Functional and structural analysis of the angucyclinone ketoreductase SimC7

By

Martin Schäfer

2016

John Innes Centre
Department of Molecular Microbiology
Norwich Research Park, Colney Lane, Norwich, NR4 7UH
United Kingdom

A thesis submitted to the University of East Anglia, Norwich, for the degree of Doctor of Philosophy

John Innes Centre

Department of Molecular Microbiology

Molecular Microbiology

Doctoral Training Partnership Programme funded by BBSRC

Supervised by

Prof. Mark J. Buttner

Department of Molecular Microbiology

John Innes Centre

Norwich, United Kingdom

Examined by

Prof. Sarah O'Connor

Department of Chemistry

John Innes Centre

Norwich, United Kingdom

Prof. Dominic Campopiano

Department of Chemistry

University of Edinburgh

Edinburgh, United Kingdom

Abstract

The natural product simocyclinone D8 (SD8) is a potent DNA gyrase inhibitor produced by *Streptomyces antibioticus* Tü6040. The biosynthetic simocyclinone (*sim*) gene cluster has been sequenced and a hypothetical biosynthetic pathway has been proposed. The tetraene linker in SD8 was suggested to be the product of a modular type I polyketide synthase working *in trans* with two monofunctional enzymes. SimC7, which belongs to the short-chain dehydrogenase/reductase (SDR) superfamily of proteins, was proposed to supply the dehydratase activity missing from two modules of the polyketide synthase.

In this study, I report the structure and function of recombinant *S. antibioticus* SimC7. Because the natural simocyclinone producer is genetically intractable, the ~72-kb *sim* cluster was isolated on a single phage artificial chromosome (PAC) clone for heterologous production in a *Streptomyces coelicolor* strain engineered for improved antibiotic production. Deletion of *simC7* resulted in production of a novel simocyclinone, 7-oxo-SD8, which carried a normal tetraene linker but was altered in the angucyclinone. I demonstrate that SimC7 is an NAD(P)H-dependent polyketide ketoreductase that catalyses the reduction of the C-7 carbonyl of the angucyclinone and the resulting hydroxyl group is essential for antibiotic activity. SimC7 shares little sequence similarity with characterized ketoreductases, suggesting it might have a distinct catalytic mechanism. To investigate this possibility, I determined the structures of SimC7 alone, in complex with $NADP^+$, and in complex with $NADP^+$ and the substrate 7-oxo-SD8. These structures show that SimC7 is distinct from previously characterized polyketide ketoreductases, lacking the conserved catalytic triad (Ser-Tyr-Lys), including the active site tyrosine that acts as central acid-base catalyst in canonical SDR proteins. Taken together with functional analyses of active site mutants, my data suggest SimC7 catalyses a substrate-assisted, two-step reaction for reduction of the C-7 carbonyl group involving intramolecular transfer of a substrate-derived proton to generate a phenolate intermediate.

Abbreviations

Apr^R	apramycin resistance
ASU	asymmetric unit
$Carb^R$	carbenicillin resistance
Cml^R	chloramphenicol resistance
CTAB	cetyltrimethylammonium bromide
DTT	dithiotreitol
GyrA	DNA gyrase subunit A
GyrB	DNA gyrase subunit B
HPLC	high-pressure liquid chromatography
Hyg^R	hygromycin resistance
IC_{50}	half maximal inhibitory concentration
Kan^R	kanamycin resistance
KR	ketoreductase
LC-MS	liquid chromatography-mass spectrometry
MIC	minimum inhibitory concentration
$NAD(P)^+$	nicotinamide adenine dinucleotide (phosphate), oxidised
NADPH	nicotinamide adenine dinucleotide (phosphate), reduced
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm wavelength
ORF	open reading frame
PAC	phage artificial chromosome
PDB	protein data bank
PKS	polyketide synthase
r.m.s.d.	root-mean-square deviation
rpm	revolutions per minute
SD8	simocyclinone D8
SDR	short-chain dehydrogenase/reductase
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SeMet	selenomethionine
$Strep^R$	streptomycin resistance
$Thio^R$	thiostrepton resistance

Acknowledgements

It was a fantastic experience to work at the John Innes Centre in Norwich. Many people have contributed to the success of my PhD in and outside the John Innes Centre and I would like to express my sincere gratitude to everyone for such a special time in my life.

First of all, I would like to thank my supervisor Mark Buttner. Mark, thank you for everything that you have done for me, for your endless support, encouragement and mentoring throughout my PhD. You have always given me the freedom to work and shape my project and helped me during exciting and challenging times throughout the past years. It was a fantastic time and I will miss you and the lab very much. I also would like to than Tung Le for developing this project together with Mark. Tung, you have done some amazing work and I learned an incredible amount during my PhD, which I could not even imagine before I started my project.

Secondly, I would like to thank my supervisor committee, Barrie Wilkinson, Mervyn Bibb and Maureen Bibb as well as Tony Maxwell, David Lawson, Stephen Bornemann and Greg Challis for constructive discussions, helpful comments and valuable advice.

Furthermore, I would like to thank Clare Stevenson and David Lawson for introducing me to protein crystallography. I deeply enjoyed this experience including the most exciting moments during ‘Diamond time’.

I would also like to thank Lionel Hill for introducing me to mass spectrometry and Sergey Nepogodiev for help with nuclear magnetic resonance spectroscopy. And I would like to thank the media kitchen staff for preparing media, buffers and special orders so that I can dedicate my time for experiments.

I would like to thank all members of the growing *Streptomyces* group for giving me such a wonderful time. A special thanks to my fantastic bench

mates Mahmoud Al-Bassam, Maureen Bibb, and Natalia Tschowri, Xiaoluo Huang and Aisling Cooke for a very cheerful and successful time in the lab. Also a great thanks to Susan Schlimpert, Matt Bush, Morgan Feeney, Kelley Gallagher, Franziska Duerr, Andrew Truman, Neil Holmes, Zhiwei Qin, Silke Alt, Juan Pablo Gomez-Escribano, Lorena Fernandez, David Widdick and Govind Chandra, Vesna Simunovic, Jean Franco Castro, Valeria Razmilic, Chiara Borsetto, Javier Aberturas and Siobhan Dorai-Raj.

I am very thankful to Jenni Rant and Anne Osbourne, who gave me a fantastic opportunity for my PIPS internship to work for the Science, Art and Writing (SAW) Trust. I enjoyed a fabulous experience in which I learned to communicate science in schools and to plan and develop various science outreach projects.

I am very grateful about the funding from the Biotechnology and Biological Science Research Council (BBSRC), the *Streptomyces* industrial account and the John Innes Foundation Hardship Fund.

Finally, I would like to thank my family and my friends for their motivation and support during the last years. Jib, I would like to thank you for your patience and you support during our wonderful time in Norwich. My boy Charlie, your endless energy and happiness are making me extremely proud and you are giving me the motivation to go through all possible challenges.

Martin Schäfer

Table of Contents

Abstract **3**

Abbreviations **4**

Acknowledgements **5**

Table of Contents **7**

Index of Tables, Figures, and Equations **12**

Publications (see Appendix) **17**

Statement of my work **18**

Chapter 1 General introduction **19**

1.1 Importance of antibiotics 20

1.2 DNA topoisomerases 21

1.2.1 DNA gyrase 23

1.3 Aminocoumarins 28

1.3.1 Biosynthesis of aminocoumarins 30

1.3.2 Resistance mechanisms against aminocoumarins 30

1.4 Simocyclinones 31

1.4.1 Biosynthesis of simocyclinones 34

1.4.2 Biological activity of simocyclinones 39

1.4.3 Resistance mechanism against simocyclinones 43

1.5 Aims of this project 44

Chapter 2 Materials and methods **45**

2.1 Enzymes 48

2.2 Bacterial strains, plasmids, and oligonucleotides 49

2.2.1 Bacterial strains 49

2.2.2 Spontaneous resistant mutants of *S. coelicolor* M1152 50
2.2.3 Plasmids and PAC clones..... 51
2.2.4 Oligonucleotides 56
2.3 Culture media, buffers and solutions..... 65
2.3.1 Culture media 65
2.3.2 Buffers and solutions 69
2.3.3 Antibiotic stock concentrations..... 73
2.3.4 Minimum inhibitory concentration (MIC) assays 74
2.4 DNA manipulations *in vivo* (Genetic manipulations)..... 74
2.4.1 Preparation of electro-competent cells of *E. coli*..... 74
2.4.2 Transformation of *E. coli*..... 75
2.4.3 Conjugation between *E. coli* and *Streptomyces* spp..... 75
2.5 DNA manipulations and *in vitro* (Cloning) 77
2.5.1 Isolation of plasmid DNA from *E. coli*..... 77
2.5.2 Isolation of chromosomal DNA from *Streptomyces* spp..... 77
2.5.3 Agarose gel electrophoresis 78
2.5.4 Purification of size-separated DNA fragments 78
2.5.5 Polymerase chain reaction..... 79
2.5.6 Cloning of DNA fragments 81
2.5.7 PAC library construction 81
2.5.8 Deletion of genes via PCR targeting and complementation.. 82
2.5.9 Site-directed mutagenesis of SimC7..... 85
2.5.10 Recombineering with single-stranded DNA oligonucleotides 85
2.5.11 Clone library from genomic DNA 86
2.5.12 Sequencing..... 86
2.6 Expression and purification of recombinant proteins..... 87
2.6.1 Protein expression and purification (His-tagged) 87

2.6.2 Protein quantification via Bradford assay 90
2.6.3 Discontinuous polyacrylamide gene electrophoresis (SDS-PAGE) 90
2.6.4 Staining and immobilisation of proteins in SDS-PAGE gels .. 92
2.7 Protein crystallography 92
2.7.1 Protein crystallisation and cryoprotection 92
2.7.2 Selenomethionine labelling of recombinant protein 94
2.7.3 X-ray data collection, structure determination and structure refinement 94
2.8 Enzyme assays *in vitro* 96
2.8.1 DNA gyrase inhibition assays 96
2.8.2 Surface plasmon resonance (SPR) 97
2.8.3 Circular dichroism (CD) spectroscopy 98
2.8.4 SimC7 oligomeric state determination by size exclusion chromatography (SEC) 98
2.8.5 Ketoreductase activity assays 99
2.9 Extraction and analysis of simocyclinones 102
2.9.1 Extraction of simocyclinones 102
2.9.2 ^{13}C-enrichment of simocyclinones 102
2.9.3 Liquid chromatography-mass spectrometry (LC-MS) 103
2.9.4 High-pressure liquid chromatography (HPLC) 104
2.9.5 Normal phase liquid chromatography (LC) 105
2.9.6 Thin layer chromatography (TLC) 105
2.9.7 Small molecule size exclusion chromatography (SEC) 105
2.9.8 Nuclear magnetic resonance (NMR) spectroscopy 106
2.10 Software for *in silico* analysis 108

Chapter 3 Heterologous expression of simocyclinones (Schäfer et al., 2015) 110

3.1 Introduction 111
3.2 Isolation of the simocyclinone (*sim*) biosynthetic gene cluster .. 112
3.3 Optimisation of the heterologous system 114
3.4 Discussion..... 118
3.5 Acknowledgements 119

Chapter 4 SimC7 is a novel NAD(P)H-dependent ketoreductase essential for the antibiotic activity of the DNA gyrase inhibitor simocyclinone (Schäfer et al., 2015) 120

4.1 Introduction 121
4.2 Structural analysis of the novel simocyclinone 7-oxo-SD8..... 122
4.3 Complementation of *simC7* mutant 146
4.4 SimC7 is an NAD(P)H-dependent ketoreductase that acts on the angucyclic polyketide 148
4.5 The C-7 hydroxyl group is required for the antibiotic activity of simocyclinone..... 150
4.6 Discussion..... 154
4.7 Acknowledgements 158

Chapter 5 Substrate-Assisted Catalysis in Polyketide Reduction: Structural Analysis of the Ketoreductase SimC7 from Simocyclinone Biosynthesis (Schäfer et al., 2016) 159

5.1 Introduction 160
5.2 Optimisation of crystal conditions for SimC7 163
5.3 Overall structure of SimC7 164
5.4 Oligomeric state of SimC7 173
5.5 Structural homologs of SimC7..... 176
5.6 Cofactor binding..... 180
5.7 Substrate binding 181

5.8 A novel catalytic mechanism for an angucyclinone ketoreductase 185
5.9 Mutagenesis of the SimC7 active site 189
5.10 Discussion 192
5.11 Acknowledgements 193

Chapter 6 Alternative approaches to manipulate simocyclinone biosynthesis for optimised heterologous production 194

6.1 Introduction 195
6.2 Addition of the adsorptive hydrophobic resin HP20 195
6.3 Spontaneous resistant mutants of *S. coelicolor* M1152 198
6.4 Alternative heterologous hosts 202
6.5 Mutational analysis of the *sim* gene cluster in *S. antibioticus* Tü6040 203
6.6 Evidence that simocyclinone intermediates are the cause of the toxicity to *S. coelicolor* M1152, and that their target is not DNA gyrase 204
6.7 Genome sequencing reveals *S. antibioticus* Tü6040 to be a 'talented producer' of secondary metabolites 205
6.8 Acknowledgements 218

Chapter 7 General discussion 219

7.1 The role of SimC7 during angucyclinone biosynthesis 220
7.2 Tetraene biosynthesis 226
7.3 Ketoreductases involved in simocyclinone biosynthesis 230
7.4 Natural resistance mechanism against simocyclinones 231

References 234

Publications 257

Index of Tables, Figures, and Equations

Chapter 1

Table 1.1 Classification and activities of DNA topoisomerases. 22
Figure 1.1 Predicted heterotetrameric structure and domain architecture of DNA gyrase. 24
Figure 1.2 Model for negative supercoiling of DNA by DNA gyrase. 26
Figure 1.3 Structures of known aminocoumarin antibiotics. 29
Figure 1.4 Classification of simocyclinones according to the main structural components. 33
Figure 1.5 Proposed biosynthetic pathway for simocyclinone D8. 35
Figure 1.6 Genomic arrangement of the simocyclinone (*sim*) gene cluster in *S. antibioticus* Tü6040. 38
Figure 1.7 Comparison of DNA gyrase inhibition by novobiocin and SD8. 41
Figure 1.8 Schematic representation of the SD8 binding site at the DNA gyrase subunit A (GyrA) dimer interface. 42

Chapter 2

Table 2.1 DNA-modifying enzymes and DNA purification kits. 48
Table 2.2 Strains of *Escherichia coli* and their genomic characteristics. 49
Table 2.3 Strains of *Streptomyces* spp. and their genomic characteristics. 50
Table 2.4 Vectors and their genetic characteristics used in this study. 51
Table 2.5 Oligonucleotides used for PCR amplification of DNA fragments. 56
Table 2.6 Antibiotics and their applications as used in this study. 73
Table 2.7 Components used for PCR amplifications. 79
Table 2.8 PCR cycling conditions. 80

Figure 2.1 Modified strategy for PCR targeting. 84
Figure 2.2 Purification of recombinant SimC7 protein from *E. coli*. 89
Equation 1 Calculation of protein concentrations via Bradford assay. ... 90
Table 2.9 Composition of stacking gel (5%) and separating gel (12%) for SDS-PAGE. 92
Equation 2 Calculation of the maximum binding response (R_{max}). 98
Figure 2.3 SD8 standard curve for determination of product conversion by (A) SimC7 and (B) SimC7 variants. 100
Figure 2.4 Detection limit for determination of product conversion by (A) SimC7 and (B) SimC7 variants. 101
Table 2.10 List of software used for bioinformatics. 108

Chapter 3

Figure 3.1 Comparison of sequenced simocyclinone (*sim*) gene cluster fragments. 113
Figure 3.2 Growth of *Streptomyces* strains. 116
Figure 3.3 Growth and SD8 production of the heterologous host. 117

Chapter 4

Figure 4.1 Analysis of simocyclinones produced by the *simC7* deletion mutant. 124
Figure 4.2 Absorbance spectra of simocyclinones. 125
Figure 4.3 High resolution tandem mass spectra for SD8 and 7-oxo-SD8. 127
Figure 4.4 Numbering of carbon atoms in 7-oxo-SD8 for NMR spectroscopy. 128
Table 4.1 ^{1}H NMR data (700 MHz) for 7-oxo-SD8 in d_6-DMSO. 129
Table 4.2 ^{13}C NMR data (700 MHz) for 7-oxo-SD8 in d_6-DMSO. 130
Figure 4.5 Selected HMBC correlations for the angucyclic polyketide moiety of SD8 and of the new simocyclinone intermediate 7-oxo-SD8. 133
Figure 4.6 Differences in 2D NMR spectra of SD8 and 7-oxo-SD8. ... 134

Figure 4.7 Isotope patters of (A) purified 7-oxo-SD8 and (B) ^{13}C-enriched 7-oxo-SD8. 136
Figure 4.8 Structural validation of ^{13}C-labelled 7-oxo-SD8. 137
Figure 4.9 ^{1}H NMR spectrum of 7-oxo-SD8. 139
Figure 4.10 Correlated spectroscopy (COSY) spectrum of 7-oxo-SD8. 140
Figure 4.11 Total correlated spectroscopy (TOCSY) spectrum of 7-oxo-SD8. 141
Figure 4.12 Heteronuclear multiple bond correlation (HMBC) spectrum of 7-oxo-SD8. 142
Figure 4.13 Heteronuclear single quantum coherence (HSQC) spectrum of 7-oxo-SD8..... 143
Figure 4.14 Attached proton test (APT) ^{13}C-NMR spectrum of 7-oxo-SD8. 144
Figure 4.15 Attached proton test (APT) ^{13}C-NMR spectrum of 7-oxo-SD8 after ^{13}C-enrichment with [1,2-$^{13}C_2$]-acetate. 145
Figure 4.16 Complementation of *simC7* deletion mutant. 147
Figure 4.17 Ketoreductase activity of SimC7. 149
Figure 4.18 Inhibition of DNA gyrase by 7-oxo-SD8. 151
Figure 4.19 Chemical structures of SD8, 7-oxo-SD8 and MGD8N2A. .. 153
Figure 4.20 The angucyclinone binding pocket in GyrA. 157

