identical concepts [73]. The advent of radiolabelling allowed the better visualisation of receptors (where they were localised in tissue) and different aspects of binding. In the 1970s, some receptors were solubilised and by the 80s and 90s they were sequenced and cloned. The first cloned human receptors were the glucocorticoid receptor [74] and the human estrogen receptor [75]. These types of studies facilitated the elucidation of the structural and functional characteristics of receptors and features such as the ligand binding region could be located and characterised.

The philosophy of the development of studies on receptors is predicated, in the first place, on the inference of receptor properties based on they interact with ligands and how the chemistry of ligands affects the biological functioning of the organism. At a molecular level, the characteristics of a ligand control how it interacts with a receptor and, on a systems level, that interaction is transducted into biological effects in a complex living system. The study of receptors and their interaction with natural and unnatural products is an example of the transactions between chemistry and biology. This is built, as indicated above, on inference. While molecular methods and techniques yield more and more evidence of how molecular interactions function, extrapolating that to a physiological endpoint remains, in many cases, problematic. It is philosophically challenging to break down a complex biological system into its component parts and to deduce from these how the complex system works and, more particularly, how an intervention at the simplest (receptor-ligand) level may perturb the whole biological system. This further suggests that the maintenance of homeostasis is not a valid objective for biological systems. They are in a continual state of flux—just as receptor-ligand interactions are not static but are evolving and transient. Looking for a snapshot of what a given compound is doing at a specific instance through, for example x-ray crystallography of its binding to a target receptor, will fail to transmit the kinetic nature of that interaction and the consequences of it on the biological system as a whole. Real time chemical:biological interactions remains an important goal for natural product chemists and those they collaborate with.

References

1. Brookes JC (2017) Quantum effects in biology: golden rule in enzymes, olfaction, photosynthesis and magnetodetection. Proc Math Phys Eng Sci 473(2201):20160822. https://doi.org/10.1098/rspa.2016.0822. Epub 2017 May 31. PMID: 28588400; PMCID: PMC5454345
2. Ben Abu N, Mason PE, Klein H et al (2021) Sweet taste of heavy water. Commun Biol 4:440. https://doi.org/10.1038/s42003-021-01964-y

3. Arodola OA, Soliman ME (2017) Quantum mechanics implementation in drug-design workflows: does it really help? Drug Des Devel Ther 11:2551–2564. https://doi.org/10.2147/DDDT.S126344. Erratum in: Drug Des Devel Ther. 2017 Nov 08;11:3205. https://doi.org/10.2147/DDDT.S155133. PMID: 28919707; PMCID: PMC5587087
4. Kulkarni PU, Shah H, Vyas VK (2022) Hybrid quantum mechanics/molecular mechanics (QM/MM) simulation: a tool for structure-based drug design and discovery. Mini Rev Med Chem 22(8):1096–1107. https://doi.org/10.2174/1389557521666211007115250. PMID: 34620049
5. Dai G (2023) Signaling by ion channels: pathways, dynamics and channelopathies. Mo Med 120(5):367–373. PMID: 37841571; PMCID: PMC10569386
6. Syrovatkina V, Alegre KO, Dey R, Huang XY (2016) Regulation, signaling, and physiological functions of G-proteins. J Mol Biol 428(19):3850–3868. https://doi.org/10.1016/j.jmb.2016.08.002. Epub 2016 Aug 8. PMID: 27515397; PMCID: PMC5023507
7. De Smet F, Christopoulos A, Carmeliet P (2014) Allosteric targeting of receptor tyrosine kinases. Nat Biotechnol 32:1113–1120. https://doi.org/10.1038/nbt.3028
8. Copeland RA (2013) Why enzymes as drug targets? Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists, 2nd edn. Wiley, Hoboken, pp 1–23
9. Paoli P (2021) Enzymatic inhibitors from natural sources: a huge collection of new potential drugs. Biomolecules 11(2):133. https://doi.org/10.3390/biom11020133. PMID: 33498517; PMCID: PMC7909571
10. Heidary Moghaddam R, Samimi Z, Asgary S, Mohammadi P, Hozeifi S, Hoseinzadeh-Chahkandak F, Xu S, Farzaei MH (2022) Natural AMPK activators in cardiovascular disease prevention. Front Pharmacol 12:738420. https://doi.org/10.3389/fphar.2021.738420. PMID: 35046800; PMCID: PMC8762275
11. Lankatillake C, Luo S, Flavel M et al (2021) Screening natural product extracts for potential enzyme inhibitors: protocols, and the standardisation of the usage of blanks in α-amylase, α-glucosidase and lipase assays. Plant Methods 17:3. https://doi.org/10.1186/s13007-020-00702-5
12. Fimognari C, Lenzi M, Hrelia P (2008) Interaction of the isothiocyanate sulforaphane with drug disposition and metabolism: pharmacological and toxicological implications. Curr Drug Metab 9:668–678
13. Asif Ali M, Khan N, Kaleem N, Ahmad W, Alharethi SH, Alharbi B, Alhassan HH, Al-Enazi MM, Razis AFA, Modu B, Calina D, Sharifi-Rad J (2023) Anticancer properties of sulforaphane: current insights at the molecular level. Front Oncol 13:1168321. https://doi.org/10.3389/fonc.2023.1168321. PMID: 37397365; PMCID: PMC10313060
14. Harris IS, Treloar AE, Inoue S, Sasaki M, Gorrini C, Lee KC, Yung KY, Brenner D, Knobbe-Thomsen CB, Cox MA et al (2015) Glutathione and thioredoxin antioxidant pathways synergize to drive cancer initiation and progression. Cancer Cell 27:211–222. https://doi.org/10.1016/j.ccell.2014.11.019
15. Kennedy L, Sandhu JK, Harper ME, Cuperlovic-Culf M (2020) Role of glutathione in cancer: from mechanisms to therapies. Biomolecules 10(10):1429. https://doi.org/10.3390/biom10101429. PMID: 33050144; PMCID: PMC7600400
16. Wu X, Zhou Qh Xu K (2009) Are isothiocyanates potential anti-cancer drugs? Acta Pharmacol Sin 30:501–512. https://doi.org/10.1038/aps.2009.50
17. Tian M, Dodds MJC, Hanley AB (2013) Allyl isothiocyanate is effective in reducing volatile sulphur compounds responsible for intrinsic oral malododor. J Breath Res 7:026001
18. Cirino G, Szabo C, Papapetropoulos A (2023) Physiological roles of hydrogen sulfide in mammalian cells, tissues, and organs. Physiol Rev 103(1):31–276. https://doi.org/10.1152/physrev.00028.2021. Epub 2022 Apr 18. PMID: 35435014
19. Dewachter L, Dénéréaz J, Liu X, de Bakker V, Costa C, Baldry M, Sirard JC, Veening JW (2022) Amoxicillin-resistant Streptococcus pneumoniae can be resensitized by targeting the mevalonate pathway as indicated by sCRilecs-seq. Elife 11:e75607. https://doi.org/10.7554/eLife.75607. PMID: 35748540; PMCID: PMC9363119

20. Bendiksen Skogvold H, Yazdani M, Sandås EM, Østeby Vassli A, Kristensen E, Haarr D, Rootwelt H, Elgstøen KBP (2022) A pioneer study on human 3-nitropropionic acid intoxication: contributions from metabolomics. J Appl Toxicol 42(5):818–829. https://doi.org/10.1002/jat.4259. Epub 2021 Nov 1. PMID: 34725838
21. Hanley AB, McBride J, Oehlschlager SL, Opara EI (1999) The use of a flow cell bioreactor as a chronic toxicity model system. Toxicol In Vitro 13:847
22. Shiraliyev R, Orman MA (2024) Metabolic disruption impairs ribosomal protein levels, resulting in enhanced aminoglycoside tolerance. eLife 13:RP94903. https://doi.org/10.7554/eLife.94903.3
23. Sharma K (2019) Cholinesterase inhibitors as Alzheimer's therapeutics Mol Med Rep 20(2):1479–1487. https://doi.org/10.3892/mmr.2019.10374. Epub 2019 Jun 11. PMID: 31257471; PMCID: PMC6625431
24. Gajendra K, Pratap GK, Poornima D, Shantaram M, Ranjita G (2024) Natural acetylcholinesterase inhibitors: a multi-targeted therapeutic potential in Alzheimer's disease. Eur J Med Chem Rep 100154
25. Vojtechova I, Machacek T, Kristofikova Z, Stuchlik A, Petrasek T (2022) Infectious origin of Alzheimer's disease: amyloid beta as a component of brain antimicrobial immunity. PLoS Pathog 18(11):e1010929. https://doi.org/10.1371/journal.ppat.1010929. PMID: 36395147; PMCID: PMC9671327
26. Cooper AM (2009) IL-17 and anti-bacterial immunity: protection versus tissue damage. Eur J Immunol 39(3):649–652. https://doi.org/10.1002/eji.200839090. PMID: 19283706; PMCID: PMC3931122
27. Ma WT, Yao XT, Peng Q, Chen DK (2019) The protective and pathogenic roles of IL-17 in viral infections: friend or foe? Open Biol 9(7):190109. https://doi.org/10.1098/rsob.190109. Epub 2019 Jul 24. PMID: 31337278; PMCID: PMC6685926
28. Yang Q, Wang J, Mi N, Zou Y (2023) Literature Overview of the Relation Between Psoriasis and Alzheimer. Neuropsychiatr Dis Treat 19:461–468. https://doi.org/10.2147/NDT.S403854. Erratum in: Neuropsychiatr Dis Treat. 2023 Nov 21;19:2519–2520. https://doi.org/10.2147/NDT.S449767. PMID: 36879948; PMCID: PMC9985423
29. Karran E, Mercken M, De Strooper B (2011) The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. Nat Rev Drug Discov 10:698–712. https://doi.org/10.1038/nrd3505
30. Lee JH, Ahn NH, Choi SB, Kwon Y, Yang SH (2021) Natural products targeting amyloid beta in Alzheimer's disease. Int J Mol Sci 22(5):2341. https://doi.org/10.3390/ijms22052341. PMID: 33652858; PMCID: PMC7956407
31. Iqbal K, Gong CX, Liu F (2014) Microtubule-associated protein tau as a therapeutic target in Alzheimer's disease. Expert Opin Ther Targets 18:307–318. https://doi.org/10.1517/14728222.2014.870156
32. Giovinazzo D, Bursac B, Sbodio JI, Nalluru S, Vignane T, Snowman AM, Albacarys LM, Sedlak TW, Torregrossa R, Whiteman M, Filipovic MR, Snyder SH, Paul BD (2021) Hydrogen sulfide is neuroprotective in Alzheimer's disease by sulfhydrating GSK3β and inhibiting Tau hyperphosphorylation. Proc Natl Acad Sci U S A 118(4):e2017225118. https://doi.org/10.1073/pnas.2017225118. PMID: 33431651; PMCID: PMC7848711
33. Kouzuki M, Ichikawa J, Shirasagi D et al (2020) Detection and recognition thresholds for five basic tastes in patients with mild cognitive impairment and Alzheimer's disease dementia. BMC Neurol 20:110. https://doi.org/10.1186/s12883-020-01691-7
34. Kinney JW, Bemiller SM, Murtishaw AS, Leisgang AM, Salazar AM, Lamb BT (2018) Inflammation as a central mechanism in Alzheimer's disease. Alzheimers Dement (N Y) 4:575–590. https://doi.org/10.1016/j.trci.2018.06.014. PMID: 30406177; PMCID: PMC6214864
35. Nguyen H, Albayay J, Höchenberger R, Bhutani S, Boesveldt S, Busch NA, Croijmans I, Cooper KW, de Groot JHB, Farruggia MC, Fjaeldstad AW, Hayes JE, Hummel T, Joseph PV, Laktionova TK, Thomas-Danguin T, Veldhuizen MG, Voznessenskaya VV, Parma V, Yanina Pepino M, Ohla K (2023) Covid-19 affects taste independent of taste–smell confusions: results from a combined chemosensory home test and online survey from a large global cohort. Chem Sens 48:bjad020. https://doi.org/10.1093/chemse/bjad020

36. Alam J, Sharma L (2019) Potential enzymatic targets in Alzheimer's: a comprehensive review. Curr Drug Targets 20(3):316–339. https://doi.org/10.2174/1389450119666180820104723. PMID: 30124150
37. Liman ER, Kinnamon SC (2021) Sour taste: receptors, cells and circuits. Curr Opin Physiol 20:8–15. https://doi.org/10.1016/j.cophys.2020.12.006. Epub 2021 Jan 13. PMID: 33709046; PMCID: PMC7943026
38. Wooding SP, Ramirez VA, Behrens M (2021) Bitter taste receptors: genes, evolution and health. Evol, Med, Public Health 9(1):431–447. https://doi.org/10.1093/emph/eoab031
39. Bufe B et al (2005) The molecular basis of individual differences in phenylthiocarbamide and propylthiouracil bitterness perception. Curr Biol 15:322–327
40. Wooding S, Gunn H, Ramos P et al (2010) Genetics and bitter taste responses to goitrin, a plant toxin found in vegetables. Chem Senses 35:685–692
41. Campbell MC et al (2014) Limited evidence for adaptive evolution and functional effect of allelic variation at rs702424 in the promoter of the TAS2R16 bitter taste receptor gene in Africa. J Hum Genet 59:349–352
42. Lu P, Zhang CH, Lifshitz LM, ZhuGe R (2017) Extraoral bitter taste receptors in health and disease. J Gen Physiol 149(2):181–197. https://doi.org/10.1085/jgp.201611637. Epub 2017 Jan 4. PMID: 28053191; PMCID: PMC5299619
43. Key TJ, Bradbury KE, Perez-Cornago A, Sinha R, Tsilidis KK, Tsugane S (2020) Diet, nutrition, and cancer risk: what do we know and what is the way forward? BMJ 368:m511. https://doi.org/10.1136/bmj.m511. Erratum in: BMJ. 2020 Mar 11;368:m996. https://doi.org/10.1136/bmj.m996. PMID: 32139373; PMCID: PMC7190379
44. Behrens M, Lang T (2022) Extra-oral taste receptors-function, disease, and perspectives. Front Nutr 9:881177. https://doi.org/10.3389/fnut.2022.881177. PMID: 35445064; PMCID: PMC9014832
45. Gerspach AC, Steinert RE, Schönenberger L, Graber-Maier A, Beglinger C (2011) The role of the gut sweet taste receptor in regulating GLP-1, PYY, and CCK release in humans. Am J Physiol-Endocrinol Metab 301:E317–E325. https://doi.org/10.1152/ajpendo.00077.2011
46. Nakamura Y, Sanematsu K, Ohta R, Shirosaki S, Koyano K, Nonaka K, Shigemura N, Ninomiya Y (2008) Diurnal variation of human sweet taste recognition thresholds is correlated with plasma leptin levels. Diabetes 57:2661–2665. https://doi.org/10.2337/db07-1103
47. Efeyan A, Comb WC, Sabatini DM (2015) Nutrient-sensing mechanisms and pathways. Nature 517(7534):302–310. https://doi.org/10.1038/nature14190. PMID: 25592535; PMCID: PMC4313349
48. Tuzim K, Korolczuk A (2021) An update on extra-oral bitter taste receptors. J Transl Med 19(1):440. https://doi.org/10.1186/s12967-021-03067-y. Erratum in: J Transl Med. 2021 Nov 26;19(1):478. https://doi.org/10.1186/s12967-021-03137-1. PMID: 34674725; PMCID: PMC8529754
49. Lee RJ, Cohen NA (2015) Role of the bitter taste receptor T2R38 in upper respiratory infection and chronic rhinosinusitis. Curr Opin Allergy Clin Immunol 15(1):14–20. https://doi.org/10.1097/ACI.0000000000000120. PMID: 25304231; PMCID: PMC5902169
50. Gaida MM, Dapunt U, Hänsch GM (2016) Sensing developing biofilms: the bitter receptor T2R38 on myeloid cells. Pathog Dis 74(3):ftw004. https://doi.org/10.1093/femspd/ftw004. Epub 2016 Jan 17. PMID: 26782143; PMCID: PMC4748947.
51. Welcome MO, Mastorakis NE (2021) The taste of neuroinflammation: molecular mechanisms linking taste sensing to neuroinflammatory responses. Pharmacol Res 167:105557. https://doi.org/10.1016/j.phrs.2021.105557. Epub 2021 Mar 15. PMID: 33737243
52. Woo JA, Castaño M, Goss A, Kim D, Lewandowski EM, Chen Y, Liggett SB (2019) Differential long-term regulation of TAS2R14 by structurally distinct agonists. FASEB J 33(11):12213–12225. https://doi.org/10.1096/fj.201802627RR. Epub 2019 Aug 20. PMID: 31430434; PMCID: PMC6902689.
53. Waterloo LAW, Löber S, Gmeiner P (2022) The bitter taste receptor TAS2R14 as a drug target. Chimia (Aarau) 76(5):418–424. https://doi.org/10.2533/chimia.2022.418. PMID: 38069713

