from Biosynthesis to total Synthesis

Strategies and Tactics for Natural Products

Edited by

ALEXANDROS L. ZOGRAFOS

WILEY

FROM BIOSYNTHESIS TO TOTAL SYNTHESIS

FROM BIOSYNTHESIS TO TOTAL SYNTHESIS

Strategies and Tactics for Natural Products

Edited by

ALEXANDROS L. ZOGRAFOS
Aristotle University of Thessaloniki, Greece

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data

Names: Zografos, Alexandros L., editor.
Title: From biosynthesis to total synthesis : strategies and tactics for natural products / edited by Alexandros L. Zografos.
Description: Hoboken, New Jersey : John Wiley & Sons, Inc., [2016] | Includes bibliographical references and index.
Identifiers: LCCN 2015037375 (print) | LCCN 2015047240 (ebook) | ISBN 9781118751732 (cloth) | ISBN 9781118753569 (Adobe PDF) | ISBN 9781118753637 (ePub)
Subjects: LCSH: Organic compounds–Synthesis. | Biosynthesis.
Classification: LCC QD262 .F76 2016 (print) | LCC QD262 (ebook) | DDC 572/.45–dc23
LC record available at http://lccn.loc.gov/2015037375

Set in 10/12pt Times by SPi Global, Pondicherry, India

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Dedicated to my mother, father and wife

CONTENTS

LIST OF CONTRIBUTORS

Elissavet E. Anagnostaki, Department of Chemistry, Laboratory of Organic Chemistry, Aristotle University of Thessaloniki, Thessaloniki, Greece and Research and Development Department, Pharmathen S.A., Thessaloniki, Greece

Stellios Arseniyadis, School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom

Louis Barriault, Department of Chemistry, University of Ottawa, Ottawa, Ontario, Canada

Erica Benedetti, Laboratoire de Chimie et Biochimie et Pharmacologiques et Toxicologiques, CNRS-Université Paris Descartes Faculté des Sciences Fondamentales et Biomédicales, Paris, France

Sebastian Brauch, Institute for Molecules and Materials, Radboud University Nijmegen, Nijmegen, The Netherlands

Rolf Breinbauer, Institute of Organic Chemistry, Technische Universität Graz, Graz, Austria

Cyril Bressy, Aix Marseille Université, Centrale Marseille, CNRS, Marseille, France

Martin Cordes, Institute for Organic Chemistry and Center of Biomolecular Drug Research (BMWZ), Leibniz Universität Hannover, Hannover, Germany and Helmholtz Center for Infection Research (HZI), Hannover, Germany

Paul E. Floreancig, Department of Chemistry, Chevron Science Center, University of Pittsburgh, Pittsburgh, PA, USA

Tobias A. M. Gulder, Department of Chemistry and Center for Integrated Protein Science Munich (CIPSM), Biosystems Chemistry, Technische Universität München, Munich, Germany

Trond Vidar Hansen, School of Pharmacy, University of Oslo, Oslo, Norway

Kazuaki Ishihara, Department of Biotechnology, Graduate School of Engineering, Nagoya University, Nagoya, Japan

Markus Kalesse, Institute for Organic Chemistry and Center of Biomolecular Drug Research (BMWZ), Leibniz Universität Hannover, Hannover, Germany and Helmholtz Center for Infection Research (HZI), Hannover, Germany

Xu-Wen Li, Shanghai Institute of Material Medica, Chinese Academy of Science, Shanghai, China

Bastien Nay, Muséum National d'Histoire Naturelle and CNRS (UMR 7245), Unité Molécules de Communication et Adaptation des Microorganismes, Paris, France

Yu Peng, State Key Laboratory of Applied Organic Chemistry, Lanzhou University, Lanzhou, China

Jakob Pletz, Institute of Organic Chemistry, Technische Universität Graz, Graz, Austria

Uwe Rinner, Institute of Organic Chemistry, Johannes Kepler University Linz, Linz, Austria and Department of Chemistry, College of Science, Sultan Qaboos University, Muscat, Oman

Floris P. J. T. Rutjes, Institute for Molecules and Materials, Radboud University Nijmegen, Nijmegen, The Netherlands

Françoise Schaefers, Department of Chemistry and Center for Integrated Protein Science Munich (CIPSM), Biosystems Chemistry, Technische Universität München, Munich, Germany

Michael Smietana, Institut des Biomolécules Max Mousseron, CNRS, Université de Montpellier, ENSCM, France

Zhen-Yu Tang, Department of Pharmaceutical Engineering, College of Chemistry and Chemical Engineering, Central South University, Changsha, China

Wouter S. Veldmate, Institute for Molecules and Materials, Radboud University Nijmegen, Nijmegen, The Netherlands

Anders Vik, School of Pharmacy, University of Oslo, Oslo, Norway

Mario Waser, Institute of Organic Chemistry, Johannes Kepler University Linz, Linz, Austria

Youwei Xie, Max-Planck-Institut für Kohlenforschung, Mülheim, Germany

Ian S. Young, Bristol-Myers Squibb Company, Chemical Development, New Brunswick, NJ, USA

Alexandros L. Zografos, Department of Chemistry, Laboratory of Organic Chemistry, Aristotle University of Thessaloniki, Thessaloniki, Greece

PREFACE

There is pleasure in the pathless woods,
there is rapture in the lonely shore,
there is society where none intrudes,
by the deep sea, and music in its roar;
I love not Man the less, but Nature more.
Lord Byron

The first time I came across with the idea of editing a book that merges selected chapters of *biosynthesis* and *total synthesis* was when I was teaching postgraduate courses of natural product synthesis at Aristotle University of Thessaloniki. This period, I realized that the best way to teach youngsters synthesis was to start from the very origin of inspiration, *nature* and its tools: *biosynthesis*.

Over the last decades, biosynthesis is filling our gaps of understanding the complex mechanisms of nature and provides useful sources of inspiration not only in the way natural products can be synthesized but also by directing synthetic chemists in developing atom-economical, efficient synthetic methods. Several are the examples that mimic biosynthetic guidelines, from modern iterative alkylations and aldol reactions to C─H oxidations that compile nowadays the modern toolbox of organic synthesis.

The handed book is constructed in the logic of presenting the parallel development of biosynthesis and organic methodology and how these can be applied in efficient syntheses of natural products. The book is divided into four sections each representing the four major biosynthetic pathways of natural products, namely, acetate, mevalonate, shikimate biosynthetic pathways, and the mixed biosynthetic pathways of alkaloids. These sections are divided into chapters that represent selected classes of natural products, for example, lipids, sesquiterpenoids, lignans, etc. Each of these chapters is further divided into three distinct subchapters: (a) biosynthesis, (b) methodological section, and (c) application of the described methodology in the total synthesis of the described family of natural products. By this way, the readers can be focused in the direct comparison between biosyntheses and the developed methodologies to construct the crucial for each class of natural product carbon bonds. Although the book, as it develops, is focused on presenting the power of biosynthesis and how this power can be applied in providing inspiration for the efficient synthesis of natural products, it was not the authors will to present only biomimetic total syntheses but rather to exploit the modern synthetic methodologies and recognize their disabilities for further improvement.

Of course this book will not have been realized without the excellent work of renowned scientists worldwide working either in the field of biosynthesis or total synthesis, who collected the existing knowledge on biosynthesis, analyzed the existing modern methodologies, and presented a bouquet of selected total syntheses. Throughout our

endeavor to complete this book, I learned many things from their expertise but I also realized that only with tight collaborations you can build long-lasting friendships. I would like to thank them all once again for their trust and effort to complete this book. We all hope that the current work will contribute to a better understanding of the current status of organic chemistry and to the discovery of novel *strategies and tactics* for the synthesis of natural products.

Alexandros L. Zografos
September 2015
Thessaloniki, Greece

1

FROM BIOSYNTHESES TO TOTAL SYNTHESES: AN INTRODUCTION

Bastien Nay[1] and Xu-Wen Li[2]

[1] *Muséum National d'Histoire Naturelle and CNRS (UMR 7245), Unité Molécules de Communication et Adaptation des Microorganismes, Paris, France*

[2] *Shanghai Institute of Material Medica, Chinese Academy of Science, Shanghai, China*

1.1 FROM PRIMARY TO SECONDARY METABOLISM: THE KEY BUILDING BLOCKS

1.1.1 Definitions

The primary and secondary metabolisms are traditionally distinguished by their distribution and utility in the living organism network. *Primary metabolites* include carbohydrates, lipids, nucleic acids, and proteins (or their amino acid constituents) and are shared by all living organisms on Earth. They are transformed by common pathways, which are studied by biochemistry (Fig. 1.1). *Secondary metabolites* are structurally diverse compounds usually produced by a limited number of organisms, which synthesize them for a special purpose, like defense or signaling, through specific biosynthetic pathways. They are studied by natural product chemistry. This distinction is not always so obvious and some compounds can be studied in the context of both primary and secondary metabolisms. This is especially true nowadays with the use of genetic and biomolecular tools, which tend to make natural product sciences more and more integrative. However, an important point to remember is that the primary metabolism furnishes key building blocks to the secondary metabolism. It would be difficult to describe in detail the full biosynthetic pathways in this section. We tried to organize the discussion as a *vade mecum*, synthetically gathering information from extremely useful sources, which will be cited at the end of this chapter.

1.1.2 Energy Supply and Carbon Storing at the Early Stage of Metabolisms

The sunlight is essential to life except in some part of the deep oceans. It provides energy for plant photosynthesis that splits molecules of water into protons and electrons and releases O_2 (Scheme 1.1). A proton gradient inside the plant chloroplasts then drags a transmembrane ATP synthase complex that produces adenosine triphosphate (ATP) while electrons released from water are transferred to the coenzyme reducer nicotinamide adenine dinucleotide phosphate hydride (NADPH). A major function of chloroplasts is to fix CO_2 as a combination to ribulose-1,5-bisphosphate (RuBP) performed by RuBP carboxylase (rubisco), forming an instable "C_6" β-ketoacid. This is cleaved into two molecules of 3-phosphoglycerate (3-PGA), which is then reduced into 3-phosphoglyceraldehyde (3-PGAL, a "C_3" triose phosphate) during the Calvin cycle. This is one of the major metabolites in the biosynthesis of carbohydrates like glucose and a biochemical mean for storing and retaining carbon atoms in the living cells.

1.1.3 Glucose as a Starting Material Toward Key Building Blocks of the Secondary Metabolism

Glucose-6-phosphate arises from the phosphorylation of glucose. It is the starting material of glycolysis, an important process of the primary metabolism, which consists in eight enzymatic reactions leading to pyruvic acid (PA)

From Biosynthesis to Total Synthesis: Strategies and Tactics for Natural Products, First Edition. Edited by Alexandros L. Zografos.

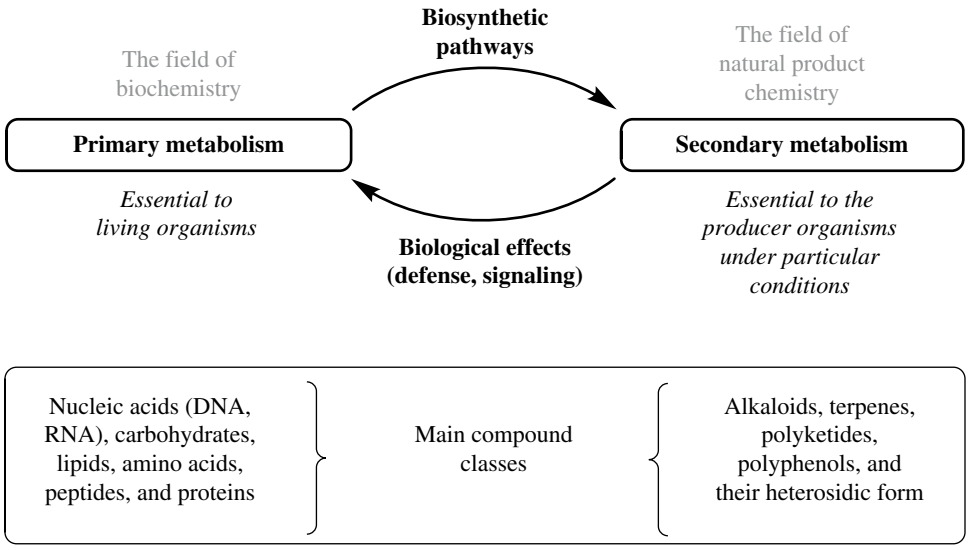
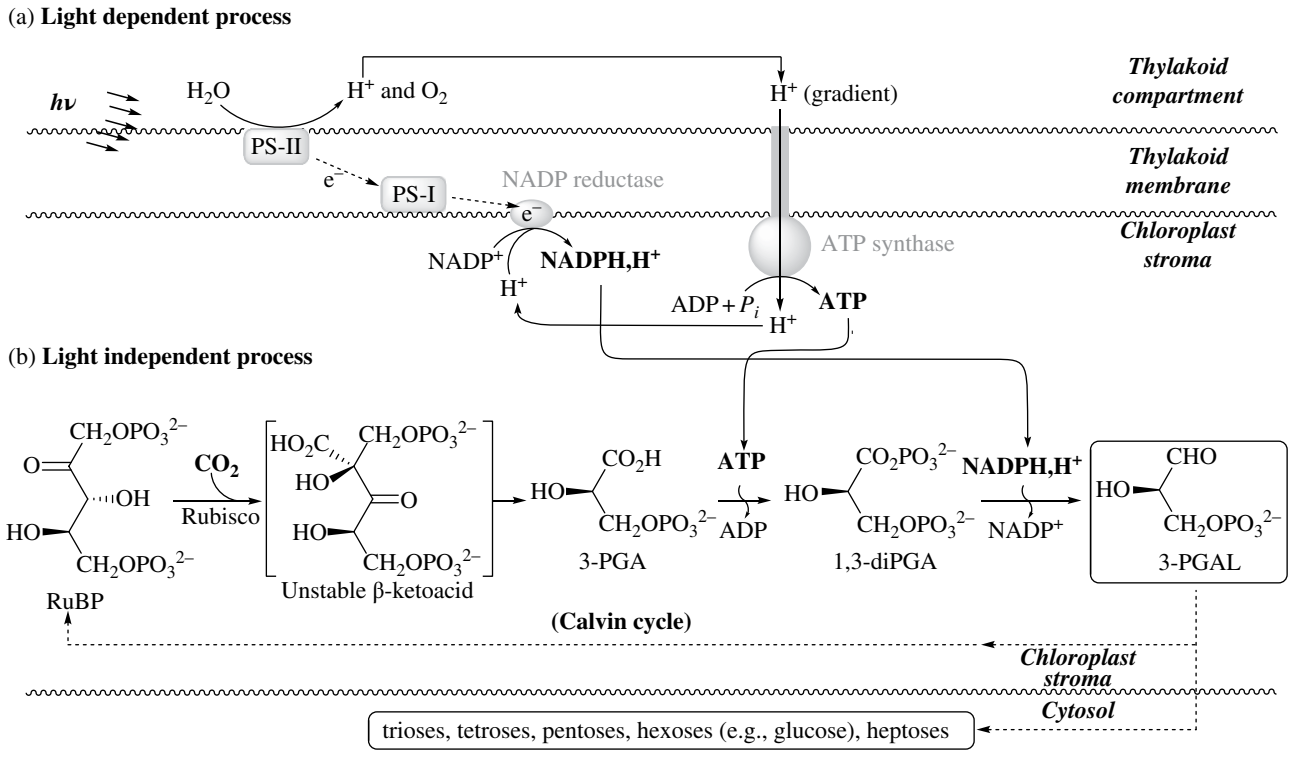
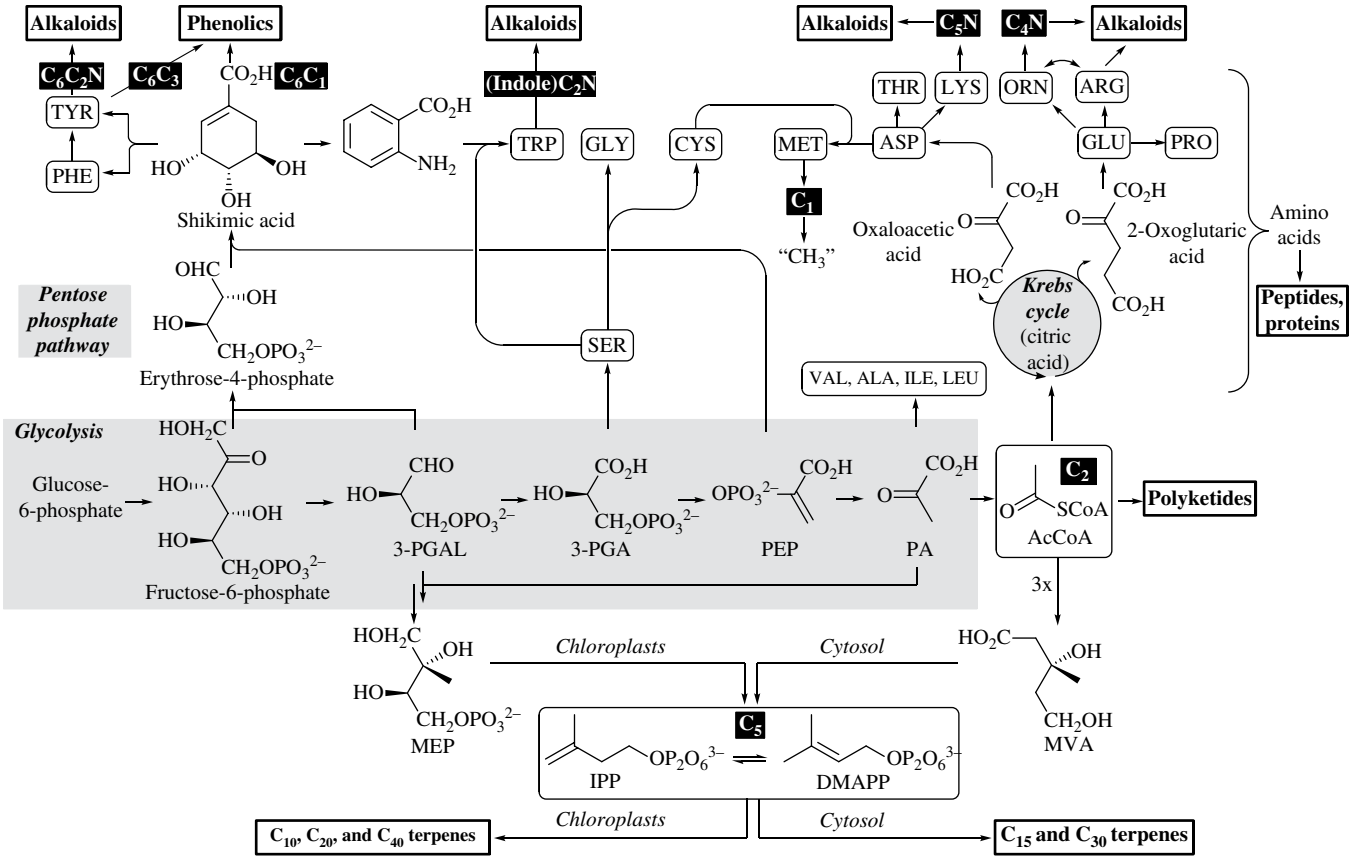


FIGURE 1.1 Primary *versus* secondary metabolisms.



SCHEME 1.1 The photosynthetic machinery (PS-I and PS-II, photosystems I and II).

(Scheme 1.2). Important intermediates for the secondary metabolism are produced during glycolysis. Glucose, glucose-6-phosphate, and fructose-6-phosphate can be converted to other hexoses and pentoses that can be oligomerized and enter in the composition of heterosides. Additionally, fructose-6-phosphate connects the pentose phosphate pathway, leading to erythrose-4-phosphate toward shikimic acid, which is a key metabolite in the biosynthesis of aromatic amino acids (phenylalanine, tyrosine, or C_6C_3 units) and C_6C_1 phenolic compounds. The next important intermediate in glycolysis is 3-PGAL, which can be redirected toward methylerythritol-4-phosphate (MEP) in the



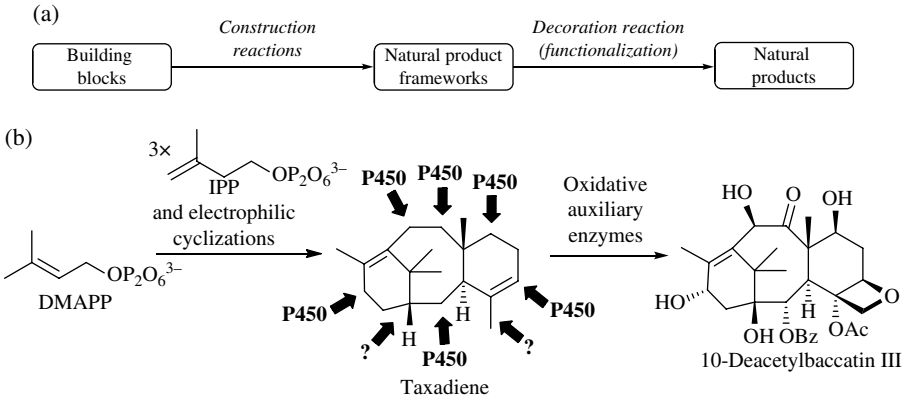
SCHEME 1.2 The building block chart, involving glycolysis, and the Krebs cycle.

chloroplast. MEP is a starting block in the biosynthesis of terpenes through C_5 isoprene units (isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP)), especially those in C_{10}, C_{20}, and C_{40} terpenes. 3-PGA is a precursor of serine and other amino acids, while phosphoenolpyruvate (PEP), the precursor of PA, is also an intermediate toward the previously mentioned shikimic acid. Lastly, PA is not only a precursor of the fundamental "C_2" acetyl coenzyme A (AcCoA) unit but also an intermediate toward aliphatic amino acids and MEP.

AcCoA is the building block of fatty acids, polyketides, and mevalonic acid (MVA), a cytosolic precursor of the C_5 isoprene units for the biosynthesis of terpenes in the C_{15} and C_{30} series (mind it is different from the MEP pathway, in product, and in cell location). Finally, AcCoA enters the citric acid or Krebs cycle, which leads to several precursors of amino acids. These are oxaloacetic acid, precursor of aspartic acid through transamination (thus toward lysine as a nitrogenated C_5N linear unit and methionine as a methyl supplier), and 2-oxoglutaric acid, precursor of glutamic acid (and subsequent derivatives such as ornithine as a nitrogenated C_4N linear unit). All these amino acids are key precursors in the biosynthesis of many alkaloids.

1.1.4 Reactions Involved in the Construction of Secondary Metabolites

Most reactions occurring in the living cells are performed by specialized enzymes, which have been classified in an international nomenclature defined by an enzyme commission (EC) number. There are six classes of enzymes depending on the biochemical reaction they catalyze: EC-1, oxidoreductases (catalyzing oxidoreduction reactions); EC-2, transferases (catalyzing the transfer of functional groups); EC-3, hydrolases (catalyzing hydrolysis); EC-4, lyases (breaking bonds through another process than hydrolysis or oxidation, leading to a new double bond or a new cycle); EC-5, isomerases (catalyzing the isomerization of a molecule); and EC-6, ligases (forming a covalent bond between two molecules). Many subclasses of these enzymes have been described, depending on the type of atoms and functional groups involved in the reaction and, if any, on the cofactor used in this reaction. For example, several cofactors can be used by dehydrogenases like NAD(P)/NAD(P)H, FAD/$FADH_2$, or FMN/$FMNH_2$. For a description of this classification, the reader can refer to specialized Internet websites like ExplorEnz [1]. What is important to realize is that most enzymes are substrate specific and have been selected during



SCHEME 1.3 (a) From building blocks to natural products and (b) the example of 10-deacetylbaccatin III.

evolution to perform specific transformations, making natural products with often and yet unknown functions.

Secondary metabolites arise from specific biosynthetic pathways, which use the previously defined building blocks. The bunch of organic reactions involved in these biosyntheses allows the construction of natural product frameworks, which are finally diversified through "decoration" steps (Scheme 1.3). It is not the purpose of this introductive chapter to describe in detail all biosynthetic pathways and the reader can refer to excellent books and articles, which have been published elsewhere [2, 3].

The reactions involved in the construction of natural product skeletons will be described later for representative classes of compounds. The identification of the building block footprint in the natural product skeleton will be emphasized as much as possible, sometimes referring to biogenetic speculations [4]. After the framework construction, the decoration steps will involve as diverse reactions as aliphatic C─H oxidations (e.g., involving a cytochrome P_{450} oxygenase) occasionally triggering a rearrangement, heteroatom alkylations (e.g., methylation by *S*-adenosylmethionine) or allylation (by DMAPP), esterifications, *heteroatom* or C-glycosylations (leading to heterosides), radical couplings (especially for phenols), alcohol oxidations or ketone reductions, amine/ketone transaminations, alkene dihydroxylations or epoxidations, oxidative halogenations, Baeyer–Villiger oxidations, and further oxygenation steps. At the end of the biosynthesis, such transformations may totally hide the primary building block origin of natural products.

1.1.5 Secondary Metabolisms

1.1.5.1 Polyketides Polyketides (or polyacetates) are issued from the oligomerization of C_2 acetate units performed by polyketide synthases (PKS) and leading to $(C_2)_n$ linear intermediates [5, 6]. If the $(C_2)_n$ intermediates arise from successive Claisen reactions performed by ketosynthase domains (KS, in nonreducing PKS), a highly reactive poly-β-ketoacyl intermediate $H{-}(CH_2C{=}O)_n{-}OH$ is formed, leading to phenolic and aromatic products through further intramolecular Claisen condensations. Furthermore, highly reducing PKSs are made of specialized enzymatic subunits working in line or iteratively to functionalize each C_2 linker bond as $CH(OH)CH_2$ (by ketoreductases (KR)), then as $HC{=}CH$ (by dehydratases (DH)), and as CH_2CH_2 (by enoyl reductases (ER)), leading to a high degree of functionalization of the final product (Fig. 1.2). By these iterative sequences, highly reduced polyketides, which can be either linear, macrocyclized, or polycyclized depending on the reactivity of the biosynthetic intermediates, can be formed [7]. With the same logic, fatty acids are also biosynthesized by fatty acid synthases.

Moreover, the PKS enzyme can be hybridized with nonribosomal peptide synthetase (NRPS) domains (see also "NRPS metabolites and peptides" in the "Alkaloids" section), leading to the acylation of an amino acid by the $(C_2)_n$ acyl intermediate. As previously, the functionalization of the acyl chain depends on the PKS enzyme, and the PKS/NRPS products are also extremely diversified (e.g., hirsutellone B; Fig. 1.2) [8].