Chapter 5

Figure 5.1 Comparison of proposed proton relay mechanisms for classic SDR ketoreductases. 162
Table 5.1 X-ray data collection and processing. 165
Table 5.2 Refinement of X-ray structures. 166
Table 5.3 Comparison of SimC7 structures. 167
Figure 5.2 Crystal structure of SimC7. 168
Figure 5.3 Structure-based sequence alignment of SimC7 with triphenylmethane reductase (TMR) from *Citrobacter* sp. and quinone oxidoreductase (QOR2) from *E. coli*. 170
Figure 5.4 Schematic representation of the highly hydrophobic substrate binding pocket of SimC7. 171

Figure 5.5 SimC7 has a very constricted active site. 172
Figure 5.6 Size exclusion chromatography of SimC7 in the presence and absence of cofactor. 174
Table 5.4 Interfaces of crystallographically observed dimers. 175
Table 5.5 Selected structural homologues of SimC7. 177
Figure 5.7 Comparison of the ternary complexes of SimC7 and LanV and the binary complex of QOR2 and respective reaction mechanisms..... 179
Figure 5.8 Simulated annealing omit electron density maps..... 183
Figure 5.9 Comparison of the 7-oxo-SD8 conformation in the ternary complex of SimC7 with SD8 conformations observed previously in other complexes..... 184
Figure 5.10 Proposed SimC7 reaction mechanism..... 187
Figure 5.11 Variations in the active site triad and substrates between SimC7, QOR2 and TMR. 188
Table 5.6 Enzymatic activities of SimC7 active site mutants. 190
Figure 5.12 Far UV circular dichroism (CD) spectra of inactive SimC7 mutants. 191

Chapter 6

Figure 6.1 Production of 7-oxo-SD8 in absence and presence of an absorptive hydrophobic resin. 197
Figure 6.2 Spontaneous mutants of *S. coelicolor* M1152 with increased resistance to SD8. 200
Table 6.1 Spontaneous mutant strains of *S. coelicolor* M1152 with increased SD8 resistance. 201
Figure 6.3 Purity and quality of isolated genomic DNA from *S. antibioticus* Tü6040. 207
Table 6.2 Summary of sequencing results for genomic DNA from *S. antibioticus* Tü6040. 208
Figure 6.4 Circular representation of the *S. antibioticus* Tü6040 genome. 212

Figure 6.5 Comparison of the SD8 binding site in the DNA gyrase subunit A (GyrA). 213
Table 6.3 Predicted biosynthetic gene clusters from *S. antibioticus* Tü6040. 214

Chapter 7

Figure 7.1 Putative biosynthetic pathway for the angucyclinone moiety of simocyclinone. 221
Figure 7.2 Potential angucyclin(on)e substrates for SimC7 and examples of angucyclin(on)es with 7-hydroxyl group. 224
Figure 7.3 Further potential SimC7 substrates with related structural motifs. 225
Figure 7.4 Potential biosynthetic pathway for the simocyclinone tetraene linker. 227

Publications (see Appendix)

Schäfer, M., T. B. Le, S. J. Hearnshaw, A. Maxwell, G. L. Challis, B. Wilkinson and M. J. Buttner (2015). "SimC7 Is a Novel NAD(P)H-Dependent Ketoreductase Essential for the Antibiotic Activity of the DNA Gyrase Inhibitor Simocyclinone." *J Mol Biol* **427**(12): 2192-2204.

Schäfer, M., C. E. M. Stevenson, D. M. Lawson, B. Wilkinson and M. J. Buttner (2016). "Substrate-Assisted Catalysis in Polyketide Reduction Proceeds via a Phenolate Intermediate." *Cell Chem Biol* **23**(9): 1019-1097

Statement of my work

This work was a collaboration project with Professor Gregory Challis, Department of Chemistry, University of Warwick, as well as Professor Anthony Maxwell, Dr. David Lawson, and Professor Barrie Wilkinson at the John Innes Centre. All experiments and data analyses have been performed by me unless otherwise stated here.

Chapter 3

For isolation of the simocyclinone biosynthetic gene cluster, a genomic library was generated and screened by Bio S&T (Montreal, Canada) using genomic DNA isolated from *Streptomyces antibioticus* Tü6040. PAC clone PAC-12I carrying the *sim* cluster was sequenced at the University of Cambridge.

Chapter 4

NMR spectra were recorded on a Bruker AVANCE-DRX 700 instrument by Dr. Ivan Prokes at the University of Warwick. SPR experiments were performed together with Stephen Hearnshaw and Clare Stevenson.

Chapter 5

Protein crystals for X-ray analysis were grown by me but harvested by Clare Stevenson. Protein structures were generated by me but final refinements before submission as well as Figures 5.2, 5.5, 5.7, 5.8 and 5.9 and Tables 5.1-5.5 were generated by David Lawson.

Chapter 6

The genome of *Streptomyces antibioticus* Tü6040 was sequenced at the Earlham Institute (formerly The Genome Analysis Centre, TGAC) (Norwich, UK).

Chapter 1
General introduction

1.1 Importance of antibiotics 20

1.2 DNA topoisomerases 21

1.2.1 DNA gyrase 23

1.3 Aminocoumarins 28

1.3.1 Biosynthesis of aminocoumarins 30

1.3.2 Resistance mechanisms against aminocoumarins 30

1.4 Simocyclinones 31

1.4.1 Biosynthesis of simocyclinones 34

1.4.2 Biological activity of simocyclinones 39

1.4.3 Resistance mechanism against simocyclinones 43

1.5 Aims of this project 44

1.1 Importance of antibiotics

Antibiotics are bioactive agents that inhibit cellular growth (bacteriostatic) or are toxic to cells (bactericidal). The name 'antimicrobials' has been commonly accepted as a synonym for antibiotics because they target mostly bacterial or fungal cells (Wright, 2007). Since the discovery of penicillin in 1928 (Fleming, 1929) and its introduction to the clinic in the 1940's (Chain *et al.*, 2005), antibiotics have become one of the most important agents for medical applications. Most antibiotic classes were discovered between 1945 and 1960, often referred to as the 'golden age' of antibiotic discovery. This trend reversed subsequently with the number of newly discovered compounds decreasing continuously (Watve *et al.*, 2001; Walsh, 2003; Wright, 2007). Since the introduction of the first antibiotics, bacterial strains have developed resistances against most clinically important antibiotics. Multiple antibiotic-resistant strains have emerged (often referred to as 'superbugs'), such as methicillin-resistant *Staphylococcus aureus* (MRSA) or vancomycin-resistant *Enterococci* (VRE). The introduction or excessive use of new antibiotics was generally followed by a wave of newly resistant bacteria (Chambers and Deleo, 2009). The time period from the introduction of an antibiotic until the first case of resistance varies but often takes less than a decade (Schmieder and Edwards, 2012). For example, the identification of MRSA was reported just one year after the introduction of methicillin in 1960 (Palumbi, 2001).

The development of new antibiotics has slowed down over the past 25 years. As an illustration, the Food and Drug Administration (FDA) approved 16 new antibiotics between 1983 and 1987, but only six between 2003 and 2007 (Spellberg *et al.*, 2008). The reason for this decline is not only reduced discovery of novel antibiotics by classic approaches, but also a strong increase in the economic costs associated with the long development process required for the production of new pharmaceuticals. The average development cost for a new antibiotic was estimated to be about $400-800 million in 2002 (DiMasi *et al.*, 2003). However, this should be measured against the costs arising from drug-resistant organisms, which probably

exceed $100 billion per year in the United States (Palumbi, 2001; Payne *et al.*, 2007). The lack of novel antibiotics and the continuous increase in multiply resistant bacteria are two major challenges that need to be addressed in the near future.

1.2 DNA topoisomerases

Topoisomerases are enzymes that control the topology (i.e. spatial structure) of DNA (Bates and Maxwell, 2007) by removing or inducing additional turns into DNA and by resolving mechanically interlocked DNA such as catenanes and knots (Schoeffler and Berger, 2008). Topoisomerases are divided into two functional groups depending on whether they transiently break one (type I) or both (type II) DNA strands during their catalytic cycle (Liu *et al.*, 1980). A summary of DNA topoisomerases and their specific functions can be found in **Table 1.1**. In most organisms DNA is negatively supercoiled. Supercoiling is not only important to compact the bacterial genome inside cells, it also plays an important role in transcription and DNA replication. The temporary separation of the DNA strands for the synthesis of DNA or RNA introduces additional supercoils next to the separated DNA segment. Topoisomerases induce compensatory positive supercoils ahead and negative supercoils behind the replication complex to relieve some of the stress in the DNA and thus maintain DNA topology around the DNA region that is unwound (Albert *et al.*, 1996; Rampakakis *et al.*, 2010).

Type II topoisomerases catalyse the interconversion of DNA topoisomers by making a transient double-stranded break in DNA and then transporting one DNA segment through another. Bacterial genomes encode two type II topoisomerases, DNA gyrase and topoisomerase IV (Topo IV), which are both heterotetramers. This is in contrast to eukaryotic type II topoisomerases, which are homodimers (Champoux, 2001). Except for an insertion of 170 amino acids near the C-terminus of GyrB, DNA gyrase and Topo IV share ~40% sequence identity. Both can relax supercoiled DNA, but

Table 1.1

Classification and activities of DNA topoisomerases.

Table adapted from Bush *et al.* (2015).

Topoisomerase	Type	Organisms	Subunit structure	Strand cleavage	Bond formed	Activities					
						Catenation/ decatenation	Knotting/ unknotting	Relaxation -ve [1)]	Relaxation +ve [1)]	Supercoiling -ve [1)]	Supercoiling +ve [1)]
Topo I	IA	Bacteria	Monomer	1	5’	Yes	Yes [2)]	Yes [2)]	No	No	No
	IB	Eukaryotes, Viruses	Monomer	1	3’	Yes	Yes [2)]	Yes [2)]	Yes	No	No
Topo II	IIA	Eukaryotes	Homo-dimer	2	5’	Yes	Yes	Yes	Yes	No	No
Topo III	IA	Bacteria, Eukaryotes	Monomer	1	5’	Yes	Yes [2)]	Yes	No	No	No
Topo IV	IIA	Bacteria	Hetero-tetramer	2	5’	Yes	Yes	Yes	Yes	No	No
Topo V	IB/ IC	Bacteria	Monomer	1	3’	unknown	unknown	Yes	Yes	No	No
Topo VI	IIB	Archaea	Hetero-tetramer	2	5’	Yes	unknown	Yes	Yes	No	No
DNA gyrase	IIA	Bacteria	Hetero-tetramer	2	5’	Yes [3)]	Yes	Yes	Yes	Yes	No
Reverse gyrase	IA	Bacteria, Archaea	Monomer	1	5’	No	No	Yes	No	No	Yes

[1)] -ve indicates negatively supercoiled DNA; +ve indicates positively supercoiled DNA.

[2)] Only if one substrate is nicked or single-stranded.

[3)] Decatenation activity of *E. coli* DNA gyrase is weak.

DNA gyrase is the only bacterial enzyme that can introduce negative supercoils into DNA (**Table 1.1**).

1.2.1 DNA gyrase

DNA gyrase is a type II DNA topoisomerase that introduces negative supercoils into DNA driven by hydrolysis of adenosine triphosphate (ATP) (Bates and Maxwell, 2007). The fact that DNA gyrase is essential in bacteria but absent from humans makes it a selective target for antimicrobials (Maxwell, 1997). This has made DNA gyrase one of the best-validated targets for antibacterial therapies (Heide, 2013).

DNA gyrase consists of two A subunits (GyrA; 97 kDa in *E. coli*) and two B subunits (GyrB; 90 kDa in *E. coli*), which form an A_2B_2 heterotetramer (Collin *et al.*, 2011). Structural studies on DNA gyrase have revealed a detailed picture of its molecular mechanism. Although there is currently no high-resolution structure of the entire enzyme complex due to its flexibility, the structure and function of various fragments of DNA gyrase have been solved separately (**Figure 1.1**). This revealed that GyrA is involved in interactions with DNA including binding and cleavage via an active-site tyrosine (Tyr122 in *E. coli*), whereas GyrB contains an ATPase domain required for the transient cleavage of DNA (**Figure 1.1**) (Collin *et al.*, 2011).

The current model for DNA supercoiling by DNA gyrase is known as the two-gate mechanism (**Figure 1.2**). During each cycle, two negative supercoils are introduced into the DNA substrate at the expense of hydrolysis of two ATP molecules (Bates and Maxwell, 2007). Three interfaces of DNA gyrase are important for conformational changes: The N-terminal domain of GyrB (N-gate), the gyrase-DNA interface (DNA-gate), and the C-terminal domain of GyrB (C-gate). First, the DNA-gate of the enzyme associates with a DNA segment, which is named the DNA gate segment (G-segment). The C-terminal domain of GyrA assists in winding the adjacent DNA around the enzyme in a right-handed supercoil of about 130 base pairs to position a second DNA segment (the DNA transfer segment or T-segment) in a positive

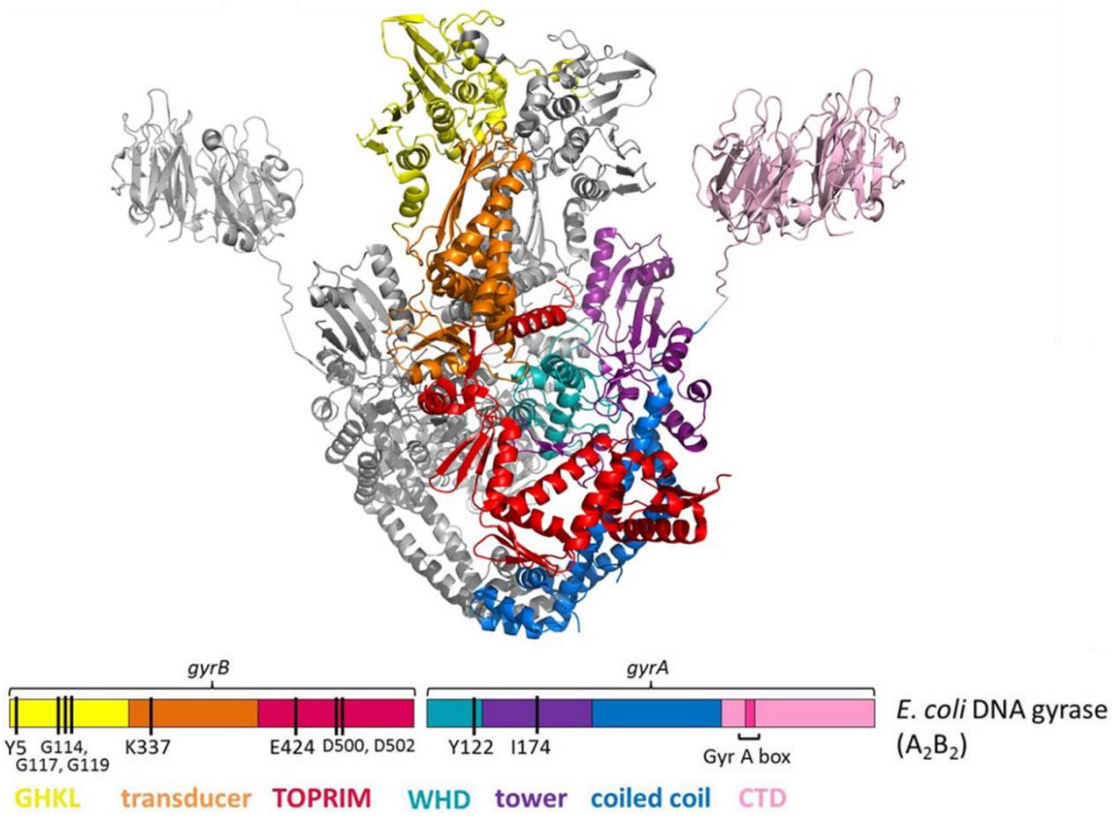


Figure 1.1

Predicted heterotetrameric structure and domain architecture of DNA gyrase.

The predicted overall structure of DNA gyrase was assembled from known structures of DNA gyrase fragments including a GyrB 43-kDa fragment (PDB accession number 1EI1), a GyrB-GyrA fusion construct (PDB accession number 3NUH), and the GyrA 35-kDa C-terminal domain (PDB accession number 3L6V). Black bars indicate catalytic residues, GHKL is the ATPase domain, TOPRIM is the topoisomerase/primase domain, WHD is the winged-helix domain and CTD is the C-terminal domain. Figure from Bush et al., 2015.

crossover above the G-segment and in between the N-gate of GyrB (**Figure 1.2**). Binding of ATP to GyrB closes the N-gate and consequently traps the T-segment to prepare for the passage of one DNA segment through the other. Cleavage of the G-segment covalently links the enzyme with the DNA substrate. The G-segment is transiently cleaved in a trans-esterification between the active site tyrosine (Tyr122) in both GyrA subunits and the phosphoryl groups in the backbone of both DNA strands to generate two DNA-phosphotyrosyl bonds. Hydrolysis of the first ATP induces a conformational change in DNA gyrase to transport the T-segment through the transiently opened G-segment and the C-gate of the enzyme. The G-segment is subsequently re-ligated and hydrolysis of the second ATP opens the N-gate to reset the enzyme for the next cycle (Nollmann *et al.*, 2007; Collin *et al.*, 2011).