54. Tokmakova A, Kim D, Guthrie B, Kim SK, Goddard WA 3rd, Liggett SB (2023) Predicted structure and cell signaling of TAS2R14 reveal receptor hyper-flexibility for detecting diverse bitter tastes. iScience 26(4):106422. https://doi.org/10.1016/j.isci.2023.106422. PMID: 37096045; PMCID: PMC10121769
55. Melis M, Errigo A, Crnjar R, Pes GM, Tomassini Barbarossa I (2019) TAS2R38 bitter taste receptor and attainment of exceptional longevity. Sci Rep 9(1):18047. https://doi.org/10.1038/s41598-019-54604-1. PMID: 31792278; PMCID: PMC6889489
56. Di Bona D, Malovini A, Accardi G, Aiello A, Candore G, Ferrario A, Ligotti ME, Maciag A, Puca AA, Caruso C (2021) Taste receptor polymorphisms and longevity: a systematic review and meta-analysis. Aging Clin Exp Res 33(9):2369–2377. https://doi.org/10.1007/s40520-020-01745-3. Epub 2020 Nov 10. PMID: 33170488; PMCID: PMC8429150
57. Campa D, De Rango F, Carrai M, Crocco P, Montesanto A, Canzian F, Rose G, Rizzato C, Passarino G, Barale R (2012) Bitter taste receptor polymorphisms and human aging. PLoS One 7(11):e45232. https://doi.org/10.1371/journal.pone.0045232. Epub 2012 Nov 2. PMID: 23133589; PMCID: PMC3487725
58. Furman D, Campisi J, Verdin E et al (2019) Chronic inflammation in the etiology of disease across the life span. Nat Med 25:1822–1832. https://doi.org/10.1038/s41591-019-0675-0
59. Nayak AP, Villalba D, Deshpande DA (2019) Bitter taste receptors: an answer to comprehensive asthma control? Curr Allergy Asthma Rep 19(10):48. https://doi.org/10.1007/s11882-019-0876-0. PMID: 31486942; PMCID: PMC6765386
60. Tuplin EW, Holahan MR (2017) Aripiprazole, a drug that displays partial agonism and functional selectivity. Curr Neuropharmacol 15(8):1192–1207. https://doi.org/10.2174/1570159X15666170413115754. PMID: 28412910; PMCID: PMC5725548
61. Pecoraro V, Moja L, Dall'Olmo L, Cappellini G, Garattini S (2014) Most appropriate animal models to study the efficacy of statins: a systematic review. Eur J Clin Invest 44(9):848–871. https://doi.org/10.1111/eci.12304. PMID: 25066257
62. Le Bras A (2024) Humanized mouse models of drug metabolism. Lab Anim 53:87. https://doi.org/10.1038/s41684-024-01357-8
63. Henderson CJ, Kapelyukh Y, Scheer N, Rode A, McLaren AW, MacLeod AK, Lin D, Wright J, Stanley LA, Wolf CR (2019) An extensively humanized mouse model to predict pathways of drug disposition and drug/drug interactions, and to facilitate design of clinical trials. Drug Metab Dispos 47(6):601–615. https://doi.org/10.1124/dmd.119.086397. Epub 2019 Mar 25. PMID: 30910785; PMCID: PMC6505380
64. Pollard TD (2010) A guide to simple and informative binding assays. Mol Biol Cell 21(23):4061–4067. https://doi.org/10.1091/mbc. E10-08-0683. PMID: 21115850; PMCID: PMC2993736
65. Beutler JA, Alvarado AB, Schaufelberger DE, Andrews P, McCloud TG (1990) Dereplication of phorbol bioactives: Lyngbya majuscula and Croton cuneatus. J Nat Prod 53(4):867–874. https://doi.org/10.1021/np50070a014. Erratum in: J Nat Prod 1991 Jul-Aug;54(4):1048. PMID: 2128889
66. Hou X, Sun M, Bao T, Xie X, Wei F, Wang S (2020) Recent advances in screening active components from natural products based on bioaffinity techniques. Acta Pharm Sin B 10(10):1800–1813. https://doi.org/10.1016/j.apsb.2020.04.016. Epub 2020 Jun 3. PMID: 33163336; PMCID: PMC7606101
67. Jacques LB, Water ET (1941) The identity and origin of the anticoagulant of anaphylactic shock in the dog. J Physiol (Lond) 99:454–466
68. Im D, Kishikawa J, Shiimura Y et al (2023) Structural insights into the agonists binding and receptor selectivity of human histamine H4 receptor. Nat Commun 14:6538. https://doi.org/10.1038/s41467-023-42260-z
69. Lim S, Jeong I, Cho J, Shin C, Kim KI, Shim BS, Ko SG, Kim B (2021) The natural products targeting on allergic rhinitis: from traditional medicine to modern drug discovery. Antioxidants (Basel) 10(10):1524. https://doi.org/10.3390/antiox10101524. PMID: 34679659; PMCID: PMC8532887

70. Agu PC, Afiukwa CA, Orji OU et al (2023) Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. Sci Rep 13:13398. https://doi.org/10.1038/s41598-023-40160-2
71. Pinzi L, Rastelli G (2019) Molecular docking: shifting paradigms in drug discovery. Int J Mol Sci 20(18):4331. https://doi.org/10.3390/ijms20184331. PMID: 31487867; PMCID: PMC6769923
72. Davis AM, Leeson PD (2023) The handbook of medicinal chemistry. In Ward SE, Davis A (eds). The Royal Society of Chemistry, 2023, Chap. 1, pp 1–39
73. Colquhoun D (1998) Binding, gating, affinity and efficacy: the interpretation of structure-activity relationships for agonists and the effects of mutating receptors. Br J Pharmacol 125:923–947
74. Hollenberg SM, Weinberger C, Ong ES, Cerelli G, Oro A, Lebo R, Thompson EB, Rosenfeld MG, Evans RM (1985) Primary structure and expression of a functional human glucocorticoid receptor cDNA. Nature 318:635–641
75. Walter P, Green S, Greene G, Krust A, Bornert JM, Jeltsch JM, Staub A, Jensen E, Scrace G, Waterfield M (1985) Cloning of the human estrogen receptor cDNA. Proc Natl Acad Sci U S A 82:7889–7893

Chapter 9
Natural Products and Microbiomes

Abstract Microbiomes can be defined in various ways but, in all cases, they are not simply the sum total of all microbial species that are present—a microbial classification definition does not reflect the complexity of a microbiome. A microbiome includes all biotic and abiotic components—including the environment in which they operate. The influence of natural products on microbiomes was initially primarily as a result of specific antimicrobial properties and hence on perturbation of the system. The link between certain microbial species and disease means that the anti-microbial properties of natural products has been a focus of attention. More recently modulation of microbial properties and behaviour (including aspects such a swarming and biofilm formation) through interaction with quorum sensing molecules has become a target. This has increased as the importance of pathogen:host interactions have been the subject of scrutiny and study. While some fungally-derived microbial natural products (e.g. penicillins) are potent antimicrobial (and, in particular, antibacterial) agents, others that are grouped under the heading of mycotoxins are toxic to mammals and prevention of their ingestion has attracted significant attention. Microbiomes exist within a world of complex interactions that include all biotic and abiotic components and their reactions.

Keywords Microbiome · Quorum sensing · Pathogen · Autoinducer · Mycotoxin

The interaction of natural products with microbial species has been at the forefront of disease prevention, food processing and preservation and, more recently, the study of the establishment of multi-component microbial communities (microbiomes). Microbiomes can be complex or simple. For example, a lichen is a symbiotic relationship between a fungus and an algae and/or a cyanobacterium. In this case, it is unclear what benefits the different contributors to the lichen gain from the association however it appears that the fungi make use of the photosynthetically produced sugars from the algae/cyanobacteria while the latter may benefit from the algal structure as growing conditions and the production of natural products by the fungi which have a number of beneficial properties. It is known, however, that the different components of the lichen contribute in a specific fashion to the community as a whole and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_9

this gives them at least some of the characteristics of a classical microbiome [1]. More complex microbiomes are found in the gut of mammals and in environments as diverse as skin, oral cavity, microwave ovens and even in oil and gas pipelines where they may contribute to the corrosion process.

9.1 Definition and Characteristics of a Microbiome

A microbiome can be defined in different ways [2]. By one definition, it is the combined biotic and abiotic constituents of a microbial system. Others emphasise the functionality of the system (actual or potential) and this leads to definitions that include the biological molecules that are produced and the metagenome. As we shall see, this definition can include not just biological macromolecules (nucleic acids and proteins) but also natural products with profound biological properties and a concomitant effect on the microorganisms that are present. Considerations about microorganisms and their role in the development of living systems originated with two particular observable aspects—the interest in deleterious effects of specific organisms (pathogens) and the use of others in support of food processing (particularly fermentation and preservation). These concepts, allied to the development of microscopy and preferential staining of certain classes of microbes—in particular, bacteria—enabled science to adopt a conventional, reductionist and classification approach. While certain species were found to be of value in virtual (or actual) monocultures such as yeasts for fermentation and lactic acid bacteria for food preservation, the emphasis on finding the causative agents for disease also followed a reductionist route. By eliminating a specific microbe, it was realised that certain diseases could be prevented or even cured. This formed the intellectual basis for Koch and his investigations into pathogenicity and for the 'Magic Bullet' approach developed by Ehrlich.

The reductionist approach was further aided by the development of more advanced, specific and sensitive techniques and methods for classification. The polymerase chain reaction (PCR) allowed microbiologists to distinguish between certain microorganisms based on specific features of their genetic sequence and this also greatly increased the sensitivity of the detection method since a very small amount of genetic material could be amplified many times over to give a detectable signal. Importantly, methods such as PCR were culture-independent. This is important since a large number of microorganisms are not able to be cultured and therefore cannot be detected or enumerated using conventional culture techniques followed by growth on particular media, staining and colony counting. The advent of molecular techniques was also greatly aided by the development of phylogenetic methods based on 16S and 18S rRNA [3]. These unique identifiers enabled the identification of individual classes of microorganisms.

The development of specific and sensitive methods not only aided the detection of individual microorganisms, it also enabled communities of species to be identified and then to be studied as a group. This facilitated the study of the ecology of

species in their natural (complex) environment and not simply the study of individual members [4]. This led to the definition of a microbiome as a 'characteristic microbial community' with defined properties. Importantly, most subsequent definitions of microbiomes included not just the microbial community present (and their interspecies interactions) but also their effect on the environment within which they are contained including the host organism. These aspects are important and any definition of microbiome which only considers the microbial component will be insufficient to describe the totality of the system. In short, a microbiome definition should include:

1. The microbial species present
2. The structural elements of the environment in which they exist (e.g. biofilm, planktonic state etc.)
3. The functional macromolecules (nucleic acids, proteins, lipids, polysaccharides etc.)
4. Microbial metabolites (sugars, natural products, molecules with defined biological properties)
5. Factors produced by the host in response to 1–4.

Finally, the microbiome should also include some realisation of the progression of the system. Biological systems are not static entities and the progression of the microbiome through different stages and factors that influence that progression should also be considered.

9.2 Natural Products and Microbiomes

The most obvious example of the effect of natural products on microbiomes is perhaps the production and impact of anti-microbial species—many of which are small molecules. The use of heavy metals (particularly mercury) in treatments for disease was part of the alchemical tradition. The story of the development of modern antimicrobial therapies is generally considered to have started in the early part of the twentieth century with the development of organoarsenic drugs such as Salvarsan. The idea that compounds could have toxic and antitoxic properties was key to the early (and later) success of a range of biologically active pharmaceuticals [5]. While salvarsan was a synthetic molecule, arsenic can also occur naturally as a cage molecule (Arsenicin A) with anti-cancer properties, isolated from the marine sponge, *Echinochalina bargibanti* [6] and therefore the potential exists for the isolation of bioactive natural products containing arsenic. Salvarsan was an effective treatment for syphilis, its use was largely discontinued with the discovery of less toxic antibiotics including beta lactams. It has, however been noted that tetracycline antibiotics appear to have been in use for many millennia. While heavy metals (including mercury and bismuth) fell out of favour for medicinal use [7] due to concerns over toxicity, they remain a treatment of choice for some conditions [8]. The increasing prevalence of antibiotic resistance has more recently limited the effectiveness of some organic antimicrobial agents. Much of the interest in antimicrobial resistance

(AMR) has arisen in recent years with the advent of multi-resistant 'superbugs' [9]. Although many scientists have pointed to the overuse of antibiotics as the reason for the development of multi-resistant strains, it is clear from phylogenetic analysis that genes coding for proteins involved in antibiotic resistance predate the significant use of antibiotics by humans [10]. One of the major classes of AMR agents are beta-lactamases and analysis suggests that these enzymes have existed for millions of years [11]. More recently, the possibility of using multiple routes to overcome AMR has been suggested and targeting inhibitors of the mevalonate pathway (which is important to bacterial cell wall synthesis) has been proposed as a way of overcoming resistance to amoxycillin [12]. In another example, the growth of aminoglycoside resistance in *E. coli* might be related to metabolic disruption and lowering ribosomal protein levels thereby making bacterial cells less sensitive [13]. If cells were simultaneously challenged with a sub-acute level of a TCA cycle inhibitor, then, in order to maintain ATP levels, the ribosome would have to function more effectively and this is likely to have the effect of increasing ribosomal protein levels, thereby increasing susceptibility to aminoglycoside antibiotics. Such a hypothesis, while not tested, is plausible.

The effect of antimicrobial species on microbial populations in a given microbiome continues to be extensively studied. In particular, the development of the concept of 'good' bacteria with beneficial properties, distinct from pathogens, has encouraged investigations into natural products that can modulate microbial populations rather than the elimination of whole classes of microbes. This has also led to a greater appreciation of the natural products derived from the commensal microbial populations in humans [14] since most individuals have a range of populations of microbes in different locations (microbiomes) and these do not cause disease. In some cases, investigations have focussed on functionally useful classes of compounds (immune modulation agents) while others have looked at the influence of small, microbially-produced molecules on systemic disease and health states [15].

While intra-microbiome microbial species and their natural products is an area of increasing interest, the effect of natural products on the structural elements of the microbiome environment and ecology (particularly factors such as biofilm formation) remains significant. Microbial species can thrive in different environments and can change the environment so as to facilitate their survival and growth. Perhaps the most significant example of this is the formation of biofilms. Biofilms exemplify the 'communal nature' of microbial populations rather than their existence as individual, free living organisms [16]. Although the study of microorganisms is largely concerned with organisms in the planktonic state (unicellular and free swimming), the majority of microorganisms actually exist in a community of heterogenous populations surrounded by a biofilm in the form of an extracellular matrix (ECM) which, in addition to protecting the microorganism from environmental stress, also enables attachment to surfaces. Microorganisms transition between free and immobilised states and each state is characterised by a particular differential gene expression. The study of single cells (and single species) in the planktonic state is part of the reductionist paradigm that has driven microbiology since the development of a means of visualising microorganisms using differential staining and microscopy and more

recently developed (and species-specific) techniques such fluorescence coupled to hybridisation [17]. The development of biofilms has been divided into three phases—initial attachment, maturation and finally decomposition of the biofilm with concomitant release of planktonic microorganisms which can then become attached elsewhere and repeat the cycle. In keeping with scientific reductionism, many studies have been directed towards single species biofilms rather than the complex communities which are more common in nature. The ECM of a biofilm is a mixture of polysaccharides and other biological macromolecules (proteins and nucleic acids). The biofilm is also highly hydrated and the water can be bound or free [18]. Cells in a biofilm are heterogenous (even if they are all of the same species, they are at varying stages of development and growth) and respond to changes in nutrient gradients and in other factors such as quorum sensing molecules. One other attribute of biofilms is that they greatly increase (by up to a thousand fold) microbial resistance to antimicrobial treatments [19]. This has important medicinal and health consequences and can make the removal of pathogenic microbial infections significantly more challenging. This has led to studies to prevent the persistence and growth of biofilms including strategies designed to impede the survival of microbial species by targeting different stages of the biofilm formation process. This includes prevention of the initial adhesion and inhibition of cell–cell signalling by quorum sensing molecules (autoinducers). Clearly biofilms are not homeostatic but are in various states of flux. In addition there is evidence for the detection of quorum sensing molecules by taste (and other) receptors in mammalian systems [20] and this suggests that there is some cross-species communication. It has been suggested that taste receptors are implicated in host immune surveillance and are present at various sites in the body. The link between bitter taste receptors and the initiation of phagocytosis [21] might suggest that detection of microbial quorum sensing molecules by bitter taste receptors might be part of a general immune response. The importance of natural products to quorum sensing will be discussed in more detail later (Sect. 9.3). If there is a mechanistic (or causative) link between bitter taste receptors and the enhancement of the immune function, then the anti-microbial activity of some natural products may be a consequence of their binding to bitter taste receptors and enhancing the innate immune system to better resist infection. A range of microbial metabolites have been demonstrated to bind to the bitter taste receptor T2R38 [22]. It has been suggested that the vagus nerve in the brain is an important aspect of control of the immune response to inflammation [23]. It is known that bitter compounds stimulate the vagus response [24] so a link between the inflammatory response, controlled by the vagus nerve and signal transduction mediated by the binding of microbial quorum sensing molecules to bitter receptors may provide a timely response to microbial infection. The binding of certain quorum sensing molecules produced by pathogens may also provide a way of differentiating between pathogenic and commensal microbes.

