Springer Theses
Recognizing Outstanding Ph.D. Research

Matthew Jenner

Using Mass Spectrometry for Biochemical Studies on Enzymatic Domains from Polyketide Synthases

Springer

Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series "Springer Theses" brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student's supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today's younger generation of scientists.

Theses are accepted into the series by invited nomination only and must fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder.
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to scientists not expert in that particular field.

More information about this series at http://www.springer.com/series/8790

Matthew Jenner

Using Mass Spectrometry for Biochemical Studies on Enzymatic Domains from Polyketide Synthases

Doctoral Thesis accepted by
the University of Nottingham, UK

Springer

Author
Dr. Matthew Jenner
School of Chemistry
University of Warwick
Coventry
UK

Supervisor
Prof. Neil Oldham
University of Nottingham
Nottingham
UK

ISSN 2190-5053 ISSN 2190-5061 (electronic)
Springer Theses
ISBN 978-3-319-32722-8 ISBN 978-3-319-32723-5 (eBook)
DOI 10.1007/978-3-319-32723-5

Library of Congress Control Number: 2016936964

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Parts of this thesis have been published in the following journal articles:

Substrate Specificity in Ketosynthase Domains from trans-AT Polyketide Synthases

***Matthew Jenner**, Sarah Frank, Annette Kampa, Christoph Kohlhaas, Geoff S. Briggs, Jörn Piel and Neil J. Oldham.*

*Angewante Chemie International Edition, 2013, **52**, 1143–1147.*

Amino Acid-Accepting Ketosynthase Domain from a trans-AT Polyketide Synthase Exhibits High Selectivity for Predicted Intermediate

*Christoph Kohlhaas,[†] **Matthew Jenner**,[†] Annette Kampa, Geoff S. Briggs, Jose P. Afonso, Jörn Piel and Neil J. Oldham.*

*Chemical Science, 2013, **4**, 3212–3217.*

A Close Look at a Ketosynthase from a trans-acyltransferase Modular Polyketide Synthase

*Darren C. Gay, Glen Gay, Abram J. Axelrod, **Matthew Jenner**, Christoph Kohlhaas, Annette Kampa Neil J. Oldham, Jörn Piel, Adrian T. Keatinge-Clay.*

*Structure, 2014, **22**, 444–451.*

Acyl Chain Elongation Drives Ketosynthase Substrate Specificity in Polyketide Biosynthesis

***Matthew Jenner**, Jose P. Afonso, Hannah R. Bailey, Annette Kampa, Sarah Frank, Jörn Piel and Neil J. Oldham*

*Angwante Chemie International Edition, 2015, **54**, 1817–1821.*

Acyl Hydrolases from trans-AT Polyketide Synthases Target Acetyl Units of Acyl Carrier Proteins

***Matthew Jenner**, Jose P. Afonso, Christoph Kohlhaas, Petra Karbaum, Sarah Frank, Jörn Piel and Neil J. Oldham Chemical Communications, 2016, **52**,* 5262–5265

Supervisor's Foreword

This thesis details outstanding work conducted by Matthew Jenner between 2010 and 2014 in the School of Chemistry, University of Nottingham under my supervision. It describes a mass spectrometric investigation of enzymatic domains from *trans*-acyl transferase polyketide synthases (*trans*-AT PKSs). These megasynthases are responsible for the biosynthesis of a wide variety of bioactive natural products such as pederin, psymberin, bacillaene, and kirromycin.

Unlike the more familiar *cis*-AT PKSs, which display collinearity between the domain architecture of the PKS and the structure of the polyketide product, *trans*-AT PKSs exhibit aberrant architecture and often incorporate novel enzymatic domains, resulting in poor biosynthetic assignments. An introduction to polyketide biosynthesis and to the differences between *cis*- and *trans*-PKSs is provided in Chap. 1 of the thesis.