DNA gyrase can be inhibited at various steps in its catalytic cycle, such as DNA binding, DNA cleavage, DNA strand passage, and ATP hydrolysis (**Figure 1.2**) (Collin *et al.*, 2011). Several classes of DNA gyrase-specific inhibitors have been discovered and described. The two most important classes of DNA gyrase inhibitors, namely the quinolones and the aminocoumarins, were discovered more than a decade before the enzyme itself was discovered in 1976 (Gellert *et al.*, 1976). The aminocoumarins will be discussed in the next chapter. The most clinically successful class of inhibitors are the quinolones, a class of synthetic antibiotics with potent activity against DNA gyrase. After the discovery of nalidixic acid in 1962 and its introduction for clinical use in 1967, the quinolones have been developed for more than five decades, giving rise to four generations of compounds with continuously improving physicochemical properties (Emmerson and Jones, 2003). The development of the fluoroquinolones in the 1970s gave rise to some of the most potent DNA gyrase inhibitors such as ciprofloxacin, with broad activity against Gram-negative bacteria. Fluoroquinolones stabilise the covalent DNA-protein complex, which leads to protein-stabilised breaks in the genome and, ultimately, cell death (Collin *et al.*, 2011). The irreversible inhibition of DNA gyrase by the fluoroquinolones explains their potency and clinical success compared to competitive and reversible inhibitors such as

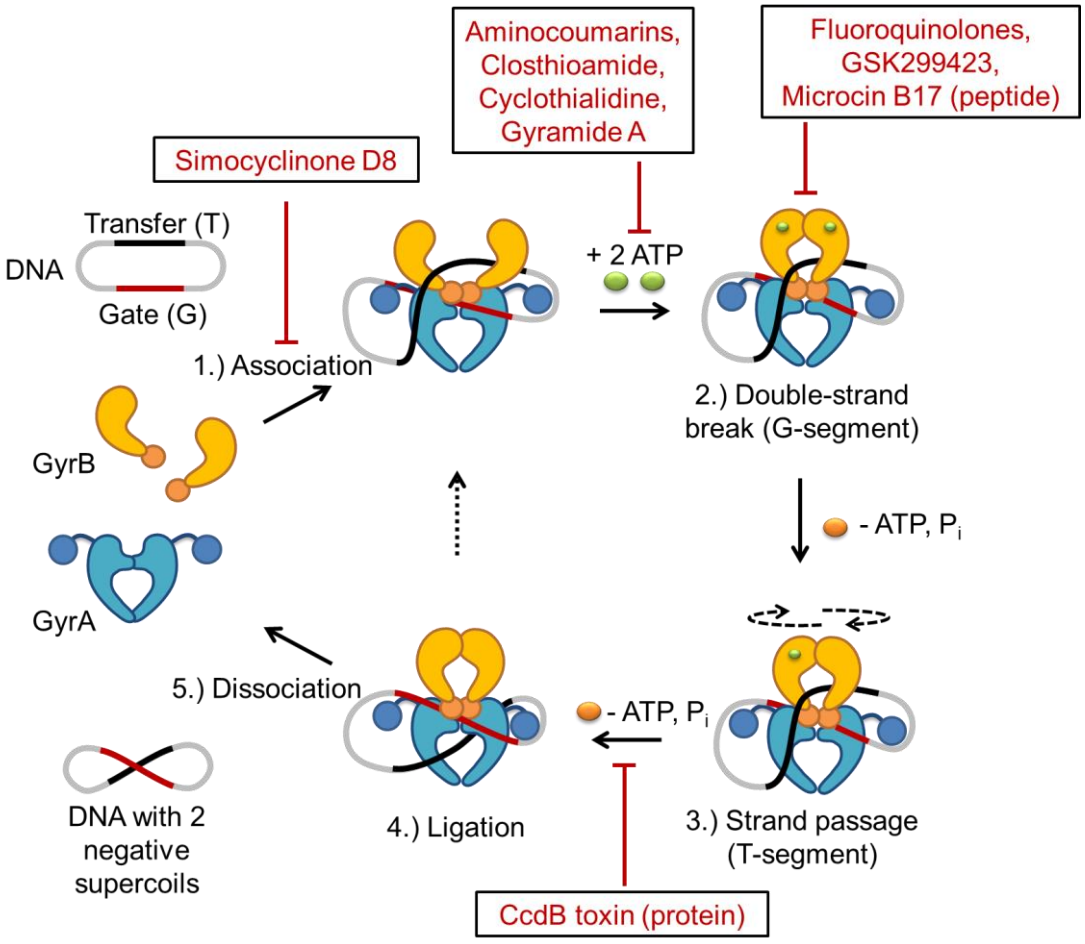


Figure 1.2

Model for negative supercoiling of DNA by DNA gyrase.

Association of the DNA gyrase subunits positions the DNA gate (G) segment above the GyrA dimer interface and the DNA transfer (T) segment in a positive crossover by wrapping the DNA around the C-terminal domain of GyrA. Binding of ATP captures the T-segment by the two GyrB subunits and the G-segment is transiently cleaved by covalently linking the 3'-phosphate of the DNA backbone to the active site tyrosine in GyrB. Hydrolysis of the first ATP initiates rotation of GyrB, which opens the DNA gate and allows transport of the T-segment through the cleaved G-segment. The G-segment is re-ligated and hydrolysis of the second ATP resets the enzyme. Inhibitors of DNA gyrase are indicated in red. The N- and C-terminal domain of GyrA are coloured in light and dark blue and the N- and C-terminal domain of GyrB are coloured in yellow and orange, respectively. The figure was modified from (Collin *et al.*, 2011; Ehmann and Lahiri, 2014).

the aminocoumarins. Fluoroquinolones are known to inhibit DNA gyrase in Gram-negative bacteria but their primary target in many pathogenic Gram-positive bacteria is topoisomerase IV (Topo IV) (Collin *et al.*, 2011). This is not surprising given the structural and sequence similarity between Topo IV and DNA gyrase. Recent crystal structures of the stabilised DNA-cleavage complex of gyrase from *S. aureus* (Bax *et al.*, 2010) and Topo IV from *Streptococcus pneumoniae* and *Acinetobacter baumannii* (Laponogov *et al.*, 2009; Wohlkonig *et al.*, 2010) have revealed that fluoroquinolones intercalate at the highly bent DNA gate. The binding site of fluoroquinolones is near the active site tyrosine and can shift slightly depending on the structure of fluoroquinolone. Although fluoroquinolones are extremely effective antibiotics, the continuous increase in bacterial resistances limits their application. Resistance against fluoroquinolones emerges from mutations in amino acid residues near the drug binding site (Laponogov *et al.*, 2009; Bax *et al.*, 2010; Wohlkonig *et al.*, 2010).

A number of other DNA gyrase inhibitors have been discovered in recent years in attempts to find alternative therapeutics. These include inhibitors that stabilise the DNA-protein intermediate (GSK299423 and the peptide Microcin B17), that inhibit ATP-hydrolysis by GyrB (closthioamide, cyclothialidine, gyramide A), or prevent DNA strand passage (CcdB protein) (**Figure 1.2**) (Goetschi *et al.*, 1993; Bahassi *et al.*, 1999; Parks *et al.*, 2007; Bax *et al.*, 2010; Foss *et al.*, 2011; Chiriac *et al.*, 2015).

1.3 Aminocoumarins

Aminocoumarins are potent inhibitors of bacterial DNA gyrase. The defining feature of the aminocoumarins is their characteristic structural core (3-amino-4,7-dihydroxycoumarin), which is usually decorated with various sugar-, acyl- or other groups (**Figure 1.3**). All the known aminocoumarins are produced by strains of the genus *Streptomyces*, although related genera have also been reported to produce coumermycin (*Actinoallomurus*) and simocyclinones (*Kitasatospora*) (Pozzi *et al.*, 2011; Bilyk *et al.*, 2016). The three 'classic' aminocoumarins are novobiocin, clorobiocin, and coumermycin A_1 (Steffensky *et al.*, 2000; Wang *et al.*, 2000; Pojer *et al.*, 2002). Their biosynthesis and mode of action has been studied in such detail that they have developed into a model system for generating new antibiotics by different approaches. These intensive genetic and biochemical studies have made novobiocin and clorobiocin two of the best studied secondary metabolites (Heide, 2013). All 'classic' aminocoumarins competitively inhibit GyrB by interfering with ATP hydrolysis because their binding site overlaps the ATP binding pocket in the ATPase domain of GyrB, as seen in the crystal structures of GyrB from *E. coli* (Holdgate *et al.*, 1997) and *Staphylococcus aureus* (Lu *et al.*, 2014) in complex with novobiocin (see below, **Figure 1.7**).

Beside the three 'classic' aminocoumarins, only two further aminocoumarins have been discovered, namely rubradirin and the simocyclinones, of which the latter will be discussed in the following section (**Figure 1.3**). In contrast to the above mentioned and well-studied 'classic' aminocoumarins, the biosynthetic pathways for the simocyclinones and rubradirin are covered by only a few experimental studies and remain largely hypothetical. Rubradirin not only lacks the deoxysugar at the 7-hydroxyl group of the aminocoumarin, which is typical for the 'classic' aminocoumarins, but it also has a completely different mode of action, inhibiting bacterial RNA polymerase instead of DNA gyrase (Kim *et al.*, 2008a).

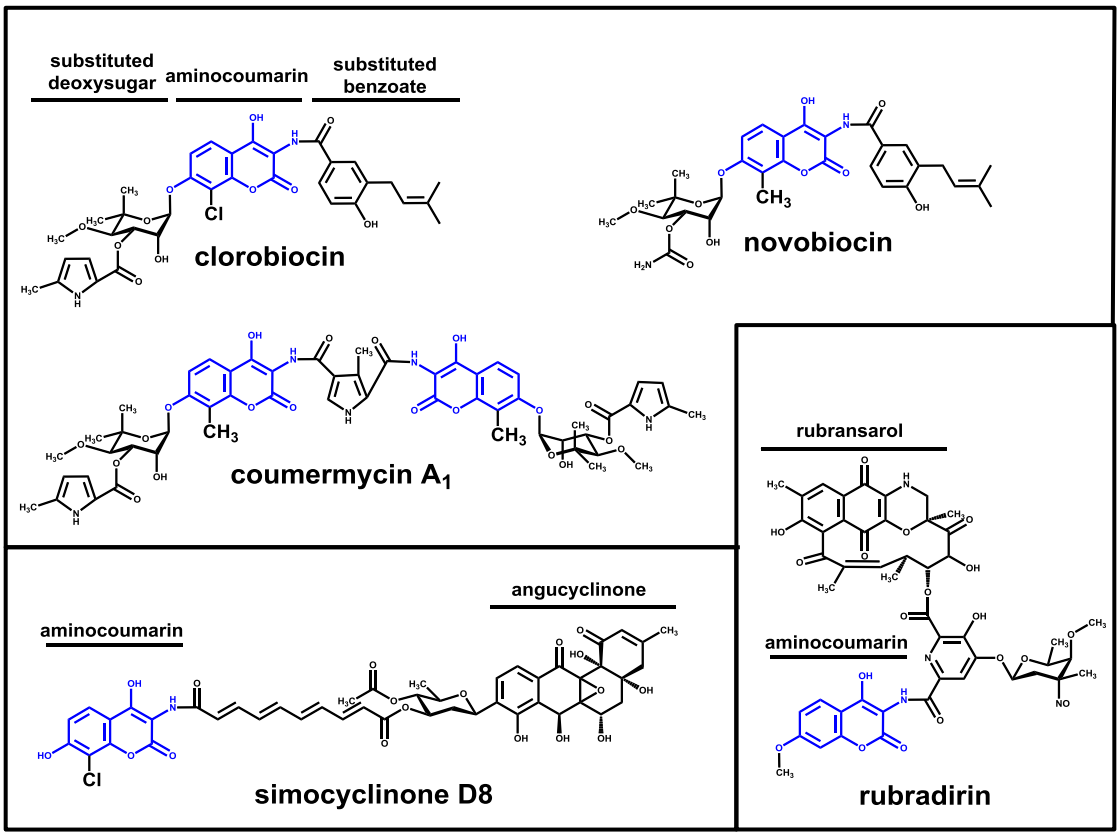


Figure 1.3

Structures of known aminocoumarin antibiotics.

Aminocoumarins are divided into 'classic' aminocoumarins (novobiocin, clorobiocin, and coumermycin A_1) and unusual aminocoumarins (simocyclinone D8 and rubradirin) based on their mode of action. Whereas the 'classic' aminocoumarins inhibit GyrB, simocyclinone inhibits GyrA by a different mode of action and rubradirin targets RNA polymerase. The aminocoumarin core (blue) is either chlorinated at position 8 (SD8 and clorobiocin), methylated (novobiocin and coumermycin A1) or unmodified (rubradirin).

1.3.1 Biosynthesis of aminocoumarins

Aminocoumarins are synthesised from L-tyrosine, which undergoes hydroxylation, dehydrogenation, and lactonisation in only a few biosynthetic steps. In novobiocin biosynthesis, NovH first activates L-tyrosine with ATP and then attaches it via a thioester to its peptidyl carrier protein (PCP) domain. The enzyme-linked tyrosine is subsequently β-hydroxylated by NovI, a cytochrome P_{450} monooxygenase, and then oxidised to β-keto-tyrosine by the heterotetrameric oxidoreductase NovJ/NovK. The last step in aminocoumarin biosynthesis, the cyclisation of the intermediate to a 3-amino-4,7-dihydroxycoumarin, remains speculative (Heide, 2009). It was initially assumed that the oxidoreductase NovC first introduces a hydroxyl group at position 2 of the aromatic ring and that this hydroxyl group makes a nucleophilic attack on the thioester bond to yield the final enol lactone of the B ring (Chen and Walsh, 2001). However, a later study revealed that *novC* was in fact not required for aminocoumarin biosynthesis and orthologs of *novC* are absent in other aminocoumarin gene clusters (Eustaquio *et al.*, 2005). The genes for novobiocin biosynthesis (*novHIJK*) are part of a single operon and their genomic arrangement is very similar to orthologs in the clorobiocin (*cloHIJK*) and coumermycin A_1 (*couHIJK*) biosynthetic gene clusters. The arrangement of the corresponding genes is slightly different in the simocyclinone (*simD6, simD1, simD2,* and *simD3*; previously named *simHIJ2K*) and the rubradirin (*rubC1-3*) biosynthetic gene clusters. An ortholog of *novK* (*simD3*) is located upstream of the remaining genes in the simocyclinone gene cluster but is absent from the rubradirin cluster, where most likely RubC3 catalyses the corresponding oxidation step (Heide, 2009).

1.3.2 Resistance mechanisms against aminocoumarins

Production of bioactive secondary metabolites requires effective resistance mechanisms to protect the host organism. Usually, specific resistance gene(s) are located within or immediately adjacent to the biosynthetic gene cluster. In the case of aminocoumarin antibiotics, the producing strains have

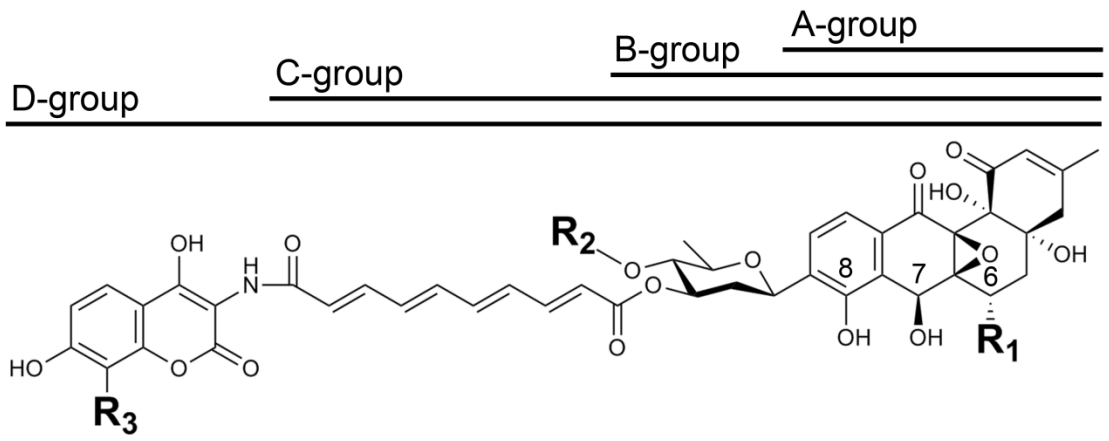
to protect their own essential DNA gyrase from inhibition. The 'classic' aminocoumarin gene clusters have resistance genes next to the biosynthetic genes. All *Streptomyces* species have the typical 'housekeeping' *gyrB*^{S} gene encoding a novobiocin-sensitive GyrB subunit. But in addition, for example, the novobiocin producer, *Streptomyces sphaeroides* NCIB 11891, has a second gene (*gyrB*^{R}) adjacent to the novobiocin biosynthetic gene cluster, which encodes a novobiocin-resistant subunit of DNA gyrase (Steffensky *et al.*, 2000). Resistance against novobiocin is mediated by increased expression of the resistant gyrase subunit, which is triggered by reduced negative supercoiling of the genome. The promoter of the novobiocin-resistant gyrase subunit responds to changes in DNA topology and is activated when negative supercoiling is reduced (Thiara and Cundliffe, 1989). Orthologs of a resistant GyrB subunit are also encoded by genes next to the clorobiocin and coumermycin A_1 biosynthetic gene clusters (Schmutz *et al.*, 2003). In addition, the clorobiocin and coumermycin A_1 clusters contain a second resistance gene (*parY*^{R}), which encodes an aminocoumarin-resistant subunit of Topo IV (Schmutz *et al.*, 2004). These observations imply that clorobiocin and coumermycin A_1 inhibit DNA gyrase and Topo IV, whereas novobiocin presumably targets only DNA gyrase.

1.4 Simocyclinones

Simocyclinones represent a new class of DNA gyrase inhibitors that are naturally produced by *Streptomyces antibioticus* Tü6040, a strain isolated from a soil sample collected in Iguaçu in Argentina as part of a search for strains producing novel secondary metabolites (Schimana *et al.*, 2000). Simocyclinone D8 (SD8) is a structurally unique hybrid antibiotic with two active ends, an angucyclinone and a chlorinated aminocoumarin, which are linked by a D-olivose deoxysugar and a tetraene linker (also referred to as an octatetraene dicarboxylate) (Schimana *et al.*, 2001; Holzenkampfer *et al.*, 2002). Thus, SD8 is an aminocoumarin, but also an angucyclinone, a class of antibiotics with more than 200 known representatives characterised by a

tetracyclic angular benz[*α*]anthracene ring system (Rohr and Thiericke, 1992; Kharel *et al.*, 2012).