1.1.5.2 Terpenes Terpenes are derived from the oligomerization of the C_5 isoprene units DMAPP and IPP. Both precursors are prompt to generate either an allylic cation (the diphosphate is a good leaving group) or a tertiary carbocation, respectively, which makes the IPP easy to react with DMAPP (Scheme 1.4). This reaction happens in the active site of a terpene synthase, which activates the departure of the diphosphate group from DMAPP, thanks to Lewis acid activation (a metal like Mg^{2+}, Mn^{2+}, or Co^{2+} is present in the enzyme active site [9]). This leads to geranyl (C_{10}, monoterpene precursor) or farnesyl (C_{15}, sesquiterpene precursor) diphosphate, depending on the location of the enzyme (chloroplast for the MEP pathway or cytosol for the MVA pathway). Geranylgeranyl (C_{20}, diterpene) and

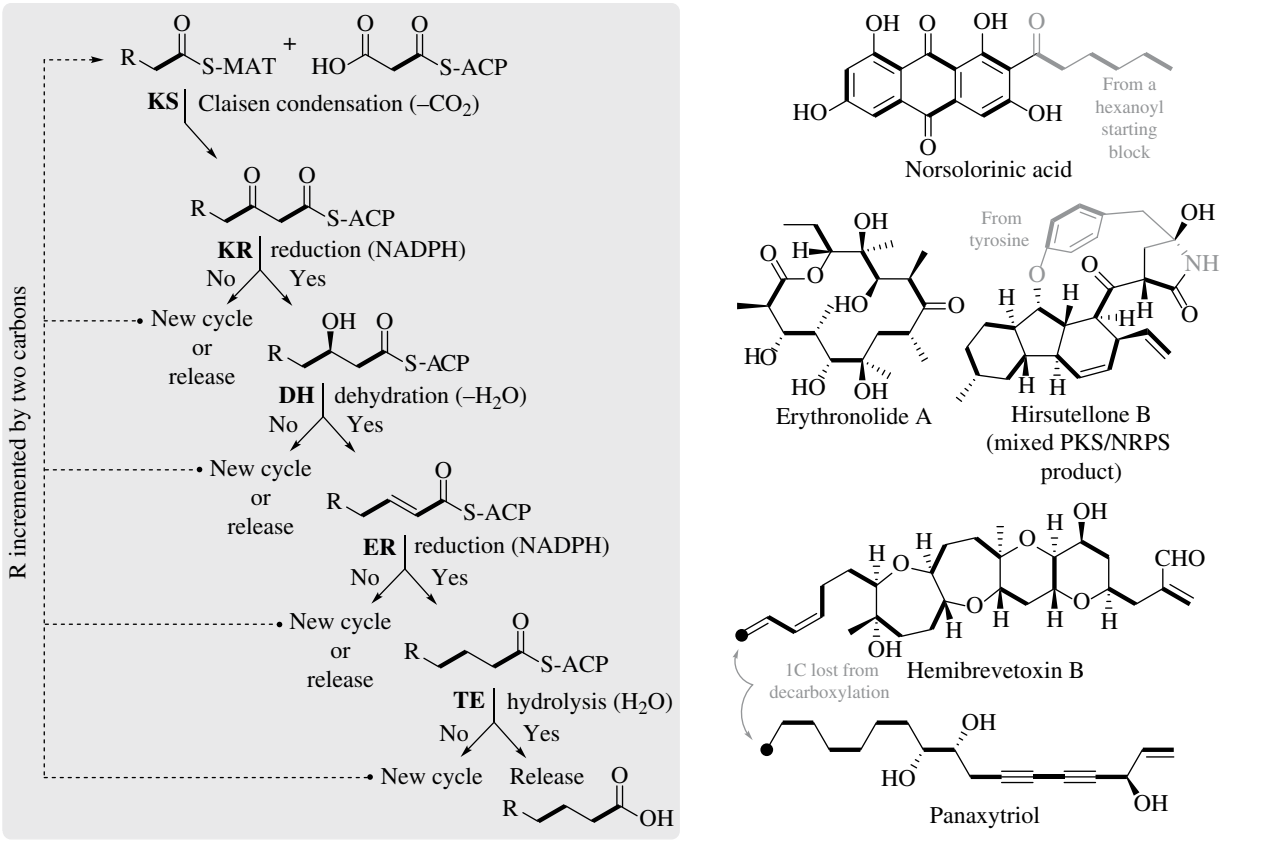
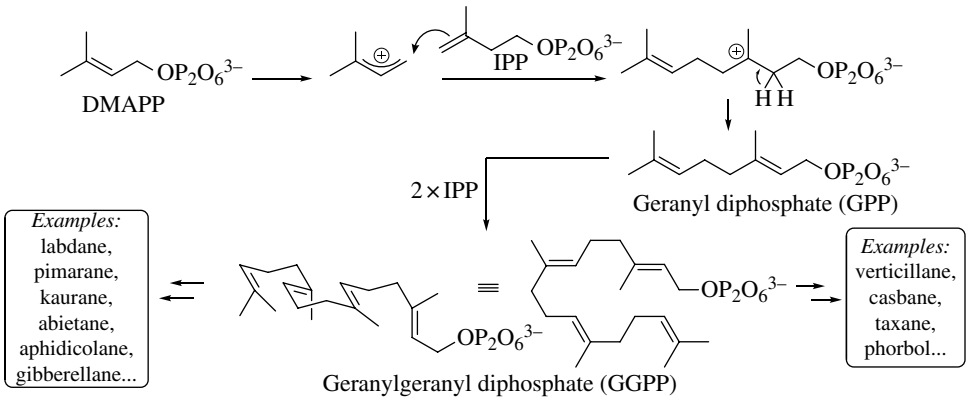


FIGURE 1.2 Chemical logic of polyketide construction leading to variable functionalization of the elongated acyl chain and examples of resulting chemical diversity. ACP, acyl carrier protein; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; KS, ketosynthase; MAT, malonyl acyl transferase; TE, thioesterase.



SCHEME 1.4 Early assembly of C_5 units in terpene biosynthesis, leading to diterpenes (C_{20}).

farnesylfarnesyl (C_{30}, triterpene precursor) diphosphates can also be obtained by further additions of IPP, leading to longer linear intermediates.

The cyclization of linear precursors is achieved by specialized cyclases, which generate a poorly functionalized natural product framework [10, 11]. Auxiliary enzymes such as oxygenases then increase the complexity and the diversity of compounds by further functionalization (Scheme 1.3b) [12]. A high degree of oxidation can be observed in compounds like thapsigargin, paclitaxel, or bilobalide (Fig. 1.3). The biosynthesis of this last compound, for example, involves a high oxygenation pattern, two Wagner–Meerwein rearrangements, and several oxidative cleavages leading to the loss of five carbons. The resulting natural products can

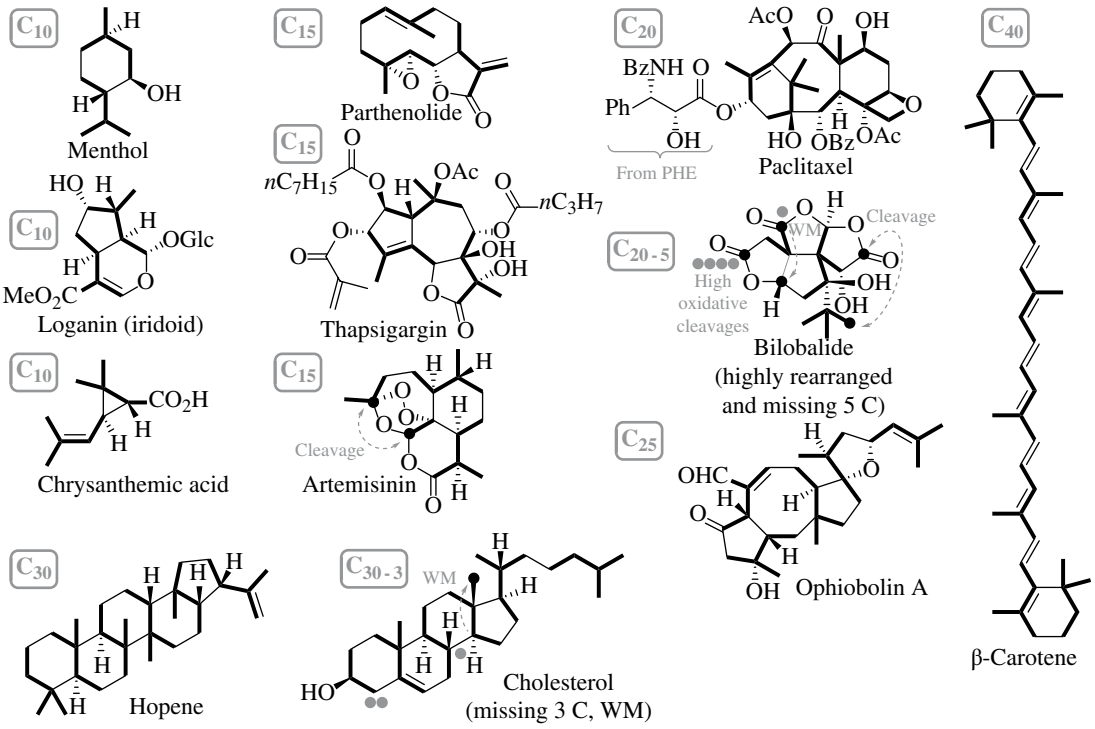


FIGURE 1.3 Chemical diversity in the terpene series (WM, Wagner–Meerwein shifts; ●, lost carbons; bold bonds are remnant of primary building blocks).

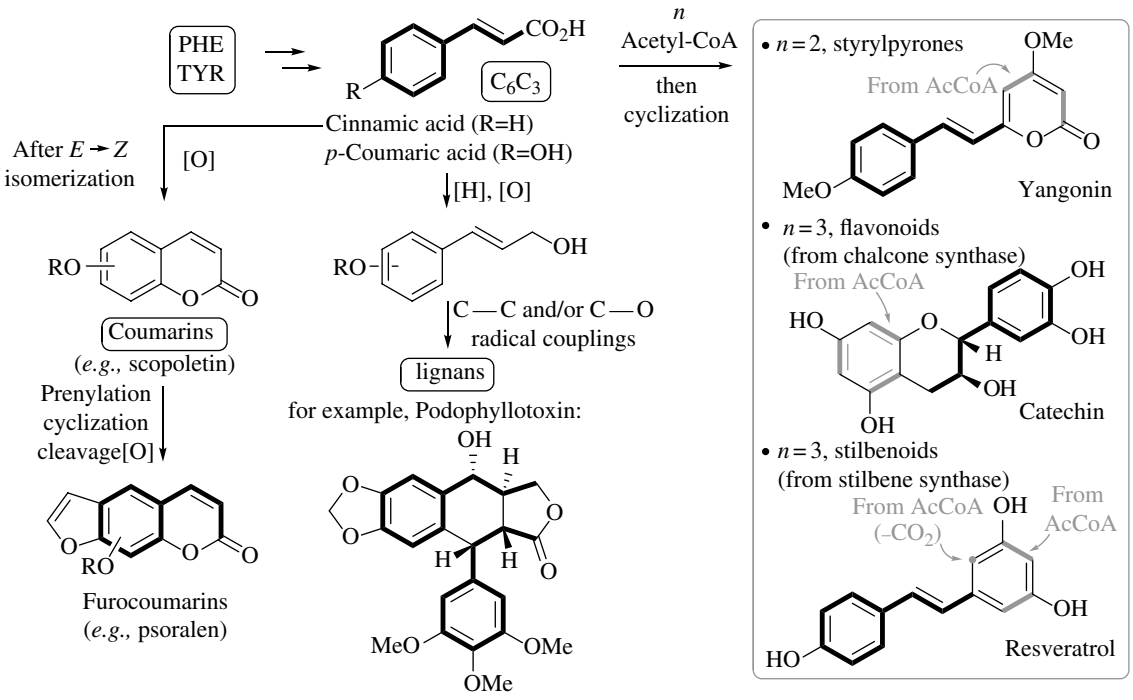
thus be extremely modified, with structures whose biogenetic origin is far from being obvious at first sight and cannot be determined without further experiments such as isotopic labeling.

1.1.5.3 Flavonoids, Resveratrols, Gallic Acids, and Further Polyphenolics We have previously discussed the polyketide origin of some phenolic compounds based on the $(C_2)_n$ motif. Other polyphenols like gallic acids are directly derived by the aromatization of shikimic acid (C_6C_1 building block; Scheme 1.2) [13]. The C_6C_3 building blocks are available from the conversion of phenylalanine and tyrosine into cinnamic and *p*-coumaric acids, respectively, and then by further hydroxylation steps (Scheme 1.5). These can dimerize into lignans (e.g., podophyllotoxin) [14, 15] through radical processes or converted to low molecular weight compounds like eugenol, coumarins, or vanillin [16]. The coenzyme A thioesters of these C_6C_3 acids can be used as initiator units by specialized ketosynthases for an elongation by two acetyl units, leading to aromatic polyketides like styrylpyrones or diarylheptanoids (e.g., curcumin) [17]. Important compounds from this metabolism are flavonoids ($C_6C_3C_6$) [18] and stilbenoids ($C_6C_2C_6$) (a decarboxylation occurs during the aryl cyclization) [19], which are synthesized by chalcone synthase and stilbene synthase, respectively. Flavonoids (e.g., catechin) and stilbenes (e.g., resveratrol) are present in large amounts in fruits and vegetables and may exert their radical scavenging properties *in vivo*.

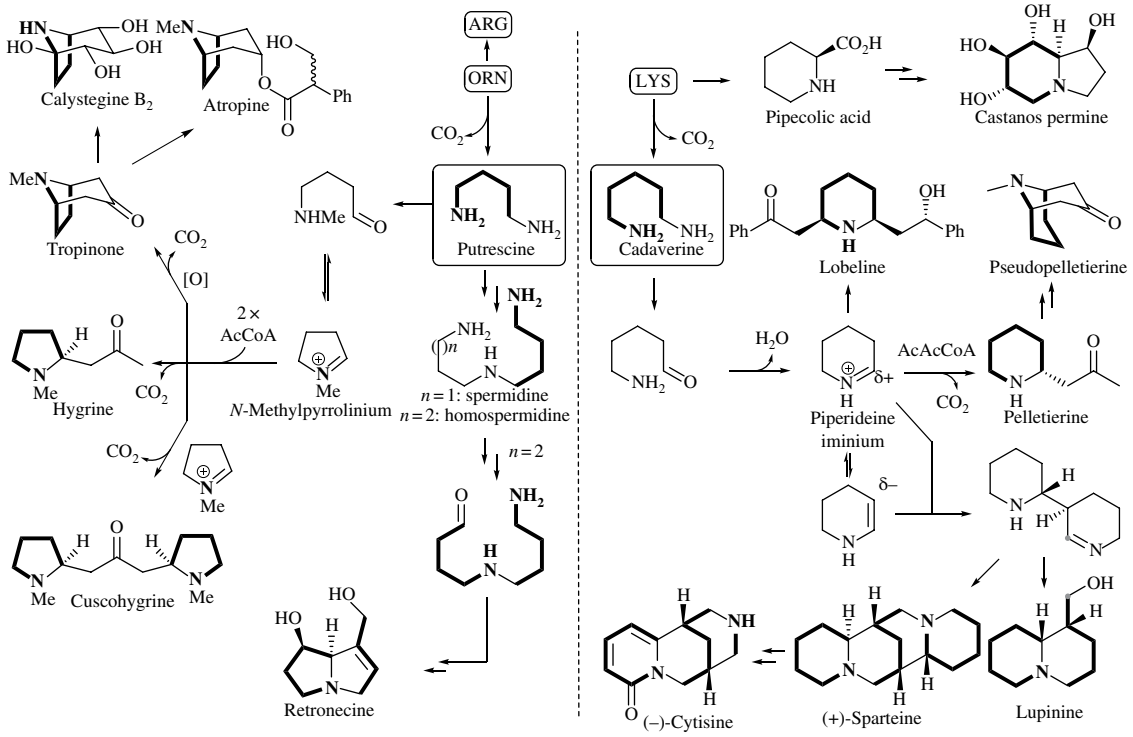
1.1.5.4 Alkaloids Alkaloids are nitrogen-containing compounds. The nitrogen(s) can be involved in an amine function, conferring basicity to the natural product (like "alkali"), or in less or nonbasic functions such as an amide, a nitrile, an isonitrile, or an ammonium salt (quaternary amines). For amines, protonation often occurs at physiological pH and may condition their biological activity. In many cases, the nitrogen is biogenetically derived from an amino acid. We will thus discuss alkaloids according to their amino acid origin.

Alkaloids Derived from the Krebs Cycle (Lysine and Ornithine Derived) As shown previously (Scheme 1.2), the Krebs cycle is a crucial metabolic process, which leads to α-ketoacids (oxaloacetic and 2-oxoglutaric acids). Their enzymatic transamination affords the two amino acids—aspartic acid and glutamic acid—which are the direct biosynthetic precursors of amino acids lysine and ornithine, respectively. These in turn produce cadaverine, a "C_5N" unit, and putrescine, a "C_4N" unit, which are major components for the biosynthesis of important alkaloids, as will be discussed later (Scheme 1.6). Additionally, ornithine is a precursor of arginine, another important amino acid.

ORNITHINE-DERIVED ALKALOIDS (INCORPORATING THE C_4N UNIT) Putrescine is derived from the decarboxylation of ornithine and is a precursor of linear polyamines like spermine. After enzymatic methylation of one amine of putrescine

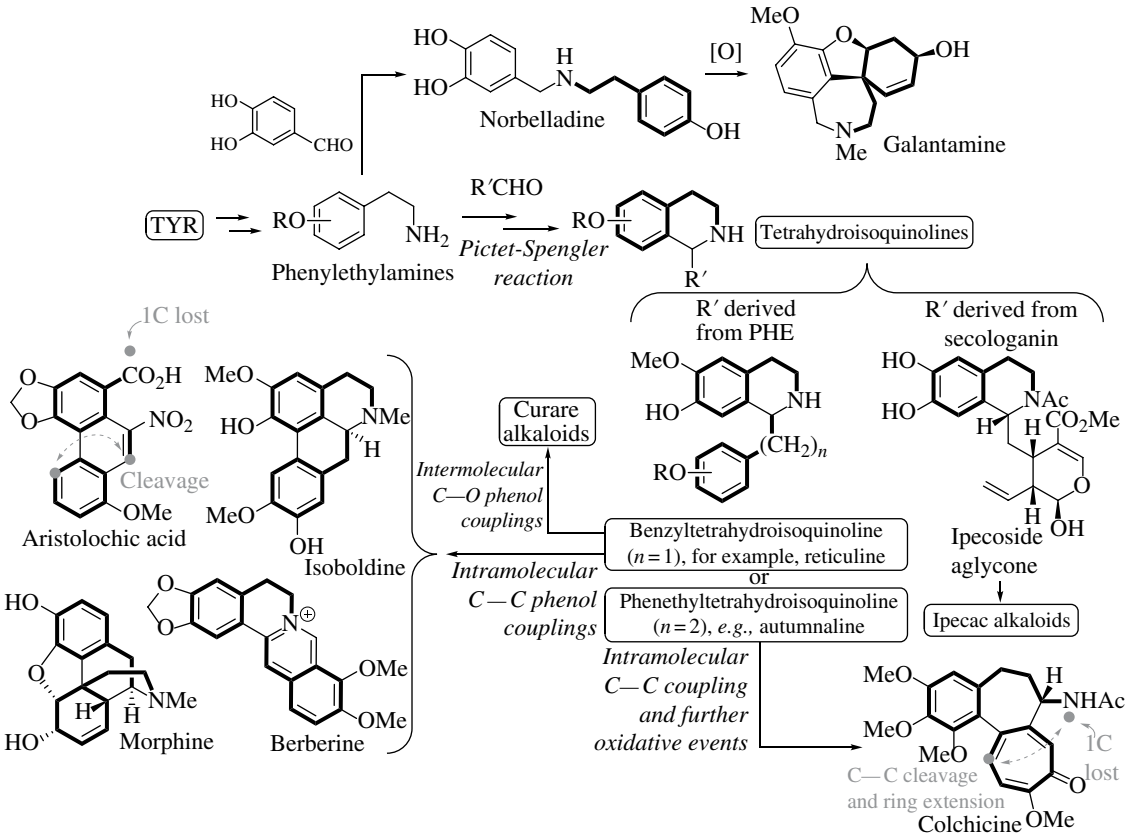


SCHEME 1.5 The phenylpropanoid biosynthetic pathways.



SCHEME 1.6 Lysine- and ornithine-derived alkaloid biosynthetic pathways (mind the structural similarities).

in the presence of *S*-adenosylmethionine, transamination of the other affords γ-(*N*-methylamino)aldehyde [20]. The resulting cyclic iminium is a key intermediate in the formation of many medicinally important alkaloids such as the plant-derived compounds cocaine, atropine, or the calystegines [21, 22]. Indeed, this iminium is a Mannich acceptor, which can react with various nucleophiles, the first of those being the carbanion of acetyl-CoA. Thus, after a stepwise



SCHEME 1.7 Tyrosine-derived alkaloid biosynthetic pathways (double head arrows figure bond cleavages during biosynthetic processes).

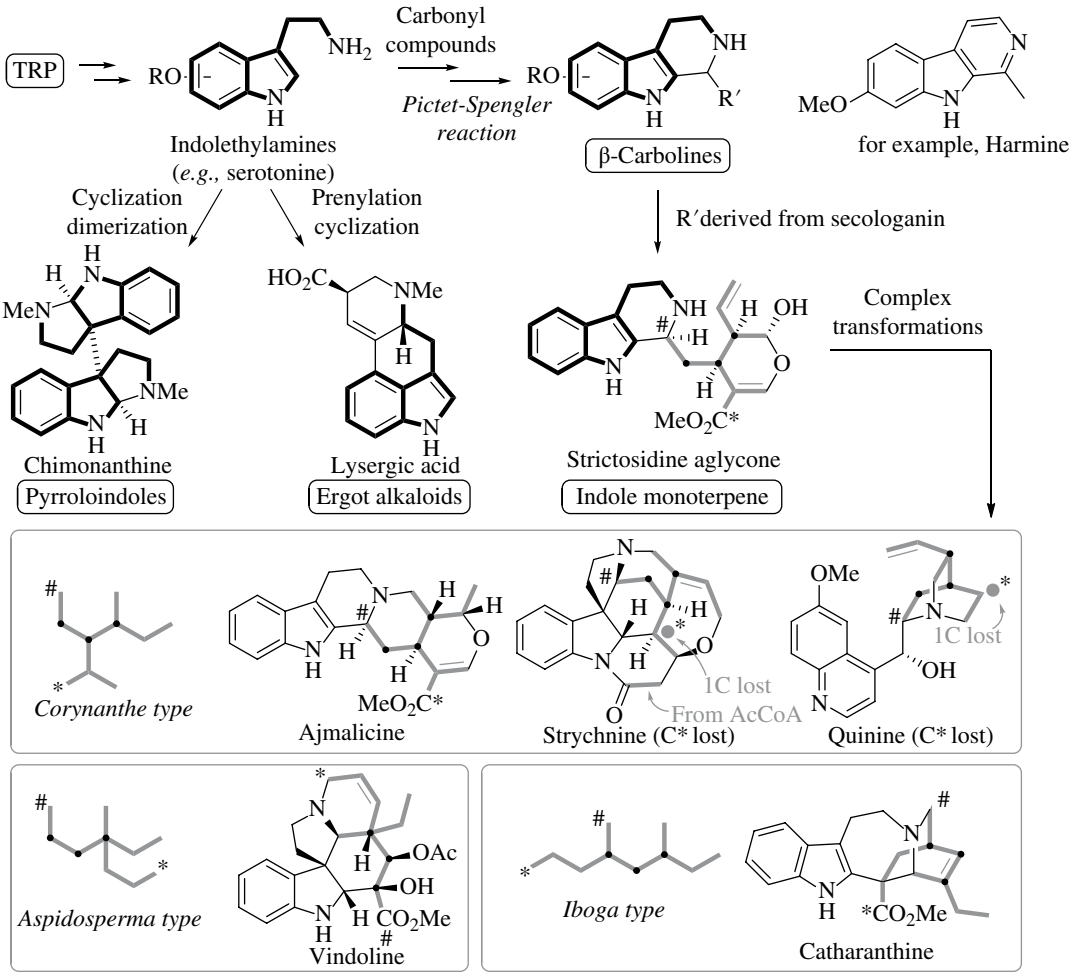
elongation by two AcCoA units, either decarboxylation can occur, leading to the acetonylpyrrolidine hygrine, or a second Mannich reaction by the intramolecular attack of the acetoacetate anion onto an oxidation-derived pyrrolinium, leading to the tropane skeleton (tropinone). The acetoacetyl-CoA intermediate can also react intermolecularly with another pyrrolinium cation, leading to cuscohygrine after decarboxylation. Finally, the pyrrolizidine alkaloids [23] are derived from homospermidine, which, when submitted to terminal oxidative deamination, leads to the bicyclic skeleton of retronecine and further *Senecio* alkaloids. We can mention herein that ornithine is a biosynthetic precursor of arginine, bearing a guanidine function, which is an intermediate toward the toxic compounds tetrodotoxin and saxitoxin (not shown).

LYSINE-DERIVED ALKALOIDS (INCORPORATING THE C_5N UNIT) From lysine to piperidine alkaloids, the biosynthetic steps parallel the one previously described from ornithine. Indeed, the oxidative deamination of cadaverine affords a δ-amino aldehyde, which cyclizes through imine formation into piperideine. Protonation results in a Mannich acceptor, which is able to react with various nucleophiles such as β-ketothioester anions. The first product of these reactions is pelletierine, which can further react through an intramolecular Mannich reaction leading to pseudopelletierine. Quinolizidines [24] can also be formed, first from the Mannich reaction of the piperideine acceptor with the corresponding enamine nucleophile and then after additional transformation steps, leading, for example, to lupinine, sparteine, or cytisine.

Indolizidine alkaloids [15] such as castanospermine and swainsonine are formed from pipecolic acid, an amino acid derived from lysine, which can be elongated by malonyl-CoA followed by ring closure. When protonated, these alkaloids are oxonium mimics strongly inhibiting glycosidases.

Tyrosine- and Phenylalanine-Derived Alkaloids Tyrosine and phenylalanine amino acids are bearing the phenylethylamine moiety of many medicinally relevant alkaloids. Further hydroxylations on the aromatic carbocycle or on the aliphatic part can be observed. Methylations can occur on phenolic oxygens and on the amine, leading to catecholamines (adrenaline, noradrenaline, dopamine). Arylethylamines are also usual to react with endogenous aldehydes through Pictet–Spengler reactions [25], leading to important biosynthetic intermediates (Scheme 1.7) like:

- Reticuline from the reaction with 4-hydroxyphenylacetaldehyde toward benzyltetrahydroisoquinoline alkaloids:



SCHEME 1.8 Tryptophan-derived alkaloid biosynthetic pathways (gray parts: monoterpenic units).

morphine, berberine, tubocurarine, isoboldine, or the highly modified aristolochic acid [26, 27]

- Automnaline from the reaction with 3-(4-hydroxyphenyl) propanal toward phenylethyltetrahydroisoquinoline alkaloids: colchicine, cephalotaxine, or schelhammericine [28, 29]
- Ipecoside from the reaction of dopamine with secologanin toward terpene tetrahydroisoquinoline alkaloids: ipecoside or emetine

Lastly, norbelladine (top of Scheme 1.7) is issued from the reductive amination of 3,4-dihydroxybenzaldehyde (derived from phenylalanine) with tyramine (derived from tyrosine) and constitutes a biosynthetic node leading to Amaryllidaceae alkaloids such as galantamine, crinine, or lycorine depending on the topology of phenolic couplings. In all these biosynthetic routes, radical phenolic couplings are key reactions for C─C and C─O bond formations and rearrangements [30, 31].