One role for natural products, in addition to binding to bitter taste receptors, could be to stimulate the production of mammalian lactonase. The production of lactonase may inhibit the effectiveness of homoserine-based quorum sensing molecules by cleaving the lactone ring thereby limiting pathogen (and biofilm) growth and affecting the balance of the microbiome (fewer Gram negative pathogens) [25]. This enzyme

(part of the PON1 complex) has additionally been implicated in a number of chronic disease states—some of which may be associated with inflammatory responses [26]. It has been suggested that a number of natural products traditionally associated with plants found in the Mediterranean diet, may potentiate PON1 expression [27] and this may be a plausible mechanism to explain some of the benefits of some natural products, although a number of different mechanisms may be operating in addition to a selective anti-microbial activity including protection against oxidative damage [28].

While AMR is an increasing challenge in the clinical world and substantial attention has been paid to the optimisation of current, known antimicrobial agents and the discovery (or synthesis) of new active species with novel mechanisms of action, the molecular arms race is one that microbes seem destined to win. Under these circumstances, control rather than elimination of pathogenic species would appear to be an approach with more long-term viability. One approach to this, using quorum sensing modulation, will be discussed in more detail in the following section.

9.3 Quorum Sensing and Natural Products

Some mention has already been made of the possible role of natural products in the modulation of microbial communication via an effect on quorum sensing. It has become clear that several GPCRs that are produced by eukaryotic species (particularly taste receptors) are able to use quorum sensing molecules as ligands however the function of microbial quorum sensing molecules is different in non-microbial cells. In particular, the immune system seems to use quorum sensing signalling molecules in their response to pathogens.

Quorum sensing molecules are used by microbial species to adjust a range of factors such as growth and virulence [29]. A number of different types of quorum sensing molecules have been reported. As mentioned in passing previously, several Gram-negative pathogens use N-acyl homoserine lactones (Fig. 9.1) as autoinducer signalling molecules.

In Gram positive bacteria, quorum sensing molecules are more usually oligopeptides and include competence stimulating peptides (CSPs). Other quorum sensing

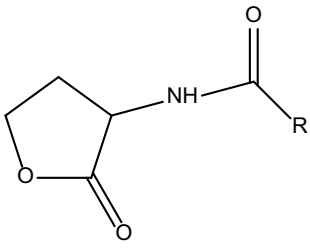


Fig. 9.1 N-acyl homoserine lactone structure

molecules include autoinducer-2, which appears to have a role in both Gram positive and Gram negative bacteria [30]. The increasing number of quorum sensing molecules with varying degrees of microbial species specificity have been reported and extensively reviewed. The ubiquity of quorum sensing mechanisms combined with their variation between different species and the role they play in as diverse microbial responses such as swarming and biofilm formation suggests that quorum sensing may be a suitable point of differentiation of pathogenic and non-pathogenic microbes and may therefore be able to be leveraged as part of a 'Magic Bullet' strategy to specifically target the inhibition of certain microbial species. Such an approach is referred to generically as 'quorum quenching'. The basis of selective quorum quenching lies in the discovery that a range of important pathogens produce different quorum sensing molecules including *Pseudomonas aeruginosa* (acyl homoserine lactones), *Streptococcus mutans* (CSPs) and *Staphylococcus aureus* (D-amino acids).

The induction of an immune response mediated by taste receptor cells in response to quorum sensing molecules may, as noted previously, be part of the oral defence against invading pathogens. In addition, the binding of compounds to taste receptors may provoke the same response in extra-oral tissue and may, therefore, pre-empt pathogen growth. In addition, there is a possibility that polymorphisms in some taste receptors may affect the host response to microbial quorum sensing molecules and this may affect susceptibility or resistance to infection [31]. Measuring the binding of natural products to taste receptors may be a useful surrogate for measuring the stimulation of the innate immune system to resist pathogen attack and may suggest novel pharmacological approaches to the use of natural products as anti-microbial agents.

Direct quorum quenching has been proposed as a way of modulating microbial pathogenicity. Molecules that may inhibit quorum sensing directly or may specifically target certain classes of quorum sensing molecules (the potential of lactonase upregulation has already been noted). Certain aspects of bacterial behaviour that is specifically modulated by quorum sensing molecules (e.g. virulence or biofilm formation) may also be potential targets. Interestingly D-amino acids, which are produced by some bacteria, can also inhibit quorum sensing and thereby inhibit biofilm formation [32] and increase the susceptibility of bacteria to anti-microbial agents. A combination of upregulating amino acid isomerase activity and anti-microbial agents may have the effect of increasing microbial susceptibility to anti-microbial therapy by preventing biofilm formation. A range of natural products that can act as quorum sensing inhibitors (QSIs) have been isolated and some of these have been used as lead compounds for the development of new QSI molecules. Since QSIs have the apparent function of blocking microbial cell–cell communication, it has been suggested that they may inhibit the build up of AMR [33]. This has led to the development of combination therapies that combine QSIs with anti-microbial agents. As the understanding of the function of various members of the AutoInducer (AI) class of molecules has grown, targets for inhibition at various stages in their pathway of activity have been identified [34]. This can occur at various stages in the process of quorum sensing including inhibition of the biosynthesis of quorum sensing molecules, enhancement of their breakdown by, for example, lactonases, blocking the receptors that

AI molecules bind to or interference with other parts of the cell–cell communication process that are dependent upon AIs. A range of natural products with different molecular structures have been found that can operate as QSIs. These operate through a range of mechanisms and affect several endpoints including reduction of virulence, prevention of biofilm formation and preventing (or limiting) levels of production of AI. All of these approaches (and others) have been used in the isolation and design of QSIs [35]. Two of the most important consequences of the use of QSIs is their synergy with conventional anti-microbials. Because QSIs affect properties such as virulence and biofilm formation rather than microbial growth directly, there is a more limited opportunity for the development of AMR. Nonetheless, because of the multiplicity of AI pathways that microbes can utilise, the development of resistance to interference in a single pathway by adoption of an alternative remains a possibility. This might suggest that optimal control of pathogen growth will not result from the complete inhibition of a given species but rather by limiting the opportunities for a single pathogenic species to become predominant. This recognises the complexity of microbial communities and the benefits of some balances of species prevalence over others. It suggests that a reductionist approach of targeting the properties of a single species (biofilm formation or virulence) may ultimately lead to the evolution of microbes that have developed resistance in the form of both AMR and resistance to (or circumvention of) QSIs. Such a development would be a most unwelcome progression and may limit the development and use of QSIs as part of a directed and specific overall microbial control strategy. The study of the ecology of sustainable microbiomes and the small molecules that affect their stability is likely to be an important direction of travel for the control (rather than elimination) of pathogens. Conceptually, prevention of pathogen-induced dysbiosis by environmental manipulation is therefore a valuable concept.

9.4 Modification of Pathogenicity and Biofilm Formation by Natural Products

The involvement of natural products in the control of microbial pathogenicity and biofilm formation are important aspects of the role of natural products in living microbial systems. Homologous natural products play an important role in the survival and growth of microbial species. In addition, microbial natural products are also implicated in the modification of the ecology and environment in which microbes operate. This is particularly relevant in the case of microbes associated with human hosts since any adaptive response is likely to be influenced, at least in part, by the microbial system, how it has been optimised and how it interacts with its associated host. In brief the interactions of natural products with pathogens and biofilms can be considered in the following contexts:

1. Microbial:microbial interactions (already considered in part by the production and use of AI and QSI molecules and the production of antibiotics).

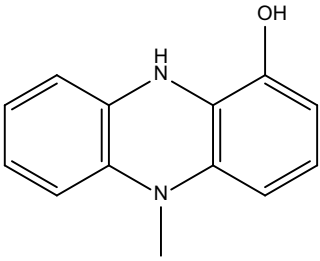
2. Microbial:host interactions—the production of microbial natural products that affect the host in which the microbes are residing (for example toxins and other bioactive constituents).
3. Host:microbial interactions—the response of the host system to a microbial presence. This has already been alluded to in the consideration of innate immune modulation, the role of molecules like AIs (acyl homoserine lactones) and their interaction with host receptors (including bitter receptors) and the production of lactonase.
4. Exogenous agent interactions—examples previously considered include the consequences of the interaction of some natural products with taste receptors and the non-sensory biological consequences of that interaction including an impact on microbial biofilm formation.

There is clearly some overlap between some of these categories however they provide a convenient framework within which a reductionist approach can be examined.

The production of compounds by microbes that inhibit the growth or even kill other microbes is a key feature of the 'golden age' of antibiotics and a defining feature of microbial:microbial interactions. The discovery of a factor in a penicillium mould that killed a bacterial strain of bacillus in 1929 was a significant event [36]. The strain initially used by Fleming was *Penicillium rubens* but this was later replaced by deep tank fermentation using *Penicillium chrysogenum.* While the antimicrobial properties of some entity produced by strains of *Penicillium* was appreciated by Fleming, it was a number of years before the active principle was isolated and characterised. Prior to the discovery of the properties of *Penicillium* species (and significantly before that principle was isolated and characterised, one of the first 'antibiotics' identified and used was pyocyanase which was isolated from *Pseudomonas aeruginosa* [37].

This active principle was originally thought to be a lytic enzyme but was later found to be (and named) pyocyanin (5-methylphenazin-1(5H)-one) (Fig. 9.2) which, in addition to inducing oxidative stress in competing species, is also an important quorum sensing molecule and virulence factor. It can, therefore facilitate the growth of *Pseudomonas aeruginosa* in different ways. Transcription of the genes responsible for the biosynthesis of pyocyanin is controlled by quorum sensing molecules. In addition to being active against other competing microbes, pyocyanin is also lethal to mammalian lung cells. *Pseudomonas aeruginosa* was found to produce the potent quorum sensing molecules, 2-alkyl-4-quinolones which are also significant anti-microbial agents [38]. The reductionist approach of linking specific biological activities with natural products can be a valuable way of optimising structure:activity relationships. The early work on penicillins (and other anti-microbial agents) relied upon observing zones of inhibition of growth of bacteria on agar plates [39]. Other more sophisticated methods have been developed [40] however rapid screening by measuring inhibition of pathogen growth remains a significant tool. Despite this, the pipeline of new antimicrobial natural products has not kept pace with the development of resistance by microbes. The use of growth and fitness modifying bioactive principles is likely to assume greater importance and this will encourage further

Fig. 9.2 Pyocyanin



investigations into how microbes interact in microbiomes and how these interactions can form the basis for the maintenance of healthy (non-pathogenic) communities.

The interest in microbes has, as we have seen, been driven at least in part by the realisation that many diseases are caused and spread by microbial entities. This is a key example of a microbe:host interaction. Some microbially-induced diseases have a distinct chemical causative agent (for example mycotoxins, marine toxins) while others cause inflammation with concomitant clinical outcomes (for example, Salmonella infection [41]. Mycotoxins are natural products derived from fungi and, in addition to being studied as vectors for human disease, many of them are also biosynthetically interesting. The fungi are often found as contaminants in foods used by humans and animals and over 300 of them have been characterised. Some of the major mycotoxins are shown in Fig. 9.3. It is noteworthy that, while they share a classification based on biological activity, they are a structurally diverse group and organisms employ a range of biosynthetic pathways. Within the general definition of mycotoxins, the toxin 3-nitropropanoic acid is one of the simplest. It is produced by *Penicillium* and *Aspergillus* species. The biosynthesis of 3-nitropropanoic acid includes an amino acid oxidase and a decarboxylase both of which are highly conserved between different fungal species and comprise the biosynthetic gene cluster in fungi [42]. While mycotoxins are most frequently characterised by their derivation from specific fungi (for example, aflatoxin from *Aspergillus flavus* and fumonisins from *Fusarium* species), 3-nitropropanoic acid is more widespread in occurrence additionally being found in some plant species and is also in bacteria and animals. When originally isolated, the source of 3-nitropropanoic acid was the tree *Hiptage mandoblata* and it was given the name hiptagenic acid [43]. This type of spread of occurrence of a natural product is relatively rare. It might be expected that the biosynthetic pathways for a natural product as simple as 3-nitropropanoic acid would share some characteristics between different producers however this appears not to be the case. The biosynthesis in fungi makes use of aspartate as a starting material while in plants, malonate is used. The mechanism of action of 3-nitropropanoic acid involves irreversible deactivation of succinate dehydrogenase and it is a plant protective agent as well as conferring some growth advantage to fungal species. Resistance to antimicrobial agents is not confined to pathogenic bacteria that affect human or animal health. Plant species that produce 3-nitropropanoic acid are protected from its effects through the action of a co-occurring oxidase enzyme that converts the toxin to malonyl semialdehyde. Bacteria that can make use of 3-nitropropanoic acid as a

sole carbon and nitrogen source have also been isolated and these detoxify the nitro acid by oxidative denitrification using an inducible monooxygenase enzyme [44]. In addition to its toxicity, 3-nitropropanoic acid can mimic the clinical symptoms of neurological conditions including Huntington's disease [45]. Other mycotoxins also have, in some cases, some useful characteristics (e.g. patulin was originally of interest because of its antibiotic properties) however the most significant members of the class are principally of interest because of their toxicity profile. These include aflatoxins, citrinin, ergot akaloids, fumonisins, ochratoxin A, patulin, trichothecenes, and zearalenone (Fig. 9.3).