About 14 different simocyclinones have been identified via nuclear magnetic resonance (NMR) or mass spectrometry (MS), which represent either intermediates or derivatives of SD8, the ultimate biosynthetic product (**Figure 1.4**). Simocyclinone production strongly depends on available carbon- and nitrogen-sources, which influence the ratio and composition of the simocyclinones produced (Schimana *et al.*, 2001). Production of SD8 is favoured by using glycerol and L-glutamine as the respective carbon and nitrogen sources in the culture medium. Production of up to 300 mg SD8 per litre was observed after about 6 days fermentation when supplemented with starch and L-glutamine. The remaining simocyclinone intermediates are only present in small amounts under these conditions. The composition of simocyclinone intermediates significantly changes when using mannitol and L-arginine as the respective carbon and nitrogen sources (Schimana *et al.*, 2001).



	A1	B1	B2	C2	C4	D4	D6	D7	D8
R_1	H	H	OH	OH	OH	OH	OH	H	OH
R_2	-	H	H	H	CO-CH_3	CO-CH_3	H	CO-CH_3	CO-CH_3
R_3	-	-	-	-	-	H	Cl	Cl	Cl

	A2	B4	C3	D2	D3	B3	C1	D1	D5
R_1	OH	OH	H	OH	H	H	H	H	H
R_2	-	CO-CH_3	CO-CH_3	H	CO-CH_3	CO-CH_3	H	H	H
R_3	-	-	-	H	H	-	-	H	Cl

Figure 1.4

Classification of simocyclinones according to the main structural components.

Four groups of simocyclinones have been classified (A to D) based on intermediates identified by mass spectrometry (yellow) or NMR (blue). Simocyclinone biosynthesis starts with the angucyclinone (aglycone, A-group), followed by attachment of the deoxysugar D-olivose (B-group), the tetraene linker (C-group), and the aminocoumarin (D-group). Hypothetical intermediates are shown in grey boxes. Figure adapted from (Theobald *et al.*, 2000).

1.4.1 Biosynthesis of simocyclinones

The simocyclinone (*sim*) biosynthetic gene cluster from *S. antibioticus* Tü6040 has been cloned and sequenced and a hypothetical pathway has been proposed (**Figure 1.5**) (Galm *et al.*, 2002; Trefzer *et al.*, 2002). The four structural elements of SD8 are synthesized separately and linked in a few enzymatic steps (**Figure 1.4** and **Figure 1.5**). SD8 is made from three primary metabolites (Galm *et al.*, 2002; Trefzer *et al.*, 2002), which explains the observed carbon and nitrogen source dependent variations in the composition of simocyclinones (Schimana *et al.*, 2001). Previous studies have shown that aminocoumarins are synthesised from L-tyrosine (Bunton *et al.*, 1963; Chen and Walsh, 2001) and D-olivose is derived from D-glucose (Wang *et al.*, 2011). Feeding experiments with stable isotopes of malonic acid revealed that the polyketide-derived angucyclinone and tetraene linker are made from acetate units although direct feeding of acetate itself was not possible because simocyclinone production was inhibited by this approach (Holzenkampfer *et al.*, 2002). Culturing *S. antibioticus* Tü6040 in an $^{18}O_2$-rich atmosphere showed that four oxygen atoms are incorporated by oxygenases from molecular oxygen (Holzenkampfer and Zeeck, 2002), consistent with a previously proposed cyclisation mechanism for novobiocin (Chen and Walsh, 2001). Cyclisation of the aminocoumarin precursor in novobiocin was predicted to arise from the introduction of a hydroxyl group at position 2 in 2,4-dihydroxy-β-keto-phenylalanine, which then attacks the thioester that links the activated aminocoumarin precursor to the peptidyl carrier protein (Chen and Walsh, 2001).

By the end of this thesis the *sim* gene cluster was sequenced three times. The three sources of the gene cluster are the original sequence from its discovery in 2002, the sequence from PAC-12I, which was subject to mutational analysis in this study (see **chapter 3.2**), and the sequence from the genome draft of *S. antibioticus* Tü6040 (see **chapter 6.7**). Comparison of the newly obtained sequences (i.e. PAC-12I and the *S. antibioticus* draft genome) with the previously published sequences of the *sim* gene cluster (Galm *et al*., 2002; Trefzer *et al.*, 2002) revealed no significant differences.

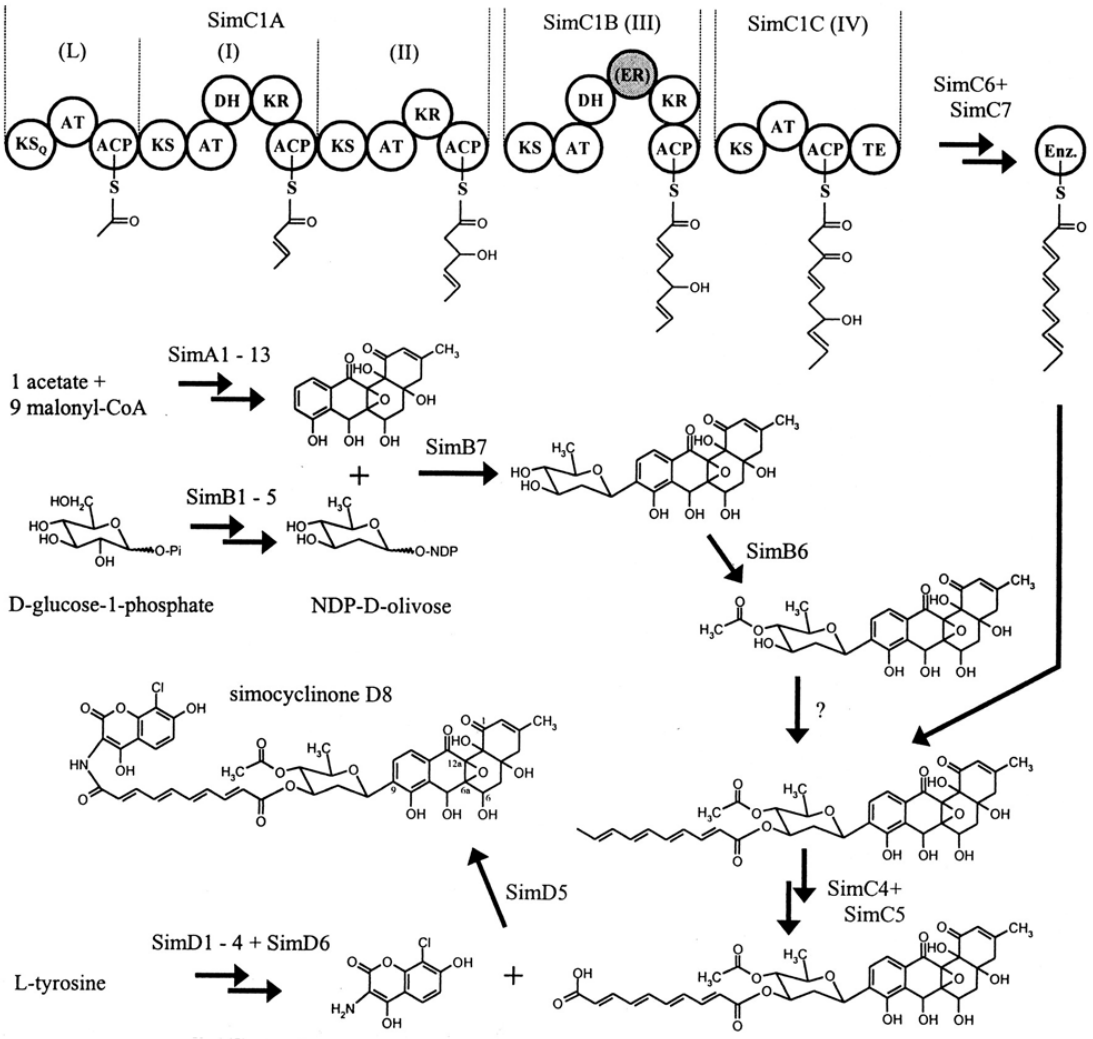


Figure 1.5

Proposed biosynthetic pathway for simocyclinone D8.

Protein domains: KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein; DH, dehydratase; KR, ketoreductase; ER, enoyl reductase; TE, thioesterase. Figure from Trefzer *et al.* (2002). KS_Q denotes an initiation factor domain similar to a KS domain with a conserved glutamine instead of an active site cysteine (Bisang *et al.*, 1999).

The proposed minimal *sim* gene cluster consists of 49 putative open reading frames, all transcribed in the same direction with exception of *simR* (**Figure 1.6**). It was proposed that 13 enzymes contribute to the biosynthesis of the angucyclinone (SimA1-13), the first detectable intermediate during simocyclinone biosynthesis. The polyketide skeleton is generated by a minimal PKS, including a ketosynthase (KS; SimA1), a ketosynthase-chain length factor (KS-CLF; SimA2) and an acyl carrier protein (ACP; SimA3). This minimal PKS resembles iterative type II PKS enzymes together with two cyclases (SimA4-5) and one ketoreductase (SimA6). Additional oxygenases (SimA7-8 and SimA13), ketoreductases (SimA9-10), and one decarboxylase (SimA12) were predicted to be involved in oxygenation, reduction and decarboxylation steps, respectively. Finally, the phosphopantetheinyl transferase SimA11 was predicted to modify the minimal PKS (Trefzer *et al.*, 2002).

The biosynthesis of the tetraene linker was predicted to involve 9 biosynthetic enzymes including the largest enzyme, the modular type I PKS (SimC1A-C). The arrangement of protein domains from SimC1A to SimC1C seems to be co-linear with the predicted pentaketide product, but note that *simC1A*, the gene encoding the first subunit of the modular PKS, is located downstream of *simC1B* and *simC1C* (**Figure 1.6**). Two additional enzymes (SimC6 and SimC7) were predicted to complement the missing ketoreductase and dehydratase functions of the type I PKS *in trans* to generate a system of four conjugated double bounds. SimC8 was proposed to be the phosphopantetheinyl transferase that post-translationally modifies the type I PKS (SimC1A-C). The octatetraene monocarboxylate intermediate would then be attached by an unknown enzyme to the 3'-hydroxyl group of D-olivose to form a carboxylic ester. A second terminal carboxylic acid would then be introduced into the tetraene linker by a dioxygenase (SimC4) and an aldehyde dehydrogenase (SimC5) (Trefzer *et al.*, 2002). The last step in simocyclinone biosynthesis is the assembly of the aminocoumarin. The genes involved (*simD1-6*) are conserved in all known aminocoumarin gene clusters, as mentioned earlier. The aminocoumarin is chlorinated at position 8 by the regio-specific halogenase SimD4 (Trefzer *et al.*, 2002).

Finally, the four structural units of SD8 are linked together in three enzymatic steps, carried out by a glycosyl-transferase (*simB7*), an unknown enzyme that connects the tetraene precursor to the deoxysugar, and an aminocoumarin ligase (SimD5, also named SimL) (Galm *et al.*, 2002; Trefzer *et al.*, 2002).

Despite the unique structure of SD8 and its complex biosynthesis, only two transcriptional regulators (SimR1 and SimR2) (Le *et al.*, 2009; Le *et al.*, 2011a; Le *et al.*, 2011b; Horbal *et al.*, 2012) and one biosynthetic enzyme (SimD5) have been characterized experimentally (Luft *et al.*, 2005; Pacholec *et al.*, 2005; Anderle *et al.*, 2007). SimD5 (SimL) is an aminocoumarin ligase that catalyses the formation of an amide bond between the aminocoumarin and the tetraene linker. SimD5 was a promising tool to develop and engineer novel antimicrobials. The narrow substrate specificity of SimD5 homologs (namely NovL, CloL, and CouL) for benzoic acid and pyrrole-carboxylic acid substrates limits the ability to generate novel aminocoumarin antibiotics (Galm *et al.*, 2004a; Galm *et al.*, 2004b). In contrast, SimD5 has a broad substrate specificity and accepts structurally diverse carboxylated acyl substrates, including polyunsaturated and saturated acyl chains and benzoic acid derivatives (Luft *et al.*, 2005; Pacholec *et al.*, 2005). Using different approaches, SimD5 turned out to be very efficient in product formation. The aminocoumarin ring from novobiocin, which is methylated at position 8, was successfully linked to simocyclinone C4, both chemo-enzymatically and through mutasynthesis (Luft *et al.*, 2005). In addition, a novel novobiocin analogue was generated *in vitro* using four enzymes. SimD5, NovM, NovP, and NovN catalysed the conversion of the aminocoumarin ring and TDP-L-noviose from novobiocin and decatetraenoic acid into a novel novobiocin analogue (Pacholec *et al.*, 2005). Finally, by combining SimD5, NovL and CouL, novel hybrid aminocoumarin antibiotics were created in a two-step mutasynthesis procedure. These analogues had antibacterial activity against Gram-positive bacteria but not against Gram-negative bacteria (*Pseudomonas aeruginosa*). The low activity of some of these analogues might be caused by a polar side chain or by the absence of a substituent at position 8 of the aminocoumarin ring system (Anderle *et al.*, 2007).

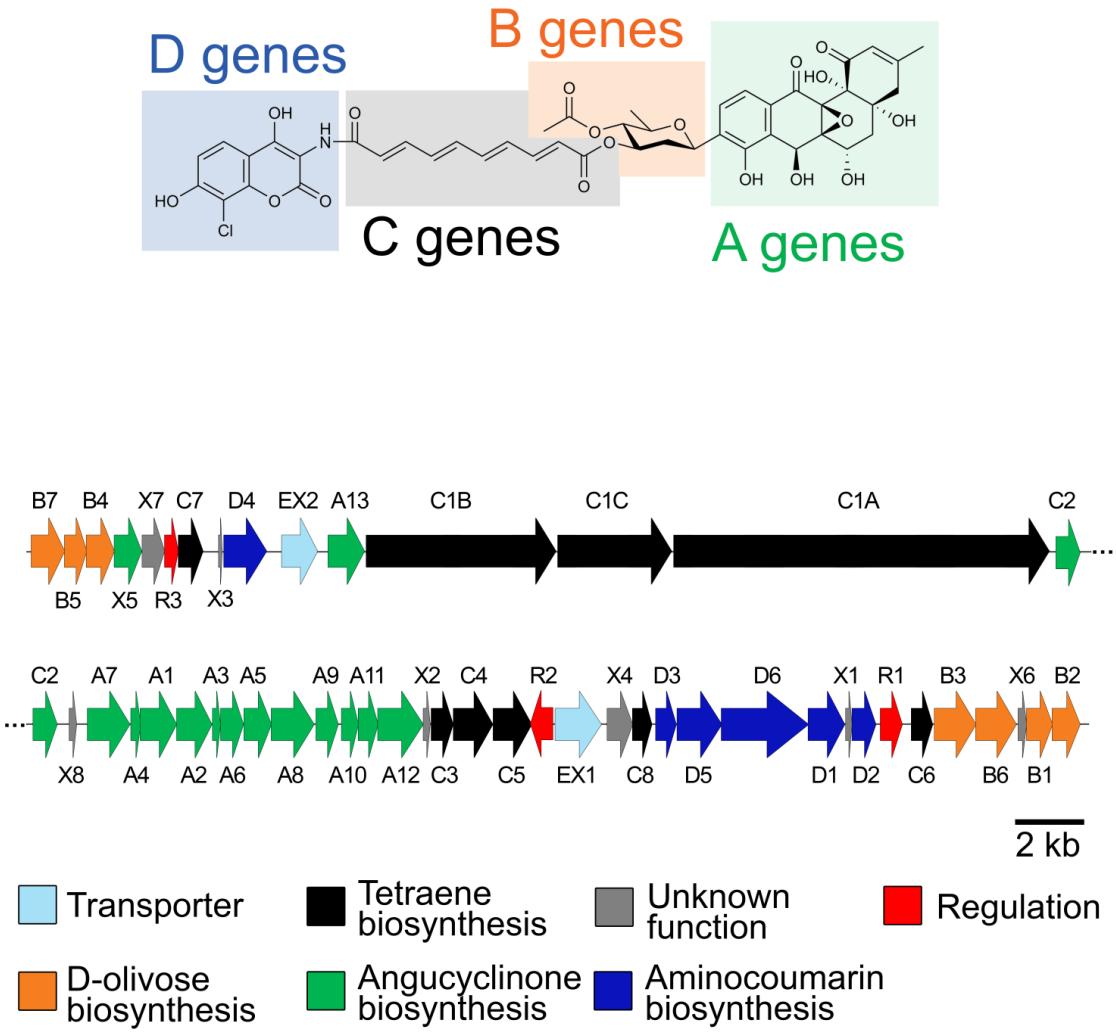


Figure 1.6

Genomic arrangement of the simocyclinone (*sim*) gene cluster in *S. antibioticus* Tü6040.

Genes are colour-coded according to their previously proposed involvement in the biosynthesis of the four structural units of simocyclinone D8 (see **Figure 1.5**) (Trefzer *et al.*, 2002).

1.4.2 Biological activity of simocyclinones

SD8 is active against Gram-positive bacteria and has cytostatic effects on different human tumour cell lines. In contrast, SD8 has no activity against Gram-negative bacteria, yeasts, or filamentous fungi at concentrations of up to 100 µg/mL (Schimana *et al.*, 2000). Although SD8 is a potent inhibitor of DNA gyrase, like the structurally and biosynthetically related 'classic' aminocoumarins, it has a completely different mode of action. SD8 is bifunctional, with an angucyclinone and an aminocoumarin at either end binding to two distinct pockets of the DNA-binding surface of GyrA (**Figure 1.7** and **Figure 1.8**) (Edwards *et al.*, 2009; Hearnshaw *et al.*, 2014). Biochemical and structural studies on the N-terminal domain of the *E. coli* GyrA revealed that SD8 binds to GyrA in a 1:1 stoichiometry (Flatman *et al.*, 2005). The interaction with GyrA shows positive cooperativity, with the high-affinity angucyclinone initiating binding, followed by binding of the low-affinity aminocoumarin, which might be facilitated by conformational changes in the protein after binding of the angucyclinone (configurational cooperativity) (Edwards *et al.*, 2011). Because of the bifunctional nature of SD8, removal of either the angucyclinone or the aminocoumarin decreases its potency as a DNA gyrase inhibitor by about 100-fold compared to SD8, which inhibits DNA supercoiling at sub-micromolar concentrations (IC_{50} ~ 0.1-0.6 µM) (Edwards *et al.*, 2009; Edwards *et al.*, 2011). This shows that both ends of SD8, and their covalent linkage, are important for effective inhibition of DNA gyrase.