Given that *trans*-AT systems constitute approximately 40 % of all bacterial multimodular PKSs, they represent a major, but poorly characterised, enzyme class that is of high relevance for drug discovery. Against this background, Matthew set out to develop tools that could be used to help understand these intriguing PKSs. Our starting point was a phylogenetic study by Nguyen et al. (*Nat. Biotechnol.* 2008), which predicted that—based on sequence analysis—ketosynthase domains from *trans*-AT PKSs should possess specificity for particular biosynthetic intermediate 'types' (e.g. acetyl starter units, β-hydroxy-, β-keto-, and enoyl-chains). The first objective was to develop a simple functional assay to confirm this proposed specificity. The initial KS acylation step of the enzyme's mechanism was selected as the first point to determine specificity, as this could be easily monitored by measuring the associated mass increase of the isolated KS domains. This work is described in the Chaps. 3 and 4: the first two results and discussion chapters of the thesis. We chose a variety of KSs from the psymberin and bacillaene PKSs, and a set of acyl *N*-acetylcysteamines (acyl SNACs) for use as substrate mimics. Plasmid vectors for the former were kindly supplied by Jörn Piel's group (University of Bonn, then ETH Zurich), as were the SNAC thioesters.

Chapter 3 describes the successful development and application of a acylation assay to investigate the tolerance of KS domains for beta-methyl branched

substrates. Selectivity for non-branched intermediates was seen in KSs that are proposed to process linear acyl chains. Matthew rationalised this observation at the amino acid level by use of homology models, and—in some very nice experiments—employed site-directed mutagenesis to modulate this selectivity and allow acceptance of branched substrates.

Chapter 4 goes on to describe the application of the acylation assay to a KS domain located immediately downstream of the non-ribosomal peptide synthase module within the PKS responsible for the biosynthesis of bacillaene. Here, selectivity for amino-acid-derived SNAC thioesters was seen, as predicted based on the knowledge of the expected substrate intermediate. Once again Matthew rationalised this result by homology modelling, and probed the role of key amino acid residues within the KS. Very pleasingly, his prediction of the importance of an asparagine residue in bonding to the amino-acid-derived substrates was recently confirmed by the Keatinge-Clay group using X-ray crystallography.

Chapter 5 presents a simple method for making acyl-ACPs for use in PKS enzyme assays. This allows the synthesis of more realistic substrate mimics, where the full phosphopantetheine linker chain tethers the acyl chain to the ACP. Matthew went on to use these products to probe the substrate specificity of the acyl hydrolase from the pederin PKS, and demonstrate that its major housekeeping role is probably in targeting unwanted acetyl-ACP, which may be derived from acetyl-CoA during initial activation of the PKS by the promiscuous phosphopantetheine transferases.

In Chap. 6, the development of an assay to measure the selectivity of the KS elongation step is described. This key process follows the initial acylation of the KS active site, which was probed in Chaps. 3 and 4. Thus, with these two assays in hand, Matthew was able to unpick the enzymology of KS domains by studying the selectivity of each step of the KS-catalysed reaction. Interestingly, the elongation step was found to be much more selective than the preceding acylation, and only those substrates that closely mimicked the natural intermediate of the KS were elongated.

In summary, this outstanding thesis represents a significant contribution to our understanding of KS specificity in this intriguing family of PKSs. The assays developed will prove very useful to researchers in the PKS field, and the findings will almost certainly find utility in future PKS engineering efforts.

I would like to thank Matthew for all the hard work he put into his Ph.D. project, for his infectious enthusiasm, and for producing such an excellent thesis.

Nottingham, UK
April 2016

Prof. Neil Oldham

Abstract

Polyketides form a group of diverse and structurally complex bioactive natural products. Their biosynthesis is directed by multi-domain polyketide megasynthases (PKSs), which extend the acyl chain by a series of condensation and optional reduction steps. Phylogenetic work has shown that, in a particular group of type I systems known as *trans*-AT PKSs, the ketosynthase (KS) domains potentially harbour specificity towards the nature of the first four carbons of the intermediate substrate (e.g. beta-hydroxy, enoyl, methyl-branched). These results suggest a close link between KS evolution and substrate specificity.