Although their classification suggests some linkage between them, as we have seen, they are structurally diverse and their association is based on biological properties (toxins) and their derivation from fungi (myco). While it may appear simple, the definition of mycotoxins can also cause some confusion. In general, the term 'mycotoxin' is reserved for compounds which are acutely toxic to vertebrates and are made by fungi. This definition typically excludes, for example, many of the toxins produced by mushrooms. As we have seen, classification by chemical structure (alkaloids, glucosinolates, beta lactams) or by biosynthetic origin (polyketides, terpenes)

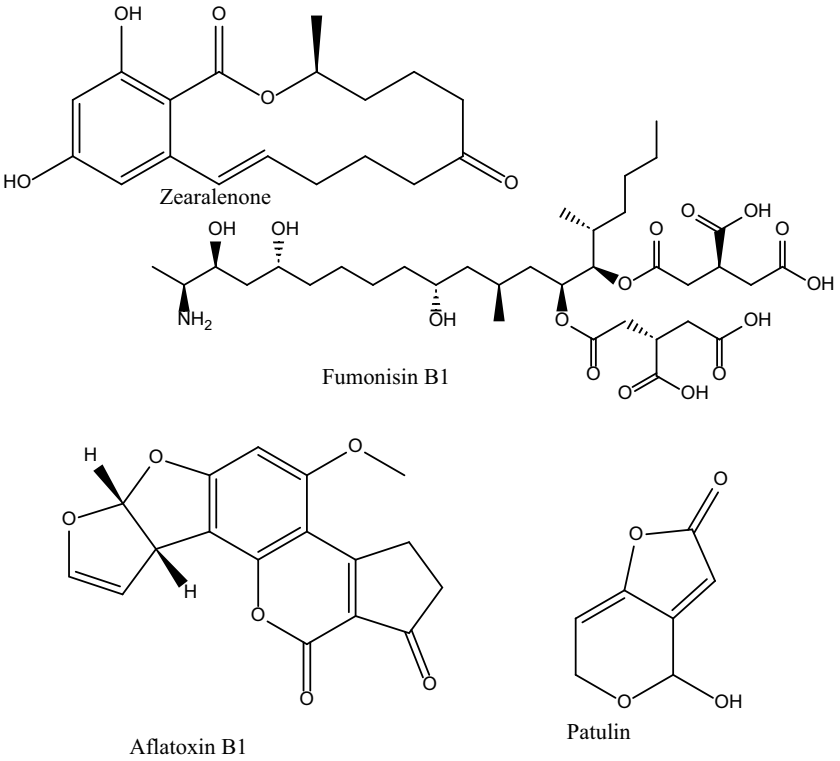


Fig. 9.3 Some representative mycotoxins

can group natural products in a logical fashion. Creating a classification based upon a broad functional description (toxin) and origin (myco) is open to confusion and misclassification. Even when functionality is based upon a more specific clinical endpoint (e.g. neurotoxins, nephrotoxins, hepatotoxins), some natural products may be classified in multiple groupings while others may have a range of modes of action. In addition, some natural products have different modes of action depending on dose ingested and the time over which exposure is experienced. For example, aflatoxin is, in acute doses, acutely toxic [46]. It is also a carcinogen and is implicated in the development of hepatocellular carcinomas [47] in cases of chronic exposure. While scientists may aspire to unify structure and effect on complex biological systems, this may not be an achievable endpoint in the classification and ordering of natural products. Marine-derived toxins are also capable of being described by multiple definitions and this may also be confusing. The principal marine toxins are tetrodotoxin, saxitoxin and conotoxin. They are derived from puffer fish, shellfish and marine snails respectively. While they can occur in different marine species, unlike mycotoxins which can affect a range of clinical targets, the mechanism of action of the principal marine toxins is to block voltage-gated sodium channels. This gives rise to a number of clinical outcomes and these are predominantly manifested in gastrointestinal and neurological effects. In the case of many shellfish-derived toxins, they are produced by phytoplankton (dinoflagellates) and then accumulated in the marine species that are later consumed. They can be further sub-divided into a classification based on the observed physiological effect after ingestion and this gives rise to the naming of Amnesic Shellfish Poisoning (ASP), Paralytic Shellfish Poisoning (PSP), Diarrheic Shellfish Poisoning (DSP) and Azaspiracid Shellfish Poisoning (AZP). These occur worldwide while Neurologic Shellfish Poisoning (NSP) is more common in the USA. This enables a definition to be incorporated into, for example, EU Regulations whereby 'Marine Biotoxins' means poisonous substances accumulated by bivalve molluscs, in particular as a result of feeding on plankton containing toxins. One of the issues that has provided a challenge for toxicologists and (in particular) regulatory authorities is that the method of detection of the compounds responsible for the observed physiological effects has relied upon a whole animal bioassay (generally a mouse) and this is both expensive and ethically undesirable. Chromatographic and biosensor techniques have been suggested as alternatives [48] and regulatory limits have been created for specific marine toxins [49].

While the classification of mycotoxins (in part) and marine toxins (in particular) are based upon host (human or animal) response, there are other examples of the interaction of microbes and/or the natural products they produce with the host organism. This is not invariably a malevolent interaction. For example, gut microbes produce short chain fatty acids (SCFAs) that are implicated in a number of beneficial effects. SCFAs play a role in the regulation of epithelial function and in the regulation of immunity. In addition to a role in the gut, SCFAs have also been implicated in effects at other organs however they appear to exert their most significant effects on the immune system [50]. The three principal physiologically important SCFAs are acetate, butyrate and propionate and they are produced from the fermentation of indigestible carbohydrates in the gut. The factors which influence the production

of SCFAs and their subsequent physiological effects have been extensively studied [51]. In addition to their physiological effects, SCFAs also act in a feedback loop with gut bacteria. The use of quorum sensing (QS) molecules to promote biofilm formation has previously been discussed in the context of the survival of pathogens. It has also been noted, however, that certain probiotic bacteria may promote the formation of gut mucosal protective biofilms and inhibit the colonisation of pathogens [52]. In addition, the role of SCFAs in host metabolism has been noted and it has also been speculated that SCFAs may inhibit the effectiveness of quorum sensing mechanisms in intestinal pathogenic microbes [53]. The complex interplay between crucial microbial metabolites and the host can be regulated by a deeper understanding of the role of quorum sensing molecules and their manipulation to optimise the gut (and other organ) microbiome and thereby to promote health.

A second series of microbial metabolites that appear to have a beneficial effect are long chain polyunsaturated fatty acids (LCPUFAs) which are produced by algae. These natural products occupy an interesting nutritional niche. Unlike essential vitamins and nutrients, there is no physiological condition associated with dietary deficiency however they have been suggested to have a role in optimal health. Two of the most studied members of the class are docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA)—both have been implicated in brain health and the development of vision in babies. Furthermore, they are considered to have an effect (although not precisely defined) on neurological benefits, in the prevention of depressive disorders and in the maintenance of cardiovascular health. While there is some limited mechanistic underpinning for some of the observed effects, there is generally a lack of substantive clinical trials in humans that would provide compelling evidence. In addition to DHA and EPA, the other important LCPUFAs are alpha-linolenic acid (ALA) and arachidonic acid (ARA). One of the principal roles of ALA is to act as a precursor for the synthesis of EPA and DHA in mammalian tissue—although the conversion rate (particularly in the case of DHA) is low (<1%) and therefore the most reliable source of EPA and DHA is through ingestion. The major dietary source of LCPUFAs is fish and other seafood however fish themselves absorb DHA and EPA through the consumption of microalgae, which are the original producers of LCPUFAs. The LCPUFAs produced by microalgae are, on the whole, free of any contaminants. Fish oil, on the other hand, can contain lipid soluble contaminants derived from the sea including PCBs (although probably not at a level at which a risk to health is likely) and mercury [54]. Microalgae can be grown in fermenters therefore biotechnological production of LCPUFAs is an option. The use of a range of microbial systems for the production of beneficial LCPUFAs has been reviewed [55] and optimisation of production is an ongoing process. One of the major algal species that is used for the production of DHA is *Schizochytrium.* This species is also able to produce astaxanthin (albeit in small amounts) and gene editing has been carried out to modify the mevalonate pathway to increase the amount of astaxanthin produced without affecting DHA levels [56].

While the study of non-essential beneficial natural products remains a challenge, the possibility of introducing multiple, non-competing biosynthetic pathways into a suitable species may increase the numbers and validity of studies and help to build

an evidence base for the study of synergies between different dietary constituents. If synthetic biology and biological engineering of pathways can be demonstrated to be specific and efficient, the colonisation of a range of human microbiomes with beneficial, engineered microbes may be a viable strategy. One area of active interest is the gut microbiome of ruminants [57]. The natural variation in the rumen microbiota and the effects of different feeding regimes makes the demonstration of cause and effect for a particular microbial species difficult to confirm. This might suggest that gene editing to introduce specific functionality (including the production of specific, bioactive natural products), may be a simpler and conceptually more robust route. This will substantially increase the possibility of the successful introduction of microbiome-based therapies [58] with targeted endpoints.

The response of the host to microbial natural products (host:microbial interaction) has been alluded to previously. In particular, the immune response and the effect of microbial quorum sensing molecules has been noted. It is important to note the difference between the innate and the adaptive immune system. The innate immune system is the general and rapid response to challenge by an external agent. It is considered to be generic and non-specific. The innate immune system orchestrates the initial inflammatory response. Phagocytosis (enveloping and digesting foreign cells) is also a function mediated by the innate immune system as is the mobilisation of a cascade of enzymes that, among other functions, digest bacterial and viral cell walls. Natural killer cells are also part of the innate immune system. Their function is to recognise abnormal cells (including those infected with viruses and tumour cells) and to then destroy the cell surface thereby leaving the abnormal cells vulnerable to other parts of the innate immune system. The adaptive immune response is slower but more specific. In addition, the adaptive immune response may 'remember' previous challenges and this means that it can be more effective in dealing with specific challenges previously encountered. There are three parts to the adaptive immune response—T cells, B cells and antibodies. In simple terms, the T cell family specifically targets cells and both destroys and remembers them so that future challenges are met in a specific (and more rapid) fashion. Some types of T cells can activate B cells (also called B lymphocytes) in a specific way based on the invading cells and the B cells replicate then produce antigen-specific antibodies. The antibodies also activate further the innate immune response and other immune system cells. In the case of commensal microbes, the immune system tolerates them however exposure to pathogens or a dysfunctional immune response can result in the triggering of a number of inflammation-related responses with clinical manifestations such as irritable bowel disease (IBD) and rheumatoid arthritis, among others. The interaction between the host and the gut microbiome is essential for the maturation and correct functioning of the immune system. Certain small molecules secreted by the microbiota have a particular role to play. For example, indole is known to be important for epithelial function. Microbial polysaccharides are also known to play a key role in the regulation of the immune response. The interactions between microbiota and the host immune system is complex and, through its manipulation, a range of beneficial outcomes may become possible. This has been particularly evident in the interest that

has developed in immune-regulated anti-cancer effects regulated by microbial interventions [59]. Immune regulation by bacteria and by the tumour microbiome has the potential to provide significant benefits [60]. Cross talk between the microbiome and cancer cells by quorum sensing peptides has been observed [61] and modulation of microbial (and immune) responses by quorum sensing natural products may provide a particular and specific benefit.

The discovery of penicillin shifted the emphasis of anti-microbial therapy from natural products derived from plants [62] to those from microbes. The increase in AMR and a growing awareness of the importance of inflammation and the immune response in chronic disease, has increased the attention paid to natural products from a range of sources, including plants [63], that can alter the properties of the gut microbiota and the metabolites excreted thereby influencing microbiome health and dysbiosis controlling pathogen growth *in situ* and impacting upon clinical endpoints related to inflammation and chronic disease. Effects have been noted on both the innate and adaptive immune systems. In addition to the effects mediated by microbes and microbial products, a number of other exogenous natural products have been investigated as potential mediators of an inflammatory response. The initial inflammatory response is the consequence of the rapid, non-specific innate immune response. One important consequence of an inappropriate immune reaction is autoimmunity. An autoimmune response is the result of a breakdown in self tolerance and activation of self-reactive T cells. In some cases, an autoimmune response is triggered by infection but the complicity of commensal organisms in the microbiome has also been suggested [64]. This is, therefore, a particular form of microbe:host:microbe interaction. In the specific case of autoimmune arthritis the inflammation can be reduced by treatment with corticosteroids however these have a number of side effects and long term use may be problematic. A range of natural products and extracts have been used and in some cases the possible molecular mechanisms have been investigated [65]. In addition, the same investigators suggested that natural products that alter the resident microbiome may have a beneficial effect through modulation of production of SCFAs. The application of resveratrol, curcumin, Boswellic acids, epigallocatechin-3-gallate, and triptolide have all been suggested to help alleviate and thereby control autoimmune inflammation [66] and are an example of the direct action of a series on exogenous agents on health. In some cases putative mechanisms of action have been suggested and these can vary depending on the target disease.

For example, in the case of triptolide (Fig. 9.4) mechanisms of action include the reduction of expression of pro-inflammatory cytokines, increased production of anti-inflammatory cytokines, interaction with a T cell surface antigen, interference with downstream inflammatory pathways and suppression of autophagy. A similar variation in activity was also observed with resveratrol and the other natural products tested in humans, in animal models and in vitro. The possibility that at least some of these activities might be mediated by effects on biofilms and other microbial characteristics has been noted previously.

The key role of biofilms in gastrointestinal health and disease has been recognised. Many of these comprise several species of microbes and the interrelationships of these can be challenging to disentangle—as is the relationship between the gut

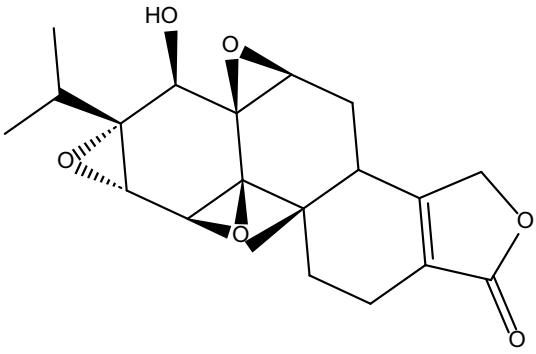


Fig. 9.4 Structure of triptolide

microbiome and the host [67]. One of the major influences on gut biofilm formation and maintenance is the presence of host and microbial proteases. The inhibition of gut proteases through the action of microbial products has been implicated in a number of inflammatory gut conditions. On the other hand, as noted previously, mucosal biofilms may prevent the attachment and adhesion of pathogenic species [68]. The use of quorum sensing molecules to modulate biofilm formation has previously been noted however host-produced hormones have also been implicated in quorum sensing and the presence of the 'taste' receptor T2R38 in the gut may inhibit gastrointestinal biofilm formation. In addition, activation of T2R38 by a bacterial quorum sensing molecule may result in activation of the innate immune system and prevention of gastrointestinal biofilm formation. This is a further example of the potential of exogenous (bitter) compounds to affect the characteristics of the host:microbiome interaction however it is not clear if the binding of bitter natural products to T2R38 will have a similar direct effect. Stimulation of the innate immune system may have both positive and negative consequences. Despite the lack of direct evidence of an effect on biofilm formation, the binding of bitter compounds to extra-oral bitter receptors has been implicated in a range of biological outcomes including hormone secretion, immunity, metabolism, and cell proliferation [69]. A further consequence of the presence of bitter taste receptors in the gut is the upregulation of the ATP-binding cassette B1 (ABCB1) [70]. This may limit the absorption of dietary bioactive compounds by activating their transport out of the cell thereby reducing their efficacy (or toxic effect). This may either limit their effectiveness as biofilm modifying agents (if limited absorption means a concomitantly limited effect on the innate immune system) or may increase residence time in the gut by reducing the ADME of a given natural product.

9.5 Natural Products—A Changed Paradigm

The traditional approach to natural products has been underpinned by the reductionist and empirical approaches of conventional chemistry. Attempting to apply the same paradigm to living systems has encouraged the development of studies designed to make use of simpler and simpler parts of the biological system and to determine which of the multiple, complex parts of a living system (apart from primary metabolism) are mechanistically of greatest significance. This has enabled natural product chemists to think in terms of single biological targets (enzyme active sites or receptors) and to avoid the consequences of the effect of natural products on whole biological systems. There remains, in many cases, a conceptual gap between the linear reasoning from a single active product to an isolated part of a biological system and the effects of natural products on whole living systems. The application of the reductionism of chemistry to the interaction between natural products and biological systems is, therefore problematic for two reasons.

1. It regards single compounds and an understanding of them as the most important key to developing knowledge about the use and potential of natural products. This ignores the possibility of synergies (or even competition) between natural products.
2. It reduces biological systems to a single essential functional component and tries to explain the effect of natural products on a complete biological system in the context of the single component.

There is a substantial history and technology associated with studies of single chemical entities. This was in part because investigations were curiosity-driven and also part of the traditional purification process adopted by alchemy. Such chemical studies evolved based upon the physicochemical characteristics of natural products. Simple extraction techniques enabled compounds with similar chemical properties (for example, alkaloids, lipids) to be preferentially extracted from a matrix containing a large number of different chemical entities. The alchemists used seven stages to effect purification. The first stage is calcination or burning. The intention is to remove extraneous matter that disguises or hides the essence of the valuable material that is the focus of the purification. The second stage is dissolution which can be considered to be analogous to solid:liquid extraction. Then follows separation which is another means of removing extraneous material. Conjunction is then used to combine the partially purified products from the first three processes. What then follows is putrefaction or fermentation. This stage is found, for example, in the development of many pigments and dyes (for example, indigo undergoes this process) [71]. It is also a feature of the release, for example, of aglycones [72]. The sixth stage in alchemical science is distillation which is a procedure that is still used in chemistry—particularly in the purification of essential oils. The final stage is referred to as coagulation or solidification. This is a unification and purification process and may be considered to be analogous to crystallisation. The crystallisation step is a way of coagulating identical molecules together from solution and removing them from other compounds

which either do not dissolve at all (and hence can be removed by filtration) or remain in solution.