Amino acid substitutions in GyrA that conferred resistance to SD8 were identified by direct mutagenesis based on the GyrA-SD8 crystal structure, and through the isolation and sequencing of spontaneous SD8-resistant mutants of *E. coli* (**Figure 1.8**) (Edwards *et al.*, 2009). Five amino acid substitutions that conferred resistance *in vivo* were identified: V44G, H45Y, H45Q, G81S, and D87Y. These mutations reduced the inhibitory activity of SD8 by up to two orders of magnitude (**Figure 1.8**). The substituted amino acid residues lie close to the binding site of SD8 seen in the crystal structure, confirming this site represents the true binding site of SD8 (Hearnshaw *et al.*, 2014). The binding sites of SD8 and ciprofloxacin on GyrA lie close to each

other, but the two antibiotics interact with different amino acid residues (Edwards *et al.*, 2009).

The potent inhibitory activity of SD8 against DNA gyrase is due to its interaction with GyrA. SD8 also binds weakly to the C-terminal domain of GyrB, but this is most likely a consequence of non-specific hydrophobic interactions and it has no effect on the ATPase activity of GyrB up to a 7:1 molar ratio of SD8 to enzyme (Flatman *et al.*, 2005; Sissi *et al.*, 2010). Despite the high structural similarity between DNA gyrase and Topo IV, SD8 is only a weak inhibitor of Topo IV with an IC_{50} of ~ 50 μM, compared to an IC_{50} of ~ 0.1-0.6 μM for DNA gyrase (Flatman *et al.*, 2005). Finally, SD8 has weak activity against eukaryotic topoisomerase II (IC_{50} ~ 5 μM), consistent with the inhibitory effect of SD8 on human tumour cell lines (Schimana *et al.*, 2000; Flatman *et al.*, 2005).

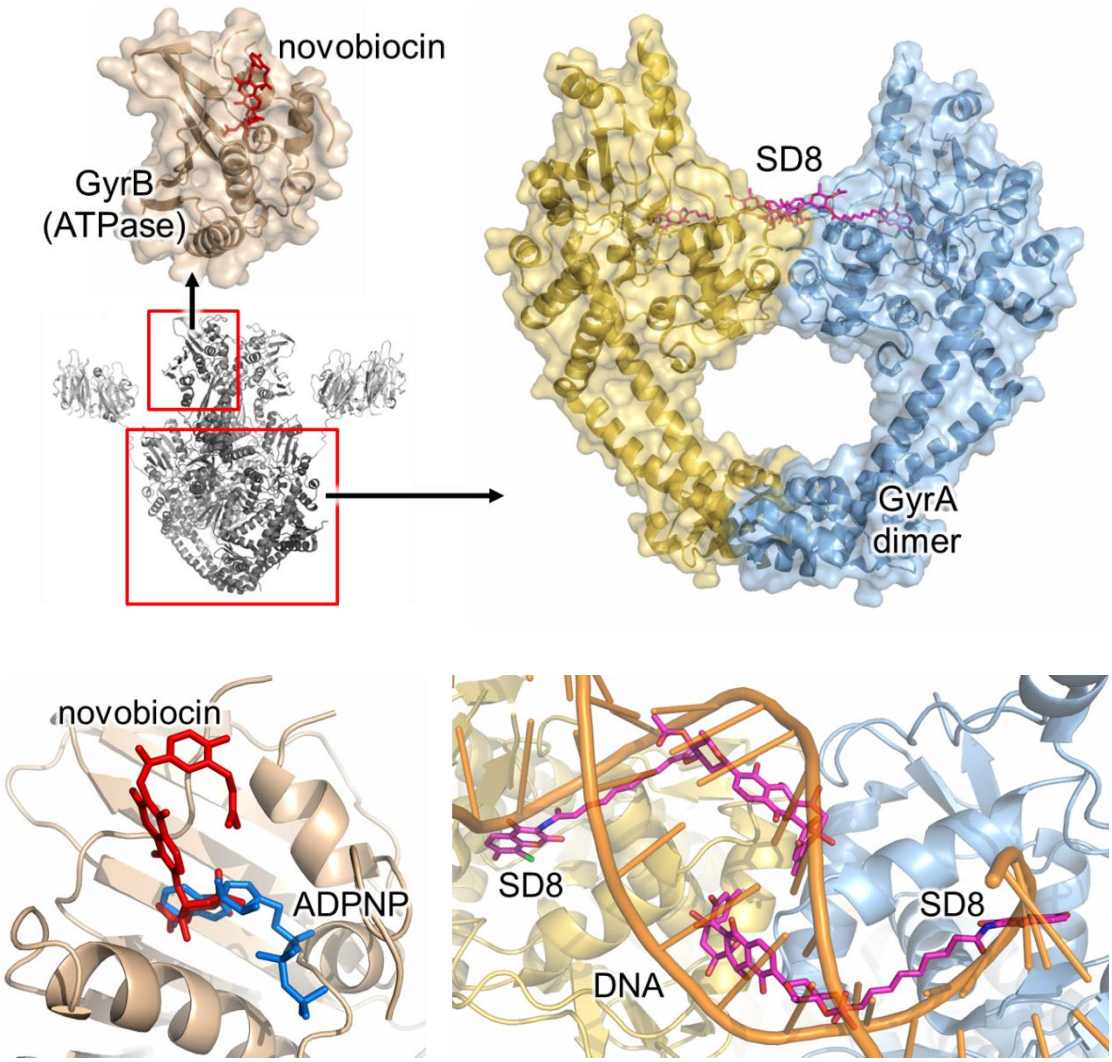


Figure 1.7

Comparison of DNA gyrase inhibition by novobiocin and SD8.

Novobiocin (red) binds to the ATPase domain of GyrB and interferes with ATP binding (blue), whereas SD8 (pink) interferes with binding of DNA (orange) to GyrA at the dimer interface. Subunits of the GyrA dimer are shown in yellow and blue. The close-up view below shows overlays with a non-hydrolysable ATP analog (ADPNP, phosphoaminophosphonic acid adenylate ester) from the ATP-GyrB complex and DNA from the DNA-GyrA complex (PDB accession numbers: 1AJ6, 1EI1, 4CKL, 2XCR).

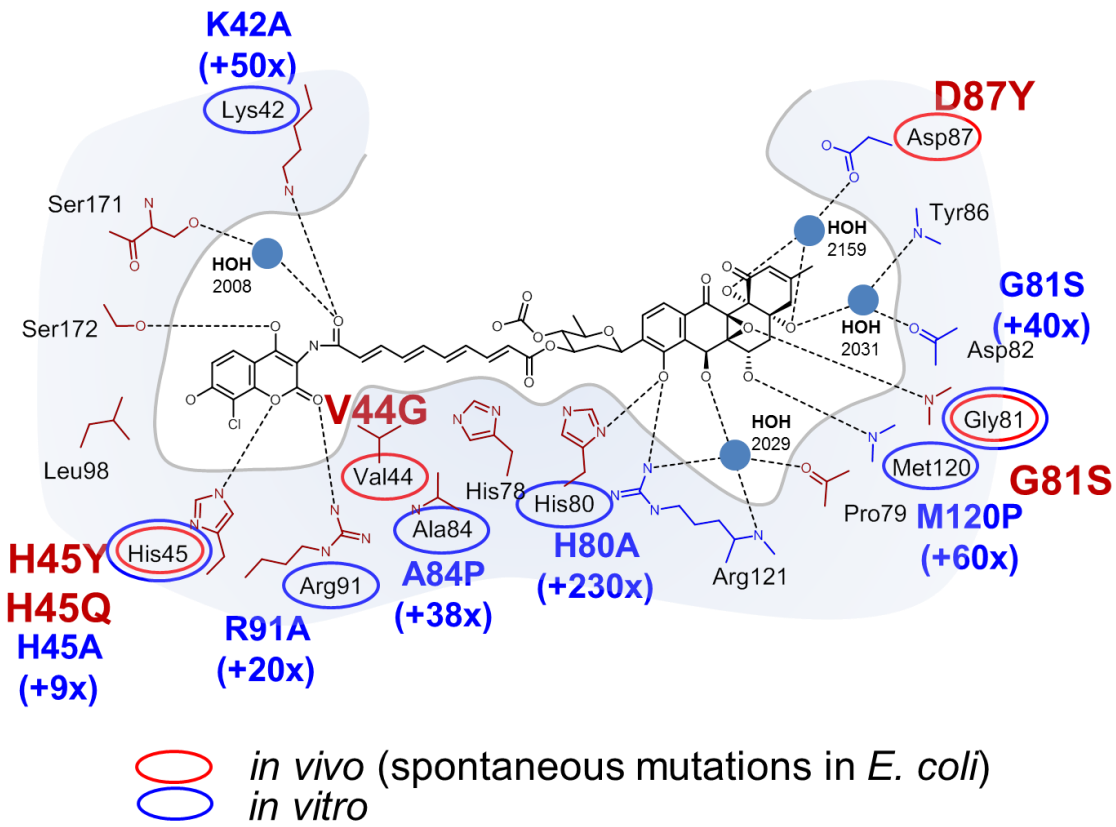


Figure 1.8

Schematic representation of the SD8 binding site at the DNA gyrase subunit A (GyrA) dimer interface.

Mutations are indicated for the *E. coli* DNA gyrase that occur spontaneously *in vivo* (red) or that were created and tested *in vitro* (blue) with fold-changes in the half-maximal inhibitory concentration (IC_{50}) indicated (Edwards *et al.*, 2009). Hydrogen bonds are shown as dashed lines and water molecules as spheres. Hydrogens have been omitted for clarity. Figure modified from Hearnshaw *et al.* (2014) (SD8-GyrA55 complex; PDB accession number: 4CKL).

1.4.3 Resistance mechanism against simocyclinones

In contrast to the biosynthetic gene clusters of 'classic' aminocoumarins, which contain genes encoding a resistant GyrB (and Topo IV) subunit, the simocyclinone cluster contains no resistance genes related to DNA gyrase or Topo IV. Instead, the *sim* cluster has two genes encoding putative transmembrane efflux pumps, *simEx1* and *simEx2* (Trefzer *et al.*, 2002). Heterologous expression of SimEx1 and SimEx2 in *Streptomyces lividans* showed that SimEx2 had no effect on resistance to SD8, whereas SimEx1 (also named SimX) increased resistance up to 30-fold when coupled to the strong consecutive *ermE** promoter (*ermEp**). The biosynthesis and export of simocyclinone are coupled through the action of a transcriptional repressor called SimR (Le *et al.*, 2009). Transcription of the gene encoding the SimX efflux pump is repressed by SimR in the absence of simocyclinone. However, simocyclinone binds to SimR and abolishes its ability to bind DNA and through this mechanism simocyclinone induces its own export (Le *et al.*, 2011a; Le *et al.*, 2011b).

1.5 Aims of this project

The complex biosynthesis of simocyclinones remains largely unexplored experimentally due to the genetic intractability of the natural producer *S. antibioticus* Tü6040. The aim of this thesis is to investigate simocyclinone biosynthesis in a heterologous system to facilitate mutagensis of the *sim* gene cluster and the analysis of the resulting simocyclinones. The principal focus of the thesis is the enzymatic function, structure and catalytic mechanism of SimC7.

In summary, the following keypoints are addressed in this thesis:

● Establishment of a heterologous system to facilitate mutagenesis of the *sim* gene cluster and analysis of the resulting simocyclinones, to investigate the simocyclinone biosynthetic pathway.

● Investigation of the true function of SimC7, a previously predicted dehydratase.

● Investigation of the catalytic mechanism of SimC7 by solving the crystal structure of the enzyme in complex with its cofactor and substrate.

Chapter 2
Materials and methods

2.1 Enzymes 48

2.2 Bacterial strains, plasmids, and oligonucleotides 49

2.2.1 Bacterial strains 49

2.2.2 Spontaneous resistant mutants of *S. coelicolor* M1152 50

2.2.3 Plasmids and PAC clones 51

2.2.4 Oligonucleotides 56

2.3 Culture media, buffers and solutions 65

2.3.1 Culture media 65

2.3.2 Buffers and solutions 69

2.3.3 Antibiotic stock concentrations 73

2.3.4 Minimum inhibitory concentration (MIC) assays 74

2.4 DNA manipulations *in vivo* (Genetic manipulations) 74

2.4.1 Preparation of electro-competent cells of *E. coli* 74

2.4.2 Transformation of *E. coli* 75

2.4.3 Conjugation between *E. coli* and *Streptomyces* spp. 75

2.5 DNA manipulations and *in vitro* (Cloning) 77

2.5.1 Isolation of plasmid DNA from *E. coli* 77

2.5.2 Isolation of chromosomal DNA from *Streptomyces* spp. 77

2.5.3 Agarose gel electrophoresis 78

2.5.4 Purification of size-separated DNA fragments 78

2.5.5 Polymerase chain reaction 79

2.5.6 Cloning of DNA fragments 81

2.5.7 PAC library construction 81

2.5.8 Deletion of genes via PCR targeting and complementation .. 82
2.5.9 Site-directed mutagenesis of SimC7 85
2.5.10 Recombineering with single-stranded DNA oligonucleotides 85
2.5.11 Clone library from genomic DNA 86
2.5.12 Sequencing 86
2.6 Expression and purification of recombinant proteins 87
2.6.1 Protein expression and purification (His-tagged) 87
2.6.2 Protein quantification via Bradford assay 90
2.6.3 Discontinuous polyacrylamide gene electrophoresis (SDS-PAGE) 90
2.6.4 Staining and immobilisation of proteins in SDS-PAGE gels .. 92
2.7 Protein crystallography 92
2.7.1 Protein crystallisation and cryoprotection 92
2.7.2 Selenomethionine labelling of recombinant protein 94
2.7.3 X-ray data collection, structure determination and structure refinement 94
2.8 Enzyme assays *in vitro* 96
2.8.1 DNA gyrase inhibition assays 96
2.8.2 Surface plasmon resonance (SPR) 97
2.8.3 Circular dichroism (CD) spectroscopy 98
2.8.4 SimC7 oligomeric state determination by size exclusion chromatography (SEC) 98
2.8.5 Ketoreductase activity assays 99
2.9 Extraction and analysis of simocyclinones 102
2.9.1 Extraction of simocyclinones 102
2.9.2 ^{13}C-enrichment of simocyclinones 102
2.9.3 Liquid chromatography-mass spectrometry (LC-MS) 103
2.9.4 High-pressure liquid chromatography (HPLC) 104

2.9.5 Normal phase liquid chromatography (LC) 105
2.9.6 Thin layer chromatography (TLC) 105
2.9.7 Small molecule size exclusion chromatography (SEC) 105
2.9.8 Nuclear magnetic resonance (NMR) spectroscopy 106
2.10 Software for *in silico* analysis 108

2.1 Enzymes

Commercially available enzymes, DNA purification kits, and size standards for DNA and protein work used in this study are listed in **Table 2.1**.

Table 2.1

DNA-modifying enzymes and DNA purification kits.

Supplier	Enzymes and kits
Roche,	*Taq* DNA polymerase
Burgess Hill, UK	*Pfu* DNA polymerase
	High Fidelity (HiFi) expand DNA polymerase
	dNTPs
	Lysozyme
	Restriction endonucleases
	T4 DNA Ligase
New England Biolabs,	Restriction endonucleases
Ipswich, UK	*Q5* High Fidelity (HiFi) DNA polymerase
	100 bp DNA ladder
	1 kb DNA ladder
	Prestained protein marker (7-175 kDa)
Qiagen,	*Taq* DNA polymerase
Hilden, Germany	Plasmid purification kit
	PCR/ Gel purification kit
Inspiralis Ltd., Norwich, UK	DNA gyrase (A_2B_2)

2.2 Bacterial strains, plasmids, and oligonucleotides

2.2.1 Bacterial strains

Bacterial organisms used in this study are listed in **Table 2.2** (*E. coli*) and **Table 2.3** (*Streptomyces* spp.).

Table 2.2

Strains of *Escherichia coli* and their genomic characteristics.

Strain	Characteristics	Reference
E. coli (B derivatives)		
Rosetta(DE3) pLysS	F^-, *ompT*, *gal*, *dcm*, $hsdS_B(rB^- mB^-)$, λ(DE3 [*lacI lacUV5-T7 gene1 ind1 sam7 nin5*]), pLysSRARE^6 (Cml^R)	Novagen (Merck, Nottingham, UK)
E. coli (K12 derivatives)		
BT340	*E. coli* DH5α carrying pCP20	(Datsenko and Wanner, 2000)
DH5α	F', *supE44*, *lacU169*, (*Φ80lacZΔM15*), Δ*hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*	(Hanahan, 1983)
DH10B	F^-, *araD139,* Δ(*ara, leu*)*7697,* Δ*lacX74, galU, galK, rpsL, deoR, Φ80dlacZ*Δ*M15, endA1, nupG, recA1, mcrA,* Δ(*mrr, hsdRMS, mcrBC*), *λ-*	(Grant *et al.*, 1990)
BW25113	Δ(*araD-araB*)*567*, Δ*lacZ4787*(*::rrnB-3)*, *lacIp-4000*(*lacI*^{Q}), *λ-*, *rpoS369*(*Am*), *rph-1*, *Δ*(*rhaD-rhaB*)*568*, *hsdR514*	(Datsenko and Wanner, 2000)
ET12567	F', *dam-13::Tn9* (Cml^R), *dcm-6*, *hsdM*, *hsdR*, *recF143, zij-202::Tn10* (Tet^R), *galK2*, *galT22*, *ara-14*, *lacY1*, *xyl-5*, *leuB6*, *thi-1*, *tonA31*, *rpsL136*, *hisG4*, *tsx-78*, *mtl-1*, *glnV44*	(MacNeil *et al.*, 1992)
TOP10	F^-, *mcrA*, Δ(*mrr-hsdRMSmcrBC*), *Φ80lacZ*Δ*M15*, Δ*lacX74*, *recA1*, *araD139*, Δ(*ara-leu*)*7697*, *galU*, *galK*, *rpsL*, *galE15*, *galK16, λ-endA1*, *nupG* (Str^R)	Invitrogen, Paisley, UK
NR698	MC4100 (F^- *araD139* Δ(*argF-lac*)*U169, rpsL150, relA1, flbB5301, deoC1, ptsF25, rbsR*), *imp4213*	(Ruiz *et al.*, 2005)
E. coli ATCC 25922™	wild type	ATCC, USA

Table 2.3

Strains of *Streptomyces* spp. and their genomic characteristics.