Tryptophan-Derived Indole and Indole Monoterpene Alkaloids As for alkaloids derived from tyrosine and phenylalanine, those derived from tryptophan are formed after decarboxylation of the amino acid (into tryptamine) and possible hydroxylation of the aromatic carbocycle (e.g., serotonin) and N-methylation (e.g., psilocin). As previously, tryptamine can also react through Pictet–Spengler reactions to form tetrahydro-β-carbolines, which can be aromatized, for example, into harmine (Scheme 1.8) [16].

When the aldehyde partner of the Pictet–Spengler reaction with tryptamine is the terpene secologanin, strictosidine is formed as an entry toward the vast monoterpene indole alkaloids [32, 33]. Hydrolysis of the glucosidic part releases the strictosidine aglycone bearing an aldehyde, while iminium formation and further cyclization and reduction can lead to ajmalicine (from oxocyclization) or yohimbine (from carbocyclization). These alkaloids are referred to as from the *Corynanthe* type, with the monoterpene carbon skeleton unmodified. Although it misses one carbon and has a very

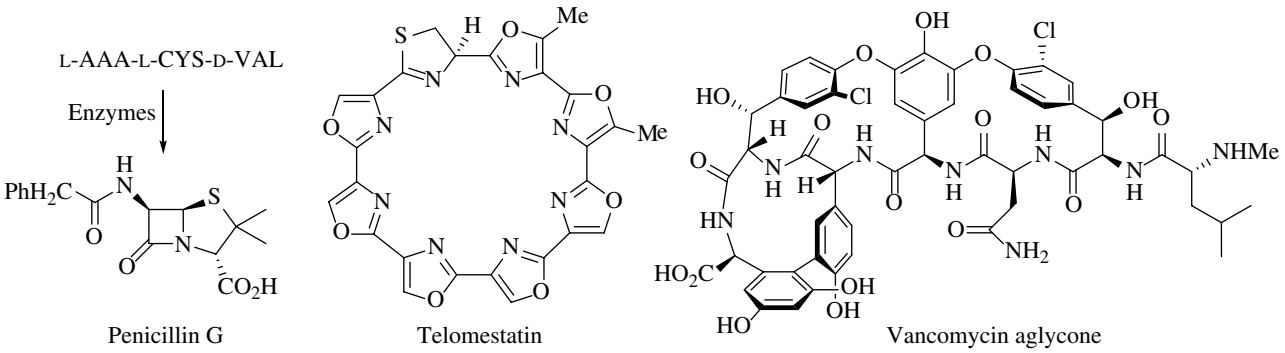


FIGURE 1.4 Structural diversity of nonribosomal peptide compounds (AAA, α-aminoadipic acid).

different structure, strychnine is related to the *Corynanthe* alkaloids, incorporating two carbons from acetyl-CoA. Highly modified monoterpene skeletons are derived from the *Corynanthe* core through C─C bond breaking and reorganization, leading to *Iboga*-type (e.g., catharanthine) and *Aspidosperma*-type (e.g., vindoline) alkaloids. The anticancer drug vinblastine is a heterodimer resulting from the nucleophilic attack of vindoline on a Mannich acceptor resulting from catharanthine, found in Madagascar periwinkle (*Catharanthus roseus*). The heteroaromatic compounds ellipticine, camptothecin, and quinine are also derived from a *Corynanthe*-type precursor, although in this case the biosynthetic relationship may not be obvious due to deep modifications of the skeleton.

Finally, two important classes of compounds have to be mentioned since they have inspired many synthetic chemists. The pyrroloindole alkaloids result from the cyclization of tryptamine, as found in physostigmine (formed by a cationic mechanism after methylation in position 3 of the indole; not shown) or in chimonanthine (presumably formed by a radical coupling mechanism; Scheme 1.8). The ergot alkaloids are derived from the 3,3-dimethylallylation on position 4 of the indole in tryptophan whose further cyclization and oxidation processes afford the natural products (e.g., lysergic acid, Scheme 1.8, and ergotamine), which have had important medical applications [34].

NRPS Metabolites and Peptides NRPS enzymes assemble amino acids, including nonproteinogenic ones, into oligopeptides. The enzymes contain several modules, and especially an adenylation domain (A), which specifically selects and activates the amino acid to be transferred as a thioester on the nearby peptidyl carrier protein (PCP) [2]. A condensation module (C) then catalyzes the formation of the peptide bonds between the newly introduced amino acyl-PCP (bearing a free amine) and the elongated peptidyl-PCP thioester. At the end of the elongation, a cyclization can occur into cyclopeptides, but the peptide can also be transferred to auxiliary enzymes like methyltransferases, glycosyltransferases, or oxidases (vancomycins are typical products of such functionalizations) [35, 36]. The formation of heterocycles is also frequently encountered in this metabolism, as in penicillins that are derived from the tripeptide α-aminoadipoyl-cysteinyl-valine or telomestatin (Fig. 1.4) [2].

Other Alkaloid Origins There are many other nitrogen sources involved in alkaloid biosyntheses, for example, nicotinic acid (originated from aspartic acid and intermediate in nicotine and anabasine biosyntheses) and anthranilic acid (originated from tryptophan and intermediate toward acridines or aurachins). The amination reaction (e.g., through transamination of carbonyl compounds) is also a way to introduce nitrogens in natural products, for example, from fatty acids, steroids (toward *Solanum* alkaloids or cyclopamines), or other terpenoids (aconitine and atisine have diterpene skeletons, while *Daphniphyllum* alkaloids are triterpene derivatives). Finally, nucleic acids can also be precursors of alkaloids like the well-known caffeine.

1.2 FROM BIOSYNTHESIS TO TOTAL SYNTHESIS: STRATEGIES TOWARD THE NATURAL PRODUCT CHEMICAL SPACE

1.2.1 The Chemical Space of Natural Products

Natural products occupy an important place in human communities as demonstrated by their vast use from ancient times to nowadays, like dyes, fibers, oils, perfumes, agrochemicals, or drugs. Broadly, both primary and secondary metabolites could be classified as natural products, while the latter, as discussed previously, are usually regarded as the "natural products" owing to their complexity and diversity arising from a variety of biosynthetic pathways. The structural chemical diversity found among

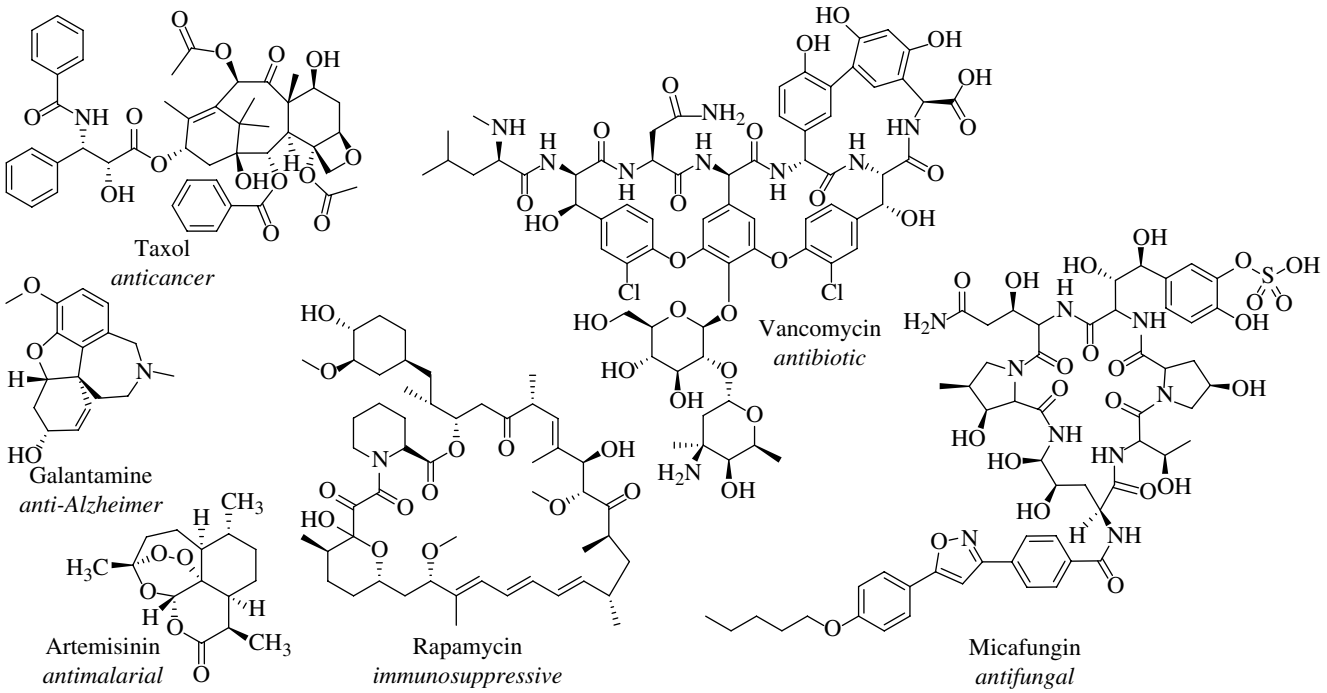
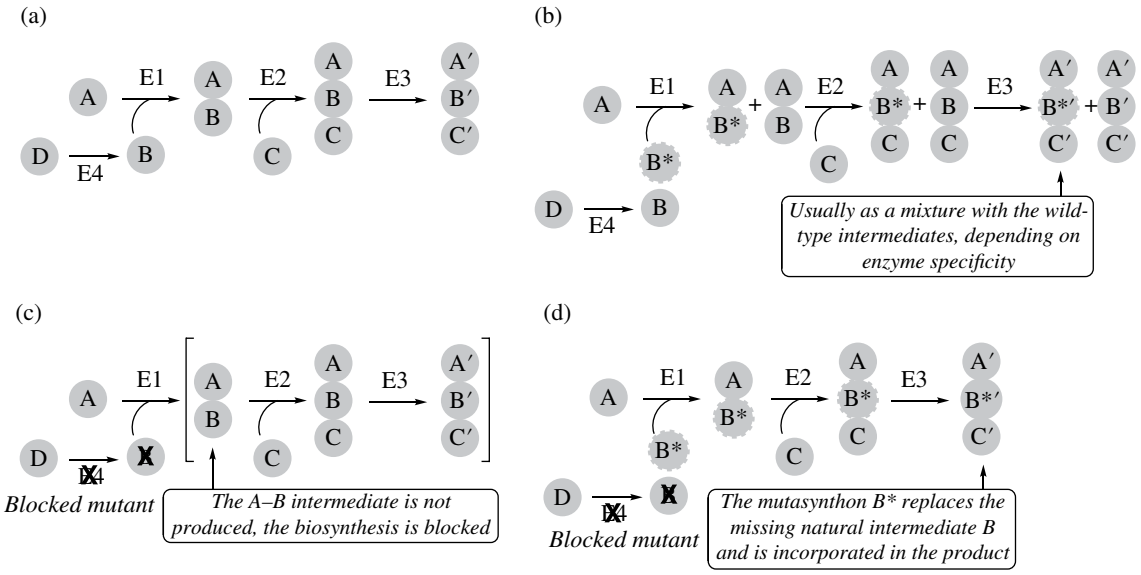


FIGURE 1.5 Some famous natural products currently used as drugs.

all living organisms, defining the chemical space of natural products [37], is the consequence of their evolution, occurring as an adaptation of organisms to their environment. It is commonly believed that secondary metabolites are produced as messengers by living organisms or as weapons against enemies, and thus they should have certain biological activities in a medicinal point of view [38]. Indeed, natural products are regarded as one of the main sources of medicines (Fig. 1.5). From the traditional medicinal extracts to every single bioactive molecule, the methods of extraction, purification, identification, and biological investigation of natural products have been well established. Their complex structures and interesting properties have attracted synthetic chemists to accomplish their total syntheses and that of medicinally relevant analogs, sometimes in the industrial context [39]. Thus, targeting the chemical space of natural products has never been more relevant than today. Although the discovery of natural products demands time and labor-consuming manipulations, it is worth to notice that the knowledge on this chemical space is still continually growing while biological advances allow for discovering and understanding potential targets. However, increasing the chemical space of human-made compounds based on natural products should benefit from transdisciplinary collaborations such as the use of coupled biosynthetic and chemical synthetic methods to design original "unnatural natural products" [40].

1.2.2 The Biosynthetic Pathways as an Inspiration for Synthetic Challenges

1.2.2.1 Precursor-Directed Biosyntheses and Mutasynthetic Strategies to Increase the Chemical Space of Natural Products As the genetics and biochemistry of natural product biosynthesis are better understood, novel biosynthetic techniques have been developed to study and generate new diversity in natural product analogs (Scheme 1.9). Precursor-directed biosynthesis (PDB) is considered as the earliest example of combining chemical and biological methods for the generation of complex natural product analogs [41, 42]. This approach, compared with the biosynthetic pathway of wild-type metabolites (Scheme 1.9a), involves the feeding of analogs of the natural biosynthetic building blocks to the living organisms (Scheme 1.9b), usually bacteria or fungi, which incorporate the modified precursors into the biosynthesized compound. Mutasynthesis, also termed as mutational biosynthesis (MBS), involves the inactivation of a key step of the biosynthesis in a mutant microorganism (Scheme 1.9c), which can then be fed by various modified or advanced building blocks (mutasynthons; Scheme 1.9d) [43]. These mutasynthons could not be incorporated by the wild type due to specificities of the enzymatic machinery. Build up on PDB, MBS eliminates the natural biosynthetic intermediate, thus generating a less complex mixture of metabolites and making the purification or yield of target products better.



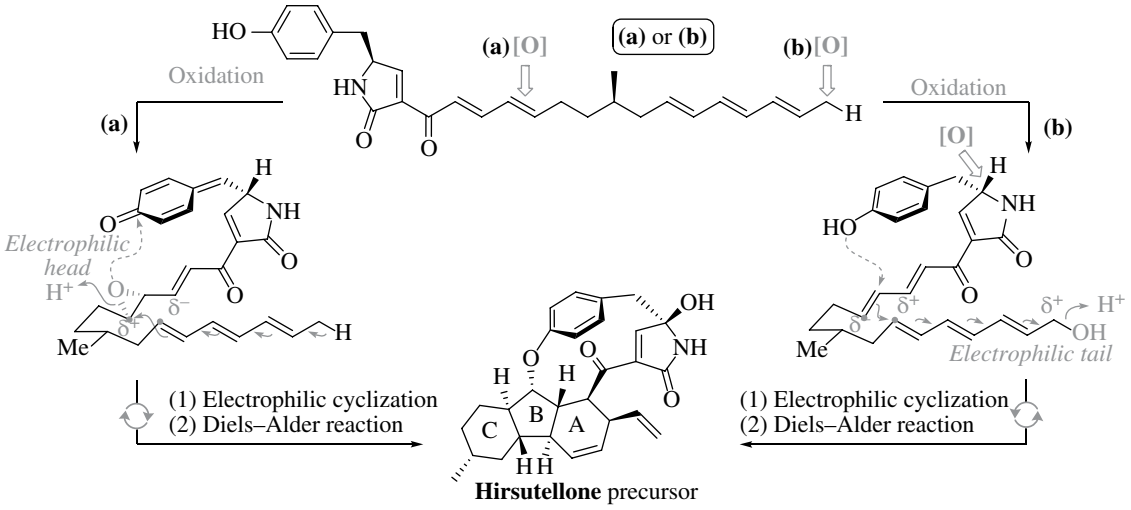
SCHEME 1.9 (a) Biosynthetic pathway of wild-type metabolites; (b) precursor-directed biosynthesis: the modified synthon B* replaces the natural synthon B; (c) biosynthetic pathway blocked by a mutation (the enzyme E4 is not functional); (d) mutasynthesis: a mutasynthon B* is introduced to replace B and is incorporated in the biosynthesis, leading to a "mutated" natural product.

Both approaches can potentially greatly increase the diversity of natural compounds.

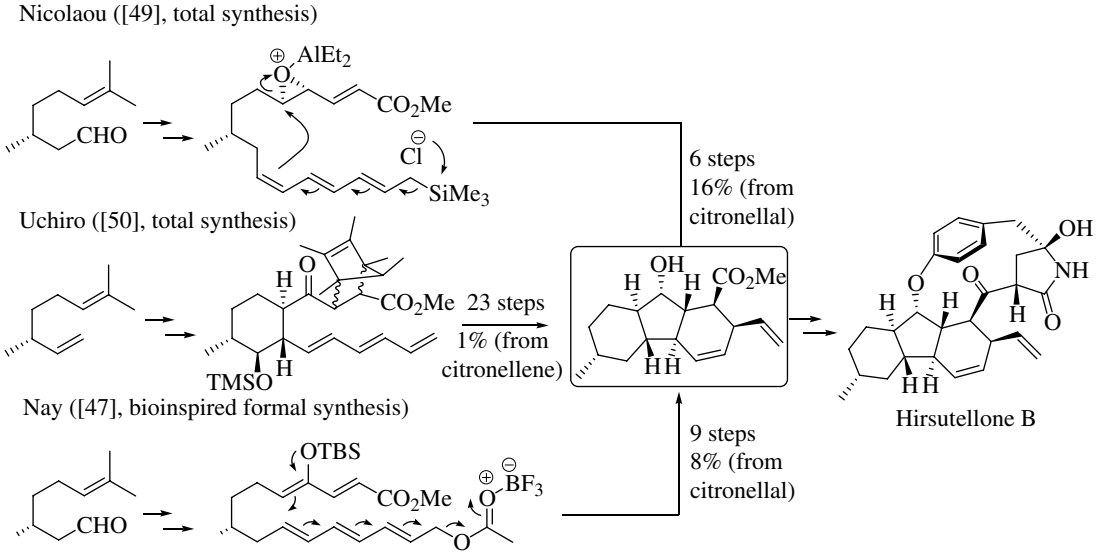
1.2.2.2 The Biomimetic Strategy: A Bridge between Biosynthesis and Total Synthesis During the past century, synthetic chemists were endeavoring to discover more efficient strategies to access complex natural products. The chemical synthesis of tropinone by Robinson in 1917 [44], one of the first biomimetic ones, is a fantastic example of an early efficient synthesis, which consisted in a multicomponent process between succinaldehyde, methylamine, and calcium acetonedicarboxylate [45]. Since then, the construction of natural products by chemical methods inspired by nature's biosynthetic pathways has attracted many synthetic chemists and participated in the progress of organic chemistry. As discussed in the book *Biomimetic Organic Synthesis* coedited by one of us (B.N.), an increasing number of total syntheses have been termed "biomimetic" or "bioinspired" during the last 20 years, meaning the use of a synthetic tactic that follows or mimics a hypothetical or proven biosynthetic pathway. Concomitantly, the biosynthesis of natural products has been more and more understood, thanks to genetic and enzymatic studies. Therefore, as a bridge between biosynthesis and total synthesis, biomimetic synthesis is able to overcome some drawbacks of conventional strategies, as it often relies on the self-assembling properties of a key reactive intermediate [46].

Tremendous works dealing with bioinspired total syntheses of secondary metabolites have thus been achieved, providing new insights in the reactivity of biomimetic precursors and occasionally leading to controversy or unresolved questions [47]. An interesting example goes to hirsutellones, a family of fungal PKS/NRPS compounds (also regarded as alkaloids due to their nitrogen) with intriguing structures and a significant antitubercular activity [7]. Their biosynthesis has been hypothesized by Oikawa who proposed a key linear precursor of the related compounds GKK1032A$_2$, made from one tyrosine, nine AcCoA, and several methylations by *S*-adenosylmethionine [48]. We applied this hypothesis to the less methylated hirsutellones (Scheme 1.10). Two different biosynthetic pathways can be proposed for the polycyclization. Pathway (a) involves the selective oxidation of one of the dienoyl double bonds (C_γ=C_δ) to generate an epoxide and of the phenol. This "electrophilic head" would then be attacked in a conjugated ene reaction involving the triene and initiating the cyclization. Formation of the bent paracyclophane would then be followed by a stereoselective intramolecular Diels–Alder (IMDA) reaction leading to the complete tricyclic core of the natural product. Pathway (b) involves the allylic oxidation at the terminal methyl group of the triene to release an allylic alcohol or cation as an "electrophilic tail." The polycyclization would then be initiated through reverse electronic activation compared to pathway (a), forming the first cyclohexane ring before the IMDA reaction occurs.

Nicolaou et al. [49] and Uchiro et al. [50] achieved the total syntheses of hirsutellone B in 2009 and 2011, respectively. We recently described a formal total synthesis by forming the key decahydrofluorene (tricyclic) core of hirsutellone in a biomimetic strategy following pathway (b) [47]. As for the synthesis of this important synthetic



SCHEME 1.10 Biosynthetic hypotheses for the biosynthesis of hirsutellones. Pathway (a) would involve the transient loss of aromaticity of the phenol and then rearomatization with bending of the paracyclophane, while pathway (b) would involve the direct attack of the phenol with concomitant formation of the bent macrocycle. The IMDA reaction would proceed lately.



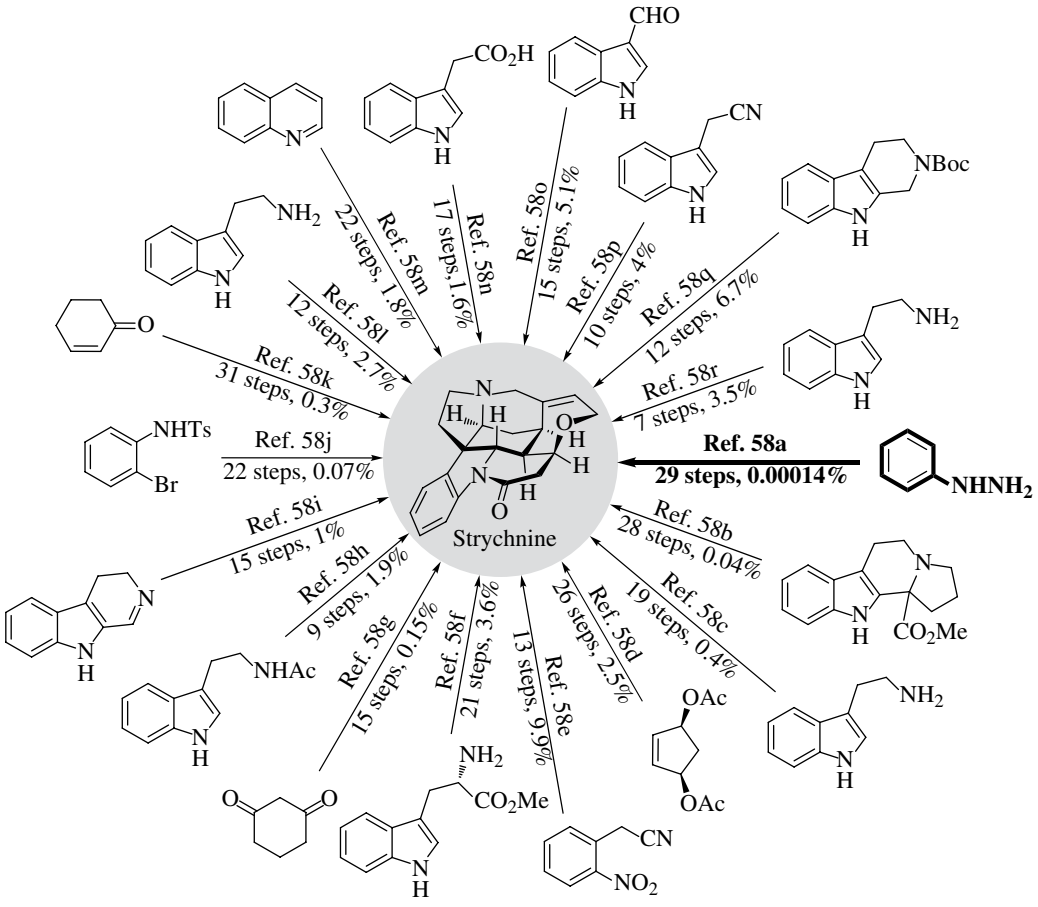
SCHEME 1.11 Biomimetic and conventional strategies toward hirsutellones: Nicolaou's total synthesis was not originally reported as biomimetic but supports the proposed tail-to-head biosynthetic pathway; Uchiro reported a conventional total synthesis; we reported a biomimetic formal synthesis leading to Nicolaou's intermediate, supporting the head-to-tail biosynthetic pathway.

intermediate with eight stereocenters, Uchiro's nonbiomimetic strategy took 23 steps from *R*-(–)-citronellene with 1% global yield (Scheme 1.11). In comparison, Nicolaou's synthetic strategy, involving an Et_2AlCl-triggered cascade cyclization, decreased the number of reaction steps to six steps starting from *R*-(+)-citronellal and with 16% overall yield. Although this was not claimed as biomimetic by the authors, this work supports the "tail-to-head" biosynthetic pathway (a) (Scheme 1.10) and nicely reveals the efficiency of biosynthetically related cascade reactions. We reported an alternative biomimetic synthesis of the tricyclic core of hirsutellones by a reverse "head-to-tail" cyclization strategy using nine steps and with 8% brsm global yield (Scheme 1.11) [47]. Interestingly, our strategy supports the biosynthetic pathway (b), thus confirming that both biosynthetic routes are possible. However, thanks to recent biosynthetic experiments using the isotopically labeled precursor (^{18}O-*phenol*)-L-tyrosine, we demonstrated that the phenolic oxygen is incorporated in analogous natural products, pyrrocidines, thus giving clues to biosynthetic pathway (b) [51].

1.2.3 The Science of Total Synthesis

1.2.3.1 The Evolution of Total Synthesis and Its Significance Today The vast utility of total synthesis and its connections with other research fields can be illustrated by a selection of key words, some of them deeply resonating with current major societal challenges: medicinal chemistry and new drugs, pharmacology, agrochemicals, biosynthetic studies, synthetic methodologies, structure determination, physical organic chemistry, catalysis, green resources, or bioinspiration. Back to the nineteenth century, the first organic synthesis of urea from ammonium cyanate, an inorganic substance, was accomplished by Wöhler in 1828 and raised the curtain of total synthesis. Total synthesis had then, for a time, played an essential role on elucidating the structure of natural products, and it is still the case nowadays when the determination of relative and absolute configurations cannot be achieved by analytic methods. With the improvement of analytical chemistry, and as chemistry and biology are better understood, the role of total synthesis slowly changed. A variety of new reactions, catalysts, and technologies have been developed for total synthesis. Most importantly, total synthesis is playing a key role for new drug discovery, chemical biology, or even material science. As introduced in the former part, a lot of natural products and derivatives were developed to provide new drugs against human diseases (Fig. 1.5), of which total synthesis enabled a larger amount of products available for further studies [52] and challenged optimized strategies for their industrial production [35]. As striking examples, we can cite Paterson's synthesis of discodermolide at the 60 g scale for anticancer clinical studies by Novartis [53], or the recent industrial production of the antimalarial drug artemisinin by Sanofi, using a semisynthetic strategy starting from a biotechnologically available advanced intermediate [54, 55].