One of the main driving forces that lies behind the development of single compound purification strategies for natural products is the need to establish proof of efficacy and dose. In addition, mixtures of compounds may be more or less active (or toxic) than the individual compounds and so an effective (but non-toxic) dose must be established. A significant advance in the isolation and purification of natural products was the isolation of morphine from the opium poppy in 1817. This was followed by the isolation of quinine from the bark of the Cinchona tree in 1820 and salicin (the precursor of Aspirin) in 1838. The history of natural products in drug discovery has been extensively reviewed [73] and the major advances documented. The other reason for isolating pure compounds is in order to determine their structures and, if possible, to modify these to create a compound with enhanced bioactivity, stability or bioavailability. Two areas in which the development of pure, single compounds has been of lesser importance has been the extraction and use of colours and of flavours. Indeed, in the case of the latter, although specific receptors have been found (including the multiple bitter receptors referred to earlier), the additive flavour properties of mixtures of individual compounds have been appreciated. This includes, in particular, the use of umami to enhance other flavours. In the case of colours, many of the colours that were commonly used were also mixtures of a number of individual compounds. This even applies to the first synthetic dye, invented by W H Perkin in the nineteenth century. Mauveine has been found to be a mixture of different compounds and this was verified by carrying out a syntho-mimetic reaction using the same materials that Perkin did [74]. It is unlikely that synergistic mixtures of bioactive compounds will be a major development in future since defining biological activity and controlling pharmacokinetics is far more complex when multiple bioactive natural products are present. One exception to this is the combination of amoxycillin and clavulanic acid which is used in certain infections since amoxycillin is a classic beta lactam antibiotic and clavulanic acid is a beta-lactamase inhibitor [75]. The possibility of using combination therapies, for example using quorum sensing inhibitors in combination with antimicrobials may be a useful strategy—particularly if selective quorum sensing inhibitors can be found among natural products [76]. This may also lead to strategies in which modulation and control of infection (and other diseases) rather than an absolute cure may be possible.

The trend towards trying to define an essential (and 'druggable') component of living systems was possibly started by the notion of the 'Magic Bullet' popularised by Ehrlich and continued through the 'Golden Age' of antibiotic research culminating in the isolation of individual receptors which could be used to screen natural products (and other compounds) for biological efficacy. One of the defining features of biological systems is their complexity. This applies intra and extra-cellularly and explains why the consequences of perturbations to even the simplest biological systems can be difficult to predict. For example, perturbations to the primary metabolism of unicellular bacteria can have a significant effect. For example, a simple *E. coli* strain which has been manipulated such that it no longer has the key TCA enzyme, succinate dehydrogenase, will change its carbon utilisation to produce excess amounts of a

required product from an engineered biosynthetic cycle. This facilitated the creation of an *E. coli* strain without a TCA cycle that can be used for production of chemicals and is therefore more biotechnologically efficient [77]. It might be interesting to see if cloning the genes for biosynthesis of 3-nitropropanoic acid, a known and irreversible inhibitor of succinate dehydrogenase, might have a similar effect. It is clear from this example that the alteration of a single intra-cellular component can have significant effects. In multicellular organisms, the opportunity for metabolic cross-talk is even more significant. However, even in unicellular organisms, there is the opportunity for extra-cellular interactions to take place. One example of this is the communication, via quorum sensing molecules, between communities of organisms in microbial communities. Microbial community communication exists for a number of reasons. One of the most significant of these is the production of biofilms which, as we have noted earlier, can both protect microbes from host-induced and exogenous antimicrobial agents and also prevent the attachment of competing species. The concept of mutual defence has led to the introduction of a concept referred to as a 'defensome' of complex microbial communities [78]. A key part of this mechanism is the transfer of genetic material between microbial species by horizontal gene transfer. While the innate and adaptive immune systems have been described for humans and occur in other mammals, bacteria can resist phage infection using a similar paradigm. The innate immune system in bacteria includes prevention of infection by phage, methylation specifically of foreign DNA and cellular destruction of the DNA before an invading phage can complete its replication cycle thereby preventing the spread of the phage infection to other members of the microbial community. The adaptive immune system employs the use of mobile genetic elements (CRISPR elements) that are incorporated into genomic DNA as spacers. Another possibility for some microorganisms is the inhibition of the growth of competing species through the use of quorum sensing signal blockers and some of these may derive from endogenous or exogenous natural products. The screening for and discovery of natural quorum quenchers is likely to be a feature of controlled colonisation of host species by benign microbial communities and may be a useful alternative to antimicrobial strategies which rely on the discovery and development of antibiotics which will be affected by the development resistance. The development of strategies that rely on maintenance of a non-pathogenic microbial community, while attractive, will require a significant advance in understanding which elements are most important in preventing pathogen outgrowth (the potential for which will still be present but unrealised) and in discovering predictive biomarkers which can provide indicators of future microbial growth. Ultimately a reductionist chemistry approach in which microbes are totally or selectively killed by chemical agents will be simpler and may be more attractive in the short term. However the history of development of anti-microbial resistance suggests that such a strategy will fail in the medium and longer term.

In the case of multi-cellular organisms, all the functionality of unicellular systems is present but this is overlaid with further levels of complexity. In an attempt to create some order and meaning to the multi-cellular system, the application of reductionist chemistry facilitated the development of specific, target-based approaches.

The premise was that some key aspects of complex living systems could be understood both on a cell-specific and a systemic level. The led to the formulation of the concept of receptors which was enunciated before there was any conclusive biological evidence for their existence. The first approach to test a receptor-based hypothesis was the development of beta-receptor blockers for the therapy of hypertension and the histamine-H2-receptor blockers for the treatment of stomach ulcers. In both cases, the concept of a receptor-mediated response was inferred. In essence, cells were hypothesised to contain an entity which responded to an exogenous chemical signalling molecule. The growth in the concept of receptors was rapid—once it was found that the receptor theory could explain how bioactive molecules might be employed in living systems [79]. One of the key concepts that underpinned receptor theory is that of specificity. While this certainly fits with Ehrlich's concept of a 'Magic Bullet', it is also consistent with previous ideas concerning the way biological systems interact with natural (and unnatural) products. Prior to the development of Ehrlich's hypothesis, the concept of 'affinities' was used to explain the biological activity of certain compounds. Some substances were thought to affect specific parts of the body (and not others). Although not a direct precursor of the 'Magic Bullet' hypothesis, it is possible to envision how this concept could lead to the receptor paradigm—a bioactive natural product binds to (has an affinity for) a specific aspect of a target organ. The second stage in the progression of this is the realisation that the 'affinity' results in a chain of events resulting in a clinical outcome. The concept of a single specific and unique receptor for each bioactive substance has subsequently been shown to be incorrect—as was demonstrated by the ubiquity of location and activities of taste receptors. Taste receptors (and studies of other receptors) demonstrate that receptor-mediated activity is not, completely reductionist since biological activity also depends on the location of the receptor. Even in the case of enzymes, it is becoming clear that, while specificity (in terms of substrate and reaction catalysed) is important in many cases, some enzymes exist on a spectrum of substrate specificity [80] and, depending on physiological conditions under which they operate, may carry out more than one type of reaction (they are polyfunctional) [81]. Another example is the olfactory receptor, OR51E2 which can bind both acetate and propionate and is involved in modulation of gut inflammation in cases of colitis [82]. It would not be unexpected if the OR51E2 receptor had a range of activities depending on where it is expressed. It is certainly involved in the initiation of a neuronal response and, if this was accompanied by anti-inflammatory activity, it may suggest a neurological link between the production of SCFAs by the gut microbiome, their transport to the olfactory system and prevention of neurological inflammatory disorders such as Alzheimer's disease [83].

Natural products have been used both as pharmacological agents and to probe the mechanism of action of biological systems. Much of the success in these investigations have resulted from the application of reductionist theories regarding individual natural products and how they interact with specific aspects of biological systems—particularly receptors. This type of paradigm has underpinned biological areas such as infection and disease and has been crucial to how chemists have treated biological systems including microbiomes and complex, multicomponent systems. The

development of an understanding about factors such as quorum sensing that affect microbial community balance has been crucial in the development of a paradigm shift whereby modulation of balanced biological systems has been seen as a viable alternative to biological reductionism. This has been accompanied by an increasing realisation that reductionism may not be the most appropriate way of considering biological systems. The pinnacle of receptor theory has had many successes however it has become clear that the site of action of a receptor may be as important as the downstream consequences of its interaction with a given agonist or antagonist. It may be time to develop better ways of dealing with the complexity of whole biological systems and their interactions with natural products rather than simply reducing that interaction to a bimolecular interaction. This is nowhere more relevant than in the complex set of relationships that underpin microbiomes and their interactions with other living systems.

References

1. Grimm M, Grube M, Schiefelbein U, Zühlke D, Bernhardt J, Riedel K (2021) The lichens' microbiota, still a mystery? Front Microbiol 30(12):623839. https://doi.org/10.3389/fmicb.2021.623839.PMID:33859626;PMCID:PMC8042158
2. Berg G, Rybakova D, Fischer D et al (2020) Microbiome definition re-visited: old concepts and new challenges. Microbiome 8:103. https://doi.org/10.1186/s40168-020-00875-0
3. Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A 74:5088–5090
4. Whipps J, Lewis K, Cooke R (1988) Mycoparasitism and plant disease control. In: Burge M (ed) Fungi biological control system. Manchester University Press, pp 161–187
5. Williams KJ (2009) The introduction of 'chemotherapy' using arsphenamine—the first magic bullet. J R Soc Med 102(8):343–348
6. Paul NP, Galván AE, Yoshinaga-Sakurai K, Rosen BP, Yoshinaga M (2023) Arsenic in medicine: past, present and future. Biometals 36(2):283–301. https://doi.org/10.1007/s10534-022-00371-y. Epub 2022 Feb 21. PMID: 35190937; PMCID: PMC8860286.
7. Gelpi A, Tucker JD (2015) After Venus, mercury: syphilis treatment in the UK before Salvarsan. Sex Transm Infect 91(1):68. https://doi.org/10.1136/sextrans-2014-051778.PMID:25609466;PMCID:PMC4317394
8. Franz KJ, Metzler-Nolte N (2019) Introduction: metals in medicine. Chem Rev 119(2):727–729. https://doi.org/10.1021/acs.chemrev.8b00685. PMID: 30990707
9. Painuli S, Semwal P, Sharma R, Akash S (2023) Superbugs or multidrug resistant microbes: a new threat to the society. Health Sci Rep 6(8):e1480. https://doi.org/10.1002/hsr2.1480.PMID:37547359;PMCID:PMC10397562
10. Aminov RI, Mackie RI (2007) Evolution and ecology of antibiotic resistance genes. FEMS Microbiol Lett 271:147–161. https://doi.org/10.1111/j.1574-6968.2007.00757.x
11. Hall BG, Barlow M (2004) Evolution of the serine beta-lactamases: past, present and future. Drug Resist Updat 7:111–123. https://doi.org/10.1016/j.drup.2004.02.003
12. Dewachter L, Dénéréaz J, Liu X, de Bakker V, Costa C, Baldry M, Sirard JC, Veening JW (2022) Amoxicillin-resistant Streptococcus pneumoniae can be resensitized by targeting the mevalonate pathway as indicated by sCRilecs-seq. Elife 24(11):e75607. https://doi.org/10.7554/eLife.75607.PMID:35748540;PMCID:PMC9363119
13. Shiraliyev R, Orman MA (2024) Metabolic disruption impairs ribosomal protein levels, resulting in enhanced aminoglycoside tolerance. eLife 13:RP94903. https://doi.org/10.7554/eLife.94903.3

14. Barber CC, Zhang W (2021) Small molecule natural products in human nasal/oral microbiota. J Ind Microbiol Biotechnol 48(3–4):kuab010. https://doi.org/10.1093/jimb/kuab010
15. Milshteyn A, Colosimo DA, Brady SF (2018) Accessing bioactive natural products from the human microbiome. Cell Host Microbe 23(6):725–736. https://doi.org/10.1016/j.chom.2018.05.013.PMID:29902438;PMCID:PMC7232905
16. Flemming HC, Wingender J, Szewzyk U et al (2016) Biofilms: an emergent form of bacterial life. Nat Rev Microbiol 14:563–575. https://doi.org/10.1038/nrmicro.2016.94
17. Goodsell DS, Haas D (2014) Visualising microorganisms from molecules to cells. FEMS Microbiol Lett 356(1):1–7. https://doi.org/10.1111/1574-6968.12485
18. Quan K, Hou J, Zhang Z, Ren Y, Peterson BW, Flemming HC, Mayer C, Busscher HJ, van der Mei HC (2022) Water in bacterial biofilms: pores and channels, storage and transport functions. Crit Rev Microbiol 48(3):283–302. https://doi.org/10.1080/1040841X.2021.1962802. Epub 2021 Aug 19 PMID: 34411498
19. Olsen I (2015) Biofilm-specific antibiotic tolerance and resistance. Eur J Clin Microbiol Infect Dis 34(5):877–886. https://doi.org/10.1007/s10096-015-2323-z. Epub 2015 Jan 29. PMID: 25630538
20. Kouakou YI, Lee RJ (2023) Interkingdom detection of bacterial quorum-sensing molecules by mammalian taste receptors. Microorganisms 11:1295. https://doi.org/10.3390/microorganisms11051295
21. Gopallawa I, Freund JR, Lee RJ (2021) Bitter taste receptors stimulate phagocytosis in human macrophages through calcium, nitric oxide, and cyclic-GMP signaling. Cell Mol Life Sci 78(1):271–286. https://doi.org/10.1007/s00018-020-03494-y. Epub 2020 Mar 14. PMID: 32172302; PMCID: PMC7492447
22. Verbeurgt C, Veithen A, Carlot S, Tarabichi M, Dumont JE, Hassid S, Chatelain P (2017) The human bitter taste receptor T2R38 is broadly tuned for bacterial compounds. PLoS ONE 12(9):e0181302. https://doi.org/10.1371/journal.pone.0181302.PMID:28902853;PMCID:PMC5597121
23. Guglielmi G (2024) Found: the dial in the brain that controls the immune system. Nature. https://doi.org/10.1038/d41586-024-01259-2. Epub ahead of print. PMID: 38693241
24. McMullen MK, Whitehouse JM, Towell A (2015) Bitters: time for a new paradigm. Evid Based Complement Alternat Med 2015:670504. https://doi.org/10.1155/2015/670504. Epub 2015 May 14. PMID: 26074998; PMCID: PMC4446506
25. Sikdar R, Elias M (2020) Quorum quenching enzymes and their effects on virulence, biofilm, and microbiomes: a review of recent advances. Expert Rev Anti Infect Ther 18(12):1221–1233. https://doi.org/10.1080/14787210.2020.1794815. Epub 2020 Aug 4. PMID: 32749905; PMCID: PMC7705441
26. Petrič B, Kunej T, Bavec A (2021) A multi-omics analysis of PON1 lactonase activity in relation to human health and disease. OMICS 25(1):38–51. https://doi.org/10.1089/omi.2020.0160. Epub 2020 Dec 11. PMID: 33306925; PMCID: PMC8045895
27. Lou-Bonafonte JM, Gabás-Rivera C, Navarro MA, Osada J (2015) PON1 and Mediterranean Diet. Nutrients 7(6):4068–4092. https://doi.org/10.3390/nu7064068.PMID:26024295;PMCID:PMC4488773
28. Sirca TB, Mureşan ME, Pallag A, Marian E, Jurca T, Vicaş LG, Tunduc IP, Manole F, Ştefan L (2024) The role of polyphenols in modulating PON1 activity regarding endothelial dysfunction and atherosclerosis. Int J Mol Sci 25(5):2962. https://doi.org/10.3390/ijms25052962. PMID: 38474211; PMCID: PMC10931629
29. Li Z, Nair SK (2012) Quorum sensing: how bacteria can coordinate activity and synchronize their response to external signals? Protein Sci 10:1403–1417
30. Pereira CS, Thompson JA, Xavier KB (2013) AI-2-mediated signalling in bacteria. FEMS Microbiol Rev 37(2):156–181. https://doi.org/10.1111/j.1574-6976.2012.00345.x
31. Lee RJ, Xiong G, Kofonow JM, Chen B, Lysenko A, Jiang P, Abraham V, Doghramji L, Adappa ND, Palmer JN et al (2012) T2R38 taste receptor polymorphisms underlie susceptibility to upper respiratory infection. J Clin Investig 122:4145–4159