Strain	Characteristics	Reference
Streptomyces albus J1074 (R^- M^-)	*S. albus* G mutant defective in *Sal*GI-mediated restriction (R^-) and modification (M^-)	(Chater and Wilde, 1976; Zaburannyi *et al.*, 2014)
Streptomyces antibioticus Tü6040	Natural producer of simocyclinone D8	(Schimana *et al.*, 2000)
Streptomyces coelicolor M145	$SCP1^-$, $SCP2^-$ (wild type derivative without plasmids)	(Kieser *et al.*, 2000)
Streptomyces coelicolor M600	$SCP1^-$, $SCP2^-$ (wild type derivative without plasmids)	(Kieser *et al.*, 2000)
Streptomyces coelicolor M1152	$SCP1^-$, $SCP2^-$ derivative of *S. coelicolor* M145 with Δ*act*, Δ*red*, Δ*cpk*, Δ*cda*, *rpoB*[*C1298T*]	(Gomez-Escribano and Bibb, 2011)
Streptomyces coelicolor M1152ex1	*S. coelicolor* M1152 with genomically integrated pIJ10480 (*ermEp*-simEx1*)	this study
Streptomyces lividans 1326	wild type ($SLP2^+$, $SLP3^+$)	(Kessler *et al.*, 1989)
Streptomyces lividans TK24	$SLP2^-$, $SLP3^-$, *str-6* ($Strep^R$)	(Hopwood *et al.*, 1983; Ruckert *et al.*, 2015)
Streptomyces venezuelae ATCC 10712™	wild type, natural producer of chloramphenicol	ATCC, USA

2.2.2 Spontaneous resistant mutants of *S. coelicolor* M1152

Spores of *Streptomyces* strains were plated on SMMS agar (5×10^5 per plate) and paper discs (6 mm diameter; GE Healthcare, Amersham, UK) were placed on top of the distributed spores. Each paper disc was prepared with 20 µL DMSO containing different amounts of SD8. Bacterial growth was monitored for 6 days by keeping the plates upright at 30 °C. Putative SD8-resistant colonies of *S. coelicolor* M1152, which grew within the inhibition zone were picked and re-grown for single colonies. Liquid cultures (5 mL of YEME and 5 mL of TSB medium) of respective putative resistant colonies

were grown for 3-4 days and were screened again for increased SD8 resistance. The screening was carried out on SMMS medium with one paper disc per plate containing larger amounts of SD8 (30 µg or 60 µg per paper disc). Resistant strains, which grew without or minimal inhibition towards the paper disc, were selected and analysed for mutations in the GyrA subunit. Therefore, a region in the *gyrA* gene was sequenced corresponding to amino acid residues Leu18 to Ile187, which covers the known binding site of SD8 in the homologous GyrA from *E. coli* (Edwards et al., 2009; Hearnshaw et al., 2014).

2.2.3 Plasmids and PAC clones

Plasmids and PAC clones used in this study are listed in **Table 2.4**.

Table 2.4

Vectors and their genetic characteristics used in this study.

Vector	Characteristics	Reference
1K3	Cosmid pKC505 (SCP2 (fertility, replication), *ori* , *aac(3)IV* (Apr^R)) with ~25kb genomic region with part of the simocyclinone gene cluster from *S. antibioticus* Tü6040 (Apr^R)	(Trefzer *et al.*, 2002)
1K3neo	1K3 with *aac(3)IV* replaced by *neo* (Kan^R)	(T. Le, unpublished)
1K3neoΔC7::Apr	1K3neo with *simC7* replaced by Apr-cassette (Kan^R, Apr^R)	this study
pCP20	*FLP*+, *λ cI857+, λ* p_R *rep*^{ts}, *bla, cat* ($Carb^R$, Cml^R)	(Cherepanov and Wackernagel, 1995; Datsenko and Wanner, 2000)
PAC-12I	pESAC13 with 85 kb *Bam*HI genomic region of the *sim* gene cluster from *S. antibioticus* Tü6040 (Kan^R, $Thio^R$)	Bio S&T, Montreal, Canada
PAC-12IΔB6	PAC-12I with in frame deletion of *simB6*	this study

Vector	Characteristics	Reference
PAC-12IΔC1C-TE	PAC-12I with in frame deletion of the thioesterase domain of *simC1C, flanked with Sna*BI sites	this study
PAC-12IΔC2	PAC-12I with in frame deletion of *simC2, flanked with Pme*I sites	this study
PAC-12IΔC3	PAC-12I with in frame deletion of *simC3, flanked with Xba*I sites	this study
PAC-12IΔC4	PAC-12I with in frame deletion of *simC4*	this study
PAC-12IΔC5	PAC-12I with in frame deletion of *simC5*	this study
PAC-12IΔC6	PAC-12I with in frame deletion of *simC6*	(T. Le and O. Shorinola, unpublished)
PAC-12IΔC7	PAC-12I with in frame deletion of *simC7*	(T. Le and O. Shorinola, unpublished)
PAC-12IΔC1C	PAC-12I with in frame deletion of *simC1C*	this study
PAC-12IΔC1ABC	PAC-12I with in frame deletion of *simC1ABC*	(T. Le and O. Shorinola, unpublished)
PAC-12IΔD6	PAC-12I with in frame deletion of simD6, flanked by *Swa*I sites	this study
PAC-12IΔLeft	PAC-12I with in frame deletion of 'left' flanking site upstream to promoter of *simB7*, flanked with *Xba*I sites	this study
PAC-12IΔRight	PAC-12I with in frame deletion of 'right' flanking site downstream to *simB2*, flanked with *Pme*I sites	this study
PAC-12IΔX7	PAC-12I with in frame deletion of *simX7*	this study
PAC-12I∅sim	PAC-12I with ~85 kb *Bam*HI deletion of the whole *sim* gene cluster and flanking regions	this study
pESAC13	ΦC31, *attP-int, sacB, oriT*, P1 replicon (Kan^R, $Thio^R$) (derivative of pPAC-S1)	Bio S&T, Montreal, Canada (Sosio *et al.*, 2000)
pET15b	Overexpression vector; T7 promoter, T7 transcription start, His-tag coding sequence (N-terminal thrombin cleavage site), MCS (*Nde*I-*Bam*HI), T7 terminator, *lacI*, *rep* (pBR322), *bla* ($Carb^R$)	Novagen (Merck, Nottingham, UK)

Vector	Characteristics	Reference
pET15b-NB-C7	pET15b with 855 bp *Nde*I-*Bam*HI fragment of *simC7* (*ptipA::simC7*) ($Carb^R$)	this study
pET15b-NB-C7-S95A	SimC7 with point mutation S95A	this study
pET15b-NB-C7-S95T	SimC7 with point mutation S95T	this study
pET15b-NB-C7-S95C	SimC7 with point mutation S95C	this study
pET15b-NB-C7-I108A	SimC7 with point mutation I108A	this study
pET15b-NB-C7-I108D	SimC7 with point mutation I108D	this study
pET15b-NB-C7-H112A	SimC7 with point mutation H112A	this study
pET15b-NB-C7-H112N	SimC7 with point mutation H112N	this study
pET15b-NB-C7-H112Q	SimC7 with point mutation H112Q	this study
pET15b-NB-C7-H112K	SimC7 with point mutation H112K	this study
pET15b-NB-C7-H112R	SimC7 with point mutation H112R	this study
pET15b-NB-C7-E132D	SimC7 with point mutation E132D	this study
pET15b-NB-C7-N137A	SimC7 with point mutation N137A	this study
pET15b-NB-C7-N137Q	SimC7 with point mutation N137Q	this study
pET15b-NB-C7-Y155F	SimC7 with point mutation Y155F	this study
pET15b-NB-C7-Y229F	SimC7 with point mutation Y229F	this study
pGM1190	*tsr*, *aac(3)IV*, *oriT*, *to* terminator, *tipAp*, RBS, *fd* terminator (Apr^R)	(Muth *et al.*, 1989)
pGM1190-NB-C7	pGM1190 with 855 bp *Nde*I-*Bam*HI fragment of *simC7* (*ptipA::simC7*)	this study
pGM1190-pB7-NB-C7	pGM1190 carrying *simC7* expressed from the *simB7* promoter (*pB7::simC7*)	this study
pGM1190-Kan-oriT	pGM1190 with *aac(3)IV* replaced by *neo* gene (Kan^R)	(N. Holmes, unpublished)
pGM1190-Kan-oriT-H-ΔC7::Apr	pGM1190-Kan-oriT with ~5.2 kb fragment of *sim* gene cluster with *simC7* replaced by Apr-cassette	this study

Vector	Characteristics	Reference
pIJ773	pBluescript KS (+) with the apramycin resistance gene *aac(3)IV* and *oriT* from plasmid RP4, flanked by FRT sites (P1-FRT-*oriT-aac(3)IV*-FRT-P2), (Apr^R)	(Gust *et al.*, 2003)
pIJ773_del_oriT	pIJ773 without *oriT* (Apr^R)	T. Le., unpublished
pIJ775	P1-(*Swa*I, I-*Sce*I)-*oriT-aac(3)VI*-(I-*Sce*I, *Swa*I)-P2, *lacZ*, b*l*a (Apr^R, $Carb^R$)	(Gust *et al.*, 2004)
pIJ790	λ-RED (*gam*, *bet*, *exo*), *cat*, *araC*, *rep101*^{ts} (Cml^R) (modified lambda-RED recombination plasmid pKD20)	(Gust *et al.*, 2003)
pIJ10257	ΦBT1-based integrative vector containing the strong, constitutive *ermE** promoter (*ermEp**) (Hyg^R)	(Hong *et al.*, 2005)
pIJ10257-NH-C7	pIJ10257 with *simC7* cloned into *Nde*I-*Hind*III site (Hyg^R)	this study
pIJ10480	pIJ10257 with *ermEp*-simEx1* (Hyg^R)	(Le *et al.*, 2009)
pJET1.2/blunt	*rep* (pMB1), *eco47IR*, P_{lacUV5}, T7 promoter, MCS, *bla* (Cml^R) (blunt end insertion site)	Fisher Scientific, Loughborough, UK
pKC505	SCP2 (fertility, replication), *ori, aac(3)IV* (Apr^R)	(Richardson *et al.*, 1987)
pR9406	self-transmissible helper plasmid derived from pUB307 ($Carb^R$)	A. Siddique and D. Figurski, unpublished
pSET152	*rep* (pMB1), *int* fC31 (ΦC31 *att-int* locus from *E. coli*), *oriT* (RK2), *ori* (pUC18), *lacZα*, *aac(3)IV* (Apr^R)	(Bierman *et al.*, 1992)
pSKB2	T7 promoter, T7 transcription start, N- and C-terminal His-tag coding sequence, N-terminal prescission pxx cleavage site, MCS (*Nco*I-*Xho*I), T7 terminator, lacI, rep (pBR322) (Kan^R)	unpublished
pSKB2-NB-C7	N-terminal 6xHis tag fused to simC7; prescission protease cleavage site to remove N-terminal His-tag (Kan^R)	this study
pUB307	RP1 (= RK2) derivative with 5.5 kb deletion of Apr^R and Tn1*; Tc-R, (oriT, oriV, Tra1, Tra2, rep, IS21, Par/Mrs)* (Kan^R, Tet^R)	(Bennett *et al.*, 1977)

Vector	Characteristics	Reference
pUZ8002	*tra*, *neo*, RP4 (Kan^R)	(Paget *et al.*, 1999)
StH18	Supercos1 with ~42 kb chromosomal fragment from *S. coelicolor* (including *gyrBA* operon)	(Bentley *et al.*, 2002)
StH18-Apr-oriT	StH18 derivative containing an Apr^R (*aac(3)IV*) and oriT cassette (from pIJ755) instead of Kan^R (*neo*)	this study
Supercos1	*rep (pUC), bla, neo,* S40 promoter, two adjacent *cos* sites and recognition sequences, T7 promoter, T3 promoter ($Carb^R$, Kan^R)	Agilent Technologies, Walbronn, Germany

rep^{ts}, temperature-sensitive origin of replication
λ p_R *,* bacteriophage λ rightward promoter
cI857, gene encoding temperature-sensitive λ repressor from pMMC6
ΦBT1 integrates intragenically into *sco4848*
ΦC31 integrates intragenically into *sco3798*

2.2.4 Oligonucleotides

DNA oligonucleotides used in this study are listed in **Table 2.5**.

Table 2.5

Oligonucleotides used for PCR amplification of DNA fragments.

Intragenic regions (bold; start and stop codons in italics and underlined, overlapping start and stop codons of genes in red), recognition sites for restriction enzymes (upper case in bold and double underlined), and FRT-sites (upper case) or point mutations (upper case and underlined) are highlighted.

Purpose	Primer	Sequence
Extension PCR		
	P1 (FRT-site)	ATTCCGGGGATCCGTCGACC
	P2 (FRT-site)	TGTAGGCTGGAGCTGCTTC
Confirmation of positive PAC clones		
right wing	right_wing_B2-F	aacgagactgcccgtgtcttcacc
	right_wing_B2-R	aggaacccgatcaccgacaggatg
left wing	left_wing_B7-F	tcaccgacgtcagcctgttcatcg
	left_wing_B7-R	gttcatcgaccgtgtgcccttcttg
simC1A	PKS-C1A-F	caactgccacctcgtactcac
	PKS-C1A-R	gtacagcgtctcatccagcag
$Thio^R$ gene	thiostrepton-F	ttatcggttggccgcgagattcc
	thiostrepton-R	atgactgagttggacaccatcg

Purpose	Primer	Sequence
left wing	left_wing_B7-new-F	gtactcatgcgaatcctgttcgt
	left_wing_B7-new-R	cagacgtcgatgaagagctcc
right wing	right_wing_B2-new-F	gtgcacatctccactgacgag
	right_wing_B2-new-R	gagtaccacgagatcgtctcc
Confirmation of simEx1 in ΦBT1 integration site		
upstream	check-simEx1-upstrm-F	ttggtaggatcgtctagaacaggag
	check-simEx1-upstrm-R	aatcaccaccatgccgatgac
downstream	check-simEx1-dwnstrm-F	ctcggtggacttgacgaactg
	check-simEx1-dwnstrm-R	cctccaacgtcatctcgttctc
Confirmation of deletions from PAC-12I		
simB6	check_del_B6-II-F	gggaagtacttctccatcatcg
	check_del_B6-II-R	gacaacgggtagtagatcatcg
simC1ABC	check-scar-simC1ABC-F	gcatgtcgtcgtacgcggcatgcag
	check-scar-simC1ABC-R	gcgacctcgacccgtcgaaggcg
simC1C (TE)	check_del_C1CTE-F	caccgatgtcctcaactatcc
	check_del_C1CTE-R	cataccttccacactcaacct
simC1C	check_scar_C1C-F	tggacgagctctatgactttgtg
	check_scar_C1C-R	gcataccttccacactcaacct

Purpose	Primer	Sequence
simC2	check_del_C2-F	agaggatgaacagttgagtcgg
	check_del_C2-R	gtattcgctgaagatgcgttgg
simC3	check_del_C3-F	tctgtccttgaaggactgacca
	check_del_C3-R	ccgatgacatgggtgttggg
simC4	check_del_C4-F	cccacttctttcccttcctcc
	check_del_C4-R	ctcctccaggttctccaacag
simC5	check_del_C5-F	gagcgatctggtcattctgtcc
	check_del_C5-R	cagttcgtcaagtccaccgag
simC6	check_del_C6-F	caggcctgtccactggaaccttgc
	check_del_C6-R	gagggacgccgacctcgtgaagcg
simC7	check_del_C7-F	ggtgctggtcctcctgcccgaggtgc
	check_del_C7-R	tcccaccctgtacgccagaaaggtgc
simD6	check_del_D6-F	agcgagaccaacatcatcag
	check_del_D6-R	ggtgaagttcgtcttgtcct
left wing	check_scar_left-F	ccgtcttcgtcgacttcaacc
	check_scar_left-R	gtgtagaagtcctggatcgtgag
right wing	check_scar_right-F	gagttggtggacccatgactg
	check_scar_right-R	gagtggagtatccggttcctg
Confirmation of strain identity		
S. coelicolor	check_dagA-F	gtgatctcatcaagtggagtg
	check_dagA-R	tagtagatgttgttgatgctgg
S. antibioticus	check_Santi-F	tctccctctgaagtcgtatctg
	check_Santi-R	gtctacaccctcaccaacctc