This thesis reports studies on the substrate specificity of crucial KS domains from *trans*-AT PKSs. Using a combination of electrospray ionisation-mass spectrometry (ESI-MS) and simple *N*-acetyl cysteamine (SNAC) substrate mimics, the substrate specificity of a range of KS domains from the bacillaene (BaeJ and BaeL) and psymberin (PsyA and PsyD) PKSs, including a KS domain immediately downstream of a non-ribosomal peptide synthase (NRPS) module, have been successfully studied with regard to the initial acylation step of KS-catalysis. In addition, the ability to alter the substrate tolerance of KS domains by simple point mutations in the active site has been successfully demonstrated. A novel method for the synthesis of acyl-acyl carrier proteins (ACP) from SNAC thioesters is also reported. A series of acyl-ACPs have been synthesised using this methodology and successfully used to probe the substrate specificity of both KS domains and the previously uncharacterised acyl hydrolase (AH) domain, PedC.

KS-catalysed chain elongation reactions have also been conducted and monitored by ESI-MS/MS. All KS domains studied exhibited higher substrate specificity at the elongation step than in the preceding acylation. Furthermore, a mechanism of reversible acylation is proposed using the PsyA ACP1-KS1 didomain. The findings reported in this thesis provide important insights into the mechanism of KS specificity and show that mutagenesis can be used to expand the repertoire of acceptable substrates for future PKS engineering efforts.

Acknowledgements

First and foremost, I would like to convey my utmost thanks to Dr. Neil Oldham for giving me the opportunity to work on this project, and for his constant support, encouragement, and guidance since joining the lab as a project student—It has been a truly enjoyable 5 years!

I also extend my gratitude to Prof. Jörn Piel and members of his group in both Bonn and Zürich for a fruitful collaboration. In particular: Sarah Frank, Christoph Kohlhaas, and Petra Pöplau for the synthesis of numerous compounds—sometimes on demand! I would like to thank Annette Kampa for all her work before and during her visit to Nottingham, and for providing numerous KS constructs. Additionally, thanks to Dr. Anna Vagstad for sending lots of plasmids with extremely detailed advice during my final year.

I would like to specially mention to Dr. Jose Afonso, who taught most of what I now know about molecular biology, and for allowing me to call him the name of any Latin football player I desired. Also, thanks go to Dr. Kleitos Sokratous who not only helped to fix numerous instrument-based issues, but was also a great friend and housemate during his time in Nottingham.

I am extremely appreciative to Prof. Panos Soultanas for allowing me to conduct all my molecular biology work in his lab. Furthermore, I extend my thanks to Dr. Geoff Briggs and Matt Green for their constant support and advice whilst working in CBS. I also want to thank all the past and present members of the Oldham, Soultanas, and Searle groups for making the lab such a great place to work, in particular to Jon Hopper, Lucy Roach, Richard Elms, Dan Scott, Lucio Manzi, Nkazi Tshuma, Liz Morris, Jenny Adlington, Alex Slater, Juliet Morgan, Jed Long, Vasillis Paschalis, and Sarah Northall.

Many thanks go to Matt, Alex, and Dan from the MonoQs, who have provided endless laughs and a welcome musical distraction from work. I would also like to give special thanks to Christina for her continuous patience, love, and reassurance throughout my studies. Huge thanks to Michael and Kathryn for making my holiday periods at home so entertaining—a time of year I always look forward to. Finally, I would like to dedicate this thesis to my parents on the year of their 30th wedding anniversary as a token of my gratitude for all their love, support, and encouragement during my studies, without which, none of this would have been possible.

Contents

Abbreviations

A-Domain	Adenylation Domain
ACP	Acyl Carrier Protein
ADP	Adenosine Diphosphate
AL	Acyl Ligase
Ala	Alanine
AMP	Adenosine Monophosphate
Arg	Arginine
Asn	Asparagine
AT	Acyltransferase
ATP	Adenosine Triphosphate
C-Domain	Condensation Domain
CID	Collision-Induced Dissociation
CoA	Coenzyme A
CR	Crotonase
CRM	Charge Residue Model
Cy	Cyclase
Cys	Cysteine
DCC	*N-N'*-Dicyclohexylcarbodiimide
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic Acid
e.q.	Equivalent
ECH	Enoyl-CoA Dehydratase
EDTA	Ethylenediaminetetraacetic Acid
ER	Enoylreductase
ESI	Electrospray Ionisation
FAS	Fatty Acid Synthase
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
GNAT	GCN5-related *N*-acetyltransferase
His	Histidine