32. Vahdati SN, Behboudi H, Navasatli SA, Tavakoli S, Safavi M (2022) New insights into the inhibitory roles and mechanisms of D-amino acids in bacterial biofilms in medicine, industry, and agriculture. Microbiol Res 263:127107. https://doi.org/10.1016/j.micres.2022.127107. Epub 2022 Jun 30. PMID: 35843196
33. Mellini M, Di Muzio E, D'Angelo F, Baldelli V, Ferrillo S, Visca P, Rampioni G (2019) In silico selection and experimental validation of FDA-approved drugs as anti-quorum sensing agents. Front Microbiol 10:2355. https://doi.org/10.3389/fmicb.2019.02355
34. Escobar-Muciño E, Arenas-Hernández MMP, Luna-Guevara ML (2022) Mechanisms of inhibition of quorum sensing as an alternative for the control of E. coli and Salmonella. Microorganisms 10(5):884. https://doi.org/10.3390/microorganisms10050884. PMID: 35630329; PMCID: PMC9143355.
35. Brackman G, Nelis HJ, Coenye T (2012) Antimicrobial drug discovery: emerging strategies. CAB eBooks; Wallingford, UK. Inhibition of quorum sensing as a novel antimicrobial strategy; pp 115–134
36. Fleming A (1929) On antibacterial action of culture of Penicillium, with special reference to their use in isolation of B. influenzae. Br J Exp Pathol 10:226–236
37. Emmerich R, Löw O (1899) Bakteriolytische enzyme als Ursache der erworbenen Immunität und die Heilung von Infectionskrankheiten durch dieselben. Z Hyg 31:1–65. https://doi.org/10.1007/BF02206499
38. Dubern JF, Diggle SP (2008) Quorum sensing by 2-alkyl-4-quinolones in Pseudomonas aeruginosa and other bacterial species. Mol Biosyst 4(9):882–888. https://doi.org/10.1039/b803796p. Epub 2008 Jun 30. PMID: 18704225
39. Balouiri M, Sadiki M, Ibnsouda SK (2016) Methods for in vitro evaluating antimicrobial activity: a review. J Pharm Anal 6(2):71–79. https://doi.org/10.1016/j.jpha.2015.11.005. Epub 2015 Dec 2. PMID: 29403965; PMCID: PMC5762448
40. Hossain TJ (2024) Methods for screening and evaluation of antimicrobial activity: a review of protocols, advantages, and limitations. Eur J Microbiol Immunol (Bp). 14(2):97–115. https://doi.org/10.1556/1886.2024.00035. PMID: 38648108; PMCID: PMC11097785
41. Galán JE (2021) Salmonella Typhimurium and inflammation: a pathogen-centric affair. Nat Rev Microbiol 19:716–725. https://doi.org/10.1038/s41579-021-00561-4
42. Johnson CW, Ohashi M, Tang Y (2024) How fungi biosynthesize 3-nitropropanoic acid: the simplest yet lethal mycotoxin. Org Lett 26(15):3158–3163. https://doi.org/10.1021/acs.orglett.4c00758. Epub 2024 Apr 8. PMID: 38588324
43. Gorter K (1920) Hiptagin, a new glucoside from Hiptage madablota. Gaertu Bull Jordin Bot Buitenzorg 2:187–202
44. Nishino SF, Shin KA, Payne RB, Spain JC (2010) Growth of bacteria on 3-nitropropionic acid as a sole source of carbon, nitrogen, and energy. Appl Environ Microbiol 76(11):3590–3598. https://doi.org/10.1128/AEM.00267-10. Epub 2010 Apr 9. PMID: 20382807; PMCID: PMC2876434
45. Hamilton BF, Gould DH, Gustine DL (2000) History of 3-nitropropionic acid. In: Sanberg PR, Nishino H, Borlongan CV (eds) Mitochondrial inhibitors and neurodegenerative disorders. Contemporary neuroscience. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-59259-692-8_2
46. Pickova D, Ostry V, Toman J, Malir F (2021) Aflatoxins: history, significant milestones, recent data on their toxicity and ways to mitigation. Toxins (Basel) 13(6):399. https://doi.org/10.3390/toxins13060399. PMID: 34205163; PMCID: PMC8227755
47. Wu HC, Santella R (2012) The role of aflatoxins in hepatocellular carcinoma. Hepat Mon 12(10 HCC):e7238. https://doi.org/10.5812/hepatmon.7238. Epub 2012 Oct 11. PMID: 23162603; PMCID: PMC3496858
48. Zhu X, Zhao Y, Wu L, Gao X, Huang H, Han Y, Zhu T (2024) Advances in biosensors for the rapid detection of marine biotoxins: current status and future perspectives. Biosensors 14:203. https://doi.org/10.3390/bios14040203
49. O'Mahony M (2018) EU regulatory risk management of marine biotoxins in the marine bivalve mollusc food-chain. Toxins (Basel) 10(3):118. https://doi.org/10.3390/toxins10030118. PMID: 29534428; PMCID: PMC5869406

50. Mann ER, Lam YK, Uhlig HH (2024) Short-chain fatty acids: linking diet, the microbiome and immunity. Nat Rev Immunol 24:577–595. https://doi.org/10.1038/s41577-024-01014-8
51. Xiong RG, Zhou DD, Wu SX, Huang SY, Saimaiti A, Yang ZJ, Shang A, Zhao CN, Gan RY, Li HB (2022) Health benefits and side effects of short-chain fatty acids. Foods 11(18):2863. https://doi.org/10.3390/foods11182863. PMID: 36140990; PMCID: PMC9498509
52. Halfvarson J, Brislawn CJ, Lamendella R, Vázquez-Baeza Y, Walters WA, Bramer LM et al (2017) Dynamics of the human gut microbiome in inflammatory bowel disease. Nat Microbiol 2:17004–17004. https://doi.org/10.1038/nmicrobiol.2017.4
53. Deng Z, Luo XM, Liu J, Wang H (2020) Quorum sensing, biofilm, and intestinal mucosal barrier: involvement the role of probiotic. Front Cell Infect Microbiol 10:538077. https://doi.org/10.3389/fcimb.2020.538077. PMID: 33102249; PMCID: PMC7546212
54. Zillioux EJ (2015) Mercury in fish: history, sources, pathways, effects, and indicator usage. In: Armon R, Hänninen O (eds) Environmental indicators. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9499-2_42
55. Abbas N, Riaz S, Mazhar S et al (2023) Microbial production of docosahexaenoic acid (DHA): biosynthetic pathways, physical parameter optimization, and health benefits. Arch Microbiol 205:321. https://doi.org/10.1007/s00203-023-03666-x
56. Liu PY, Li G, Lin CB, Wu JJ, Jiang S, Huang FH et al (2022) Modulating DHA-producing Schizochytrium sp. toward astaxanthin biosynthesis via a seamless genome editing system. ACS Synth Biol 11:4171–4183. https://doi.org/10.1021/acssynbio.2c00490
57. Belanche A, Patra AK, Morgavi DP, Suen G, Newbold CJ, Yáñez-Ruiz DR (2021) Editorial: Gut microbiome modulation in ruminants: enhancing advantages and minimizing drawbacks. Front Microbiol 11:622002. https://doi.org/10.3389/fmicb.2020.622002. PMID: 33505388; PMCID: PMC7829182
58. Gayetri R, David B (2019) Editing the microbiome the CRISPR way. Phil Trans R Soc B37420180103
59. Gupta KH, Nowicki C, Giurini EF, Marzo AL, Zloza A (2021) Bacterial-based cancer therapy (BBCT): recent advances, current challenges, and future prospects for cancer immunotherapy. Vaccines (Basel) 9(12):1497. https://doi.org/10.3390/vaccines9121497. PMID: 34960243; PMCID: PMC8707929
60. Yang L, Wang Q, He L, Sun X (2024) The critical role of tumor microbiome in cancer immunotherapy. Cancer Biol Therapy 25(1). https://doi.org/10.1080/15384047.2024.2301801
61. Wynendaele E, Verbeke F, D'Hondt M, Hendrix A, Van De Wiele C, Burvenich C, Peremans K, De Wever O, Bracke M, De Spiegeleer B (2015) Crosstalk between the microbiome and cancer cells by quorum sensing peptides. Peptides 64:40–48. https://doi.org/10.1016/j.peptides.2014.12.009. Epub 2015 Jan 3. PMID: 25559405
62. Petrovska BB (2012) Historical review of medicinal plants' usage. Pharmacogn Rev 6:1. https://doi.org/10.4103/0973-7847.95849
63. Abdallah EM, Alhatlani BY, de Paula Menezes R, Martins CHG (2023) Back to nature: medicinal plants as promising sources for antibacterial drugs in the post-antibiotic era. Plants (Basel) 12(17):3077. https://doi.org/10.3390/plants12173077. PMID: 37687324; PMCID: PMC10490416
64. Mathis D, Benoist C (2011) Microbiota and autoimmune disease: the hosted self. Cell Host Microbe 10(4):297–301. https://doi.org/10.1016/j.chom.2011.09.007. PMID: 22018229
65. Dudics S, Langan D, Meka RR, Venkatesha SH, Berman BM, Che CT, Moudgil KD (2018) Natural products for the treatment of autoimmune arthritis: their mechanisms of action, targeted delivery, and interplay with the host microbiome. Int J Mol Sci 19(9):2508. https://doi.org/10.3390/ijms19092508. PMID: 30149545; PMCID: PMC6164747
66. Moudgil KD, Venkatesha SH (2022) The anti-inflammatory and immunomodulatory activities of natural products to control autoimmune inflammation. Int J Mol Sci 24(1):95. https://doi.org/10.3390/ijms24010095. PMID: 36613560; PMCID: PMC9820125
67. Motta JP, Wallace JL, Buret AG et al (2021) Gastrointestinal biofilms in health and disease. Nat Rev Gastroenterol Hepatol 18:314–334

68. Stecher B, Hardt WD (2011) Mechanisms controlling pathogen colonization of the gut. Curr Opin Microbiol 14:82–91
69. Zhou X et al (2024) Role of bitter contributors and bitter taste receptors: a comprehensive review of their sources, functions and future development. Food Sci Human Wellness 13(4):1806–1824
70. Jeon TI, Seo YK, Osborne TF (2011) Gut bitter taste receptor signalling induces ABCB1 through a mechanism involving CCK. Biochem J 438(1):33–37. https://doi.org/10.1042/BJ20110009. PMID: 21592089; PMCID: PMC3319131
71. Aino K, Hirota K, Okamoto T, Tu Z, Matsuyama H, Yumoto I (2018) Microbial communities associated with indigo fermentation that thrive in anaerobic alkaline environments. Front Microbiol 9:2196. https://doi.org/10.3389/fmicb.2018.02196. PMID: 30279681; PMCID: PMC6153312
72. Marazza JA, Garro MS, de Giori GS (2008) Aglycone production by Lactobacillus rhamnosus CRL981 during soymilk fermentation. Food Microbiol 26(3):333–339. https://doi.org/10.1016/j.fm.2008.11.004. Epub 2008 Dec 16. PMID: 19269578
73. Dias DA, Urban S, Roessner U. A historical overview of natural products in drug discovery. Metabolites 2(2):303–336. https://doi.org/10.3390/metabo2020303. PMID: 24957513; PMCID: PMC3901206
74. Cova TFGG, Pais AACC, Seixas de Melo JS (2017) Reconstructing the historical synthesis of mauveine from Perkin and Caro: procedure and details. Sci Rep 7:6806. https://doi.org/10.1038/s41598-017-07239-z
75. Bryan J (2011) Still going strong at 30: co-amoxiclav. Pharm J 286:762
76. Lu L, Li M, Yi G, Liao L, Cheng Q, Zhu J, Zhang B, Wang Y, Chen Y, Zeng M (2022) Screening strategies for quorum sensing inhibitors in combating bacterial infections. J Pharm Anal 12(1):1–14. https://doi.org/10.1016/j.jpha.2021.03.009. Epub 2021 Apr 2. PMID: 35573879; PMCID: PMC9073242
77. Zhou H, Zhang Y, Long CP et al (2024) A citric acid cycle-deficient Escherichia coli as an efficient chassis for aerobic fermentations. Nat Commun 15:2372. https://doi.org/10.1038/s41467-024-46655-4
78. Beavogui A, Lacroix A, Wiart N et al (2024) The defensome of complex bacterial communities. Nat Commun 15:2146. https://doi.org/10.1038/s41467-024-46489-0
79. Prüll C-R, Maehle A-H, Halliwell RF (2009) A short history of the drug receptor concept. Palgrave Macmillan, Basingstoke, pp 146–157
80. Guengerich FP (2022) Roles of cytochrome P450 enzymes in pharmacology and toxicology: Past, present, and future. Adv Pharmacol 95:1–47. https://doi.org/10.1016/bs.apha.2021.12.001. Epub 2022 Jul 18. PMID: 35953152; PMCID: PMC9869358
81. Gillam EM (2005) Exploring the potential of xenobiotic-metabolising enzymes as biocatalysts: evolving designer catalysts from polyfunctional cytochrome P450 enzymes. Clin Exp Pharmacol Physiol 32(3):147–152. https://doi.org/10.1111/j.1440-1681.2005.04165.x. PMID: 15743395
82. Kotlo K, Anbazhagan AN, Priyamvada S, Jayawardena D, Kumar A, Chen Y, Xia Y, Finn PW, Perkins DL, Dudeja PK, Layden BT (2020) The olfactory G protein-coupled receptor (Olfr-78/OR51E2) modulates the intestinal response to colitis. Am J Physiol Cell Physiol 318(3):C502–C513. https://doi.org/10.1152/ajpcell.00454.2019. Epub 2020 Jan 8. PMID: 31913697; PMCID: PMC7099522
83. Kinney JW, Bemiller SM, Murtishaw AS, Leisgang AM, Salazar AM, Lamb BT (2018) Inflammation as a central mechanism in Alzheimer's disease. Alzheimers Dement (N Y) 4:575–590. https://doi.org/10.1016/j.trci.2018.06.014. PMID: 30406177; PMCID: PMC6214864

Chapter 10
The Future of Natural Product Research. New Horizons

Abstract Natural products exist in a world of reductionism and complexity. Large complex systems are composed of simpler components although it is not clear that additive rules apply. The example of the similarity between light and electrons (wave:particle duality) suggests that as yet undiscovered convergences between components may exist. The biosynthesis of natural products can provide clues as to how living systems evolved and how control of living systems works on a cellular level. The concept of Dark Matter can be applied to biological systems and to natural products and may explain why metagenomic analysis suggests that the numbers of natural product biosynthetic genes seems to exceed the numbers of products isolated and characterised. Future directions will include further progress in the study of bioactive peptides and the discovery of new species and structures. The Nagoya protocol will help to provide a regulatory framework for natural product discovery and exploitation. Further research on extremophiles will uncover new natural products and pathways. The development of new techniques will pave the way for the real time investigation of the production of natural products and their interaction with living systems.

Keywords Natural product research · Priorities · New paradigms · Opportunities and challenges

Natural product research has made a significant contribution to the study of chemistry over many years. Nonetheless, there remains a number of problems and significant opportunities.

10.1 Reductionism and Complexity

The principles of reductionism in chemistry have been established for a considerable time and have their roots in alchemy. While the principles of mathematical reductionism were a key principle of the nineteenth and twentieth centuries, biological

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

B. Hanley, *Natural and Unnatural Product Chemistry*,
https://doi.org/10.1007/978-3-031-81698-7_10

reductionism has grown in importance and relevance as the molecular mechanisms of living systems (and the means to probe them) has become better understood however attempts to create an additive, natural world where every aspect can be reduced to a level where inspection and understanding becomes possible, has not been invariably successful. The paradigm of the 'reducibility' of living systems and are subject to a mechanical explanation was prominent in the seventeenth century. Studies of anatomy and, particularly, of the circulatory system gave added weight to the concept of a mechanical, rather than a vitalistic view of living systems. In essence, this form of reductionism suggests that all systems are equivalent to the sum of their parts. Depending on which scientific viewpoint being proposed, the individual parts can be examined with differing levels of granularity and this provides different levels of insight. For example, a social scientist may regard society as being composed of interacting individuals. If all such interactions can be quantified then society can be fully understood. A biologist may regard an individual living system as being composed of interacting cells and organs. Understanding these (and how they interact), will provide a complete picture of the whole living system. At this level, the system is a complex series of interrelated reactions that impact upon each other. In particle physics, the way nature behaves is postulated to be able to be understood in the context of fundamental particles and forces. This has led to the development of concepts like 'supersymmetry' to try to explain why things are the way they are and why, indeed, the universe exists. From the standpoint of chemistry, systems are made up of reactions and molecular interactions and this suggests that the key to understanding living systems is by understanding the chemical components and how they interact. The definition of natural products encompasses this concept. Traditionally natural products have been generally considered to be organic molecules produced by living systems. While this simplistic view is suitable in most cases, it does not adequately describe the contribution of inorganic natural products. Advances in molecular biology and the development of tools to modify and transfer genes between species has also added a further layer of complexity. If a natural product is present in a living system as a result of the transfer of a heterologous biosynthetic pathway from another organism then should that product be considered 'natural' or 'unnatural'. Similarly, if a gene is altered or constructed in a way that does not occur naturally, is the protein produced 'natural' or 'unnatural'. Clearly the demarcation line between natural and unnatural products is indistinct.