Purpose	Primer	Sequence
Confirmation of simC7 flanking site in 1K3neo		
	Check-C7-1K3-F	gtaccagcgactcttcactc
	Check-C7-1K3-R1	acaaagtagacgggcatcag
	Check-C7-1K3-R2	gtcatagaggtcattggcgt
Deletions in PAC-12I (PCR targeting)		
simB6	del_simB6-F	gagcatgctgaccggcctgcacaccagctg***gtga*****cccac**ATTCCGGGGATCCGTCGACC
	del_simB6-R	gcgatggcacgaaaggcaggggtgccgtggtcaagt***tca***TGTAGGCTGGAGCTGCTTC
simC1ABC	del_simC1ABC_F	aacgacgatgacaagcttcggcagtatctgaagaag***gtg***ATTCCGGGGATCCGTCGACC
	del_simC1ABC_R	gactcaactgttcatcctctgaaccggatcgacgcg***tca***TGTAGGCTGGAGCTGCTTC
simC1C (TE domain)	del_C1CTE-F-SnaBI	tcccggctgacgacggggaccgccgggctgccgctg***gtg*****TACGTA**ATTCCGGGGATCCGTCGACC
	del_C1CTE-R-SnaBI	gaccgatgcgtaggaccggtagga***cta*****cagaacctcggc****TACGTA**TGTAGGCTGGAGCTGCTTC
simC1C	del_simC1C-F	cgcccttaccaagtcgatccctggaagcggtgccgt***atg***ATTCCGGGGATCCGTCGACC
	del_simC1C-R	gacccgtacggcgaccgatgcgtaggaccggtagga***cta***TGTAGGCTGGAGCTGCTTC
simC2	del_C2_F-PmeI	ccagttgtccgctggttcgtcggccggtcctgacca***gtg***c**GTTTAAAC**ATTCCGGGGATCCGTCGAC C
	del_C2_R-PmeI	tgccggtactcccccggcaggcgcagtatcggggaa**tca****GTTTAAAC**gTGTAGGCTGGAGCTGCTT C
simC3	del_C3_F-XbaI	ccgcctgtccctctgtccttgaaggactgaccactc***atg*****TCTAGA**ATTCCGGGGATCCGTCGACC
	del_C3_RII-XbaI [1)]	gacgcctggtgc***tca*****tcggcttgctccgctcgtcgtgct****TCTAGA**TGTAGGCTGGAGCTGCTTC
simC4	del_simC4_F	gacccggagcacgacgagcggagcaagccg***atg*****agcacc**ATTCCGGGGATCCGTCGACC
	del_simC4_R	ggagcgggtacggggtgcatggaacctaccgttcag***tca***TGTAGGCTGGAGCTGCTTC
simC5	del_simC5_F	ggatcgccgatccggcctgactgaacggtaggttcc***atg***ATTCCGGGGATCCGTCGACC
	del_simC5_R	tcggccgcatcggaacgtccagcactggct***tga*****ggtggc**TGTAGGCTGGAGCTGCTTC

Purpose	Primer	Sequence
simC6	del_simC6-F	ggatcgaccgggaatgacgatatccggagttgcgtg***atg***ATTCCGGGGATCCGTCGACC
	del_simC6-R	tgaacaggtcggcccgcggcaggccgggtcagaccg***tca***TGTAGGCTGGAGCTGCTTC
simC7	del_simC7-F	ccggcttcgcccagacgggatgagaggacctgaacg***atg***ATTCCGGGGATCCGTCGACC
	del_simC7-R	cgctgaacacccagggctgactgaacggcgctggga***cta***TGTAGGCTGGAGCTGCTTC
simD6	del_simD6_F-SwaI	tgcgcgcggacgcgatggcgcgtcgggaggccaccc***atg***c**ATTTAAAT**ATTCCGGGGATCCGTCGACC
	del_simD6_R-SwaI	gcgagagcatcggtcgcgggttcatcatcgtcccct***tca*****ATTTAAAT**gTGTAGGCTGGAGCTGCTTC
left flanking region	del_left_side-F-XbaI	gtcaggccgtcttcgtcgacttcaacccctcgccgggca**TCTAGA**ATTCCGGGGATCCGTCGACC
	del_left_side-R-XbaI	atg***tca***ggcccgcggctgatgcctgcccagaacgacgac**TCTAGA**TGTAGGCTGGAGCTGCTTC
right flanking region	del_right_side-F-PmeI	ttccacttccggggtgaggctgagcaaaggcgtggtcac**GTTTAAAC**ATTCCGGGGATCCGTCGACC
	del_right_side-R-PmeI	ggtaagcggttcggtgagtggagtatccggttcctgaac**GTTTAAAC**TGTAGGCTGGAGCTGCTTC
Cloning of DNA fragments		
simC7 gene	simC7-NdeI-F	GCA***cat*****ATG**aagattcttgtgaccggagc
	simC7-BamHI-R	CAG**AGGCTT*****cta***aatgctctgagctgcc
	simC7-HindIII-R	GCA**GGATCC*****cta***aatgctctgagctgcc
fragment with *simC7*	Cloning-C7scar-HindIII-F	GCG**AAGCTT**gggctcgaccggatgcagtatgacc
	Cloning-C7scar-HindIII-R	GCG**AAGCTT**cgcggtgatctcaccgaacgactgc
simB7 promoter	pB7-NdeI-F	tgaCATATGtcgttctgggcaggcatcag
	pB7-NdeI-R	gca**CAT**ATGtactccccattcagtggagg

Purpose	Primer	Sequence
Sequencing		
cosmid	Seq-C7-1-F	gggctcgaccggatgcagtatgac
fragment	Seq-C7-2-F	gcgatgttcgtggcgatcgg
with simC7	Seq-C7-3-F	cgcagtgggagagcgcctatcc
	Seq-C7-4-F	gtgcctgacggcccgatagc
	Seq-C7-5-F	gtttgtgactcaggtaatcg
	Seq-C7-6-F	ccgggaattcggctcgctcg
	Seq-C7-1-R	cgcggtgatctcaccgaacgactg
	Seq-C7-2-R	aagtcggcccgcgagatttgtttgg
	Seq-C7-3-R	tttcactaggttgcagcgcg
	Seq-C7-4-R	ctcgggcaggaggaccagca
	Seq-C7-5-R	gcgcagtcgaccaccgcgatcag
	Seq-C7-6-R	cgcctcccgcatcacctcga
Apr^R gene	Apr-F	atgtcatcagcggtggagtgcaatg
	Apr-R	tcagccaatcgactggcgag
gyrA fragment	check_SCO3873-F	ctcgagacggagatgcagcg
	check_SCO3873-R	gatgttggtcgccatgccga
pESAC13	pESAC13-F	cgagttgcatgataaagaagac
	pESAC13-R	cgaattgactagtgggtagg
pET15b insert	T7-promoter-F	taatacgactcactataggg
	T7-terminator-R	gctagttattgctcagcgg
pGM1190	pGM1190-F	ctgaggtcattactggaccg
	pGM1190-R	tcgcaattcctttagttgttcc

Purpose	Primer	Sequence
Recombineering of DNA gyrase in S. coelicolor		
	Recomb-GyrA-H45A-F (cac->gcc)	gctgccggacgtccgcgacggcctcaagcccgtg**GCC**cgccgcgtgctgtacgccatgtacgacggcggc
gyrA (*sco3873*)	Recomb-GyrA-H45Y-F (cac->tac)	gctgccggacgtccgcgacggcctcaagcccgtg**TAC**cgccgcgtgctgtacgccatgtacgacggcggc
	Recomb-GyrA-H45Q-F (cac->cag)	gctgccggacgtccgcgacggcctcaagcccgtg**CAG**cgccgcgtgctgtacgccatgtacgacggcggc
	Recomb-GyrA-H80A-F (cac->gcc)	cgtcggcgacgtcatgggcaactaccacccg**GCC**ggcgactcctccatctacgacgccctggtccgcctc
	Recomb-GyrA-G81S-F (ggc->tcg)	cgtcggcgacgtcatgggcaactaccacccgcac**TCG**gactcctccatctacgacgccctggtccgcctc
	RecGyrA-H80A-G81S-F (cacggc->gcctcg)	cgtcggcgacgtcatgggcaactaccacccg**GCCTCG**gactcctccatctacgacgccctggtccgcctc
Control for galK	Oligo-100 (tag->tat)	aagtcgcggtcggaaccgtattgcagcagctt**TAT**catctgccgctggacggcgcacaaatcgcgcttaa
***Confirmation of* mutations in DNA gyrase via recombineering**		
	Check-gyrA-H45-F	cggcctcaagcccgtgcac
	Check-gyrA-H80-F	cgacgccctggtccgcctc
	Check-gyrA-H45A-F	cggcctcaagcccgtg**GCC**
	Check-gyrA-H45Y-F	cggcctcaagcccgtg**TAC**
	Check-gyrA-H45Q-F	cggcctcaagcccgtg**CAG**
	Check-gyrA-H80A-F	ctacgacgccctggtc**GCC**
	Check-gyrA-G81S-F	cgacgccctggtccgc**TCG**

Purpose	Primer	Sequence
	Check-gyrA-H80A-G81S-F	cgacgccctggtc**GCCTCG**
***Targeting of neo* in StH18**		
	StH18-Apr-oriT-F	tcaagatctgatcaagagacaggatgaggatcgtttcgcATTCCGGGGATCCGTCGACC
	StH18-Apr-oriT-R	gcgtcgcttggtcggtcatttcgaaccccagagtcccgcTGTAGGCTGGAGCTGCTTC
SimC7 mutants		
S95A	S95Anew-F	tgctctccGCCatcgccgtgacctggc
S95C	S95C-F	tgctctccTGCatcgccgtgacctggc
S95T	S95Tnew-F	tgctctccACCatcgccgtgacctggc
S95A/T/C	S95AT-R	gcaccacatgccggaccc
I108A	I108A-F	gacccgGCGggccggatgcacctc
I108D	I108D-F	gacccgGATggccggatgcacctc
I108A/ D	I108A-R	ccggtccgcgtccggccaggtc
H112A	H112A-F	ggatgGCCctcgccgtcga
H112N	H112N-F	ggatgAACctcgccgtcgagcg
H112Q	H112Q-F	ggatgCAGctcgccgtcgagcg
H112K	H112K-F	ggatgAAActcgccgtcgagcg
H112R	H112R-F	ggatgCGTctcgccgtcga
H112A/N/Q/K/R	H112R-R	ggccgatcgggtcccgg
E132D	E132D-F	ggcccGATgcgctcgccac
	E132D-R	ggacgaaggtccagccaagcccg
N137A	N137Anew-F	gccaccGCCgcgctgggctgggctc
N137Q	N137Qnew-F	gccaccCAGgcgctgggctgggctc
N137A/ Q	N137Q-R	gagcgcctcgggccggac
Y155F	Y155F-F	ctgcgccTTCccgggcgc

Purpose	Primer	Sequence
	Y155F-R	cgcaccatgtcgccgccg
Y229F	Y229F-F	ggctcTTCccggccgaggtg
	Y229F-R	cctccagcacggcgcggg

[1)] Putative RBS of *simC4* (GGAG) is underlined.

2.3 Culture media, buffers and solutions

2.3.1 Culture media

Respective liquid media were prepared without addition of agar.

DNA medium (Difco nutrient agar)

Difco™ Nutrient Agar (4.6 g) was added to 200 mL distilled water in 250 mL Erlenmeyer flasks before autoclaving.

DNB medium (Difco nutrient broth)

Difco™ Nutrient Broth (4.6 g) was added to 200 mL distilled water in 250 mL Erlenmeyer flasks before autoclaving.

IMA agar (Instant Mashed Potato)

Potato Smash™	20 g/L	
LabM agar	20 g/L	

L agar (Lennox broth)

Agar	10 g/L	
Difco™ Bacto tryptone	10 g/L	
NaCl	5 g/L	85.6 mM
Glucose	1 g/L	5.6 mM

LB medium (Luria Bertani Broth)

Agar	10 g/L	
Difco™ Bacto tryptone	10 g/L	
Difco™ yeast extract	5 g/L	
NaCl	5 g/L	85.6 mM
Glucose	1 g/L	5.6 mM

MacConkey galactose indicator agar

Difco™ MacConkey agar	40 g/L	
D-(+)-galactose	10 g/L	55.5 mM

The medium was made fresh on the day of use.

Müller Hinton Broth (M-H Broth)

Beef infusion solids	2 g/L

MYM medium (Maltose yeast malt)

Difco™ Bacto Agar	18 g/L	
Maltose	4 g/L	11.7 mM
Yeast extract	4 g/L	
Malt extract	10 g/L	

Prior inoculation, the solidified medium was melted and trace element solution (2 mL/L) was added.

R2 medium (Okanishi *et al.*, 1974; Hopwood and Wright, 1978)

The medium was used to select and grow *Streptomyces* strains after conjugation.

Sucrose	103 g/L	300 mM
K_2SO_4	0.25 g/L	1.4 mM
$MgCl_2 \times 6H_2O$	10.1 g/L	49.8 mM
Glucose	10 g/L	5.6 mM
Difco™ casaminoacids	0.1 g/L	

Components for 1 litre medium were dissolved in 800 mL distilled water and the solution was autoclaved in 250 mL Erlenmeyer flasks (80 mL each) containing 2.2 g Difco Bacto agar. Prior inoculation, the solidified medium was melted and following autoclaved or sterile-filtrated solutions were added per flask:

KH_2PO_4 (0.5% w/v in dH_2O)	1 mL
$CaCl_2 \times 2H_2O$ (3.68% w/v)	8 mL
L-proline (20% w/v)	1.5 mL

TES buffer (5.73% w/v, pH 7.2)	10 mL
Trace element solution	0.2 mL
NaOH (1 M) (unsterile)	0.5 mL
Required growth factors for auxotrophs	0.75 mL

R5 medium (Kieser *et al.*, 2000)

Sucrose	103 g/L	300 mM
K_2SO_4	0.25 g/L	1.4 mM
$MgCl_2 \times 6H_2O$	10.1 g/L	49.8 mM
Glucose	10 g/L	5.6 mM
TES buffer	5.7 g/L	25 mM
Difco™ yeast extract	5 g/L	
Difco™ casaminoacids	0.1 g/L	
Trace element solution	2 mL/L	

The solution was autoclaved in 250 mL Erlenmeyer flasks (100 mL each) containing 2.2 g Difco Bacto agar. Prior inoculation, the solidified medium was melted and following autoclaved or sterile-filtrated solutions were added per flask:

KH_2PO_4 (0.5% w/v in dH_2O)	1 mL
$CaCl_2 \times 2H_2O$ (5 M)	0.4 mL
L-proline (20% w/v)	1.5 mL
NaOH (1 M) (unsterile)	0.7 mL
Required growth factors for auxotrophs	0.75 mL

SFM medium (Soya flour mannitol)

The medium was used to prepare spores of *Streptomyces* strains.

Agar	20 g/L	
Mannitol	20 g/L	109.8 mM
Soya flour	20 g/L	

SOC medium (Super optimal catabolite repression)

Tryptone	20 g/L

Yeast extract	5 g/L	
Glucose	3.6 g/L	20 mM
$MgSO_4$ x7H_2O	1.2 g/L	4.9 mM
$MgCl_2$ x6H_2O	952 mg/L	4.7 mM
NaCl	584 mg/L	10 mM
KCl	186 mg/L	2.5 mM

Supplemented minimal medium, solid (SMMS)

The medium was used for antibiotic assays with *Streptomyces* strains.

Agar	15 g/L	
Dicfo casaminoacids	2 g/L	
TES buffer	5.7 g/L	25 mM

The solution except agar was adjusted to pH 7.2 with 5N NaOH and 200 mL was poured into 250 mL Erlenmeyer flasks containing 3 g of Lab M agar. At time of use, following compounds were added:

NaH_2PO_4 (50 mM)	2 mL
K_2HPO_4 (50 mM)	2 mL
MgSO4 (1 M)	1 mL
Glucose (50% w/v)	3.6 mL
Trace element solution	0.2 mL
Growth factors for auxotrophs	1.5 mL

TSB (Tryptone soya broth)

Oxoid tryptone soya broth powder (CM129)	30 g/L

YEME (Yeast extract-malt extract medium)

Difco™ Bacto peptone	5 g/L	
Difco™ yeast extract	3 g/L	
Oxoid malt extract	3 g/L	
Glucose	10 g/L	55.5 mM
Sucrose	340 g/L	993.3 mM

After autoclaving, magnesium chloride was added:

$MgCl_2 \times 6\ H_2O$ (2.5 M)	1.02 g/ L	5 mM

2x YT medium (double concentrated yeast terrific broth medium)

Difco Bacto tryptone	16 g/L
Difco Bacto yeast extract	10 g/L
NaCl	5 g/L

2.3.2 Buffers and solutions

Gyrase assay dilution buffer (pH 7.5)

Tris/HCl (pH 7.5)	6.1 g/L	50 mM
KCl	7.5 g/L	100 mM
DTT	309 mg/L	2 mM
EDTA (250 mM, pH 8.0)	292 mg/L	1 mM
Glycerol	500 g/L	50% (v/v)

Gyrase assay relaxation buffer (pH 7.5)

Tris/HCl (pH 7.5)	4.2 g/L	35 mM
KCl	1.8 g/L	24 mM
$MgCl_2$ x6H_2O	813 mg/L	4 mM
DTT	309 mg/L	2 mM
Spermidine	261 mg/L	1.8 mM
Glycerol (50%)	65 g/L	6.5% (v/v)
Albumin	0.1 g/L	

Gyrase assay supercoiling buffer (pH 7.5)

Tris/HCl (pH 7.5)	4.2 g/L	35 mM
KCl	1.8 g/L	24 mM
$MgCl_2$ x6H_2O	813 mg/L	4 mM
DTT	309 mg/L	2 mM
Spermidine	261 mg/L	1.8 mM
ATP	507 mg/L	1 mM

Glycerol (50%)	65 g/L	6.5% (v/v)
Albumin	0.1 g/L	

Elution buffer I (protein extraction) (pH 8.0)

Tris/HCl (1 M, pH 8.0)	6.1 g/L	50 mM
NaCl (5 M)	14.6 g/L	250 mM
Glycerol (50%)	50 g/L	5% v/v
Imidazole (1 M, pH 8.0)	23.8 g/L	350 mM

Elution buffer III (protein extraction for crystallography) (pH 8.0)

HEPES (1 M, pH 8.0)	4.8 g/L	20 mM
NaCl (5 M)	2.9 g/L	50 mM
Glycerol (50%)	200 g/L	20% v/v
Imidazole (1 M, pH 8.0)	20.4 g/L	300 mM

HSB-EP+ (pH 7.4)

HEPES (pH 7.4)	2.4 g/L	10 mM
NaCl (5 M)	8.8 g/L	150 mM
EDTA (250 mM, pH 8.0)	877 g/L	3 mM
Surfactant P-20	0.5 mL/L	0.05% (v/v)

Loading dye for DNA samples (6-fold concentrated)

Bromophenol blue	2.5 g/L	3.7 mM
Xylene cyanol FF	2.5 g/L	4.6 mM
Glycerol (50%)	330 mL/L	33% (v/v)

Loading dye for protein samples (4-fold concentrated)