HMGS	3-hydroxy-3-methylglutaryl-CoA
IEM	Ion Evaporation Model
IPTG	Isopropylthio-β-galactoside
KR	Ketoreductase
KS	Ketosynthase
LB	Luria Broth
LC	Liquid Chromatography
Leu	Leucine
Lys	Lysine
Met	Methionine
MS/MS	Tandem Mass Spectrometry
MS	Mass Spectrometry
MT	Methyltransferase
NMR	Nuclear Magnetic Resonance
NRPS	Non-Ribosomal Peptide Synthase
Phe	Phenylalanine
PKS	Polyketide Synthase
PPant	Phosphopantetheine
PPTase	4'-Phosphopantetheinyl Transferase
Pro	Proline
Ser	Serine
SNAC	*N*-Acetyl Cysteamine
TE	Thioesterase
TFA	Trifluoroacetic Acid
Thr	Threonine
TOF	Time of Flight
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

Chapter 1
Introduction

1.1 Natural Products

The realm of natural products and their derivatives has provided the most successful source of bioactive drug molecules for generations [1]. The medicinal use of natural products can be dated back to 2600 B.C. in Mesopotamia, where the earliest records of natural products are documented on clay tablets, including simple oils such as *Commiphora myrrha* (myrrh) to treat colds and inflammation [2]. Since then, many bioactive compounds have been discovered including salicin (the precursor to aspirin (**1**), from *Salix alba*), morphine (**2**) (from *Papaver somniferum*) and the anti-malarial quinine (**3**) (from *Cinchona succirubra*). However, arguably the most notable natural product discovery is that of penicillin (**4**) from the fungus *Penicillium notatum*, which was credited to Alexander Fleming in 1929, but with major contributions from Howard Florey, Edward Penley Abraham, and Ernst Chain. Since this breakthrough, the introduction of antibiotics such as tetracycline (**5**) and kanamycin (**6**) has decreased the mortality rate from bacteria-induced diseases drastically (Fig 1.1) [3].

Despite the positive impact antibiotics have had on human health, their misuse combined with global travel has allowed strains of pathogenic bacteria to acquire resistance to these compounds, requiring the need for novel antibiotic agents [4]. Since the early 80's natural products or their derivatives have comprised ~40 % of new chemical entities, and represent 70 % of current antibiotic compounds [5]. Despite this, in recent years pharmaceutical drug discovery programmes have tended to focus upon combinatorial libraries of small fragment molecules. It is estimated that <10 % of the worlds biodiversity has been investigated to date for potential biological activity, therefore a metaphorical chemical vault of natural products remains undiscovered. Accessing this chemical diversity remains the current challenge for researchers.

M. Jenner, *Using Mass Spectrometry for Biochemical Studies on Enzymatic Domains from Polyketide Synthases*, Springer Theses,
DOI 10.1007/978-3-319-32723-5_1