Within each of the world views noted previously is the conviction that large and complex systems are composed of smaller components. The question for each scientific discipline is what is the most relevant and useful scale to work at? This must, in part, be a pragmatic choice that takes into account how much information is required to develop meaningful paradigms. While the ultimate reductionist view is that the world can be expressed purely in terms of equations and mathematical interpretations, may be ultimately correct, it is, in fact, nothing more than a philosophical artifice and, for most cases, is insoluble. One of the other arguments against this ultra-reductionist world view is that it implies that determinism is paramount. If everything can be reduced to equations (or even to fundamental particles) and these equations can be solved then nothing truly novel or new can ever arise. This then frames an

argument against a purely and absolutely reductionist view—interactions (and the consequences and results that arise from them) cannot be completely quantified. This view suggests that new concepts 'emerge' from interactions within complex systems in ways that cannot be predicted. One way to consider such systems—without resorting to suggesting that there is something uniquely special about living systems (vitalism)—is to use the analogy of pointillism. The principle of pointillism is that a picture can be created from small dots of colour. At the level of the individual dots, the picture is not revealed. By standing further away, the picture is able to be viewed. This analogy, although limited, suggests that, rather than a perpetual quest for reductionism and to consider things as ever diminishing entities, there can be an optimal scale at which the properties of a living system are revealed more effectively. The idea that systems can be fully defined and precise is contradicted by much of what we know and understand—not least from concepts developed to support the quantum nature of the universe. The wave:particle duality of light has been demonstrated, for example [1] and, what is more, the nature of light depends upon the experimental conditions under which it is measured—it can either be quantised and act as a particle or act as a wave and, for example, demonstrate interference. This is the opposite of a completely defined and characterised system. Light is crucial to the functioning of biological systems—particularly those that employ photosynthesis. Being unable to absolutely define light except in an ambiguous fashion, means that biological systems will be partly obscured. In another example, much of what we understand about how chemical reactions function is predicated on the properties of electrons and how these properties affect reactivity, bonding and how molecules are created. Indeed, simple chemical bonding has been described as the sharing of electrons. Rather than electrons being discrete particles, atomic orbitals are thought of as wavefunctions that describe the probability of the position of an electron. If electrons are thought of in terms of a probability distribution then the interaction of electrons to form a bond is also a probability distribution. A chemical reaction is not the simple sharing of electrons from two (or more) different atoms but is a merging of probability distributions.

Given the difficulty in conceptualising chemical reactions in absolute terms or of understanding the nature of light and the details of precisely how it might interact with living systems, the concept of suitability of scale is valid. In typical mechanistic chemistry, the interaction of atoms to form a bond is conceptualised in terms of the movement and sharing of electrons. Using this model, electrons can be envisaged as discrete entities rather than a probability function. In a similar way, while molecular interactions and chemical reactions can be thought of as the basis for every biological interaction, trying to reduce complex living systems to a series of interlinked reactions may not be the most effective way of visualising the effects of natural products. It is important to recognise both the importance of molecular interactions (for example the binding of products to receptors) and the systemic consequences that occur as a result and may not be predictable even if the initial state and some details of the consequences of the initial interaction are known. Just as living systems can be examined using a reductionist approach (for example by using techniques like CRISPR to selectively silence certain genes), by determining which parts of a natural

product are important for a given observed activity, it is possible to probe further how complex living systems actually operate. This is an important aspect of the use of structures that do not occur naturally (unnatural product chemistry). In contrast, the structures of natural products and the consequences of the attributes of those structures have built up in conjunction with living systems. By determining which structures occur and which ones do not occur, it is possible to determine the lineage of evolution of living systems based on the types of natural products they produce. For example 3-nitropropanoic acid is a fungal natural product. The corresponding 4-nitrobutanoic acid does not occur. This may be a consequence of the absence of suitable starting materials (aspartate is the fungal precursor and the use of just two enzymes—an amine oxidase and a decarboxylase are sufficient for the biosynthesis of the product. However, while the homologue of aspartic acid—glutamic acid—occurs naturally, it is not converted to 4-nitrobutanoic acid), or the specificity of the biosynthetic enzymes. Additionally, unlike 3-nitropropanoic acid, 4-nitrobutanoic acid cannot operate as an inhibitor of succinate dehydrogenase so there is no growth or survival advantage in the production of 4-nitrobutanoic acid. It is a biological cul de sac. The evolutionary approach has also been used to design 'pseudo natural products'. These are defined as the product of fragment-based molecular design to derive new compounds from the arrangement of natural product fragments in novel ways [2]. This approach addresses two important constraints that limit the realisation of the potential of natural products. The first of these is the plethora of compounds that have not, as yet, been discovered. The second is that the evolution of novel natural product structures is a slow process and attempting to accelerate this has been challenging. A range of approaches have been used to create natural product compound libraries in vitro however these all have limitations both in the numbers and uniqueness of the structures that can be created. Nonetheless, by setting specific structural properties, it is possible to build a paradigm that will create some new structures. The 'pseudo natural product' approach starts by defining 2000 natural product fragment groups [3] and then to combine those fragments in novel ways to create unnatural products with new structures and properties. Not only might these properties be relevant to known biological targets, they may also open up new biological space for interactions. Although such constructed libraries have met with some success, a range of novel structures of mixed biosynthetic origin are also known in nature—for example terpene alkaloids [4] and polyketides of mixed biosynthetic origin [5] have both been described. Such molecules introduce additional complexity into natural product chemistry but, unlike the 'pseudo natural products' described above, they are fully derived from nature and hence are completely natural products.

10.2 Heterogeneity of Biosynthesis of Natural Products

There are a number of aspects to the heterogenous production of natural products. As discussed previously, for example, the biosynthetic pathway that leads to production of glucosinolates is sufficiently non-specific that a range of final products occur

without the necessity of a complete and unique set of biosynthetic enzymes for each one. Another example of heterogeneity is the production of a natural product in one part of a species and its subsequent transportation to another part. This occurs, for example, with the biosynthesis of theanine in tea (*Camellia sinensis*) occurring in the roots then being transported to leaves. The control of the biosynthesis is at the cellular level [6]. In addition, this means that the expression of the biosynthetic genes and the controlling transcription factors does not need to occur within the same gene cluster.

The concept of making use of novel substrates to develop new natural products relies upon a lack of substrate specificity in biosynthetic enzymes. This may at first seem unexpected since enzymes often demonstrate exquisite specificity giving rise, for example, to specific enantiomers. The manipulation of enzymes has been noted (Chap. 7) and the flexibility of some biosynthetic enzymes and pathways has been demonstrated by transgenic expression of genes and the accumulation of natural products. One of the most important and successful approaches to the optimisation of natural product formation has been the use of heterologous expression systems and non-natural hosts. The process begins with the identification of biosynthetic gene clusters (BGCs)—some of which may be silent (not expressed). Silent or cryptic BGCs can be thought of as one of the natural product equivalents of Dark Matter. Dark Matter which exists but only by inference, is thought to make up most of the matter in the universe [7]. In an analogous fashion, natural product Dark Matter can be thought of as BGCs whose products cannot be detected because the gene clusters themselves are inactive and therefore do not create any product. The implication is that it is not sufficient to be able to find and quantify the BGCs—their products must also be able to be found and quantified. The BGCs can be cloned into a suitable host and then activated to produce the BGC product (proteins comprising the biosynthetic pathway) and, eventually, desired natural product in optimal yield [8]. A key aspect of a successful procedure is that the host strain should be metabolically and genetically characterised. In addition the host strain should be capable of rapid growth, have an adequate supply of the required biosynthetic precursor and be able to be genetically manipulated (be able to be transfected with the BGC of interest and to express its gene products). It is also preferable if the background levels of natural products in the host species is low so the unnatural product is produced without significant other metabolites being present which may compete for cellular resources or present purification challenges. In some cases, competing biosynthetic pathways are removed through manipulation of competing BGCs using, for example, CRISPR technology. In summary, the development of heterologous host strains is a major advance in the optimisation of production of natural products in an unnatural production system.

The use of BGCs using novel substrates to produce unnatural products is an alternative way of looking at the heterogeneity of biosynthesis of unnatural products. In order to realise the potential of unrecognised BGCs, it is important to quantify the BGCs that are present over a number of species [9]. Such studies may also suggest taxonomic niches with the greatest potential for novel natural product formation. It is noteworthy that, according to data mining, there are more BGCs than there are known natural products and this also suggests that there is considerable potential for

the discovery of novel chemical structures. It also indicates that microbial species have the potential to produce considerable chemical diversity and this may be realised both by discovering new BGCs and by capitalising on the lower substrate specificity of many natural product BGCs thereby enabling the biosynthesis of novel molecules [10].

10.3 Natural Products and Biological Systems

The reductionist approach to natural products has provided many insights into how and why these non-essential compounds are synthesised by living systems. There is a considerable history of biological systems being used for the synthesis of natural products that are then put to a diverse number of uses including, perhaps most significantly, as beneficial bioactive agents [11]. While synthetic chemistry has evolved and increasingly sophisticated methods have been used to prepare a range of bioactive agents, recent advances in systems biology has led to an increase in the use of natural systems as 'cellular factories' for the production of biopharmaceuticals and other products [12]. This has, in large part, derived from the development of 'combinatorial biosynthesis' [13] in which enzymes and pathways are manipulated in order to facilitate the production of natural products with desired properties.

The innate functions of natural products and their synthesis is not necessarily in line with (or even related to) their function as useful, isolated entities. The chemical evolution of natural products has been studied at length and a number of hypotheses leading to paradigms have been suggested [14]. A number of theories have arisen regarding the evolution of natural products. One concept is that biosynthetic enzymes for natural products are natural mutations of enzymes that are involved in primary metabolism or arose from the modification of chemical structures by core enzymes involved in primary metabolism. This may, in part, answer the question of how natural products were formed. The next question is why did particular products persist? Most evidence suggests that natural products act as messengers to the environment that organisms find themselves in and that they confer an advantage (for example, as feeding deterrents). This will create a degree of selective advantage for the resource and energy expenditure that the organism must commit to the production of a given natural product. The biological action of natural products does not correlate with their effects on human physiology. Any medicinal or pharmacological benefit to humans (or other consuming species) is, therefore, purely coincidental. Natural products are, therefore, compounds that convey a specific message from the producing organism to the environment and may also operate within the producing organism. For example, quorum sensing molecules are produced by many organisms and have an endogenous use. In addition, the production of some molecules is important for the primary functioning of cells. Purines and pyrimidines are essential components of nucleic acids. Some sugars and polysaccharides (prebiotics) are able to be used as a metabolizable carbon source by gut bacteria. Microbes in various locations in living systems (microbiomes) have an effect on the immune system of the host and have an effect on

health and disease [15] and on other functions. The complexity of inter-organismal interactions can mean that discerning the effects of individual natural products can be challenging. It is clear, however, that the impact and role of natural products is different in different species and they are part of the ecology of living systems.

10.4 The Future Directions of Natural Products

The future of investigations is likely to involve three, partially overlapping areas—the discovery of new structures, the delineation of new biological activities and novel interactions with living systems.

Structure determination (and biosynthesis) is part of the tradition of natural product chemistry. As has been discussed, the isolation of natural products remains challenging. This is in part because of the large number of structures present and the presence of comparatively small amounts of potential target compounds. Unlike primary metabolites, many secondary natural products are only present in very small amounts. In addition, part of the impetus driving the discovery of new compounds is the opportunity to find novel biological activity. A significant number of isolated compounds can be defined as 'biosimilars' and are of limited commercial value. Nonetheless at least some of these can be used as starting points for the partial synthesis of novel (or more physiologically stable) bioactive entities. Computational methods (including metagenomics) have also been used to investigate natural product 'Dark Matter' and to probe unexpressed biosynthetic pathways.

One area of natural product chemistry that has been comparatively less exploited is that of bioactive peptides and proteins [16]. Three of the most important classes of bioactive peptides are antimicrobial peptides, ACE-inhibitory peptides, and antioxidative peptides and all three have been the target for isolation and purification. Two of the major obstacles to the use of protein-based therapeutics are their comparatively low oral bioavailability and their lack of stability to endogenous proteases. Smaller, peptide bioactive molecules are more bio-effective. A range of bioactive peptides have been isolated from natural sources—these include several antimicrobial compounds and post-translationally modified compounds such as lantibiotics). In some cases the peptides are found intact in the natural source while in other cases they are released after hydrolysis. The study of the library of bioactive peptides has been termed 'peptidomics'—by analogy with proteomics. In common with other natural products, bioactive peptides are often present in very small amounts and activity guided fractionation coupled to identification using chromatographic methods linked to mass spectrometry is commonly used. The suite of specific endo or exopeptidases are used to hydrolyse natural material to release bioactive peptides or other compounds with desired biological or physicochemical properties. The discovery, use and development of protein and peptides requires a range of specialised techniques and these are frequently used in conjunction with oneanother.

One example of a dipeptide of importance to human health is the mycotoxin, Ochratoxin A. Ochratoxin A is the most commonly found member of a family of

similar mycotoxins produced by *Aspergillus* and *Penicillium* species which grow on some foodstuffs. Its target organ is the kidney and it is, therefore, significantly nephrotoxic. It is also classified as a potential human carcinogen. It was first discovered in corn meal in 1965 [17] and has been extensively studied since then. Chemically, ochratoxin A is a dimer of L-phenylalanine and isocoumarin and the two molecules are linked by an amide bond. Ochratoxin A appears to have a number of modes of toxicity and these have been investigated primarily in animal models. While the simplest (and best) method of preventing Ochratoxin A contamination is by ensuring that crops are harvested and stored in such a way that fungal growth does not occur (ensuring grain crops are kept dry is an effective method of preventing fungal contamination), irradiation can be effective in preventing formation or degrading Ochratoxin A as can heat treatment (although this may affect the properties of the crop with a concomitant deleterious affect on quality or processing parameters), adsorption, chemical and biological (including the use of amidases) methods have all been used—with varying degrees of success [18]. The possibility of using control of quorum sensing to modulate Ochratoxin A production has been suggested [19]. Such an approach has been suggested as a more general control system for mycotoxin production, either directly through the degradation of specific mycotoxins such as patulin by lactonase secreted by bacteria [20], or through interruption of quorum sensing pathways [21] or other related control methods [22].

In addition to the discovery (and sometimes reclassification) of new natural products, another driver will be the discovery of novel biological activities that can be ascribed to natural products. One of the challenges to the successful commercialisation of natural products is the Nagoya protocol [23]. This seeks to prevent the unfair exploitation of the countries' of origin of biological material. Certain pathways will be favoured as targets for the discovery of novel biologically active natural products. This choice will be driven by the availability of test and screening methods and the size of the potential market. This will mean that biological activity associated with wealthier countries is likely to be prioritised over, for example, antimicrobial compounds or bioactive approaches to diseases that affect poorer populations. Nonetheless, there are some initiatives that focus on the potential of natural products as sources of novel bioactive molecules [24]. One such area, which is a focus in many areas of the world due to population demographics and the increased numbers of ageing members of society, is age-related cognitive decline and, in particular, Alzheimer's disease (AD). Two areas of particular interest are anti-oxidants (to prevent neural oxidative damage) and natural acetyl cholinesterase inhibitors such as galantamine. Much of the most promising work in this area relies upon ethnopharmacological approaches and traditional medicines [25]. An alternative approach treats AD as a multi-factorial disease with inflammation being a key component [26, 27] and solutions including the use of dithiocarbamates as prodrugs for the administration of hydrogen sulphide have been suggested [28].

The idea of known natural products being a source of novel biological activity has been investigated [29]. The use of advanced techniques driven by Artificial Intelligence has enabled new uses for existing natural products to be more effectively explored [30].

10.5 Synthetic Genes

The development of chemical and biological methods for nucleic acid manipulation has made the de novo synthesis of genes a viable possibility. Gene synthesis has a number of advantages. First of all, unlike the biological synthesis of DNA, it is not necessary to have a template of single strand DNA to work from. In addition, precise sequences can be created without the errors that occur in biological systems such as in the use of Taq and other polymerases [31]. The synthesis of DNA can be a crucial aspect of the development of synthetic biology [32] and to the creation of new biosynthetic pathways. The synthetic manipulation of silent natural product biosynthetic gene clusters such that they are expressed is a key aspect of synthetic biology as applied to natural products [33]. Ultimately synthetic genes may be used in the creation of wholly artificial cells with designed properties [34] and for the synthesis of unnatural products.