1.2.3.2 Strychnine as a Case Study: A Classic among the Classics Herein, we would like to illustrate the evolution of total synthesis by one of the most famous natural products, strychnine (Scheme 1.12). For decades, strychnine was regarded as one of the most challenging natural products to be synthesized [56]. The correct structure of strychnine was determined by Woodward and Brehm in 1948, one century after its discovery [57]. Since then, this remarkable natural product witnessed the evolution of total synthesis.



SCHEME 1.12 The 18 total syntheses of strychnine (1954–2011) [58a–r].

The landmark synthesis of strychnine was reported by Woodward and coworkers in 1954, 6 years after its structure determination [58].

Since then, many synthetic chemists have been confronted with strychnine, which is a classic among the classics of total synthesis. Overall, 18 total syntheses of strychnine have been reported so far [58a–r], the shortest one in only 7 linear steps from tryptamine by Vanderwal [58r], to be compared with the earliest total synthesis of Woodward in 29 linear steps from phenylhydrazine [53a]. The efficiency of these works can be evaluated by looking at the overall yields, from 0.00014% [58a] to nearly 10% yield [58e]. For sure, these improvements not only took benefits from Corey's retrosynthetic "Logic of Chemical Synthesis" but also from those famous chemists's creativity and from new achievements in synthetic and catalytic methodologies. Indeed, new methodological concepts have arisen by the last 20 years, such as those of ideal synthesis [59], atom economy, step economy, redox economy, and sustainable approaches [60].

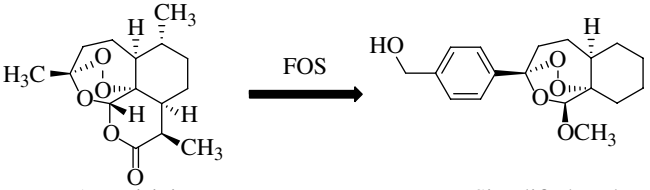
The efficiency of total synthesis should then benefit to the growing research efforts in chemical biology and drug discovery in the future, in connection with recently designed strategies like diversity-oriented synthesis (DOS) and function-oriented synthesis (FOS).

1.2.3.3 DOS and FOS: Two Strategies to Optimize Biological Hits and Synthetic Efficacy Classical combinatorial chemistry has allowed for the synthesis of vast amounts of products, yet poorly overlapping the chemical space of natural products, essentially due to their limited structural diversity and drug-likeness. Chemists are thus searching for ever more efficient ways to generate rapidly more complex and diverse functional compounds. As discussed before, precursor-directed biosynthesis and mutasynthesis have been developed by biochemists and exemplify a biological mean to diversify structures in a natural product series. In addition, organic chemists have designed new strategies for this purpose, such as DOS and FOS.

DOS and Divergent Total Synthesis DOS, often compared with the classical target-oriented synthesis (TOS), is using forward synthetic analysis with the aim of transforming various building blocks, through planned reactions, to efficiently generate complex and diverse compounds matching with a large portion of the chemical space. To some extent, this is the opposite way as the well-established retrosynthetic analysis of TOS. The strategy of DOS mainly relies on the variation of three parameters [61]: (1) the building blocks, to introduce a vast number of functional groups in the skeleton; (2) the stereochemistry, which can be introduced by various stereoselective reagents; and (3) the molecular skeleton, which could achieve the highest level of structural complexity and diversity by using different synthetic methods, such as multicomponent reactions, combinational synthesis, folding pathway, and branched pathway [62]. In any case, DOS greatly increases the chemical space to enable more biological and pharmaceutical investigation.

Using an analogous strategy, diverse natural products were synthesized through collective natural product synthesis or divergent total synthesis. This powerful concept was applied by MacMillan and coworkers to the asymmetric synthesis of six monoterpene indole alkaloids using organocascade catalysis (strychnine, aspidospermidine, vincadifformine, akuammicine, kopsanone, and kopsinine) [63]. Dai and coworkers exploited the combination of a biosynthetically inspired strategy with such a divergent approach for the synthesis of seven monoterpene indole alkaloids (mersicarpine, leuconodines B and D, leuconoxine, melodinine E, leuconolam, and rhazinilam) [64]. In the taiwaniaquinoid series, four natural products were synthesized by Li and coworkers after two to three steps from a common intermediate prepared on the gram scale [65]. Such collective strategies are more and more encountered in the literature, taking advantage of a common synthetic route leading to key intermediates to access entire families of compounds rather than a sole natural product target.

FOS The common problem encountered with total synthesis is the high complexity of natural products, which often takes many steps to be achieved and lowers overall yields. One way to solve this problem is, as discussed before, to think and design efficient synthetic strategies, for example, using redox or step economy, to shorter the route. Another approach is to design less complex synthetic targets by maintaining or improving selected functions involved in the biological activity, which is the so-called FOS. FOS is based on the study of complex targeted molecules with the aim of shortening the synthetic work to develop diverse simplified but still functional targets, keeping key structural features to effect biological functions [66]. FOS is thus deeply related to drug discovery. Many simplified small compounds can indeed be proposed from structure–activity relationship studies. For example, the famous antimalarial artemisinin gave simplified but functional analogs with potent *in vitro* antimalarial activities in the same range of IC_{50} as that of the natural product (Scheme 1.13) [67]. Other than DOS, which focuses on structure complexity and diversity, FOS concentrates more



Artemisinin
IC_{50} = 9.2 nM

Simplified analog
IC_{50} = 15 nM

SCHEME 1.13 The FOS strategy toward a simplified but still potent analog of the antimalarial artemisinin.

on the functional groups involved in the biological functions, while both of them are somehow inspired by total synthesis.

1.2.4 Conclusion: A Journey in the Future of Total Synthesis

The future of total synthesis is written in our laboratory notebooks. It will not only be conditioned by new synthetic achievements and new methodologies and technologies improving the efficacy of experiments but also by their applications to answer questions from new horizons. All of us will agree, as it was said by others, that total synthesis is marked by beauty and it has sometimes been compared with art. Not so many fields can respond to such criteria, and it is due to the free creativity we are able to exert. In theory, total synthesis could provide any compound, from the simplest to the most complex ones. But can we provide enough material for deepened studies in other research fields [52]? Indeed, our products, once achieved, are not to be stored indefinitely in tiny flasks. They should lead to new projects, new questions, and new answers.

Thus, how studying in depth the biological, the chemical, and the physical properties of a natural product when its natural source is rare, low producing, and sometimes no more available? This question is in the hand of two scientific communities: the biotechnological and the synthetic chemist ones. Let's bet that we will still answer many of such questions by continuing to improve qualitatively and quantitatively our productivity by making our syntheses simpler and faster (and thus, as we may say, more "elegant") and by being the driving forces in building strong interdisciplinary bridges.

Further Reading on Total Synthesis and Biosynthesis

- J.-N. Bruneton, *Pharmacognosie, phytochimie, plantes medicinales*, Tec & Doc Lavoisier: Paris (2009)
- E. J. Corey and X.-M. Cheng, *The Logic of Chemical Synthesis*, Wiley: New York (1989)
- J. Cossy and S. Arseniyadis, *Modern tools for the synthesis of complex bioactive molecules*, Wiley: Hoboken (2012)
- P. M. Dewick, *Medicinal natural products, a biosynthetic approach*, Wiley: Chichester (2009)
- T. Hudlicky and J. W. Reed, *The way of synthesis*, Wiley-VCH: Weinheim (2007)
- K. C. Nicolaou and E. J. Sorensen, *Classics in total synthesis*, Wiley-VCH: Weinheim (1996)
- K. C. Nicolaou and S. A. Snyder, *Classics in total synthesis II*, Wiley-VCH: Weinheim (2003)
- K. C. Nicolaou and J. S. Chen, *Classics in total synthesis III*, Wiley-VCH: Weinheim (2011)
- E. Poupon and B. Nay (Eds), *Biomimetic Organic Synthesis*, Wiley-VCH: Weinheim (2011)

The reader interested in biosynthetic pathways can also refer to the interactive KEGG atlas (biosynthetic pathways) available on Internet: http://www.kegg.jp/kegg/atlas/.

REFERENCES

1. http://www.enzyme-database.org/class.php (accessed September 28, 2015).
2. T. D. H. Bugg, *Nat. Prod. Rep.* 2001, **18**, 465–493.
3. P. M. Dewick, *Medicinal natural products, a biosynthetic approach*, Wiley: Chichester (2009): refer to Chapter 2 for a description of the construction mechanisms.
4. R. Thomas, *Nat. Prod. Rep.* 2004, **21**, 224–248.
5. C. Hertweck, *Angew. Chem. Int. Ed.* 2009, **48**, 4688–4716.
6. M. A. Fischbach, C. T. Walsh, *Chem. Rev.* 2006, **106**, 3468–3496.
7. X.-W. Li, A. Ear, B. Nay, *Nat. Prod. Rep.* 2013, **30**, 765–782.
8. D. Boettger, C. Hertweck, *ChemBioChem* 2013, **14**, 28–42.
9. S. Frick, R. Nagel, A. Schmidt, R. Bodemann, P. Rahfeld, G. Pauls, W. Brandt, J. Gershenzon, W. Boland, A. Burse, *Proc. Natl. Acad. Sci. U.S.A* 2013, **110**, 4194–4199.
10. Example of the trichodiene synthase: M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *Proc. Natl. Acad. Sci. U.S.A* 2001, **98**, 13543–13548.
11. Example of taxadiene synthase: R. Croteau, R. E. B. Ketchum, R. M. Long, R. Kaspera, M. R. Wildung, *Phytochem. Rev.* 2006, **5**, 75–97.
12. L. M. Podust, D. H. Sherman, *Nat. Prod. Rep.* 2012, **29**, 1251–1266.
13. L. Pouységu, D. Deffieux, G. Malik, A. Natangelo, S. Quideau, *Nat. Prod. Rep.* 2011, **28**, 853–874.
14. T. Umezawa, *Phytochem. Rev.* 2003, **2**, 371–390.
15. L. B. Davin, N. G. Lewis, *Phytochem. Rev.* 2003, **2**, 257–288.
16. F. Bourgaud, A. Hehn, R. Larbat, S. Doerper, E. Gontier, S. Kellner, U. Matern, *Phytochem. Rev.* 2006, **5**, 293–308.
17. D. Yu, F. Xu, J. Zeng, J. Zhan, *IUBMB Life* 2012, **64**, 285–295.
18. K. Saito, K. Yonekura-Sakakibara, R. Nakabayashi, Y. Higashi, M. Yamazaki, T. Tohge, A. R. Fernie, *Plant Physiol. Biochem.* 2013, **72**, 21–34.
19. J. Chong, A. Poutaraud, P. Hugueney, *Plant Sci.* 2009, **177**, 143–155.
20. S. Biastoff, W. Brandt, B. Dräger, *Phytochemistry* 2009, **70**, 1708–1718.
21. A. J. Humphrey, D. O'Hagan, *Nat. Prod. Rep.* 2001, **18**, 494–502.
22. B. Dräger, *Nat. Prod. Rep.* 2004, **21**, 211–223.
23. D. Ober, E. Kaltenegger, *Phytochemistry* 2009, **70**, 1687–1695.
24. D. S. Seigler, *in Plant secondary metabolism*, Springer US, Boston (1999), pp. 546–567.
25. J. Stöckigt, A. P. Antonchick, F. Wu, H. Waldmann, *Angew. Chem. Int. Ed.* 2011, **50**, 8538–8564.

26. J. Ziegler, P. J. Facchini, R. Geißler, J. Schmidt, C. Ammera, R Kramell, S. Voigtländer, A. Gesell, S. Pienkny, W. Brandt, *Phytochemistry* 2009, **70**, 1696–1707.
27. V. Sharma, S. Jain, D. S. Bhakuni, R. S. Kapil, *J. Chem. Soc. Perkin* 1982, **1**, 1153–1155.
28. A. R. Battersby, R. B. Herbert, E. McDonald, R. Ramage, J. H. Clements, *J. Chem. Soc. Perkin* 1972, **1**, 1741–1746.
29. H. Abdelkafi, B. Nay, *Nat. Prod. Rep.* 2012, **29**, 845–869.
30. J. Eichhorn, T. Takada, Y. Kita, M. H. Zenk, *Phytochemistry* 1998, **49**, 1037–1047.
31. A. M. Takos, F. Rook, *Int. J. Mol. Sci.* 2013, **14**, 11713–11741.
32. S. E. O'Connor, J. J. Maresh, *Nat. Prod. Rep.* 2006, **23**, 532–547.
33. J. Stöckigt, S. Panjikar, *Nat. Prod. Rep.* 2007, **24**, 1382–1400.
34. C. L. Schardl, D. G. Panaccione, P. Tudzynski, *in The Alkaloids: Chemistry and Biology*, Vol. **63**, G. A. Cordell (Ed.), Elsevier: Amsterdam (2006), 45–86.
35. G. Yim, M. N. Thaker, K. Koteva, G. Wright, *J. Antibiot.* 2014, **67**, 31–41.
36. B. K. Hubbard, C. T. Walsh, *Angew. Chem. Int. Ed.* 2003, **42**, 730–765.
37. C. M. Dobson, *Nature* 2004, **432**, 824–828.
38. L. H. Caporale, *Proc. Natl. Acad. Sci. U.S.A* 1995, **92**, 75–82.
39. A. Bauer, M. Brönstrup, *Nat. Prod. Rep.* 2013, **31**, 35–60.
40. T. Liu, Y.-M. Chiang, A. D. Somoza, B. R. Oakley, C. C. C. Wang, *J. Am. Chem. Soc.* 2011, **133**, 13314–13316.
41. D. E. Cane, F. Kudo, K. Kinoshita, C. Khosla, *Chem. Biol.* 2002, **9**, 131–142.
42. C. J. B. Harvey, J. D. Puglisi, V. S. Pande, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* 2012, **134**, 12259–12265.
43. (a) A. Kirschning, F. Hahn, *Angew. Chem. Int. Ed.* 2012, **51**, 4012–4022; (b) S. Weist, R. D. Süssmuth, *Appl. Microbiol. Biotechnol.* 2005, **68**, 141–150; (c) K. Weissman, *Trends Biotechnol.* 2007, **25**, 139–142.
44. R. Robinson, *J. Chem. Soc. Trans.*, 1917, **111**, 762–768.
45. J. W. Medley, M. Movassaghi, *Chem. Commun.* 2013, **49**, 10775–10777.
46. E. Gravel, E. Poupon, *Eur. J. Org. Chem.* 2008, 27–42.
47. X.-W. Li, A. Ear, L. Roger, N. Riache, A. Deville, B. Nay, *Chem. Eur. J.* 2013, 16389–16393.
48. H. Oikawa, *J. Org. Chem.* 2003, **68**, 3552–3557.
49. K. C. Nicolaou, D. Sarlah, T. R. Wu, W. Zhan, *Angew. Chem. Int. Ed.* 2009, **48**, 6870–6874.
50. H. Uchiro, R. Kato, Y. Arai, M. Hasegawa, Y. Kobayakawa, *Org. Lett.* 2011, **13**, 6268–6271.
51. A. Ear, S. Amand, F. Blanchard, A. Blond, L. Dubost, D. Buisson, B. Nay, *Org. Biomol. Chem.* 2015, **13**, 3662–3666.
52. C. A. Kuttruff, M. D. Eastgate, P. S. Baran, *Nat. Prod. Rep.* 2014, **31**, 419–432.
53. (a) S. J. Mickel, G. H. Sedelmeier, D. Niederer et al., *Org. Proc. Res. Dev.* 2004, **8**, 92–100 (Part 1); (b) S. J. Mickel, G. H. Sedelmeier, D. Niederer et al., *Org. Proc. Res. Dev.* 2004, **8**, 101–106 (Part 2); (c) S. J. Mickel, G. H. Sedelmeier, D. Niederer et al., *Org. Proc. Res. Dev.* 2004, **8**, 107–112 (Part 3); (d) S. J. Mickel, G. H. Sedelmeier, D. Niederer et al., *Org. Proc. Res. Dev.* 2004, **8**, 113–121 (Part 4); (e) S. J. Mickel, D. Niederer, R. Daeffler et al., *Org. Proc. Res. Dev.* 2004, **8**, 122–130 (Part 5).
54. C. J. Paddon, P. J. Westfall, D. J. Pitera et al., *Nature* 2013, **496**, 528–532.
55. (a) J. Dhainaut, A. Dlubala, R. Guevel, A. Medard, G. Oddon, N. Raymond, J. Turconi (Sanofi –Aventis, Fr.) WO 2011/026865 (2011); (b) V. Kraft, G. Kretzschmar, K. Rossen (Sanofi –Aventis, Fr.) WO 2011/030223 (2011).
56. J. S. Cannon, L. E. Overman, *Angew. Chem. Int. Ed.* 2012, **51**, 4288–4311.
57. R. B. Woodward, W. J. Brehm, *J. Am. Chem. Soc.* 1948, **70**, 2107–2115.
58. (a) R. B. Woodward, M. P. Cava, W. D. Ollis, A. Hunger, K. Schenker, *J. Am. Chem. Soc.* 1954, **76**, 4749–4761; (b) P. Magnus, M. Giles, R. Bonnert, C. S. Kim, L. McQuire, A. Merritt, N. Vicker, *J. Am. Chem. Soc.* 1992, **114**, 4403–4405; (c) M. E. Kuehne, F. Xu, *J. Org. Chem.* 1993, **58**, 7490–7497; (d) S. D. Knight, L. E. Overman, G. Pairaudeau, *J. Am. Chem. Soc.* 1993, **115**, 9293–9294; (e) V. H. Rawal, S. Iwasa, *J. Org. Chem.* 1994, **59**, 2685–2686; (f) M. E. Kuehne, F. Xu, *J. Org. Chem.* 1998, **63**, 9427–9433; (g) D. Solé, J. Bonjoch, S. García-Rubio, E. Peidró, J. Bosch, *Angew. Chem. Int. Ed.* 1999, **38**, 395–397; (h) M. J. Eichberg, R. L. Dorta, K. Lamottke, K. P. C. Vollhardt, *Org. Lett.* 2000, **2**, 2479–2481; (i) M. Ito, C. W. Clark, M. Mortimore, J. B. Goh, S. F. Martin, *J. Am. Chem. Soc.* 2001, **123**, 8003–8010; (j) M. Nakanishi, M. Mori, *Angew. Chem. Int. Ed.* 2002, **41**, 1934–1936; (k) T. Ohshima, Y. Xu, R. Takita, S. Shimizu, D. Zhong, M. Shibasaki, *J. Am. Chem. Soc.* 2002, **124**, 14546–14547; (l) G. J. Bodwell, J. Li, *Angew. Chem. Int. Ed.* 2002, **41**, 3261–3262; (m) Y. Kaburagi, H. Tokuyama, T. Fukuyama, *J. Am. Chem. Soc.* 2004, **126**, 10246–10247; (n) H. Zhang, J. Boonsombat, A. Padwa, *Org. Lett.* 2007, **9** 279–282; (o) G. Sirasani, T. Paul, W. Dougherty Jr., S. Kassel, R. B. Andrade, *J. Org. Chem.* 2010, **75**, 3529–3532; (p) C. Beemelmanns, H.-U. Reissig, *Angew. Chem. Int. Ed.* 2010, **49**, 8021–8025; (q) S. B. Jones, B. Simmons, A. Mastracchio, D. W. C. MacMillan, *Nature* 2011, **475**, 183–188; (r) D. B. C. Martin, C. D. Vanderwal, *Chem. Sci.* 2011, **2**, 649–651.
59. (a) P. A. Wender, *Nat. Prod. Rep.* 2014, **31**, 433–440; (b) T. Gaich, P. S. Baran, *J. Org. Chem.* 2010, **75**, 4657–4673.
60. T. Newhouse, P. S. Baran, R. W. Hoffmann, *Chem. Soc. Rev.* 2009, **38**, 3010–3021.
61. (a) M. D. Burke, G. Lalic, *Chem. Biol.* 2002, **9**, 535–541; (b) D. R. Spring, *Org. Biomol. Chem.* 2003, **1**, 3867–3870.
62. C. Cordier, D. Morton, S. Murrison, A. Nelson, C. O'Leary-Steele, *Nat. Prod. Rep.* 2008, **25**, 719–737.
63. S. B. Jones, B. Simmons, A. Mastracchio, D. W. C. MacMillan, *Nature* 2011, **475**, 183–188.
64. Y. Yang, Y. Bai, S. Sun, M. Dai, *Org. Lett.* 2014, **16**, 6216–6219.
65. J. Deng, R. Li, Y. Luo, J. Li, S. Zhou, Y. Li, J. Hu, A. Li, *Org. Lett.* 2013, **15**, 2022–2025.
66. P. A. Wender, V. A. Verma, T. J. Paxton, T. H. Pillow, *Acc. Chem. Res.* 2008, **41**, 40–49.
67. G. H. Posner, P. M. O'Neill, *Acc. Chem. Res.* 2004, **37**, 397–404.

SECTION I

ACETATE BIOSYNTHETIC PATHWAY

2

POLYKETIDES

FRANÇOISE SCHAEFERS AND TOBIAS A. M. GULDER

Department of Chemistry and Center for Integrated Protein Science Munich (CIPSM), Biosystems Chemistry, Technische Universität München, Munich, Germany

2.1 POLYKETIDE BIOSYNTHESIS

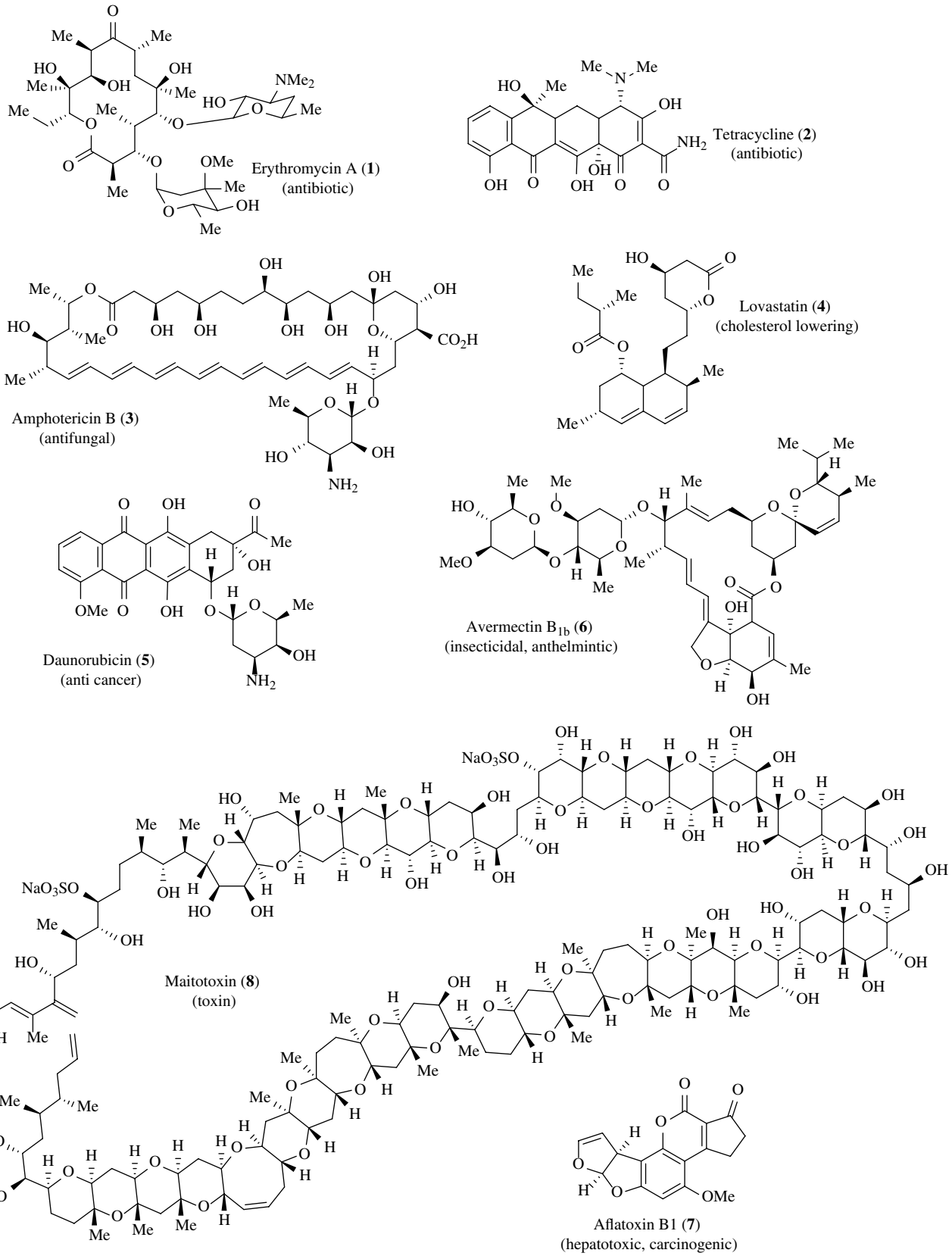
2.1.1 Introduction

Polyketides are a fascinating class of natural products with highly diverse, often complex structures combined with a likewise impressive breath of strong and selective biological activities. Because of these extraordinary properties, polyketides continue to fascinate and inspire medicinal, synthetic, and natural product chemists and biochemists alike. This appeal of polyketides can easily be illustrated by some representative examples shown in Scheme 2.1. From a structural point of view, this limited selection already comprises a tremendous diversity, from small aromatic and aliphatic polycyclic compounds, to densely functionalized and stereochemically challenging macrolides, up to strikingly large and complex polyethers. In addition, these compounds exhibit a multitude of different biological properties with many important clinical applications: erythromycin A (**1**) [1] and tetracyclines [2], for example **2**, are widely applied antibiotics. The polyene amphotericin B (**3**) is an effective agent against various systemic fungal infections and shows antiprotozoal activity [3]. Lovastatin (**4**) is an effective cholesterol-lowering agent and has served as a blueprint for the development of a large set of synthetic statins [4]. The DNA intercalating agent daunorubicin (**5**) is a potent anticancer chemotherapeutic [5]. The avermectin family of antiparasitics, exemplified by **6**, have been instrumental for the treatment of severe maladies caused by parasitic worms, in particular onchocerciasis (river blindness) [6]. The biomedical significance of this small selection of polyketides is underlined by the fact that they—and/or synthetic derivatives inspired by them—can all be found on the World Health Organization's List of Essential Medicines [7].