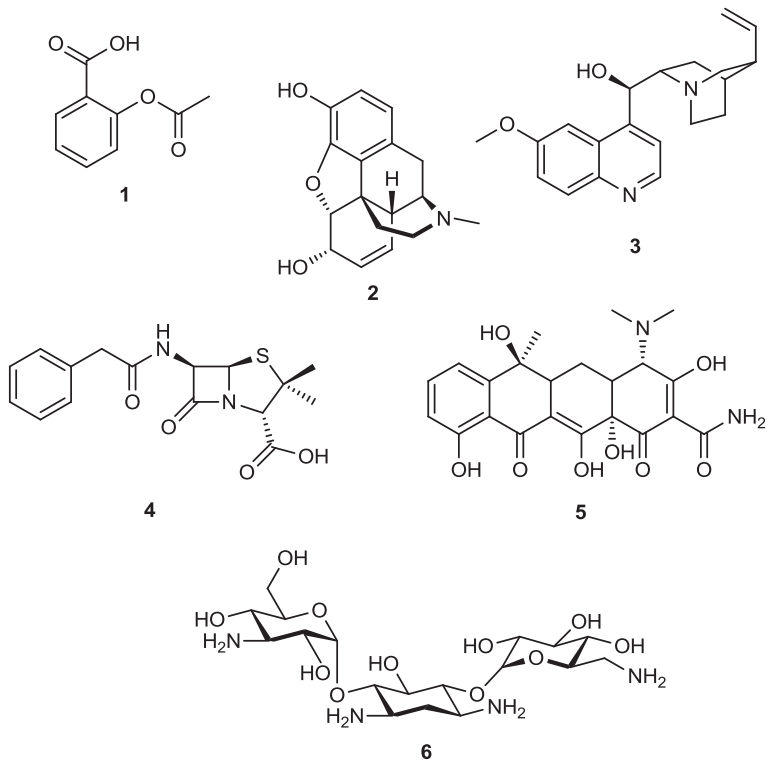


Fig. 1.1 Aspirin (**1**), Morphine (**2**), Quinine (**3**), Penicillin G (**4**), Tetracycline (**5**), Kanamycin (**6**)

1.1.1 Polyketides and Non-ribosomal Peptides

Polyketides (PK), non-ribosomal polypeptides (NRP) and their hybrids (PK-NRP) represent a diverse class of natural products that exhibit a wide range of pharmaceutical activities (Fig. 1.2). Such biological activity includes: antibiotic (Erythromycin A (**7**)), immunosuppressant (FK-506 (**8**)), antifungal (Amphotericin B (**9**)), antitumour (Geldanmycin (**10**)) and hypolipidemic agents (Lovastatin (**11**)). With such an extensive range of therapeutic biological activities, natural product research is constantly working to discover novel polyketide structures capable of enhancing human health.

Although many of these highly sought-after natural products are, in principle, available from sources in nature, obtaining sufficient quantities to characterise their potential therapeutic activity is often difficult [6]. Moreover, efficient chemical synthesis of these complex compounds is extremely challenging. Investigating the biosynthetic mechanisms responsible for such attractive compounds has allowed the chemistry of polyketide-producing enzymes to be harnessed for the production of novel, biochemically bioactive natural products [7, 8].

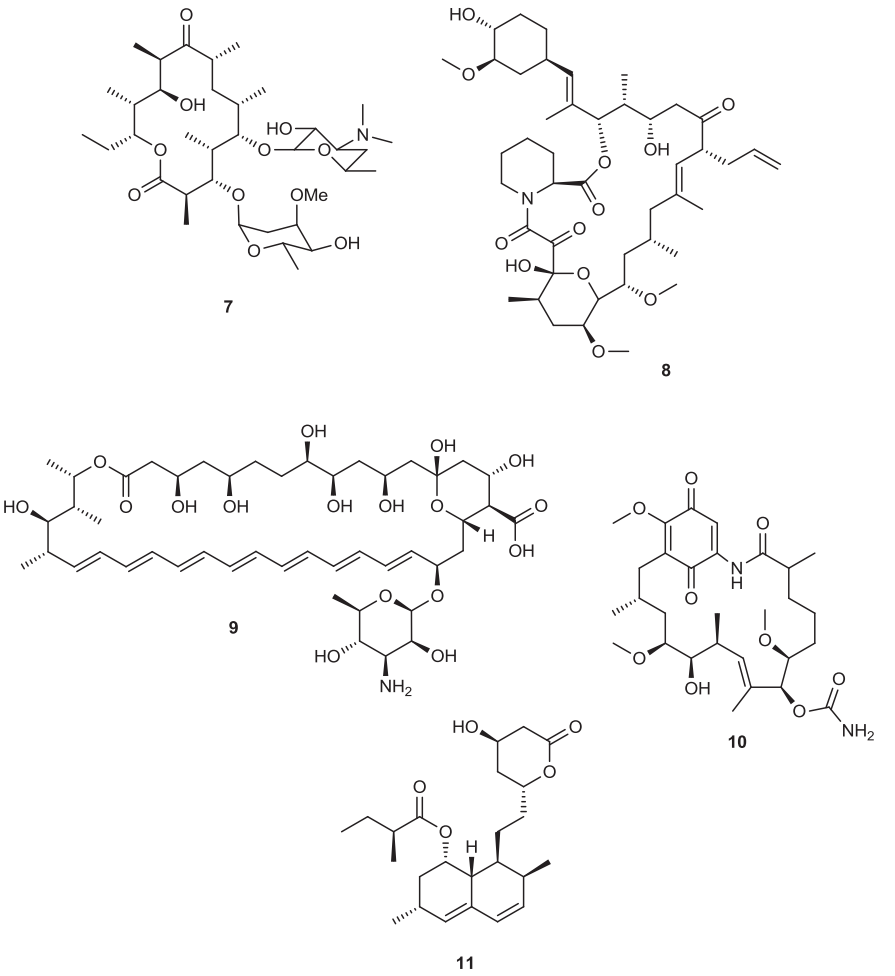
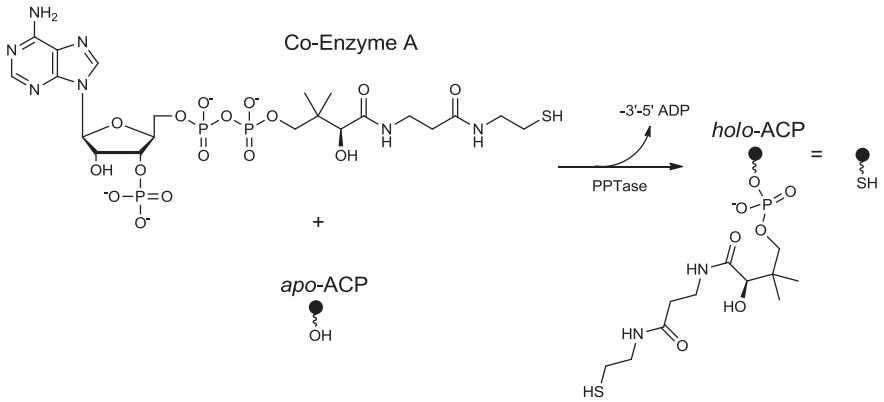