10.6 Extremophiles

Extremophiles are organisms that survive (and in some cases, thrive) in extreme environments. These include high pressure, high or low pH, high levels of salt (or other ionic species), radiation and both high and low temperatures [35]. Among the most significant products from extremophiles are the thermostable nucleic acid polymerases. While the use of enzymes such as Taq polymerase has revolutionised molecular biology and forensic science (and made considerable amounts of money), other thermostable enzymes have been used in a range of other areas including biofuel production and in other areas of biotechnology [36]. The stability of extremophiles to more robust biotechnological conditions may make them an increasingly important resource for synthetic biology and biotechnology. In terms of mechanisms to be able to survive under extreme conditions, thermostable enzymes make use of adjunct molecules such as heat shock proteins and chaperones to maintain their structure, stability and activity. In addition extremophiles of different types are able to produce a range of molecules to enable them to survive in extreme conditions including ice nucleating proteins, pigments, extremozymes and exopolysaccharides.

In addition to maintaining their cellular integrity and macromolecular integrity, some extremophiles produce natural products with novel properties. For example, the cyclic peptides, mixirins A, B and C were found in a *Bacillus* strain found in the Arctic. Some extremophiles are found in ecological niches that have become contaminated. Rather than isolate compounds from individual species found in extreme environments, metagenomic analysis of samples of whole communities has facilitated the discovery of extremozymes and of DNA libraries that can be expressed in host organisms and the properties of the proteins investigated in more detail. Among the targets for extremophile natural products are anti-viral and anti-cancer compounds and those with novel neurological activity.

While extremophile microbial species (including fungi) have been extensively studied and a range of compounds isolated including polyketides and others with novel biological activity, extremophile plant species have been less well characterised. Of these, the most commonly examined are halophytes and these have been found to contain a range of natural products including, in particular phenolics [37]. While secondary natural products are not usually essential to plant growth, they are sometimes produced in higher amounts in response to stress [38] so it might be expected that both the amounts and types of secondary natural products found in plants would be elevated in plants grown under stress conditions.

10.7 New Techniques and Methods

The development of natural product chemistry has both led and followed technological developments. While the direction of technological progress may be difficult to predict, some areas of possible progress in natural product chemistry are possible to discern.

One of the areas that has attracted growing interest is understanding, in real time, how living systems operate. Many studies have relied upon either taking the cellular components (receptors, enzymes etc.) and measuring their interaction with natural products or measuring the cellular consequences (production of specific components such as inflammatory cytokines, stress response factors etc.) of the interactions of natural product. There have been some biomarker and some DNA-based techniques developed [39] and these have been proposed as being able to measure cellular events in real time. Despite the development of these and other techniques, end point rather than real time assays remain more common and technological advances remain to be implemented.

The second area that contains significant potential for advances is in the characterisation of unknown living species and unknown (unexpressed) biosynthetic pathways. In the case of the latter, genome mining for unknown products and pathways have been developed [40]. The development of in silico methods will facilitate the identification of putative structures and potential products [41]. This is likely to be an area of significant advances. In the case of unknown species, the majority of microbial species are not culturable and will be accessible by genome mining (including in ecological niches that harbour extremophiles [42]. A number of techniques are available for genome mining to identify unknown species and, as more rapid and accurate methods arise, these will be applied in this area. In the case of plants and fungi, it has been estimated that over 2000 new species of each are identified and characterised each year and some of these will contain novel natural products. The discovery of new species should be seen in the context of there being over 350,000 identified plant species so new members are relatively rare. Nonetheless, at least some of these will likely be a valuable resource for natural product chemists.

The application of quantum methods in the area of natural products has been noted. Quantum mechanical calculations have been used in the study of natural product

biosynthesis. In addition computational methods have been used for the discovery and structure elucidation of natural products [43]. Quantum biology (coupled to chronobiology) has the potential to develop into an interesting new area [44] and the development of quantum techniques in living systems will facilitate new ways of looking at the interactions between natural products and complex living systems [45].

10.8 Conclusion

Natural products (and their unnatural analogues) have played a pivotal role in the development of chemistry, science and health. The philosophy of reductionism which has underpinned approaches to natural products has given chemists a view of how molecules are related to each other and to the living world. Attempts to apply the same principles of reductionism to the study of how natural products interact with living systems have demonstrated the limitations of an approach that relies upon simplification. While there have been significant advances including studies of specific molecules (Chap. 3), biosynthesis (Chap. 4), bioengineering (Chaps. 5 and 6), enzymes and enzyme systems (Chap. 7), receptors (Chap. 8) and microbial communities (Chap. 9), it has been clear that complex biological systems remain, in some ways, opaque to investigation.

While traditional natural product chemistry has, after a period of explosive growth, become less well studied, there are still both traditional (discovery chemistry) and novel (real time interactions with living systems and quantum biology) approaches that will increase both knowledge about the living systems on the planet and require the development of better and more appropriate methodologies. In the future, natural product chemistry will remain driven by curiosity about the world and the development of greater understanding about how different components of living systems interact and enhance each other.

References

1. Young T (1804) Phil Trans R Soc 94:1–16
2. Grigalunas M, Brakmann S, Waldmann H (2022) Chemical evolution of natural product structure. J Am Chem Soc 144(8):3314–3329. https://doi.org/10.1021/jacs.1c11270. Epub 2022 Feb 21. PMID: 35188375; PMCID: PMC8895405
3. Over B, Wetzel S, Grütter C, Nakai Y, Renner S, Rauh D, Waldmann H (2013) Natural-product-derived fragments for fragment-based ligand discovery. Nat Chem 5:21. https://doi.org/10.1038/nchem.1506
4. Cherney EC, Baran PS (2011) Terpenoid-alkaloids: their biosynthetic twist of fate and total synthesis. Isr J Chem 51(3–4):391–405. https://doi.org/10.1002/ijch.201100005. PMID: 26207071; PMCID: PMC4508874

5. Contigli C, Valle MS, Oloris SCS, Pimenta LPS, Takahashi JA (2023) Polyketides from fungi. In: Carocho M, Heleno SA, Barros L (eds) Natural secondary metabolites. Springer, Cham. https://doi.org/10.1007/978-3-031-18587-8_17
6. Lin S, Zhang Y, Zhang S, Wei Y, Han M, Deng Y, Guo J, Zhu B, Yang T, Xia E, Wan X, Lucas WJ, Zhang Z (2024) Root-specific theanine metabolism and regulation at the single-cell level in tea plants (Camellia sinensis). Elife 13:RP95891. https://doi.org/10.7554/eLife.95891. PMID: 39401074; PMCID: PMC11473105
7. Randall L (2018) What is Dark Matter? Nature 557(7704):S6–S7. https://doi.org/10.1038/d41586-018-05096-y. PMID: 29743708
8. Kang HS, Kim ES (2021) Recent advances in heterologous expression of natural product biosynthetic gene clusters in Streptomyces hosts. Curr Opin Biotechnol 69:118–127. https://doi.org/10.1016/j.copbio.2020.12.016. Epub 2021 Jan 11. PMID: 33445072
9. Wei B, Du AQ, Zhou ZY, Lai C, Yu WC, Yu JB, Yu YL, Chen JW, Zhang HW, Xu XW, Wang H (2021) An atlas of bacterial secondary metabolite biosynthesis gene clusters. Environ Microbiol 23(11):6981–6992. https://doi.org/10.1111/1462-2920.15761. Epub 2021 Sep 21. PMID: 34490968
10. Rokas A, Mead ME, Steenwyk JL, Raja HA, Oberlies NH (2020) Biosynthetic gene clusters and the evolution of fungal chemodiversity. Nat Prod Rep 37(7):868–878. https://doi.org/10.1039/c9np00045c. Epub 2020 Jan 3. PMID: 31898704; PMCID: PMC7332410
11. Nasim N, Sandeep IS, Mohanty S (2022) Plant-derived natural products for drug discovery: current approaches and prospects. Nucleus (Calcutta) 65(3):399–411. https://doi.org/10.1007/s13237-022-00405-3. Epub. PMID: 36276225; PMCID: PMC9579558
12. Dzobo K (2022) The role of natural products as sources of therapeutic agents for innovative drug discovery. In: Comprehensive pharmacology, pp 408–422. https://doi.org/10.1016/B978-0-12-820472-6.00041-4. Epub 2022 Jun 9. PMCID: PMC8016209
13. Kim E, Moore B, Yoon Y (2015) Reinvigorating natural product combinatorial biosynthesis with synthetic biology. Nat Chem Biol 11:649–659. https://doi.org/10.1038/nchembio.1893
14. Katz L, Baltz RH (2016) Natural product discovery: past, present, and future. J Ind Microbiol Biotechnol 43(2–3):155–176. https://doi.org/10.1007/s10295-015-1723-5
15. Zheng D, Liwinski T, Elinav E (2020) Interaction between microbiota and immunity in health and disease. Cell Res 30:492–506. https://doi.org/10.1038/s41422-020-0332-7
16. Purohit K, Reddy N, Sunna A (2024) Exploring the potential of bioactive peptides: from natural sources to therapeutics. Int J Mol Sci 25(3):1391. https://doi.org/10.3390/ijms25031391. PMID: 38338676; PMCID: PMC10855437
17. van der Merwe KJ, Steyn PS, Fourie L, Scott DB, Theron JJ (1965) Ochratoxin A, a toxic metabolite produced by Aspergillus ochraceus Wilh. Nature 205(976):1112–1113. https://doi.org/10.1038/2051112a0. PMID: 5833211
18. Ding L, Han M, Wang X, Guo Y (2023) Ochratoxin A: overview of prevention, removal, and detoxification methods. Toxins (Basel) 15(9):565. https://doi.org/10.3390/toxins15090565. PMID: 37755991; PMCID: PMC10534725
19. Gao J, Liu H, Jin Y et al (2024) Glucose and HODEs regulate Aspergillus ochraceus quorum sensing through the GprC-AcyA pathway. Cell Mol Life Sci 81:241. https://doi.org/10.1007/s00018-024-05160-z
20. Dor S, Nudel K, Eagan JL, Cohen R, Hull CM, Keller NP, Prusky D, Afriat-Jurnou L (2024) Bacterial-fungal crosstalk is defined by a fungal lactone mycotoxin and its degradation by a bacterial lactonase. Appl Environ Microbiol 90:e00299-e324
21. Affeldt KJ, Brodhagen M, Keller NP (2012) Aspergillus oxylipin signaling and quorum sensing pathways depend on g protein-coupled receptors. Toxins (Basel) 4(9):695–717. https://doi.org/10.3390/toxins4090695. Epub 2012 Sep 18. PMID: 23105976; PMCID: PMC3475224
22. Bacon CW, Hinton DM, Mitchell TR (2017) Is quorum signaling by mycotoxins a new risk-mitigating strategy for bacterial biocontrol of Fusarium verticillioides and other endophytic fungal species? J Agric Food Chem 65(33):7071–7080
23. Burton G, Evans-Illidge EA (2014) Emerging R and D law: the Nagoya Protocol and its implications for researchers. ACS Chem Biol 9(3):588–591. https://doi.org/10.1021/cb500045t. PMID: 24650394

24. Atanasov AG, Zotchev SB, Dirsch VM (2021) International Natural Product Sciences Taskforce; Supuran CT. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov 20(3):200–216. https://doi.org/10.1038/s41573-020-00114-z. Epub 2021 Jan 28. PMID: 33510482; PMCID: PMC7841765
25. Bordoloi S, Pathak K, Devi M et al (2024) Some promising medicinal plants used in Alzheimer's disease: an ethnopharmacological perspective. Discov Appl Sci 6:215. https://doi.org/10.1007/s42452-024-05811-7
26. Hanley AB (2021) Alzheimer's disease—a clear and future epidemic. Sci Parliament 77(1):27
27. Olajide OA, Sarker SD (2020) Alzheimer's disease: natural products as inhibitors of neuroinflammation. Inflammopharmacol 28:1439–1455. https://doi.org/10.1007/s10787-020-00751-1
28. The protective effects of dietary isothiocyanates against cancer—a delivery mechanism for hydrogen sulfide. 2020. White Paper. Biophys Ltd. Hanley A B
29. Hanley AB, Graz M (2023) Drug repurposing—new wine in old bottles. Sci Parliament 79:9
30. Thomford NE, Senthebane DA, Rowe A, Munro D, Seele P, Maroyi A, Dzobo K (2018) Natural products for drug discovery in the 21st century: innovations for novel drug discovery. Int J Mol Sci 19(6):1578. https://doi.org/10.3390/ijms19061578. PMID: 29799486; PMCID: PMC6032166
31. Potapov V, Ong JL (2017) Examining sources of error in PCR by single-molecule sequencing. PLoS One 12(1):e0169774. https://doi.org/10.1371/journal.pone.0169774. Erratum in: PLoS One. 2017 Jul 6;12(7):e0181128. https://doi.org/10.1371/journal.pone.0181128. PMID: 28060945; PMCID: PMC5218489
32. Hughes RA, Ellington AD (2017) Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. Cold Spring Harb Perspect Biol 9(1):a023812. https://doi.org/10.1101/cshperspect.a023812. PMID: 28049645; PMCID: PMC5204324
33. Tang X (2023) Synthetic biology to revive microbial natural product discovery. mLife 2(2):123–125. https://doi.org/10.1002/mlf2.12071. PMID: 38817621; PMCID: PMC10989979
34. Xu C, Hu S, Chen X (2016) Artificial cells: from basic science to applications. Mater Today (Kidlington) 19(9):516–532. https://doi.org/10.1016/j.mattod.2016.02.020. PMID: 28077925; PMCID: PMC5222523
35. Rekadwad BN, Li WJ, Gonzalez JM, Punchappady Devasya R, Ananthapadmanabha Bhagwath A, Urana R, Parwez K. Extremophiles: the species that evolve and survive under hostile conditions. 3 Biotech 13(9):316. https://doi.org/10.1007/s13205-023-03733-6. Epub 2023 Aug 25. PMID: 37637002; PMCID: PMC10457277
36. Hatti-Kaul R, Abouhmad A (2020) Extremophiles: a promising source of novel natural products. In: Lee NM (ed) Biotechnological applications of extremophilic microorganisms. (Life in Extreme Environments; vol 6). De Gruyter, pp 1–42. https://doi.org/10.1515/9783110424331-001
37. Loukas A, Kappas I, Abatzopoulos TJ (2018) HaloDom: a new database of halophiles across all life domains. J of Biol Res-Thessaloniki 25:2. https://doi.org/10.1186/s40709-017-0072-0
38. Salam U, Ullah S, Tang ZH, Elateeq AA, Khan Y, Khan J, Khan A, Ali S (2023) Plant metabolomics: an overview of the role of primary and secondary metabolites against different environmental stress factors. Life (Basel) 13(3):706. https://doi.org/10.3390/life13030706. PMID: 36983860; PMCID: PMC10051737
39. Sheth RU, Wang HH (2018) DNA-based memory devices for recording cellular events. Nat Rev Genet 19:718–732. https://doi.org/10.1038/s41576-018-0052-8
40. Yee DA, Niwa K, Perlatti B et al (2023) Genome mining for unknown–unknown natural products. Nat Chem Biol 19:633–640. https://doi.org/10.1038/s41589-022-01246-6
41. Lee YY, Guler M, Chigumba DN et al (2023) HypoRiPPAtlas as an Atlas of hypothetical natural products for mass spectrometry database search. Nat Commun 14:4219. https://doi.org/10.1038/s41467-023-39905-4
42. Chen YH, Chiang PW, Rogozin DY et al (2021) Salvaging high-quality genomes of microbial species from a meromictic lake using a hybrid sequencing approach. Commun Biol 4:996. https://doi.org/10.1038/s42003-021-02510-6

43. Nugroho AE, Morita H (2019) Computationally-assisted discovery and structure elucidation of natural products. J Nat Med 73:687–695. https://doi.org/10.1007/s11418-019-01321-8
44. Mazzoccoli G (2022) Chronobiology meets quantum biology: a new paradigm overlooking the horizon? Front Physiol 13:892582. https://doi.org/10.3389/fphys.2022.892582. PMID: 35874510; PMCID: PMC9296773
45. Tantillo DJ (2013) Walking in the woods with quantum chemistry--applications of quantum chemical calculations in natural products research. Nat Prod Rep 30(8):1079–1086. https://doi.org/10.1039/c3np70028c. Epub 2013 Jun 24. PMID: 23793561