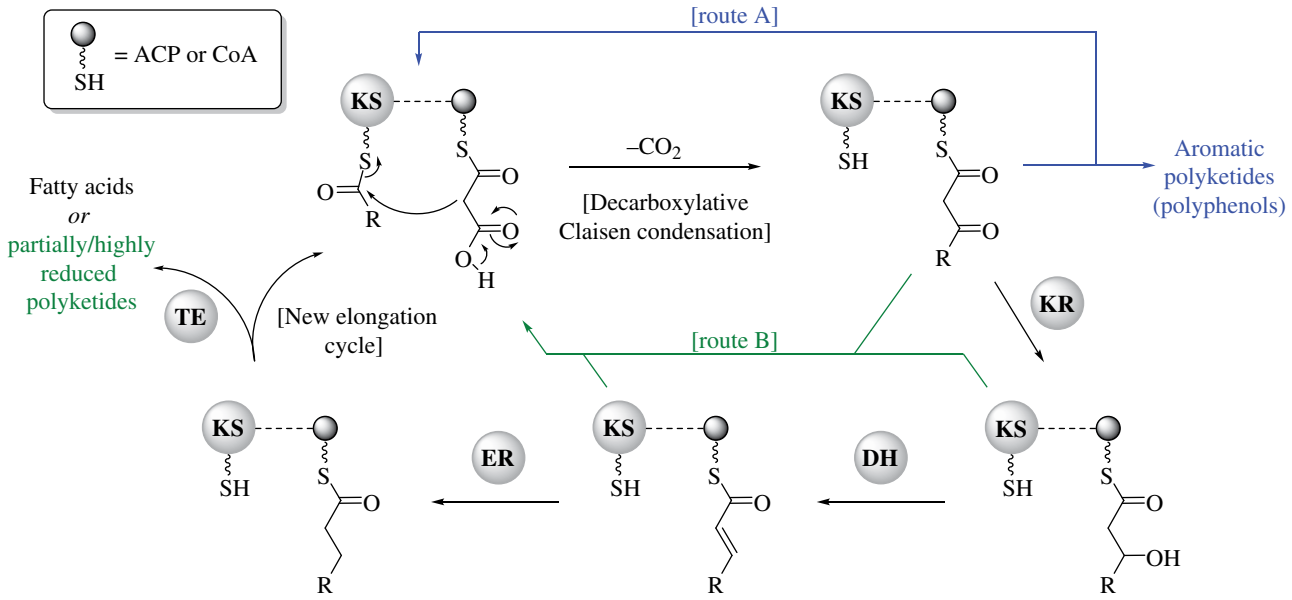
Besides these desirable biomedical properties, the impact of the biological activity of polyketides on mankind also extends to detrimental effects. Prominent examples are the aflatoxins, such as B1 (**7**), produced by *Aspergillus* sp. [8]. This group of compounds exhibits acute hepatotoxic effects and strongly increases the risk of developing liver cancer upon chronic exposure [9]. The compounds can primarily be detected on dried goods, such as corn, nuts, fruits, and spices, but also find their way into fresh produce, such as meat or milk. Aflatoxin levels in food are thus tightly regulated (with typical tolerance levels in between 0 and 30 μg/kg). Another impressive example is maitotoxin (**8**), the largest nonpolymeric, nonpeptidic organic molecule isolated from nature until today [10]. This giant polyether is produced by dinoflagellates (e.g., *Gambierdiscus toxicus*) and moves up the food chain to ultimately enrich in fish, for example, in *Ctenochaetus striatus* (called maito in Tahiti) from which **8** was initially discovered. Ingestion of such contaminated food leads to the severe illness ciguatera. This effect is caused by the tremendous toxicity of **8** (LD_{50} of 50 ng/kg in mice).

Equally impressive as the polyketide structural and functional diversity are the enzymatic machineries that evolved to produce these fascinating secondary metabolites,

From Biosynthesis to Total Synthesis: Strategies and Tactics for Natural Products, First Edition. Edited by Alexandros L. Zografos.



SCHEME 2.1 Selection of structurally and functionally diverse polyketide natural products **1–8**.



SCHEME 2.2 Biosynthetic routes to fatty acids and polyketides.

almost exclusively employing simple acetate and malonate building blocks for core structure assembly. Within this chapter, we discuss the logic of biosynthetic polyketide assembly. This diverse topic has been comprehensively reviewed in excellent general articles in recent years [11–15]. We herein focus on the introduction of the basic mechanisms and concepts of polyketide core structure assembly using a number of well-studied representative examples, thus allowing to compare the biosynthetic versus chemical synthetic strategies (see following chapters) to assemble such beautiful small molecules.

2.1.2 Assembly of Acetate/Malonate-Derived Metabolites

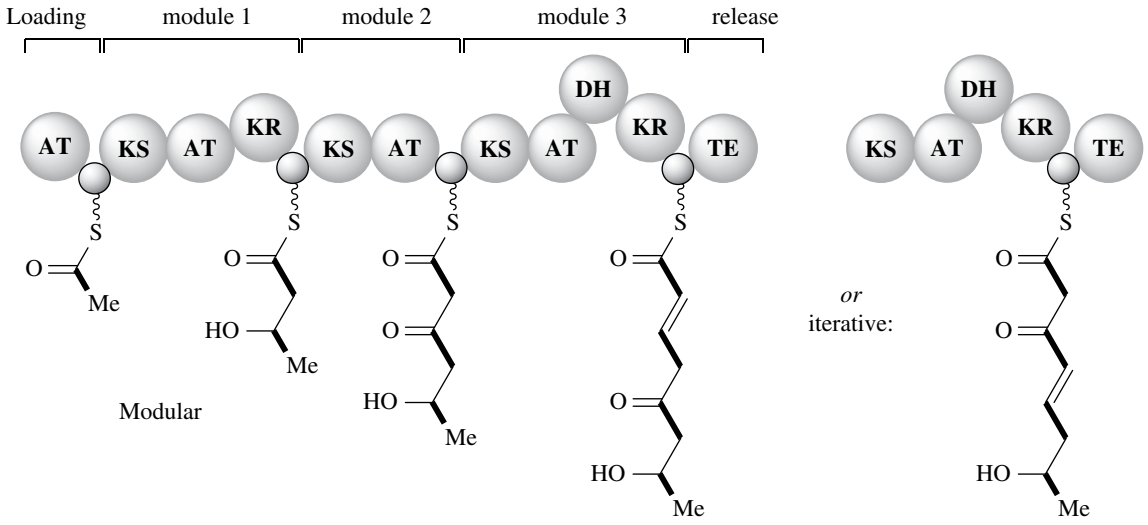
The basic principle of polyketide assembly is highly related to that of fatty acid biosynthesis [14, 16]. In both biosynthetic systems, an acyl-primed ketosynthase (KS) catalyzes chain extension by decarboxylative Claisen condensation with malonate activated by its attachment to coenzyme A or an acyl carrier protein (ACP) via a thioester bond (Scheme 2.2). In fatty acid synthases (FASs), the resulting ketone is reduced to the corresponding alcohol by a ketoreductase (KR), dehydrated by action of a dehydratase (DH) to give the alkene with subsequent double-bond reduction by an enoyl reductase (ER) yielding the saturated system (cf. Section 3.2). The latter can then be transferred onto the KS domain and enter the next cycle of chain extension and complete reduction. This homologation process facilitates the assembly of long-chain saturated fatty acids, for example, palmitic acid, after seven cycles, which will ultimately be released from the catalytic system by saponification of the linkage to the thiol activating group, for example, catalyzed by a thioesterase (TE) domain [16].

Remarkably, polyketide synthases (PKSs) utilize the identical, small set of biosynthetic reactions to generate the tremendous structural diversity observed in the polyketide natural product family. The diversification of core structures accessible by PKSs when compared to FASs is achieved by flexible application of the reductive transformations processing the ß-keto groups resulting from chain extension. The class of aromatic polyphenols thereby can be accessed by omitting reductions (Scheme 2.2, route A). Alternatively, the oxidation state at the ß-carbon can individually be controlled after each elongation step to keep ketones, alcohols, or double bonds in the respective positions (Scheme 2.2, route B), leading to highly or partially reduced PKS products. Together with an increased flexibility concerning the acyl and malonyl building blocks combined with further post-PKS modifications—such as hydroxylation, halogenation, or glycosylation tailoring reactions—this strategy gives rise to a virtually inexhaustible diversity of possible product structures.

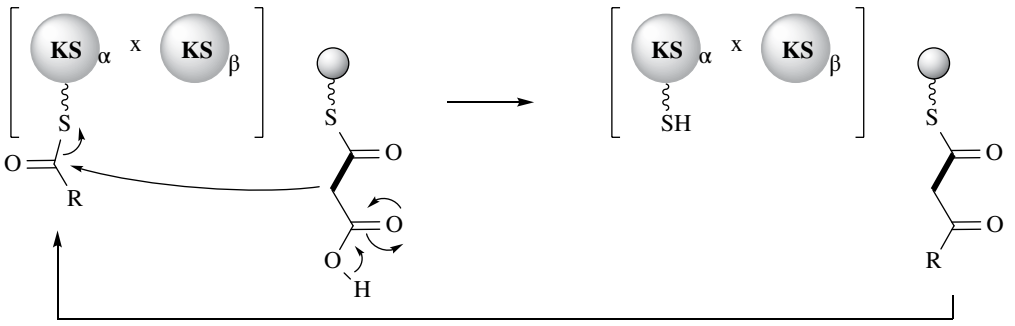
2.1.3 Classification of Polyketide Biosynthetic Machineries

Depending on the overall structural organization of the individual catalytic units, their exact function, and precursor requirements, PKSs can roughly be classified into three substantially different types (Scheme 2.3) [11–13]. In analogy to the FAS classification, type I PKSs are multifunctional megaenzymes bearing all catalytic units on a large polypeptide [14, 17–22]. Type II PKSs are dissociable complexes

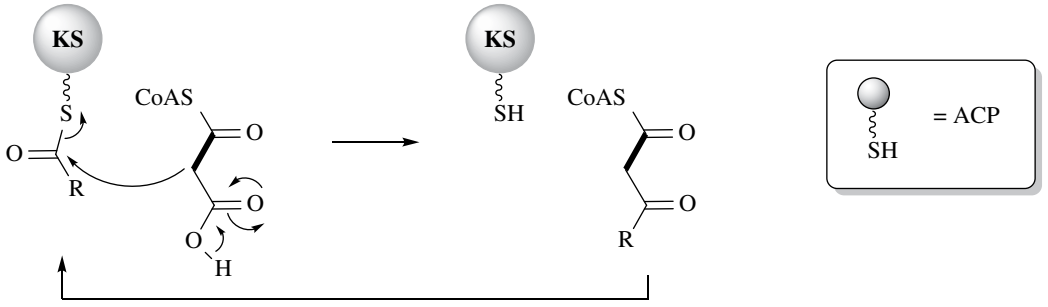
(a)



(b)



(c)



SCHEME 2.3 Classification of PKSs. (a) Type I PKS. (b) Type II PKS. (c) Type III PKS.

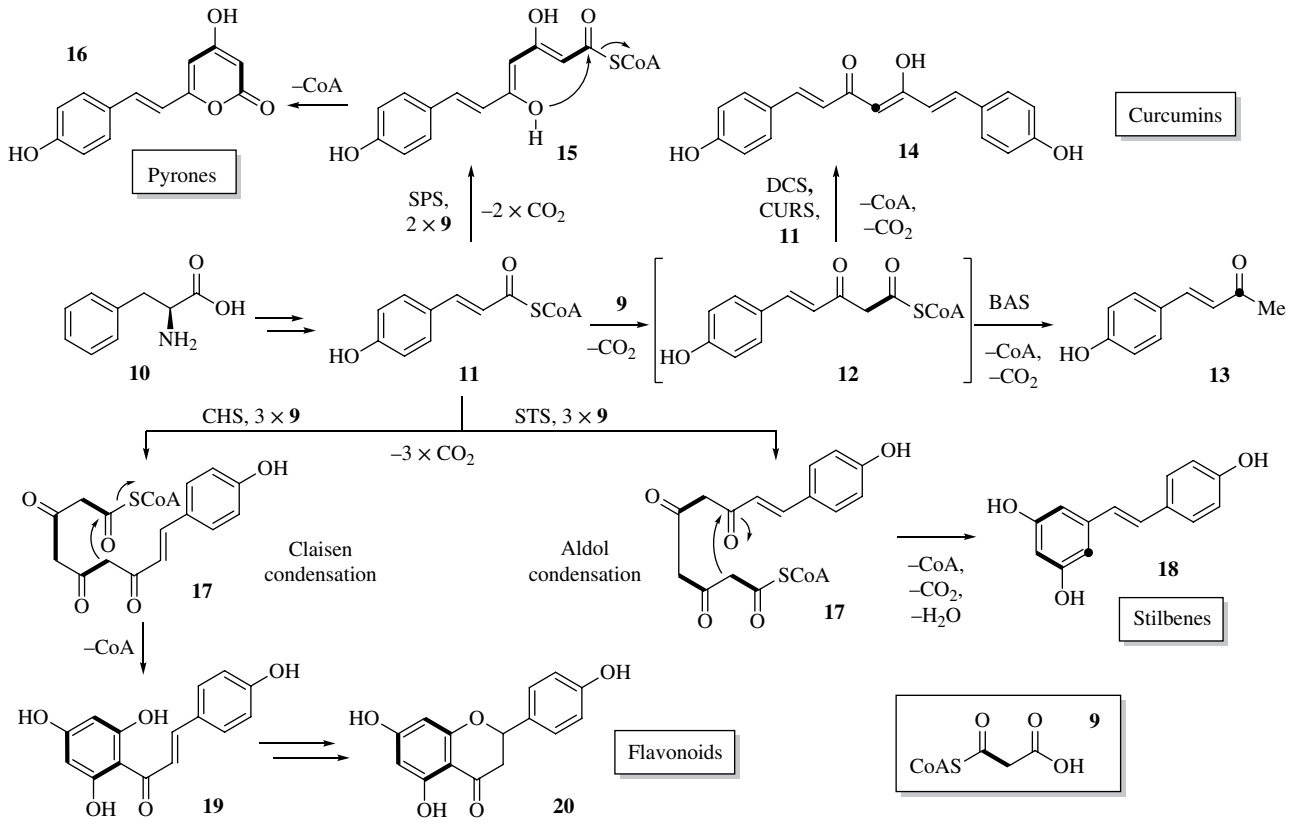
of individual proteins with single functions [23–26]. Both these systems rely on thiotemplate mechanisms in which the substrates and growing polyketide intermediates are covalently attached to the biosynthetic proteins via thioester bonds. The additional type III PKSs, by contrast, only require a single KS active site and utilize free, that is, coenzyme A-bound rather than ACP-bound building blocks for polyketide assembly [27–31].

The three types of PKSs are not equally distributed across the kingdoms of life and, owing to their structural and functional differences, generate distinct types of products [13]. The type III (mainly found in plants, some in bacteria and fungi) and type II (only bacteria) PKSs produce aromatic natural products by iterative use of their respective catalytic units with no to little reductive alterations. Fungal iterative type I PKSs are capable of crafting both aromatic and structurally diverse reduced polyketides. These PKSs are often further classified as nonreducing (NR), partially reducing (PR), or highly reducing (HR), depending on the degree of reductive processing during chain assembly [20]. Despite the iterative use of the overall PKS system, each individual catalytic domain can optionally be employed in each elongation cycle, thus facilitating biosynthesis of highly complex products. Bacterial iterative type I PKSs have long thought to be rare, but a number of interesting examples encoding structurally challenging molecules, such as enediynes [32, 33] or polycyclic tetramate macrolactams [34, 35], have been identified in recent years. Finally, modular type I PKSs that have so far only been found in bacteria and few protists exclusively generate reduced, often highly elaborate natural products [17]. In these systems, each PKS module contains all catalytic domains responsible for the incorporation and the desired reductive processing of a single C_2-building block. This so-called colinearity principle facilitates a direct translation of PKS module organization into expected product structure and *vice versa* [17], a convenient correlation not currently possible for any of the other types of PKSs. In recent years, however, a new subtype of modular type I PKSs emerged. The most prominent difference to previous textbook PKSs is the lack of individual AT domains in each module but instead the presence of a single AT domain outside the modular PKS structure that is acting on all ACPs in *trans* [36]. Many of these *trans*-AT PKSs feature module skipping or repeated utilization of certain modules as well as domain architectures deviating from the predictable, long-known *cis*-AT PKS standards. While this complicates product prediction significantly, recent detailed insights into precursor selection mechanisms are beginning to demystify *trans*-AT PKSs.

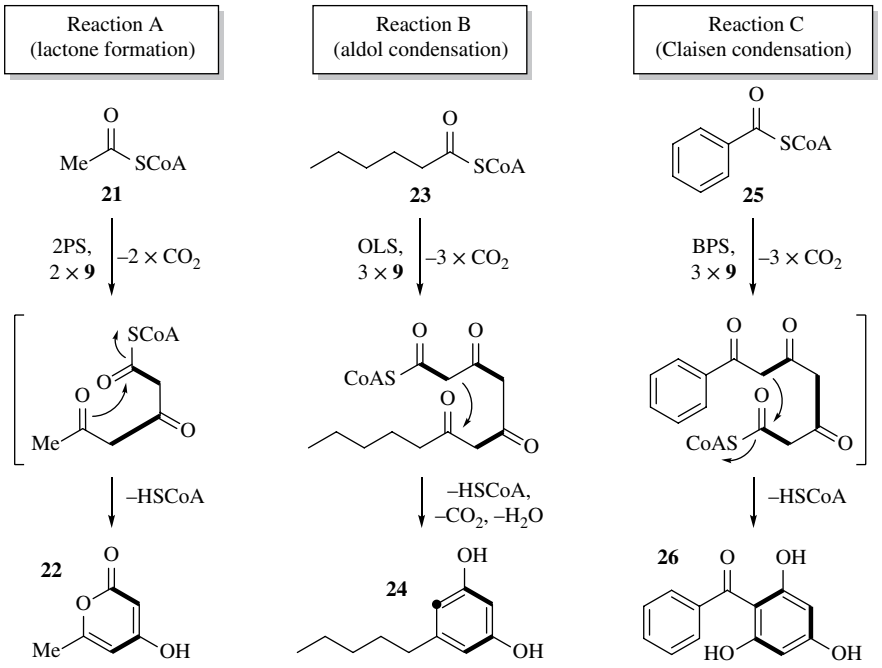
Although the organization of PKSs into types I to III will continue to be challenged by newly discovered systems that do not match the scheme [37, 38], it is extremely useful to explain fundamental differences in polyketide biosynthetic diversity. We will thus adhere to this classification and will begin to introduce the seemingly simplest type III PKS systems.

2.1.3.1 Type III PKSs in Plants, Bacteria, and Fungi Type III PKSs are proteins of about 40–45 kDa size forming homodimers that bear two catalytically active sites, each facilitating product formation by decarboxylative condensation of malonyl-CoA (**9**) with a starter unit (cf. Scheme 2.3c) [27–31]. The broad product spectrum of such systems is a result of their high flexibility concerning the number of elongation steps, the final chain termination/cyclization chemistry, and starter unit selection. Among the best-studied type III PKSs are those involved in the biosynthesis of phenylpropanoids in plants. Starting from L-phenylalanine (**10**), deamination by phenylalanine ammonia lyase (PAL), *p*-hydroxylation by cinnamic acid 4-hydroxylase (C4H), and activation of the carbonyl by 4-coumaroyl:CoA ligase (4CL) by thioester formation with coenzyme A yield the central precursor *p*-coumaroyl-CoA (**11**) [39]. Alternatively, **11** is accessible by direct deamination of L-tyrosine with tyrosine ammonia lyase (TAL) in some microorganisms and grasses [40–42]. Compound **11** now serves as the starter unit for many plant-derived type III PKSs (Scheme 2.4). Decarboxylative elongation with one unit of **9** catalyzed by benzalacetone synthase (BAS) to **12** with concomitant hydrolysis and decarboxylation yields *p*-hydroxybenzalacetone (**13**) [43]. Similar elongation of **11** with **9** catalyzed by diketide-CoA synthase (DCS) in the presence of curcumin synthase (CURS) results in the formation of the curcuminoid carbon skeleton, for example, of bisdemethoxycurcumin (**14**) [44, 45]. Notably, curcumins are also directly accessible from cinnamoyl-CoA precursors by action of a single enzyme, curcuminoid synthase (CUS) [46]. Elongation of **11** by two units of **9** catalyzed by styrolpyrone synthase (2PS) gives **15**, which is directly cyclized by intramolecular lactonization to give pyrone **16** [47]. Such pyrone formation is also catalyzed by 4-coumaroyltriacetic acid synthase (CTAS) [48]. The latter, however, uses three units of **9** for chain extension, thus resulting in the longer-chain *p*-coumaroyl triacetic acid lactone. Both chalcone synthase (CHS) and stilbene synthase (STS) likewise use three units of **9** to give intermediate **17** [49]. In STS, this is further processed by Aldol condensation to give stilbenes, such as resveratrol (**18**) [50]. Alternatively, CHS catalyzes Claisen condensation producing chalcone **19**, the precursor of flavonoids, such as naringenin (**20**) [51].

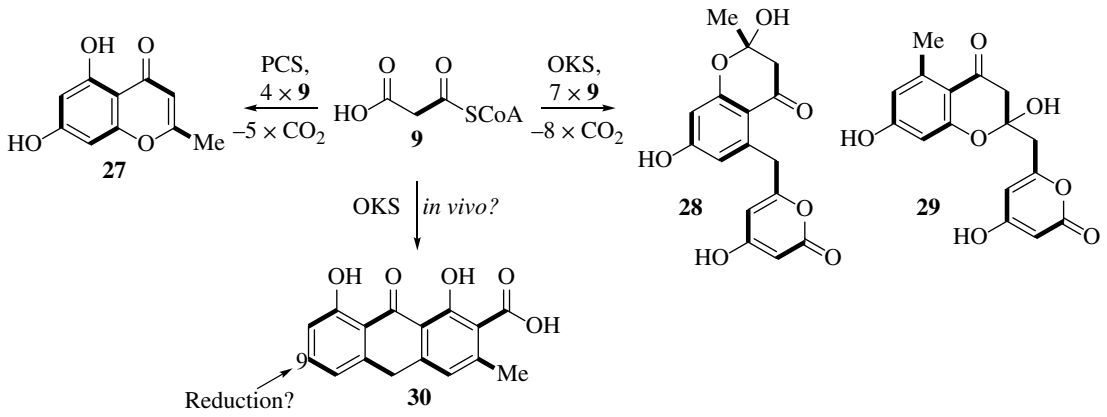
Although Scheme 2.4 presents only a small selection of currently characterized plant type III PKSs, it contains all important basic reactions of this class of enzymes. Besides sole elongation chemistry, here leading to **13** or **14**, cyclization reactions allow further diversification by either (A) lactonizations to give pyrones, (B) aldol condensations to resorcinols or resorcylic acids, or (C) Claisen condensations yielding phloroglucinols (Scheme 2.5). Using



SCHEME 2.4 Type III PKS involved in phenylpropanoid biosynthesis in plants.



SCHEME 2.5 Selected examples of plant type III PKSs utilizing diverse starter units.



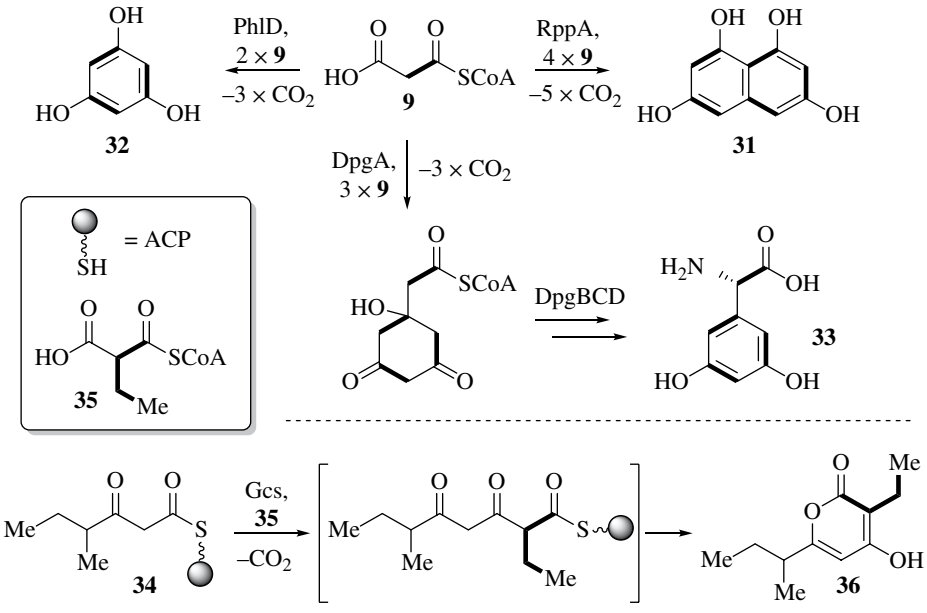
SCHEME 2.6 Recent examples of plant type III PKSs producing long-chain products.

this set of reactions A–C, variations in starter units facilitate access to further chemical diversity. For example, extension of acetyl-CoA (**21**) with two units of **9** catalyzed by 2-pyrone synthase (2PS) gives pyrone **22** [47]. Threefold extension of hexanoyl-CoA (**23**) with **9** catalyzed by olivetol synthase (OLS) produces olivetol (**24**) [52], a central building block in cannabinoid biosynthesis. Benzophenone synthase (BPS) extends benzoyl-CoA (**25**) by three C_2 units leading to benzophenone **26** [53].

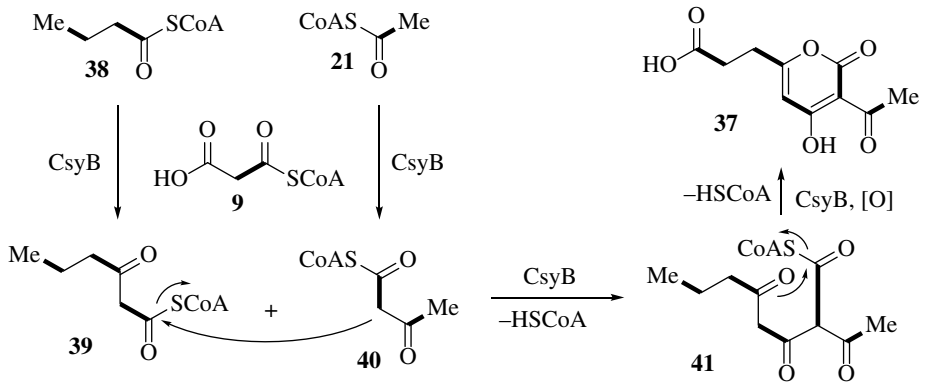
As suggested by the aforementioned examples, most type III PKSs in plants catalyze two or three chain-extension cycles. However, currently characterized systems range from one to eight elongation reactions. Two interesting enzymes for long-chain assembly were discovered in *Aloe* sp., pentaketide chromone synthase (PCS) [54], and oktaketide synthase (OKS) [55]. *In vitro*, PCS was shown to utilize five units of **9** to form chromones **27** (Scheme 2.6). By homologation of eight units of **9**, OKS unexpectedly yielded SEK4 (**28**) and SEK4b (**29**), two compounds previously known as shunt products from bacterial type II PKSs [56]. While chromones, comparable to **27**, could indeed also be PKS products *in vivo*, for example, leading to aloesin-type secondary metabolites known from *Aloe* sp., OKS is thought to require an enzymatic partner *in vivo* [30]. A KR might be involved in the reduction of the carbonyl function at C-9 of a theoretical, linear octaketide precursor, which subsequently would allow to funnel OKS biosynthesis toward chrysophanol-related natural products, such as **30**. Similar KR activity-mediated product control can be found in many bacterial type II PKS systems (cf. Section 2.1.3.2).