Fig. 1.2 Erythromycin A (**7**), FK-506 (**8**), Amphotericin B (**9**), Geldanmycin (**10**), Lovastatin (**11**)

This thesis reports efforts to better understand the enzymology of key proteins involved in polyketide biosynthesis, and how understanding the specificity of these enzymes can assist the discovery of novel natural products.

1.2 Polyketide Synthases

Polyketide synthases (PKSs) are large protein complexes responsible for the biosynthesis of an array of complex, biologically active compounds, many of which are employed in medicine. Several classes of PKSs exist, each with an architectural variation on a common mechanistic theme.



Scheme 1.1 Post-translational attachment of phosphopantetheine from co-enzyme A to a conserved serine residue on an *apo*-ACP, catalysed by a PPTase

1.2.1 Polyketide Biosynthesis

Despite the structural diversity of polyketides, the building blocks of these compounds are simple acyl-CoAs, and the biosynthetic logic is closely associated with that of fatty acid synthases (FASs) [9, 10]. In both polyketide and fatty acid biosynthesis, the growing chain is covalently tethered to an acyl carrier protein (ACP) via a phosphopantetheine (PPant) moiety. The PPant chain, derived from CoA, is attached post-translationally to a conserved serine residue on the *apo*-ACP by a phosphopantetheinyl transferase (PPTase), yielding the mature *holo*-ACP (Scheme 1.1) [11, 12].

The *holo*-ACP, in conjunction with an acyltransferase (AT) and a ketosynthase (KS), forms the minimal set of enzymatic domains required for a single chain elongation reaction. This set of enzymatic domains is often termed a module. Polyketide biosynthesis is initiated by the loading of a starter unit, derived from an acyl-CoA, onto the PPant arm of the ACP [12, 13]. In fatty acid biosynthesis the starter unit is typically limited to acetyl, whereas PKSs can utilise a variety of primer units including acetyl-, propionyl-, butyryl and other variants [10]. The role of the AT domain is to catalyse the transfer of a malonyl-derived extender unit from a CoA-thioester onto the thiol of the ACP phophopantetheinyl moiety (Fig. 1.3).

The elongating carbon-carbon bond formation is then achieved by Claisen condensation chemistry catalysed by the KS domain. Further optional modification at the β-keto position is directed by the presence of ketoreductase (KR), dehydratase (DH) and enoylreductase (ER) domains, generating β-hydroxyl, olefinic and fully saturated intermediates respectively (Fig. 1.4) [12, 14, 15].