Type III PKSs have long been thought to exclusively occur in plants. This changed with the discovery of the 1,3,6,8-tetrahydroxynaphthalene (**31**, THN) synthase RppA from *Streptomyces griseus* in 1999 [57, 58], which meanwhile was identified in a number of other actino- and proteobacteria [59–62]. RppA catalyzes the formation of **31** by decarboxylative fusion of five malonyl-CoAs (**9**) and cyclization of the resulting linear precursor (Scheme 2.7). Since the discovery of RppA, type III PKS biochemistry in bacteria has been extended by a number of other enzymes. This included phloroglucinol (**32**) synthase PhlD from *Pseudomonas fluorescens* [63] and DpgA from *Amycolatopsis orientalis* [64], which, together with DpgBCD, assembles 3,5-dihydroxyphenylglycine (**33**), an important amino acid building block in glycopeptide antibiotics [65]. Further examples include type III PKSs involved in bacterial alkylresorcinol/-quinone (e.g., ArsB, SrsA, FtpA) [66–68], and -pyrone (e.g., ArsC, BpsA, PKS18) [69–71] biosynthesis. Some of these enzymes exhibit properties unknown for plant-derived congeners: for example, ArsB and ArsC derive their long-chain alkyl starter units from the type I FAS ArsA by direct interaction with the respective ACP [66], and SrsA and FtpA do not only accept **9** but also methylmalonyl-CoA as extender unit [67, 68]. These unusual features are combined in the type III PKS Gcs from *Streptomyces coelicolor* that not only preferentially accepts an ACP-bound, branched starter unit **34** [72] but is also capable of extending it with ethylmalonyl-CoA (**35**) to ultimately yield germicidin C (**36**) [73].

Only in very recent years, the first type III PKSs from fungi have been identified and functionally studied. CsyA from *Aspergillus oryzae* [74] and 2′-oxoalkylresorcylic acid synthase (ORAS) from *Neurospora crassa* [75, 76] were shown to produce tri- and tetraketide pyrones from long-chain fatty acid acyl-CoAs using malonyl-CoA (**9**) extender units. ORAS furthermore is capable of assembling resorcylic acids and resorcinols from linear tetra- and pentaketide precursors. AnPKS from *Aspergillus niger* likewise crafts tri- and tetraketide pyrones from fatty acid acyl-CoAs, but additionally is capable of accepting benzoyl-CoA starter units [77]. CsyB from *A. oryzae* exhibits a very uncommon reactivity (Scheme 2.8). It is the key enzyme in the biosynthesis of csypyrone B1 (**37**) [78]. CsyB-mediated formation of **37** is achieved by unprecedented fusion of two ß-ketoacyl-CoA units **39** and **40** derived from decarboxylative extension of a butyryl-CoA (**38**) and acetyl-CoA (**21**), respectively



SCHEME 2.7 Selected examples of type III PKS systems in bacteria.



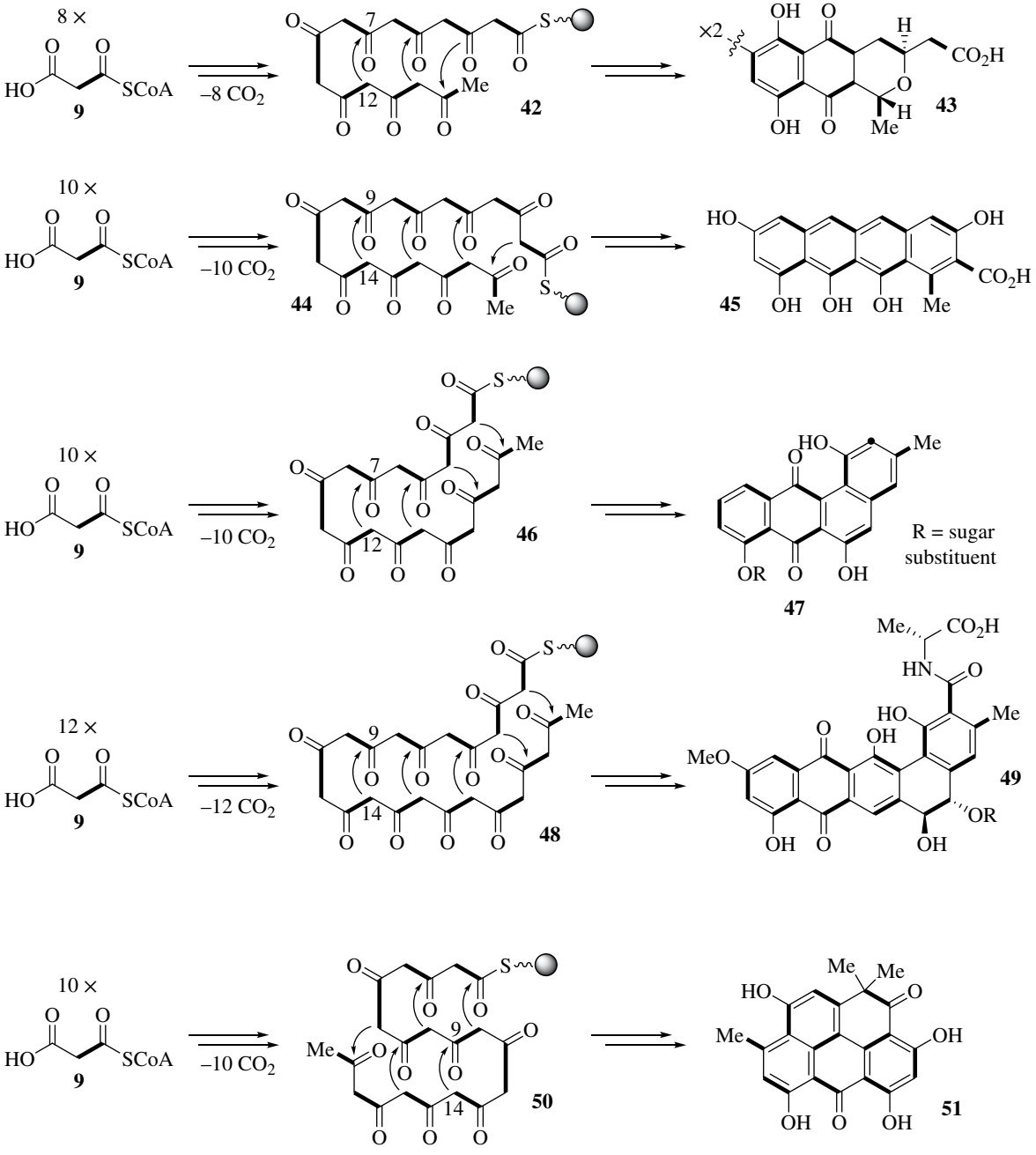
SCHEME 2.8 Csypyrone B1 (**37**) biosynthesis catalyzed by CsyB, a fungal type III PKSs.

[79, 80]. Subsequent cyclization of intermediate **41** and oxidation delivers **37** (Scheme 2.8).

As these few examples of novel bacterial and fungal systems clearly indicate, many seemingly firm rules in type III PKS biosynthesis will continue to be challenged with an increasing number of such PKSs from diverse organisms being identified and functionally studied.

2.1.3.2 Type II PKS Systems in Bacteria Many biomedically important aromatic natural products are derived from bacterial type II PKSs [23]. Within these biosynthetic machineries, individual enzymes form a functional complex, the so-called minimal PKS. This is composed of a KSα/KSβ heterodimer and a dissociated ACP that serves as product tether (cf. Scheme 2.3b). ACP loading with building blocks is catalyzed by a malonyl-CoA:ACP transacylase (MAT), an activity that can be provided by FabD from fatty acid biosynthesis [81], or might alternatively also be achieved by self-acylation [82, 83]. While KSα catalyzes iterative, decarboxylative chain extension of a starter unit exclusively utilizing malonyl-CoA (**9**), KSβ lacks the necessary active site cysteine residue to perform this function. Thus, it is not directly involved in chain elongation chemistry, but rather in chain initiation [84] and in control of the length of the final polyketide intermediate (hence its additional name CLF = chain length factor) [85–87].

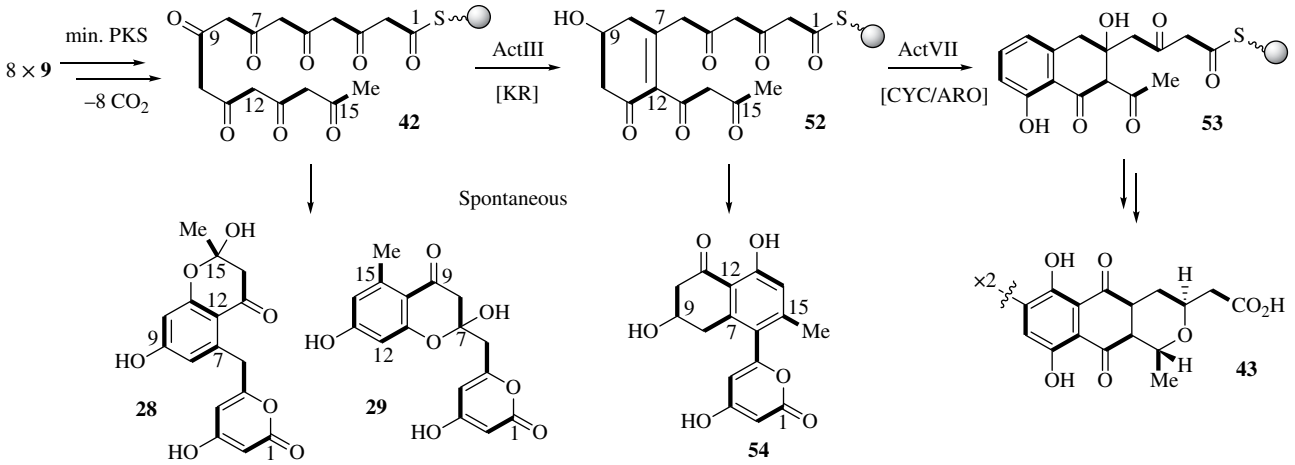
In most bacterial type II PKS products, the polyketide precursor is composed of eight (octaketides, e.g., precursor **42**, toward benzoisochromanes such as actinorhodin (**43**) [88]), ten (decaketides, e.g., precursors **44** and **46** toward tetracenomycins **45** [89] and angucyclines such as landomycin (**47**) [90]), or twelve (dodecaketides, e.g., precursor **48** leading to pentangular polyphenols such as pradimicin (**49**) [91]) C_2-building blocks (Scheme 2.9), with the largest



SCHEME 2.9 Typical type II PKS intermediates (R group indicates further glycosylation).

known example to date being the pentadecaketide fredericamycin [92]. The linear polyketide precursor molecules seem to generally be preorganized in a U-shaped manner [13]. Cyclization by aldol condensation thus leads to linear or angular polycyclic products with no more than two sides of any ring being annealed to another ring system. An interesting exception to this general rule is resistomycin (**51**), which results from cyclization of an S-shaped decaketide precursor **50** [93]. The predominant lack of reductive processing of the intermediate polyketide chain results in the characteristic alternating 1,3-oxygenation pattern in type II PKS products, which in some cases is obscured to some extent by post-PKS oxygenation.

As indicated in Scheme 2.9, the cyclization of the first aromatic ring predominantly occurs between C-12 and C-7 or C-14 and C-9. Furthermore, in each type II PKS pathway, all subsequent cyclization reactions of the highly reactive polyketide intermediate need tight control. Besides significant contributions of the minimal PKS in precursor protection and potentially also in control of the first ring cyclization,

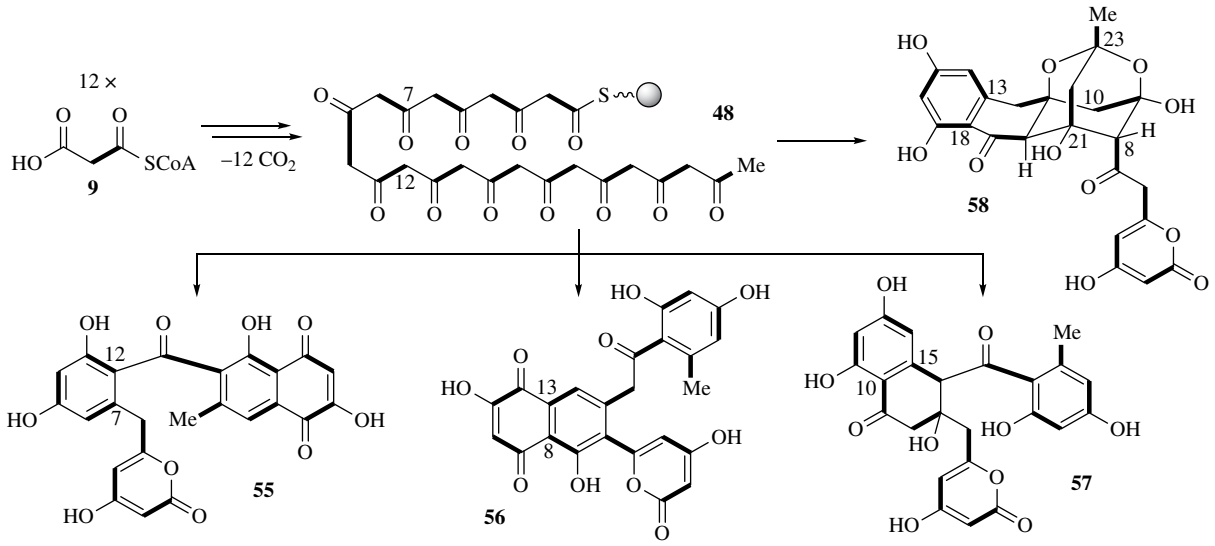


SCHEME 2.10 Control of the regiochemistry in ring formation in the biosynthesis of actinorhodin (**43**).

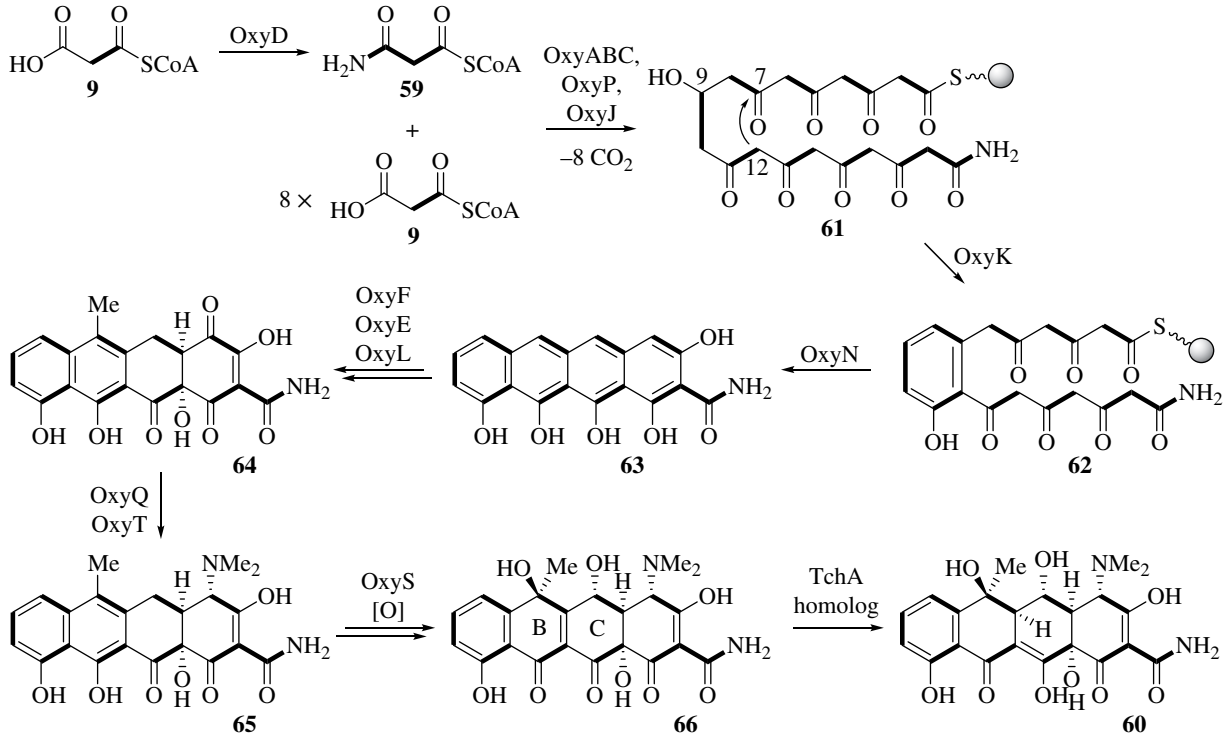
correct folding and thus cyclization of the polyketide are further orchestrated by cyclases (CYC), aromatases (ARO), and a primary KR, the latter almost exclusively reducing precursor chains at C-9. The impact of these important chaperone-like accessory biocatalysts was comprehensively summarized in an excellent recent article by Tang et al. [94]. We thus herein focus on two instructive examples, starting with the exceptionally well-studied biosynthesis of actinorhodin (**43**) [24]. If control over first ring cyclization regioselectivity toward **43** is assumed for the *act* minimal PKS alone, exclusive formation of C-12 to C-7 cyclized PKS products would have to be expected (Scheme 2.9). However, investigations into the product spectrum of the minimal PKS revealed the formation of SEK4 (**28**), with the expected C-12 to C-7 cyclization, together with SEK 4b (**29**) with a C-10 to C-15 connectivity [56] (Scheme 2.10). This clearly rules out an exclusive control of first ring cyclization by the minimal PKS and rather points toward a mere temporal protection of the growing polyketo chain against undesired cyclization within the KS complex. If the KR ActIII is employed in addition, selective reduction at C-9 leads to an exclusive C-12 to C-7 cyclization as shown in intermediate **52**, ultimately leading to mutactin (**54**) [95] (Scheme 2.10). The interplay of the minimal PKS with the KR and the altered structure and conformation of the C-9-reduced substrate thus apparently facilitate strict control over first ring cyclization outcome. The second ring is selectively formed by action of the CYC/ARO Act VII to give **53**, which is ultimately converted into actinorhodin (**43**) by action of the remaining pathway-specific enzymes [24].

The lack of cyclization control by minimal PKSs alone was also impressively demonstrated by expression of the type II PKS *whiE* from *Streptomyces coelicolor* [96]. This led to the formation of more than 30 structurally highly diverse polyketides, with first ring cyclization ranging from C-12 to C-7 as in TW93e (**55**), C-8 to C-13 as in TW93f (**56**), C-10 to C-15 as in TW93g (**57**), and C-13 to C-18 leading to the unprecedented 2,4-dioxaadamantane structure TW93h (**58**) (Scheme 2.11).

In addition to shaping diverse structures by flexible control of cyclization regiochemistry with the aforementioned enzymes, type II PKSs are also capable of incorporating diverse starter units, for example, short-chain fatty acids or benzoate, by interaction with ACPs from type III PKS systems loaded with the respective molecule (which in turn can be derived from FAS or type I PKSs) or by dedicated transferases [97]. An interesting example of an unusual starter unit is malonamate (**59**) found in the class of tetracycline natural products, such as oxytetracycline (**60**) [98, 99]. The amide functionality in **59** is installed by action of the amidotransferase OxyD, either using malonyl-CoA (**9**) or ACP-bound malonate as the substrate [100] (Scheme 2.12). Enhanced priming of the minimal PKS OxyABC with **59** versus acetate is supported by the thiolase OxyP that shows preferential activity toward acetyl-ACP hydrolysis, thus policing starter unit selection [101]. Elongation of the starter unit with eight building blocks **9** and selective reduction at C-9 by KR OxyJ lead to theoretical intermediate **61**. Cyclization by the first ring CYC/ARO OxyK [102] results in **62**, which gets elaborated to the tetracyclic core structure pretetramid (**63**) by CYC OxyN [103]. Further post-PKS tailoring of the structure by C-methylation (OxyF) and oxygenation (OxyE and OxyL) yields 4-keto-anhydrotetracycline (**64**), which upon reductive amination (OxyQ) and di-*N*-methylation (OxyT) gives anhydrotetracycline (**65**) [104]. Oxygenation of the B-ring by OxyS (also named OtcC) and of the C-ring by an additional oxygenase delivers **66** [105]. In analogy to the chlortetracycline pathway, the last reduction step to **60** is likely catalyzed by a homolog of the reductase TchA [106]. This series of methylation, oxygenation, and reduction tailoring steps thus establishes the highly functionalized, bioactive oxytetracycline (**60**) structure [98, 99].



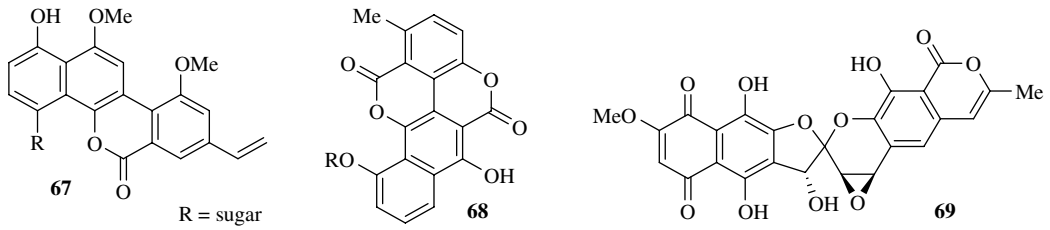
SCHEME 2.11 Products of the minimal type II PKS *whiE* formed by spontaneous cyclization.



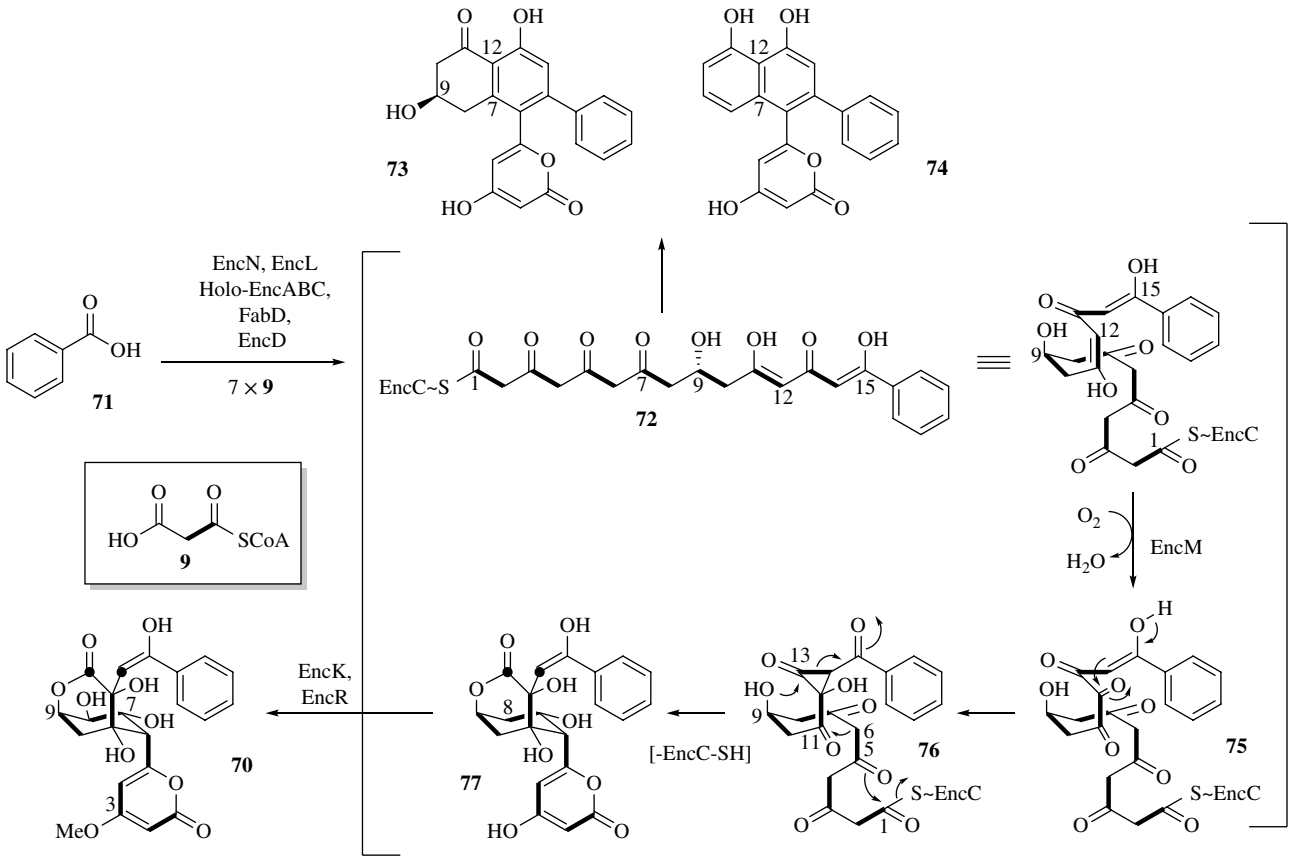
SCHEME 2.12 Biosynthesis of oxytetracycline (**60**).

Besides rather conservative oxidative functionalization of type II PKS products by tailoring enzymes, PKS core structures are in some pathways heavily modified by extensive oxidative rearrangement reactions. Examples for such processes are found in the biosynthesis of the gilvocarcins, such as **67**, resulting from the rearrangement of an angucyclic precursor of type **47** [107] (cf. Scheme 2.9); of chartreusin (**68**), derived from linear tetracyclic systems comparable to **45** [108]; or of griseorhodin A (**69**) [109], derived from a hexacyclic angular aromatic precursor molecule (Scheme 2.13) [13, 23].

This small selection nicely illustrates the potential of oxidative post-PKS tailoring steps to restructure type II PKS architectures [13, 23]. Nonetheless, the resulting products remain to be aromatic polyketides. The single most impressive deviation from this typical type II PKS product

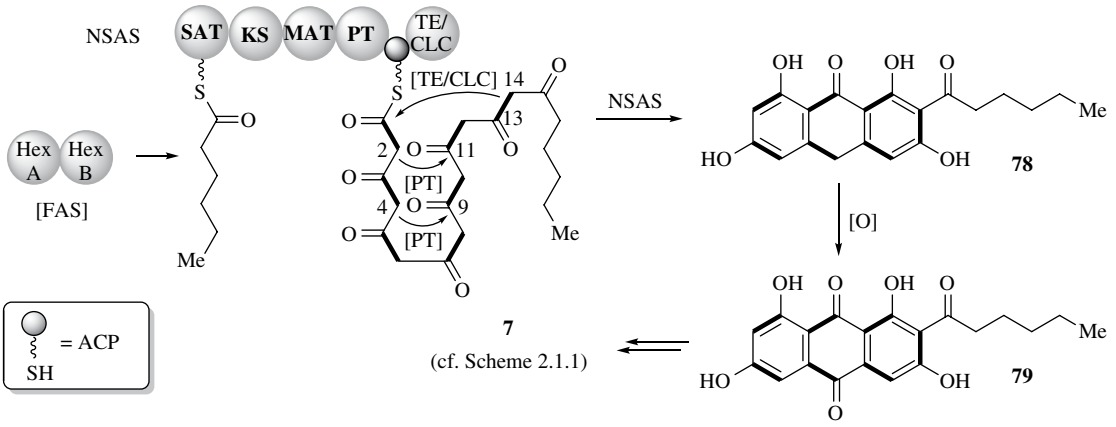


SCHEME 2.13 Selection of unusual type II polyketides derived from oxidative core structure rearrangements.



SCHEME 2.14 Biosynthesis of the highly unusual type II PKS product enterocin (**70**).

space is found in the biosynthesis of enterocin (**70**) in *Streptomyces maritimus* [110] (Scheme 2.14). In this case, the benzoate:ACP ligase EncN transfers a benzoate starter unit **71** to the ACP EncC of the enterocin minimal PKS EncABC [111]. Efficient loading of this precursor is controlled by the type II TE EncL that selectively cleaves misprimed starter units, for example, acetate, from EncC [112], similar to the function of OxyP in the biosynthesis of **60** earlier [101]. Subsequent attachment of elongating **9** onto EncC is catalyzed by a homolog of the MAT FabD [113]. Interestingly, this set of enzymes alone does not produce any isolable, benzoate-primed polyketides. With the additional KR EncD, however, selective C-9 reduction leads to theoretical intermediate **72**. This suggests that EncD serves an important catalytic role that helps to direct product assembly and that it likely reduces the nascent polyketide chain during elongation rather than after its complete assembly [114]. Without any further enzymes, **72** spontaneously cyclizes to wailupemycins, for example, by the common C-12 to C-7 cyclization pattern yielding wailupemycins F (**73**) and G (**74**) and thus leading to typical aromatic type II PKS products [113]. In the presence of the flavoprotein EncM, however, oxidation at C-12 to the corresponding ketone **75** precedes the formation of cyclopropanone **76**. Subsequent lactone formation between the alcohol function at C-9 with concomitant cleavage of the C–C bond between C-13 and C-14 and two further cyclization reactions (from the oxygen at C-5 to C-1 and from C-6 to C-11) furnishes intermediate



SCHEME 2.15 Biosynthesis of norsolorinic acid (**79**) by a fungal type I iPKS.

77 [115, 116]. The oxygenation reaction at C-12 by EncM, which induces this Favorskii-type rearrangement sequence, remarkably proceeds via a flavin-*N*-5-oxide [117]. The catalytic activity of this single enzyme thereby reroutes the cyclization outcome, leading to the highly unusual enterocin core structure **77**. The biosynthetic assembly of enterocin (**70**) is completed by action of the SAM-dependent methyltransferase (MT) EncK, which methylates the hydroxyl group at C-3, and the cytochrome P450 hydroxylase EncR that mediates C-8 hydroxylation [113].

As the biosynthesis of enterocin (**70**) impressively demonstrates, type II PKSs are not limited to the biosynthesis of polycyclic, aromatic structures. Undoubtedly, future research on this type of PKSs will continue to reveal unprecedented biosynthetic features.

2.1.3.3 Type I PKS Systems in Fungi and Bacteria Type I PKSs are multidomain monomodular (iterative type I PKS, iPKS) or multimodular (modular type I PKS) systems in which the catalytic domains are located and covalently linked within large multifunctional enzymes (cf. Scheme 2.3a) [13]. These proteins can thus reach enormous sizes and are therefore also often referred to as megasynthases. With only a few exceptions, fungi typically harbor monomodular type I iPKSs [21, 22]. Fungal PKSs are further classified according to the degree of reductive processing of the assembled polyketide chain as NR, PR, and HR [20]. Owing to the PKS domain organization in a single protein and the programmed, iterative use of these catalytic units, biochemical and mechanistic investigations of these systems have been extremely challenging. It is thus not surprising that the elucidation of many important fungal type I PKS functions was mostly achieved within the last two decades. In the following section, we will briefly introduce some of the key features of fungal type I PKSs, exemplarily using the norsolorinic acid (NR) and lovastatin (HR) synthases.

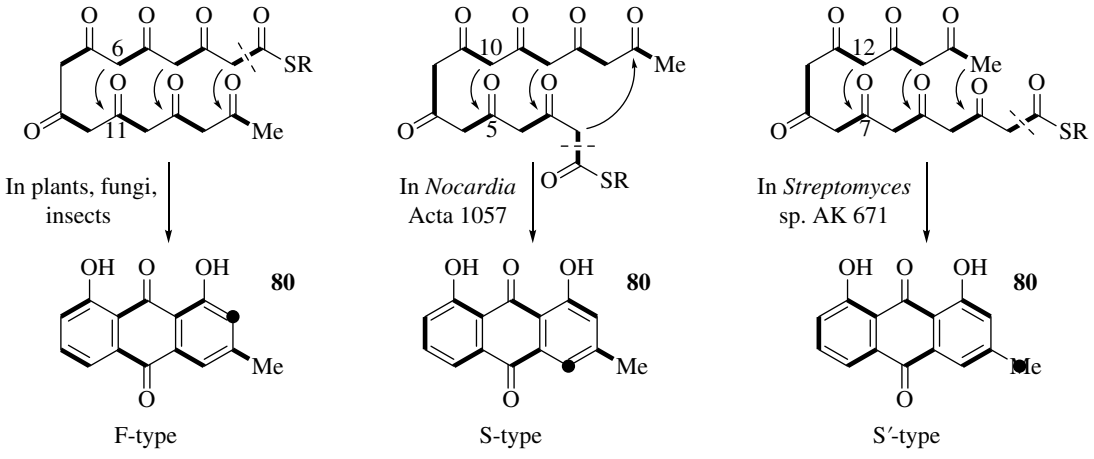
Fungal NR type I iPKSs produce aromatic compounds, structurally highly related to bacterial type II PKS structures [21]. One of the best-studied examples is norsolorinic acid synthase (NSAS) that encodes **79**, the first isolable precursor in sterigmatocystin and aflatoxin biosynthesis [118] (Scheme 2.15). Early isotope labeling and genetic experiments suggested the starter unit to be hexanoate [119, 120]. This building block is provided by HexAB, homologs of fungal α- and β-FAS subunits clustered with NSAS [121, 122]. The exact domain organization of NSAS was elucidated by bioinformatic analysis using the Udwary–Merski algorithm (UMA) [123]. This revealed the presence of six discrete domains, four of which were identified by their conserved sequences to be KS, MAT, ACP, and a terminal TE–Claisen cyclase (TE/CLC) domain. In addition, an amino-terminal domain was predicted, which was later identified to be responsible for hexanoate loading and thus named starter unit transacylase (SAT) [124]. Remarkably, SAT utilizes an active site cysteine residue for activating hexanoate as a thioester, in contrast to the common serine-mediated oxoester activation found in MAT domains. Finally, a product template (PT) domain was postulated, which was thought to be responsible for chain length control and stabilization of the growing polyketide chain [123]. In fact, PT sequences were shown to cluster according to the structures of their respective products [125]. The exact roles of these individual domains were recently firmly established in an impressive deconstruction approach [126]. After homologation of the hexanoate starter to the ACP-bound octaketide precursor, PT mediates regioselective Aldol condensation reactions between C-4/C-9 and C-2/C-11 [127]. As shown before for WA PKS [128], TE/CLC subsequently catalyzes a C-14 to C-1 Claisen reaction to form the third ring system while cleaving **78** from the synthase [129]. In the absence of TE/CLC, the control of the selectivity of this off-loading mechanism is lost. This leads to the formation of **78** along with the respective pyrone resulting from nucleophilic attack of the C-13 oxygen function onto the thioester bond (not shown). Compound **78** gets transformed into norsolorinic acid (**79**) by oxidation. By a complex

series of reactions, **79** is ultimately restructured to yield aflatoxins, such as **7** [8] (cf. Scheme 2.1).

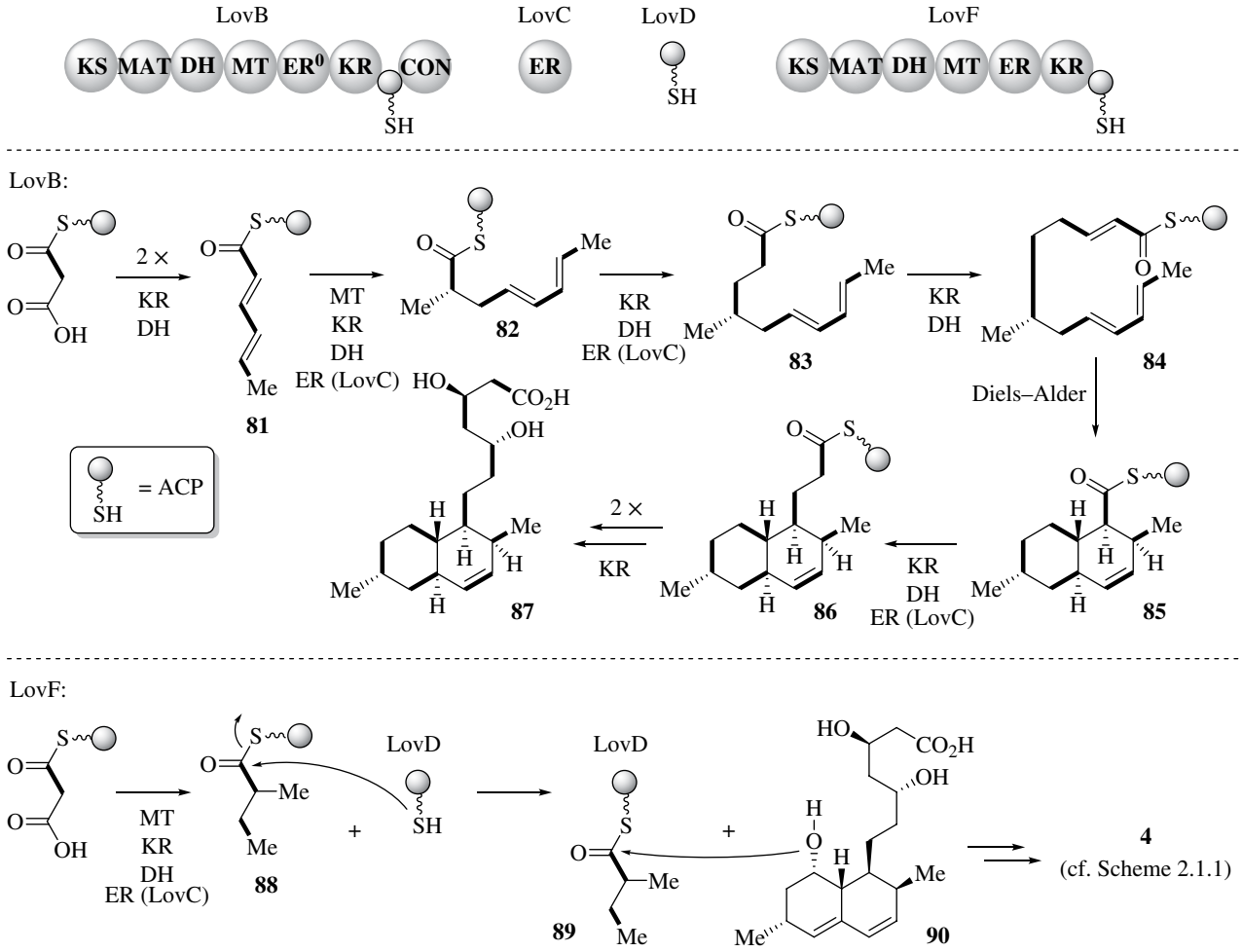
As exemplified by **79**, the products formed by fungal iPKSs strongly resemble bacterial type II PKS products. However, differences in folding of the linear polyketide backbone result in different cyclization regiochemistry. This almost exclusively leads to the incorporation of three intact C_2 units in the ring system first cyclized in bacteria, referred to as S-type folding (from ***S****treptomyces*), whereas only two C_2 units are found in the first ring in case of eukaryotes (F-type folding from **f**ungi) [130]. This phenomenon was illustrated by comparative isotope labeling studies elucidating the folding pattern of the ubiquitous octaketide chrysophanol (**80**) in higher plants, fungi, and insects when compared to that observed in the bacterial strains *Streptomyces* sp. AK 671 and *Nocardia* Acta 1057 [131] (Scheme 2.16). Indeed, in all eukaryotic producers of **80**, the expected F-type folding was observed [132]. *Nocardia* Acta 1057 and *Streptomyces* sp. AK 671 even provided two distinguishable S-type folding patterns, likely resulting from different first ring cyclization events, that is, C-10 to C-5 and C-12 to C-7, respectively [132, 133].

The natural products derived from fungal HR type I PKSs open an entirely different structural space. A prototypical example of such a system is the biosynthetic machinery encoding lovastatin (**4**). The HR PKS LovB (lovastatin nonaketide synthase (LNKS)) of this pathway contains functional KS, MAT, DH, KR, ACP, and C-MT domains in addition to a terminal CON domain resembling NRPS condensation domains and a nonfunctional ER (ER^0) [134, 135] (Scheme 2.17). Heterologous expression experiments in *Aspergillus nidulans* [136] and gene deletion and chemical complementation studies in *Aspergillus terreus* [137] revealed that the free-standing ER LovC is additionally required for the production of the first isolable intermediate of lovastatin biosynthesis, dihydromonacolin L (**87**). It is acting as *trans*-ER, a concept also found in other fungal HR type I PKSs [22]. The complex programming of the activity of the individual catalytic domains of LovB during each elongation reaction was beautifully illuminated by complete reconstitution of the LovB/C enzymatic machinery [138]: after priming, LovB catalyzes two rounds of chain extension with concomitant action of KR and DH to give enzyme-bound sorbic acid (**81**). Without NADPH in the enzymatic assays, reduction cannot take place and the non-reduced triketide precursor does not get further elaborated but rather gets cleaved off the PKS by nonenzymatic pyrone formation (not shown). Further extension and C-methylation at the newly introduced α-C-atom are followed by KR, DH, and ER (LovC) activity to give **82**. If the enzymatic system lacks SAM, therefore preventing C-methylation, also no ER activity is observed, neither at this nor at any later stage of polyketide assembly. This consequently leads to polyunsaturated pyrone or ketone shunt products (not shown) [135, 138]. The activity of the MT thus controls downstream ER processing and is therefore decisive for the selective formation of **87**. Further elongation with complete reduction to **83**, followed by extension and partial reduction to **84**, sets the stage for a Diels–Alder reaction to furnish decalin **85**. The Diels–Alder reaction is controlled by LovB [139], potentially involving the biochemically yet uncharacterized CON domain [22]. Further extension of **85** with full reduction gives **86**, which gets elaborated into **87** by two additional elongation steps with reductive processing by the KR only. This example clearly demonstrates the high degree of complexity in programming the activity of the individual catalytic domains of LovB during a particular elongation step and the utility of LovC acting as *trans*-ER. The recent crystal structure of LovC provides first explanations into the molecular interactions of LovB and LovC during polyketide assembly [140].

In addition to the decalin core structure **87**, lovastatin (**4**) bears a short, branched side chain. This building block is assembled by LovF, a second HR type I PKS [135]. The



SCHEME 2.16 Different folding patterns in the biosynthesis of chrysophanol (**80**) in eukaryotes and prokaryotes.

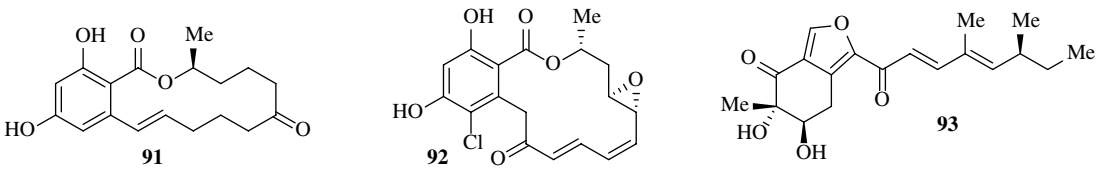


SCHEME 2.17 HR fungal type I PKS machinery encoding lovastatin (**4**).

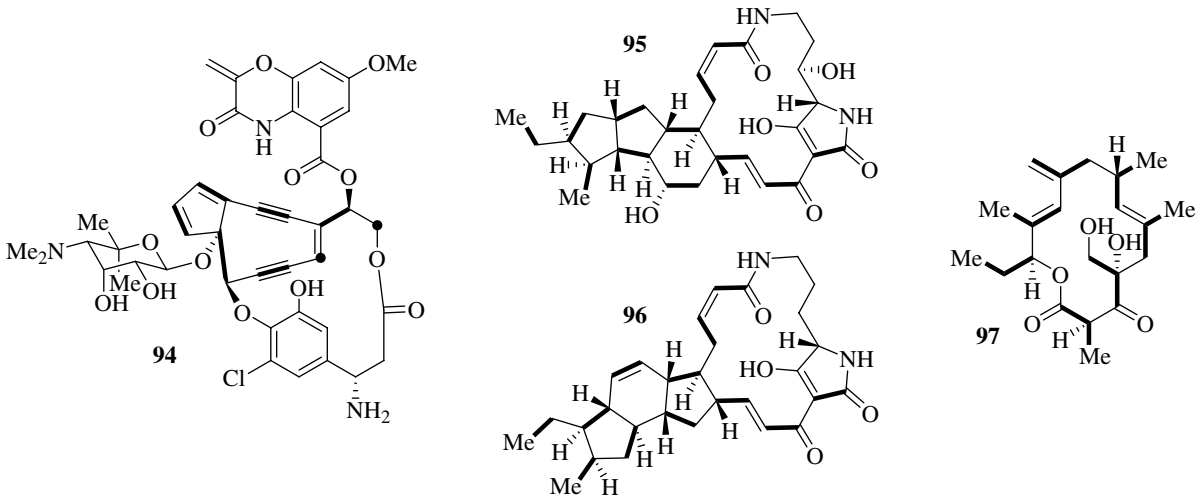
domain architecture of LovF is almost identical to that of LovB (Scheme 2.17). However, it is devoid of a terminal CON domain and contains an active ER. After priming, LovF catalyzes a single elongation cycle with all processing domains active, thus leading to the C-methylated intermediate **88**. The free-standing ACP LovD intercepts this precursor to give **89** [141]. Nucleophilic attack of the hydroxyl group located at the decalin core of **90**, which in turn is derived from **87** by oxidative tailoring to the thioester functionality in **89** followed by intramolecular lactone formation, concludes lovastatin (**4**) biosynthesis.

The interplay of two PKSs is a common theme in fungal polyketide biosynthesis [22]. This strategy allows to expand the structural space covered by such PKSs without further complicating the programmed teamwork of domains within a single iterative PKS. Lovastatin (**4**) biosynthesis thereby represents an example of convergent PKS interactions, where two independently acting PKSs assemble their products that finally get fused by action of an additional enzyme, in this case the ACP LovD. However, there are also many examples of sequential interplay of PKSs, where one synthase provides a precursor that gets further elaborated by a second PKS. In the biosynthesis of norsolinic acid (**79**; cf. Scheme 2.15), for example, the hexanoate starter unit is assembled by the FAS HexAB, transferred to NSAS, and further processed to ultimately yield **79** [122]. The transfer of the building block in such cases is thought to be catalyzed by direct interactions of the ACP of the first PKS with the SAT of the second [142]. Such dual PKS systems also facilitate the biosynthesis of molecules bearing structural features of both, NR and HR PKS products. Typical examples are the resorcylic acid lactones (RALs), such as zearalenone (**91**) [143, 144] or radicicol (**92**) [145] (Scheme 2.18). In some cases, the NR PKS-derived aromatic residue furthermore gets heavily modified by post-PKS processing, as, for example, in asperfuranone (**93**) biosynthesis [146].

Iterative type I PKSs were thought to almost exclusively occur in fungi or, in the rare bacterial cases, to provide rather simple building blocks, for example, for typical modular type I bacterial PKSs [13]. However, more and more evidence



SCHEME 2.18 Fungal natural products originating from dual HR/NR type I PKS systems.

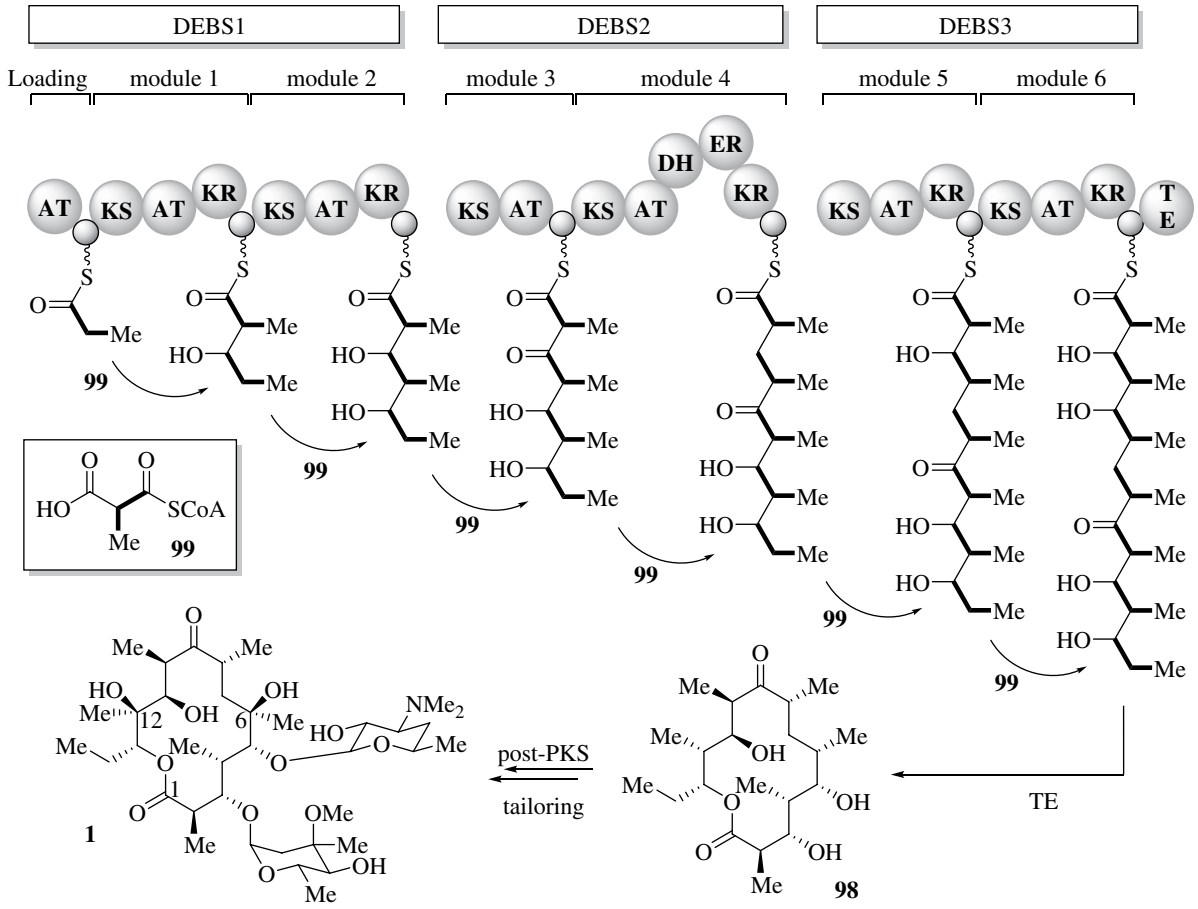


SCHEME 2.19 Selection of natural products derived from uncommon bacterial type I iterative PKSs.

underlining the importance of iterative type I PKSs in bacterial secondary metabolism is uncovered. For example, the polyketide core structure of enediynes, such as C-1027 (**94**) (Scheme 2.19) [32] or calicheamicin (not shown) [33], is assembled by such synthases. In addition, a single iterative PKS module is also involved in the biosynthesis of polycyclic tetramate macrolactams, such as HSAF (**95**) [147] or ikarugamycin (**96**) [34, 35]. Very recently, even the complex programming of different oxygenation states in the iterative PKS machinery encoding the 14-membered galbonolide macrolides, for example, galbonolide B (**97**) in *Streptomyces* spp., has been discovered [148, 149]. These emerging examples clearly indicate that further unexpected pathways involving iterative type I PKSs await being discovered, not only in fungi but also in bacteria and likely in other organisms.

The vast majority of bacterial type I PKSs, however, is encoded on large, multimodular megaenzymes that closely resemble assembly line processes [17]. In these systems, each elongation step is catalyzed by an individual module that is also equipped with the catalytic domains necessary for the desired reductive processing of the β-carbonyl of the elongated intermediate to either the alcohol (KR), the alkyne (KR, DH), or the alkene (KR, DH, ER). The nascent polyketide precursor thereby moves across the PKS during assembly, ultimately reaching the terminal ACP domain. Typically, a TE domain catalyzes final cleavage of the PKS-bound intermediate by either hydrolysis or macrocyclization [150].

The prototypical example of such modular type I PKSs is DEBS that assembles 6-deoxyerythronolide B (**98**), the PKS precursor of erythromycin (**1**) [151–153]. The landmark discovery of this giant enzymatic machinery (total size of the PKS > 1 MDa), independently by Katz et al. [154] and Leadlay et al. [155], also largely contributed to a dramatic shift of biosynthetic research, from elaborate feeding experiments to modern biomolecular approaches. DEBS is composed of 28 domains distributed over 7 modules on three large PKS proteins DEBS 1–3 (Scheme 2.20). Each of the six elongation modules catalyzes the incorporation of propionate by decarboxylative Claisen-type condensation utilizing methylmalonate (**99**) as building block. Modules 1, 2, 5, and 6 all only contain KR, leading to secondary alcohols at the β-positions in the intermediates bound to the respective ACP domains. Module 3 is devoid of any reductive domain, thus retaining the β-carbonyl. The presence of KR, DH, and ER in module 4 leads to full reduction of the respective β-carbon of the intermediate attached to the ACP of this module. Upon full assembly of the polyketide chain, product release is achieved by TE-catalyzed macrolactone formation to yield **98**. The observed absolute configuration in **98** is installed by perfect stereocontrol during the KR-mediated reduction events [153]. Further post-PKS tailoring reactions, that is, stereoselective C-6 and C-12 hydroxylation and two glycosylations, conclude the biosynthesis of erythromycin (**1**) [151, 152].

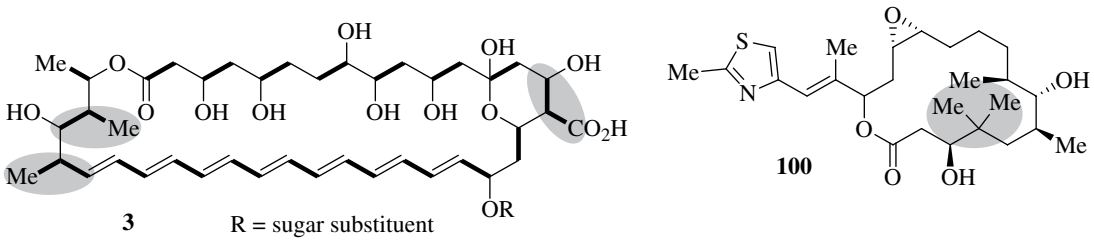


SCHEME 2.20 DEBS, a prototypical example of a bacterial type I modular PKS.

While the modular structure of such type I PKSs causes high metabolic costs because of the resulting huge enzymatic machineries that need to be assembled, it also facilitates biosynthesis of highly complex natural product scaffolds. The structural space accessible by these PKSs is further enlarged by their ability to incorporate different α-substituted building blocks by employing the respective 2-substituted malonate precursors, thus leading to alkyl-branched products. This flexibility is a result of the individual AT domains in each module that can selectively activate different extender units. An example for the selective incorporation of methyl branches can be found in amphotericin biosynthesis, where 16 malonates and 3 methylmalonates (gray label) are incorporated into the final structure, for example, **3** [156] (Scheme 2.21). Furthermore, optional MT domains are capable of introducing methyl or even dimethyl branches, for example, in the biosynthesis of epothilones (**100**, gray label) [157].

The biosynthesis of many complex bacterial PKS products largely follows the aforementioned assembly line logic [17], although exceptions exist [158]. The observed collinearity makes these systems ideal targets for the controlled generation of novel "unnatural" natural products by modifications of the overall PKS modular architecture [159] or individual catalytic functions. In the case of DEBS, this has led to, that is, the generation of a large set of derivatives of **98** by altering functions of individual catalytic domains within the megasynthase [160]; the reconstitution of mono-, bi-, or trimodular PKS substructures [161–164]; and the remarkable complete reconstitution of the entire DEBS machinery with protein extracts [165] or even purified proteins [166]. Such experiments have strongly contributed to our current understanding of PKS assembly lines. In addition, extensive structural studies on these PKSs are beginning to provide deep mechanistic insights into their functions [153, 167–170], recently culminating in high-resolution structures of a complete pikromycin PKS module [171, 172].

The aforementioned textbook knowledge on type I bacterial PKSs is mainly based on investigations targeting systems from actinomycete bacteria. These organisms have long been studied because of their proven track record to biosynthesize numerous bioactive small molecules. In recent years, however, natural product researchers have started to interrogate unusual bacterial taxa, for example, bacteria from remote habitats, symbiotic bacteria, or pathogens.

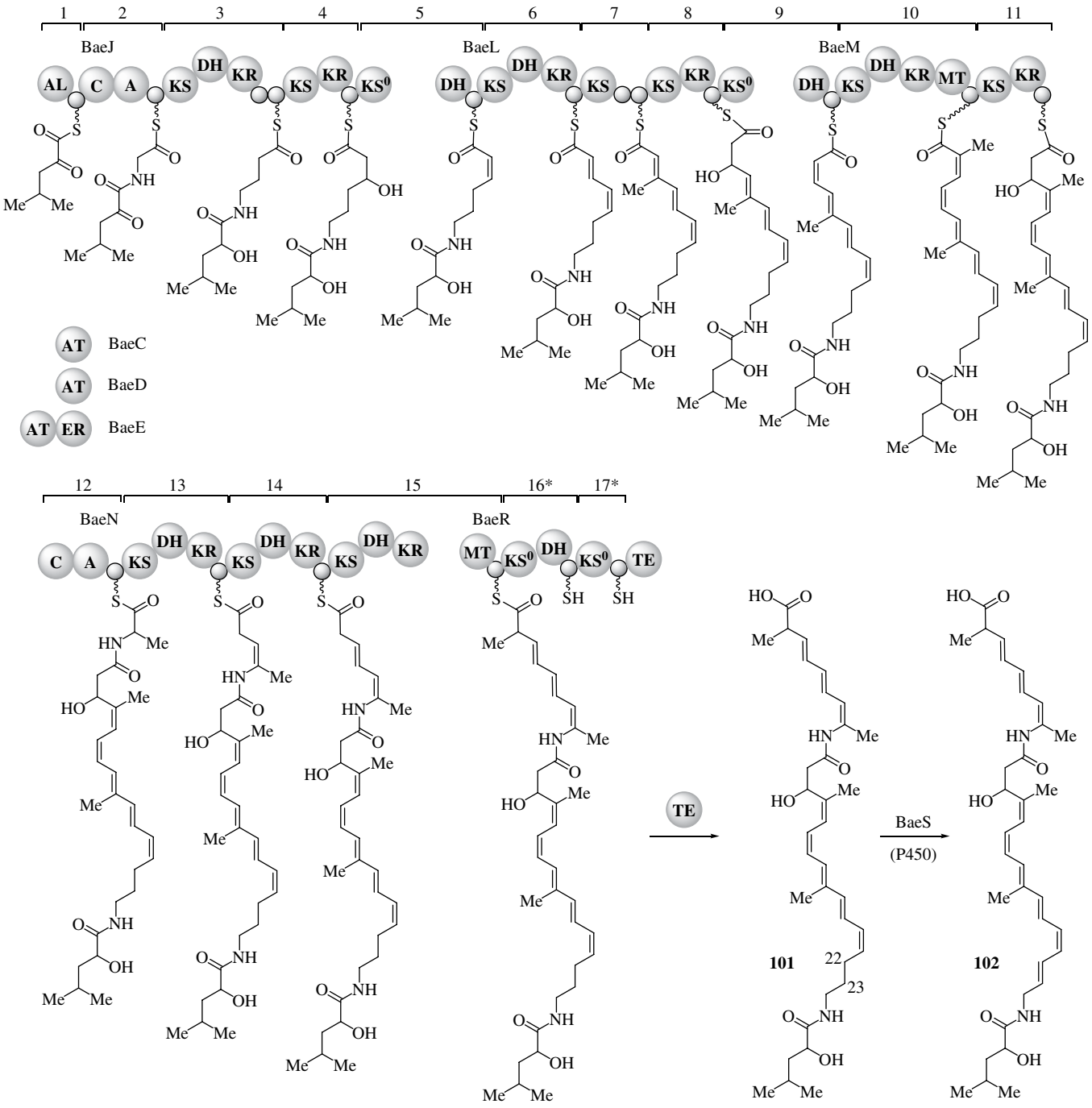


SCHEME 2.21 Incorporation of alkyl branches into amphotericin B (**3**) and epothilone A (**100**).

This has led to the discovery of a new class of multimodular type I PKSs [173, 174]. In these PKSs, all individual modules lack the AT domains. The PKS building blocks are instead provided by a free-standing AT acting in *trans*. These systems were consequently named *trans*-AT PKSs (or "AT-less" PKSs), as opposed to the classical *cis*-AT PKSs described in the preceding text [36]. Interestingly, phylogenetic analyses of multimodular PKSs showed that *cis*- and *trans*-PKSs have evolved independently from FAS [175] following two different modes of evolution. While *cis*-PKSs predominantly evolve by duplication of individual modules with subsequent alterations in domain composition [176], *trans*-AT PKSs derive from multiple genetic parts by extensive horizontal gene transfer [177].

Importantly, *trans*-AT PKSs do not adhere to the collinearity principle of their *cis*-AT counterparts [36]. For example, individual modules might be used more than once or might be completely omitted. Furthermore, while *cis*-PKSs only comprise eight different module architectures, depending on the presence/absence of the optional domains (none, KR, KR/DH, KR/DH/ER, each with or without MT), more than 50 variations can be found in *trans*-AT PKSs. This diversity in module structures gives rise to a large number of catalytic features, including ß-branching cassettes, dehydrating bimodules leading to *cis*-configured double bonds, etc. [178]. A detailed discussion of all features of this remarkable class of PKSs would thus clearly exceed the scope of this article. We instead selected a single, representative example to briefly introduce some intriguing features of *trans*-AT PKS systems, the biosynthetic machinery encoding bacillaene (**102**). It was initially discovered as *pksX* in *Bacillus subtilis* 168 as the first *trans*-AT PKS to be identified [179, 180]. An almost identical cluster (*bae*, Scheme 2.22) was later found in *Bacillus amyloliquefaciens* and firmly linked to the production of **101** and **102** by gene deletion experiments [181]. Within *bae* (*pksX*), the *trans*-AT BaeC (PksC) loads all ACPs with malonyl-CoA [182]. The acyl-CoA ligase (AL) in module 1 attaches the α-ketoisocaproate starter unit to the first ACP [183]. This building block gets elongated by a nonribosomal peptide synthetase (NRPS) module (module 2) using glycine. Due to the modularity of both, bacterial PKS type I and NRPS biosynthetic enzymes, hybrid PKS/NRPS products encoded by a combination of PKS and NRPS modules are a common theme [17], in particular in *trans*-AT PKSs [36]. Two noncanonical reactions are catalyzed within module 3: the DH not only reduces the glycine-derived β-carbonyl after elongation, as expected, but remarkably also the remote ketone function of the starter unit [183]. In addition, full reduction to the alkane is observed, although no ER is located in module 3. This catalytic activity is instead provided in *trans* by BaeE [184]. The next step is likewise unusual. The dehydrating bimodule 4/5 catalyzes the formation of a *cis*-configured double bond [185]. The same catalytic activity is conferred by bimodule 8/9 further downstream. The next unusual modification is catalyzed by module 7, a so-called β-branching module, which introduces a methyl branch *via* a mechanism resembling exomethylene bond formation in mevalonate-dependent isoprenoid biosynthesis [182]. Another intriguing feature of this PKS is the introduction of an enamine structural portion in module 13 by double-bond generation in β,γ-position [186]. The last two modules, 16* and 17*, seem to be superfluous for the production of **101** and **102**. Remarkably, virtually all intermediates of the complex biosynthetic pathway have been structurally characterized. This was facilitated by the deletion of the terminal TE domain, which unexpectedly led to the hydrolytic cleavage of pathway intermediates from all modules [185]. These results were not only instrumental for understanding bacillaene biosynthesis in detail but also led to the discovery of a proofreading enzyme that cleaves stalled intermediates from *trans*-AT PKSs. This activity is provided by enzymes previously annotated as AT, in this case BaeD, that rather possesses TE-like activity [187]. After final cleavage of **101** off the synthase by the terminal TE, the biosynthesis of bacillaene (**102**) is concluded by oxidation at C-22/C-23 catalyzed by the cytochrome P450 BaeS [185].

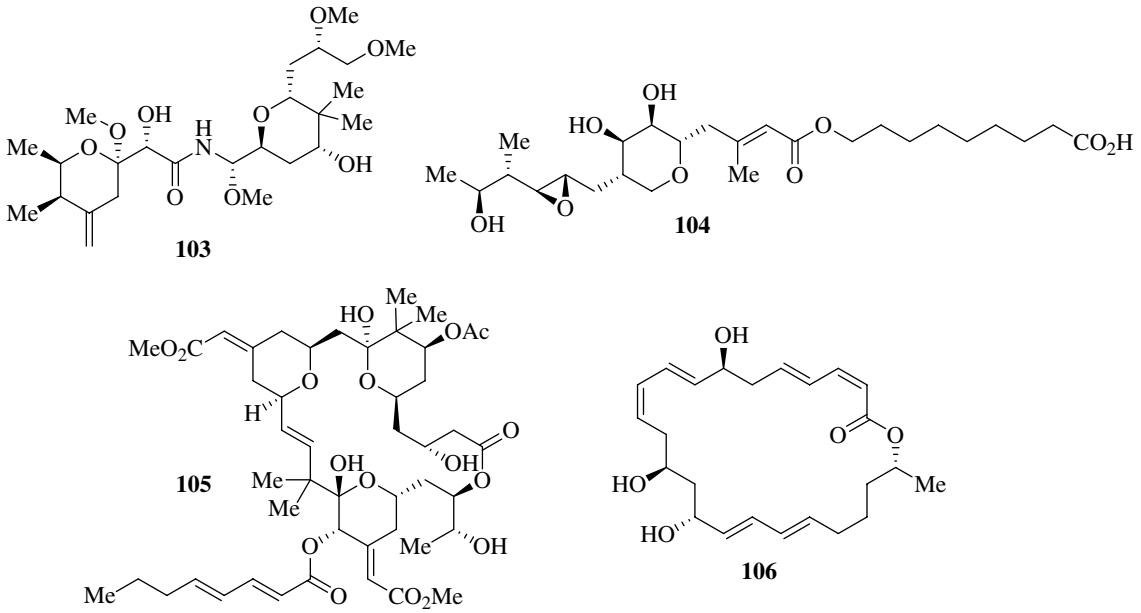
The aforementioned example nicely showcases the catalytic versatility of *trans*-At PKSs. It is thus not surprising that many structurally complex, biologically highly active secondary metabolites have meanwhile been attributed to *trans*-AT PKSs. Examples include pederin (**103**) [173], mupirocin (**104**) [188], bryostatins (**105**) [189, 190], and macrolactins (**106**) [181, 191] (Scheme 2.23).



SCHEME 2.22 The bacterial type I modular *trans*-AT PKS encoding bacillaene (**101**).

2.1.4 Conclusion

Within this chapter, we strived to introduce the mechanistic principles of polyketide core structure assembly. Despite the structural diversity of polyketide natural products, there are only few different catalytic domains facilitating a small number of basic reactions—common to all types of PKSs—that are crucial for product biosynthesis. By smart recombination of these catalytic units, nature has succeeded in generating a combinatorial library of bioactive small molecules with a structural and biological diversity that is virtually inexhaustible. The structural and functional space covered by these molecules is further expanded by post-PKS functionalization reactions catalyzed by so-called tailoring enzymes. This includes reorganization of the entire PKS



SCHEME 2.23 Representative examples of natural products derived from *trans*-AT PKSs: pederin (**103**), mupirocin (**104**), bryostatins (e.g., **105**), and macrolactins (e.g., **106**).

core (cf. Scheme 2.23) [13] but also subtle changes by group transfer reactions (e.g., glycosylation, alkylation) [192, 193] or oxidative transformations (e.g., hydroxylation, halogenation, phenol-oxidative cross coupling) [194–198]. The final steps in the biosynthesis of PKS-derived polyethers are in impressive example of oxidative tailoring discussed in Section 4.2. Investigations into the chemistry and biochemistry of these fascinating natural products will undoubtedly continue to reveal fascinating new chemical structures, biological activities, and biocatalytic potential for decades of natural product research.

REFERENCES

1. S. Pal, *Tetrahedron* 2006, **62**, 3171–3200.
2. I. Chopra, M. Roberts, *Microbiol. Mol. Biol. Rev.* 2001, **65**, 232–260.
3. A. C. Mesa-Arango, L. Scorzoni, O. Zaragoza, *Front. Microbiol.* 2012, **3**, 1–10.
4. J. A. Tobert, *Nat. Rev. Drug Discov.* 2003, **2**, 517–526.
5. G. Aubel-Sadron, D. Londos-Gagliardi, *Biochimie* 1984, **66**, 333–352.
6. W. C. Campbell, *Curr. Pharm. Biotechnol.* 2012, **13**, 853–865.
7. For WHO Model List of Essential Medicines, see here: http://www.who.int/medicines/publications/essentialmedicines/en/ (accessed on October 22, 2015).
8. M. Abrar, F. M. Anjum, M. S. Butt, I. Pasha, M. A. Randhawa, F. Saeed, K. Waqas, *Crit. Rev. Food. Sci. Nutr.* 2013, **53**, 862–874.
9. T. W. Kensler, B. D. Roebuck, G. N. Wogans, J. D. Groopman, *Toxicol. Sci.* 2011, **120**, S28–S48.
10. M. Murata, T. Yasumoto, *Nat. Prod. Rep.* 2000, **17**, 293–314.
11. D. A. Hopwood, *Chem. Rev.* 1997, **97**, 2465–2498.
12. J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* 2001, **18**, 380–416.
13. C. Hertweck, *Angew. Chem. Int. Ed. Engl.* 2009, **48**, 4688–4716.
14. S. Smith, S. C. Tsai, *Nat. Prod. Rep.* 2007, **24**, 1041–1072.
15. C. D. Reeves, *Crit. Rev. Biotechnol.* 2003, **23**, 95–147.
16. B. J. Rawlings, *Nat. Prod. Rep.* 1998, **15**, 275–308.
17. M. A. Fischbach, C. T. Walsh, *Chem. Rev.* 2006, **106**, 3468–3496.
18. B. J. Rawlings, *Nat. Prod. Rep.* 2001, **18**, 231–281.
19. B. J. Rawlings, *Nat. Prod. Rep.* 2001, **18**, 190–227.
20. R. J. Cox, *Org. Biomol. Chem.* 2007, **5**, 2010–2026.
21. R. J. Cox, T. J. Simpson, *Meth. Enzymol.* 2009, **459**, 49–78.
22. Y.-H. Chooi, Y. Tang, *J. Org. Chem.* 2012, **77**, 9933–9953.
23. C. Hertweck, A. Luzhetskyy, Y. Rebets, A. Bechthold, *Nat. Prod. Rep.* 2007, **24**, 162–190.
24. A. Das, C. Khosla, *Acc. Chem. Res.* 2009, **42**, 631–639.
25. B. J. Rawlings, *Nat. Prod. Rep.* 1999, **16**, 425–484.
26. B. Shen, *Top. Curr. Chem.* 2000, **209**, 1–51.
27. M. B. Austin, J. P. Noel, *Nat. Prod. Rep.* 2003, **20**, 79–110.
28. K. Watanabe, A. P. Praseuth, C. C. Wang, *Curr. Opin. Chem. Biol.* 2007, **11**, 279–286.
29. J.-L. Ferrer, M. B. Austin Jr. C. Stewart, J. P. Noel, *Plant Physiol. Biochem.* 2008, **46**, 356–370.
30. I. Abe, H. Morita, *Nat. Prod. Rep.* 2010, **27**, 809–838.
31. D. Yu, F. Xu, J. Zeng, J. Zhan, *IUBMB Life* 2012, **64**, 285–295.

32. W. Liu, S. D. Christenson, S. Standage, B. Shen, *Science* 2002, **297**, 1170–1173.
33. J. Ahlert, E. Shepard, N. Lomovskaya, E. Zazopoulos, A. Staffa, B. O. Bachmann, K. Huang, L. Fonstein, A. Czisny, R. E. Whitwam, C. M. Farnet, J. S. Thorson, *Science* 2002, **297**, 1173–1176.
34. J. Antosch, F. Schaefers, T. A. Gulder, *Angew. Chem. Int. Ed. Engl.* 2014, **53**, 3011–3014.
35. G. Zhang, W. Zhang, Q. Zhang, T. Shi, L. Ma, Y. Zhu, S. Li, H. Zhang, Y. L. Zhao, R. Shi, C. Zhang, *Angew. Chem. Int. Ed. Engl.* 2014, **53**, 4840–4844.
36. J. Piel, *Nat. Prod. Rep.* 2010, **27**, 996–1047.
37. R. Müller, *Chem. Biol.* 2004, **11**, 4–6.
38. B. Shen, *Curr. Opin. Chem. Biol.* 2003, **7**, 285–295.
39. M. Kaneko, E. I. Hwang, Y. Ohnishi, S. Horinouchi, *J. Ind. Microbiol. Biotechnol.* 2003, **30**, 456–461.
40. D. A. Scott, P. M. Hammond, G. M. Brearley, C. P. Price, *J. Chromatogr.* 1992, **573**, 309–312.
41. J. Rösler, F. Krekel, N. Amrhein, J. Schmid, *Plant. Physiol.* 1997, **113**, 175–179.
42. J. A. Kyndt, T. E. Meyer, M. A. Cusanovich, J. J. Van Beeumen, *FEBS Lett.* 2002, **512**, 240–244.
43. I. Abe, Y. Takahashi, H. Morita, H. Noguchi, *Eur. J. Biochem.* 2001, **268**, 3354–3359.
44. Y. Katsuyama, T. Kita, N. Funa, S. Horinouchi, *J. Biol. Chem.* 2009, **284**, 11160–11170.
45. Y. Katsuyama, T. Kita, S. Horinouchi, *FEBS Lett.* 2009, **583**, 2799–2803.
46. Y. Katsuyama, M. Matsuzawa, N. Funa, S. Horinouchi, *J. Biol. Chem.* 2007, **282**, 37702–37709.
47. S. Eckermann, G. Schröder, J. Schmidt, D. Strack, R. A. Edrada, Y. Helariutta, P. Elomaa, M. Kotilainen, I. Kilpeläinen, P. Proksch, T. H. Teeri, J. Schröder, *Nature* 1998, **396**, 387–390.
48. T. Akiyama, M. Shibuya, H. M. Liu, Y. Ebizuka, *Eur. J. Biochem.* 1999, **263**, 834–839.
49. J. Schröder, G. Schröder, *Z. Naturforsch. [C]* 1990, **45**, 1–8.
50. M. B. Austin, M. E. Bowman, J. L. Ferrer, J. Schroder, J. P. Noel, *Chem. Biol.* 2004, **11**, 1179–1194.
51. J. L. Ferrer, J. M. Jez, M. E. Bowman, R. A. Dixon, J. P. Noel, *Nat. Struct. Biol.* 1999, **6**, 775–784.
52. F. Taura, S. Tanaka, C. Taguchi, T. Fukamizu, H. Tanaka, Y. Shoyama, S. Morimoto, *FEBS Lett.* 2009, **583**, 2061–2066.
53. L. Beerhues, *FEBS Lett.* 1996, **383**, 264–266.
54. I. Abe, Y. Utsumi, S. Oguro, H. Morita, Y. Sano, H. Noguchi, *J. Am. Chem. Soc.* 2005, **127**, 1362–1363.
55. I. Abe, S. Oguro, Y. Utsumi, Y. Sano, H. Noguchi, *J. Am. Chem. Soc.* 2005, **127**, 12709–12716.
56. H. Fu, D. A. Hopwood, C. Khosla, *Chem. Biol.* 1994, **1**, 205–210.
57. N. Funa, Y. Ohnishi, I. Fujii, M. Shibuya, Y. Ebizuka, S. Horinouchi, *Nature* 1999, **400**, 897–899.
58. N. Funa, Y. Ohnishi, Y. Ebizuka, S. Horinouchi, *J. Biol. Chem.* 2002, **277**, 4628–4635.
59. J. Cortés, J. Velasco, G. Foster, A. P. Blackaby, B. A. Rudd, B. Wilkinson, *Mol. Microbiol.* 2002, **44**, 1213–1224.
60. S. Li, S. Grüschow, J. S. Dordick, D. H. Sherman, *J. Biol. Chem.* 2007, **282**, 12765–12772.
61. J. Zeng, R. Decker, J. Zhan, *Appl. Biochem. Biotechnol.* 2012, **166**, 1020–1033.
62. F. Gross, N. Luniak, O. Perlova, N. Gaitatzis, H. Jenke-Kodama, K. Gerth, D. Gottschalk, E. Dittmann, R. Müller, *Arch. Microbiol.* 2006, **185**, 28–38.
63. W. Zha, S. B. Rubin-Pitel, H. Zhao, *J. Biol. Chem.* 2006, **281**, 32036–32047.
64. C. C. Tseng, S. M. McLoughlin, N. L. Kelleher, C. T. Walsh, *Biochemistry* 2004, **43**, 970–980.
65. H. Chen, C. C. Tseng, B. K. Hubbard, C. T. Walsh, *Proc. Natl. Acad. Sci. U S A* 2001, **98**, 14901–14906.
66. A. Miyanaga, N. Funa, T. Awakawa, S. Horinouchi, *Proc. Natl. Acad. Sci. U S A* 2008, **105**, 871–876.
67. M. Funabashi, N. Funa, S. Horinouchi, *J. Biol. Chem.* 2008, **283**, 13983–13991.
68. T. Hayashi, Y. Kitamura, N. Funa, Y. Ohnishi, S. Horinouchi, *Chembiochem* 2011, **12**, 2166–2176.
69. N. Funa, H. Ozawa, A. Hirata, S. Horinouchi, *Proc. Natl. Acad. Sci. U S A* 2006, **103**, 6356–6361.
70. C. Nakano, H. Ozawa, G. Akanuma, N. Funa, S. Horinouchi, *J. Bacteriol.* 2009, **191**, 4916–4923.
71. P. Saxena, G. Yadav, D. Mohanty, R. S. Gokhale, *J. Biol. Chem.* 2003, **278**, 44780–44790.
72. J. A. Chemler, T. J. Buchholz, T. W. Geders, D. L. Akey, C. M. Rath, G. E. Chlipala, J. L. Smith, D. H. Sherman, *J. Am. Chem. Soc.* 2012, **134**, 7359–7366.
73. L. Song, F. Barona-Gomez, C. Corre, L. Xiang, D. W. Udwary, M. B. Austin, J. P. Noel, B. S. Moore, G. L. Challis, *J. Am. Chem. Soc.* 2006, **128**, 14754–14755.
74. D. Yu, J. Zeng, D. Chen, J. Zhan, *Enzyme Microb. Technol.* 2010, **46**, 575–580.
75. N. Funa, T. Awakawa, S. Horinouchi, *J. Biol. Chem.* 2007, **282**, 14476–14481.
76. S. B. Rubin-Pitel, H. Zhang, T. Vu, J. S. Brunzelle, H. Zhao, S. K. Nair, *Chem. Biol.* 2008, **15**, 1079–1090.
77. J. Li, Y. Luo, J. K. Lee, H. Zhao, *Bioorg. Med. Chem. Lett.* 2011, **21**, 6085–6089.
78. Y. Seshime, P. R. Juvvadi, K. Kitamoto, Y. Ebizuka, I. Fujii, *Bioorg. Med. Chem.* 2010, **18**, 4542–4546.
79. M. Hashimoto, T. Koen, H. Takahashi, C. Suda, K. Kitamoto, I. Fujii, *J. Biol. Chem.* 2014, **289**, 19976–19984.
80. T. Mori, D. Yang, T. Matsui, M. Hashimoto, H. Morita, I. Fujii, I. Abe, *J. Biol. Chem.* 2015, **290**, 5214–5225.
81. W. P. Revill, M. J. Bibb, D. A. Hopwood, *J. Bacteriol.* 1995, **177**, 3946–3952.
82. T. S. Hitchman, J. Crosby, K. J. Byrom, R. J. Cox, T. J. Simpson, *Chem. Biol.* 1998, **5**, 35–47.