Glycerol	400 mL/L	40% (v/v)
Tris/HCl (1 M, pH 8.0)	29.1 g/L	240 mM
SDS (25%)	80 g/L	8% (w/v)
Bromophenol blue	0.4 g/L	0.04% (w/v)
β-mercaptoethanol	40 mL/L	4% (v/v)
EDTA (250 mM, pH 8.0	14.6 g/L	50 mM

Lysis buffer I (protein extraction) (pH 8.0)

NaCl (5 M)	8.8 g/L	150 mM
Triton X100	1 g/L	0.1% (w/v)
Lysozyme	1 g/L	0.1% (w/v)
Protease inhibitor (EDTA-free, Roche)		1 tablet per 10 mL

Lysis buffer IV (protein extraction for crystallography) (pH 8.0)

HEPES (1 M)	4.8 g/L	20 mM
NaCl (5 M)	2.9 g/L	50 mM
Glycerol (50%)	100 mL/L	10% (v/v)
Imidazole (1 M, pH 8.0)	681 mg/L	10 mM
Protease inhibitor (EDTA-free, Roche)		1 tablet per 10 mL

Running buffer (1x concentrated) (SDS-PAGE) (pH 8.4 adj. with HCl)

Tris/HCl (1 M, pH 8.0)	3 g/L	25 mM
Glycine	14.4 g/L	192 mM
SDS	1 g/L	0.1 % (w/v)

Running buffer III (pH 8.0)

HEPES (1 M)	4.8 g/L	20 mM
NaCl (5 M)	8.8 g/L	150 mM
Glycerol (50%)	50 g/L	5% (v/v)

SET buffer with lysozyme (pH 8.0)

Sucrose (103 g/L)	1.4 g/L	4.1 mM
EDTA (250 mM, pH 8.0)	7.3 g/L	25 mM
Tris/HCl (1 M, pH 8.0)	3 g/L	25 mM
Lysozyme (from chicken egg)	2 g/ L	0.2% (w/v)

Solution 2 (NaOH-SDS)

NaOH	2 g/L	50 mM
SDS	10 g/L	1% (w/v)

STEB buffer (pH 8.0)

Sucrose	400 g/L	1.17 M
Tris/HCl (1 M, pH 8.0)	12.1 g/L	100 mM
EDTA (250 mM, pH 8.0)	29.2 g/L	100 mM
Bromophenol blue	0.5 g/L	0.3 mM

Storage buffer I (protein extraction) (pH 8.0)

Tris/HCl (1 M, pH 8.0)	2.4 g/L	20 mM
NaCl (5 M)	8.8 g/L	150 mM
Glycerol (50%)	50 g/L	5% (v/v)
$MgCl_2$ x6H_2O	2 g/L	10 mM
EDTA (250 mM, pH 8.0)	584 mg/L	2 mM
DTT	154 mg/L	1 mM

Storage buffer X (protein extraction for crystallography) (pH 8.0)

HEPES (1 M, pH 8.0)	4.8 g/L	20 mM
Glycerol (50%)	100 g/L	10% (v/v)
TCEP (0.5 M, pH 7.0)	625 mg/L	5 mM

TBE buffer (Tris-Borate-EDTA buffer, pH 8.3)

Trizma base	10.8 g/L	89.2 mM
Boric acid	5.5 g/L	89 mM
EDTA	930 mg/L	3.2 mM

TE buffer (Tris-EDTA buffer, pH 8.0)

Tris/HCl (1 M, pH 8.0)	1.2 g/L	10 mM
EDTA (250 mM, pH 8.0)	292 mg/L	1 mM

Trace element solution

$ZnCl_2$	40 mg/L	0.29 mM
$FeCl_3 \times 6\ H_2O$	200 mg/L	0.74 mM
$CuCl_2 \times 2\ H_2O$	10 mg/L	0.06 mM

$MnCl_2 \times 4\ H_2O$	10 mg/L	0.05 mM
$Na_2B_4O_7 \times 10\ H_2O$	10 mg/L	0.03 mM
$(NH_4)_6Mo_7O_{24} \times 4\ H_2O$	10 mg/L	0.01 mM

Washing buffer I (protein extraction) (pH 8.0)

NaCl (5 M)	14.61 g/L	250 mM
Tris/HCl (1 M, pH 8.0)	6.1 g/L	50 mM
Glycerol (50% v/v)	50 mL/L	5% (v/v)

Washing buffer III (protein extraction for crystallography) (pH 8.0)

HEPES (1 M, pH 8.0)	4.8 g/L	20 mM
NaCl (5 M)	2.9 g/L	50 mM
Glycerol (50% v/v)	100 g/L	10% (v/v)
Imidazole (1 M, pH 8.0)	146.9 g/L	10 mM

2.3.3 Antibiotic stock concentrations

The antibiotics used in this study are listed in **Table 2.6**. For growth in liquid medium, the concentration of respective antibiotics was reduced by 50%.

Table 2.6

Antibiotics and their applications as used in this study.

Antibiotic	**Stock (mg/mL)**	**Final concentration (µg/mL)**	
		***Streptomyces* spp. DNA, MYM, R2, R5, SFM, Tu2**	***E. coli* L, LB**
Carbenicillin	100	-	100
Chloramphenicol	25 [1)]	-	25
Kanamycin	50	-	50
Nalidixic acid	50 [2)]	25	-
Apramycin	50	50	50
Hygromycin	40 (54)	50	25 [4)]
Thiostrepton	50 [3)]	50	50

[1)] in ethanol; [2)] in 0.3M NaOH; [3)] in DMSO; [4)] on DNA medium.

2.3.4 Minimum inhibitory concentration (MIC) assays

The minimum inhibitory concentration (MIC) was determined for simocyclinones against bacterial strains on microtiter plates as described previously (Wiegand *et al.*, 2008). Culture medium (200 µL) was applied on a flat-bottom 96-well microtiter plate (Fisher Scientific, Loughborough, UK) and was mixed with 1% (v/v) of HPLC-purified simocyclinones in DMSO to a final concentration of up to 100 µg/mL. Müller Hinton broth was inoculated to 5×10^5 colony forming units (CFU) per mL with an *E. coli* reference strain for antibiotic susceptibility testing (ATCC 25922™) or an *E. coli* strain with increased membrane permeability (NR689) and bacterial growth was analysed by measuring the OD_{600} after 16-20 hours at 37 °C with an EON™ microplate spectrophotometer (BioTek Instruments, Winooski, USA) and evaluated with the Gen5™ data analysis software. The assay was modified for *Streptomyces* species by inoculating SMMS medium with 1×10^5 spores per well, incubation for 6 days at 30 °C, and manual inspection of bacterial growth. The MIC was defined as the minimum concentration of a compound required for inhibiting bacterial growth.

2.4 DNA manipulations *in vivo* (Genetic manipulations)

2.4.1 Preparation of electro-competent cells of *E. coli*

A single colony or glycerol stock of the respective *E. coli* strain was used to inoculate 10 mL of LB medium with respective antibiotics. Pre-cultures were grown overnight at 37 °C while shaking (250 rpm) and 1% inoculum was used for 50 mL of LB medium. The main culture was grown at 37 °C while shaking to an OD_{600} of 0.3-0.6. Cultures with temperature-sensitive plasmids were incubated at 30 °C throughout. The cultures were cooled on ice for 10-30 min and were then continuously kept on ice. The cells were harvested and washed twice by repeated centrifugation (5 min at 1,825*g* and 4 °C) and carefully dissolved in 10 mL of ice-cold 10% (v/v) glycerol. After a final

centrifugation (5 min at 1,825*g* and 4 °C), the cells were dissolved in about 500 µL of the remaining liquid.

2.4.2 Transformation of *E. coli*

Cultures of *E. coli* were transformed with isolated vector DNA or PCR-amplified fragments of DNA via heat-shock or electroporation. For heat-shock transformation, respective isolated DNA was added to 100 µL of chemically competent cells. The cells were incubated on ice for 30 min, heat-shocked for 45 seconds, and kept on ice for 2 min before addition of 800 µL of SOC medium. For electroporation, cultures were made electro-competent as described (see **chapter 2.4.1**) and 200-300 µL of competent cells were carefully mixed with 2-5 µL of DNA solution (about 300 ng purified PCR product or 2-5 µg total nucleic acids from DNA isolation). The cell solution was transferred to pre-cooled 2 mm electroporation cuvettes (Isogen Life Science, De Meern, Netherlands) and electroporation was carried out using a Bio-Rad Gene Pulser I (Bio-Rad, Hertfordshire, UK) with a resistance of 200 Ω, a capacitance of 25 µF, an electrical potential of 2.5 kV, and an expected time constant of 4.5-4.9 ms. Immediately after electro-transformation, 750 µL of liquid SOC medium was added. All cultures were allowed to recover for 1 hour at 37 °C or 2 hours at 30 °C while shaking (250 rpm) before selection of transformed cells on LB agar medium containing respective antibiotics. Transformed cells were transferred onto new LB agar medium with respective antibiotics and clones were verified via PCR and DNA restriction digests.

2.4.3 Conjugation between *E. coli* and *Streptomyces* spp.

Because some *Streptomyces* have evolved a highly efficient system to detect and degrade foreign methylated DNA, a methylation-deficient *E. coli* strain was used to introduce DNA into *Streptomyces* strains.

Respective DNA constructs were first conjugated from *E. coli* DH5α or DH10B (donor) to the methylation-deficient *E. coli* strain ET12567 (recipient) via triparental conjugation using *E. coli* TOP10/pR9406 (self-transmissible helper plasmid). Cultures were grown shaking at 37 °C in 50 mL LB medium with respective antibiotics till OD_{600} ~0.6 using 1% inoculum from a 10 mL overnight pre-culture. Cultures were washed twice with 10 mL LB medium to remove antibiotics from the medium and were dissolved in 0.5 mL LB medium. The cultures were mixed by adding 10 µL of each culture on the same spot on LB agar. After triparental mating overnight at 37 °C, the mixed cultures were collected and transferred on LB agar with respective antibiotics to select for exconjugants of *E. coli* ET12567 carrying the respective vector construct and the helper plasmid. Successful conjugational transfer of DNA was confirmed via PCR.

One single exconjugant colony of *E. coli* ET12567 with helper plasmid and confirmed DNA construct was grown overnight in 10 mL LB medium with respective antibiotics and 1% inoculum was used to grow 50 mL of main culture till OD_{600} ~0.6. Cells were harvested by centrifugation (5 min at 1,026*g*), washed twice with 10 mL of LB to remove antibiotics from the medium, and dissolved in 0.5 mL LB medium. Meanwhile, about 1×10^7 spores were heat-shocked for 10 min at 50 °C in 10 µL dH_2O. The spores were dissolved in 0.5 mL 2xYT medium and were mixed with the washed *E. coli* colony. Cells and spores were pelleted by centrifugation (1 min at 13,200*g*) and dissolved in 100 µL remaining liquid. A 10-fold dilution series was prepared with 100 µL of dH_2O and plated on R2-S medium for conjugation of PAC clones or SFM medium with 10 mM $MgCl_2$. After 16-22 hours incubation at 30 °C, the plates were overlaid with 1 mL dH_2O containing nalidixic acid and respective antibiotics to inhibit growth of *E. coli* colonies and to select for the integrated DNA construct, respectively. Single *Streptomyces* exconjugants were transferred at least twice on new medium with nalidixic acid and respective antibiotics to avoid contamination with *E. coli* colonies. Exconjugants were confirmed via PCR for presence of respective DNA constructs.

2.5 DNA manipulations and *in vitro* (Cloning)

2.5.1 Isolation of plasmid DNA from *E. coli*

Phage artificial chromosome (PAC) derivatives were isolated from *E. coli* by growing single colonies overnight at 37 °C in 10 mL LB medium in closed 25 mL glass flasks. Cells were harvested by centrifugation (5 min at 1,026*g*), resuspended in 300 µL of solution I (50 mM Tris/HCl; 10 mM EDTA; pH 8.0). Cells were lysed and proteins were denatured by immediately adding 450 µL of solution II (200 mM NaOH; 1% SDS) and mixing the solution by inverting the tubes 10 times. The alkaline lysis was immediately stopped by addition of 450 µL solution III (3 M potassium acetate; pH 5.5) and inverting the tubes 10 times. Lysed cells and denatured proteins were pelleted by centrifugation (5 min at 13,200*g*) and the supernatant was immediately extracted with 400 µL of PCI (phenol : chloroform : isoamyl alcohol, 25:24:1, v/v, saturated with 10 mM Tris/HCl; pH 8.0). The extraction was repeated with 400 µL chloroform to remove phenol residues. The upper aqueous phase was mixed with 600 µL of isopropanol and was incubated on ice for 10-20 min to precipitate the DNA, followed by centrifugation (1 min at 13,200*g*). The DNA pellets were washed with 300 µL of 70% ethanol, dried, and dissolved overnight in 50 µL of TE buffer (10 mM Tris/HCl; 1 mM EDTA; pH 8.0). For isolation of plasmid DNA, the silicon-based column system QIAprep Spin Miniprep kit (Qiagen, Manchester, United Kingdom) was used according to the manufacturer's manual.

2.5.2 Isolation of chromosomal DNA from *Streptomyces* spp.

Cultures of *Streptomyces* strains were grown in a mixture of 5 mL YEME containing 34% (w/v) sucrose and 5 mL TSB. After 1-6 days, mycelial cultures from 1-2 mL medium were harvested by centrifugation (1 min at 13,200*g*), washed with 1 mL dH_2O, resuspended in 0.5 mL lysozyme buffer, and incubated for about 30 min at 37 °C to weaken the bacterial cell wall. Proteins were denatured by adding 250 µL of 2% (w/v) SDS solution and vigorous mixing for 1-2 min until the viscosity of the solution decreased

noticeably. Total DNA was extracted with 250 µL of PCI (phenol : chloroform : isoamylalcohol, 25:25:1, v/v, saturated with 10 mM Tris/HCl; pH 8.0). The solution was mixed vigorously for 1-2 min, phases were separated by centrifugation (2 min at 13,200*g*), and the upper aqueous phase was transferred to new 1.5 mL vials. The solution was mixed with 0.1 volume of 3 M sodium acetate solution (pH 4.8) and then with 1 volume of isopropanol. The DNA was alcohol-precipitated for 5 min at room temperature and then pelleted by centrifugation (5 min at 13,200*g*). The DNA was washed in 500 µL of 70% ethanol, dried, and dissolved in 100 µL of TE buffer (10 mM Tris/HCl; 1 mM EDTA; pH 8.0).

2.5.3 Agarose gel electrophoresis

DNA fragments were analysed via agarose gel electrophoresis. Samples were mixed with 0.1 volume of 6-fold concentrated loading dye prior loading and were separated in 1-1.5% agarose gels using an electric potential of 80-120 Volt. The gel was subsequently stained in ethidium bromide solution (about 1-5 µg/L EtBr final concentration). The DNA was visualised on a UV table and documented with an attached camera.

For high molecular weight DNA, samples were prepared as described and separated by pulsed-field gel electrophoresis (PFGE) using a CHEF DR II pulse field gel apparatus (Bio-Rad). DNA fragments were separated in 1% agarose gels using an alternating electrical field of 6 V/cm for 16 hours.

2.5.4 Purification of size-separated DNA fragments

DNA fragments of respective length were cut out from agarose gels with a razor blade. The gel slice was cut into small fragments, 300 µL of PCI was added, and the mix was frozen in liquid nitrogen for 30 seconds. The vial was centrifuged (15 min at 13,200*g*), 300 µL of TE buffer was added, and the vial was centrifuged again. The aqueous phase was transferred into a new 1.5 mL vial and purified from phenol residues by addition of 300 µL of

chloroform. After centrifugation (10 min at 13,200*g*), the upper phase was transferred to a new vial and the DNA was precipitated on ice for 10 min after addition of 0.1 volume of 3 M potassium acetate solution (pH 5.5), 1 volume isopropanol, and 1 µL of glycogen. The DNA was pelleted (10 min at 13,200*g* and 4 °C) and washed in 300 µL of 70% ethanol before dissolving in 25 µL of water or TE buffer.

2.5.5 Polymerase chain reaction

DNA fragments were amplified via polymerase chain reaction (PCR). Components were mixed as listed in **Table 2.7** and PCR cycling was carried out as shown in **Table 2.8**.

Table 2.7

Components used for PCR amplifications.

Standard DNA polymerase (*Taq*) was used for verification of DNA fragment lengths, high-fidelity DNA polymerases were used for sequencing (*Pfu* and *Q5*) and extension PCR (*Expand HiFi* or *Q5*). High GC-enhancer (5-fold; 5 µL) was added instead of DMSO for *Q5* DNA polymerase.

Compound	Volume	Final concentration
Forward/ reverse primer (25 µM)	1 µL (each)	25 pmol (each)
dNTPs (10 mM)	0.5 µL	50 µM (each)
DMSO	1.25 µL	5 %
Polymerase buffer (5-/10-fold)	5 µL/ 2.5 µL	1-fold
DNA Polymerase [1)]		
Taq (Qiagen, 5 U/µL) or	0.125 µL	0.625 U
Pfu (Roche, 3.5 U/µL) or	(1 µL)	(3.5 U)
Q5 HiFi (NEB, 2 U/µL) or	(1 µL)	(2 U)
Expand HiFi (Roche, 3.5 U/µL)	(1 µL)	(3.5 U)
Template DNA	1 µL	~100 ng
dH_2O	up to 25 µL	

Table 2.8

PCR cycling conditions.

PCR conditions for *Taq* and *Pfu* DNA polymerases (A), extension PCR using *Expand HiFi* DNA polymerase (B), and high-fidelity *Q5* DNA polymerase (C). Conditions to proof gene deletions are given in parenthesis. Elongation time was 1 min per 1 kb DNA template or 30 seconds for *Q5* DNA polymerase.

A)

Cycle	Temperature	Time	Repeat
Initial denaturation	94 °C	4 min	1x
Denaturation	94 °C	45 s	30x
Annealing	55-60 °C	45 s	
Elongation	72 °C	25 s to 2 min	
Final elongation	72 °C	5 min	1x
Cooling	4 °C	1 min	1x

B)

Cycle	Temperature	Time	Repeat
Initial denaturation	94 °C	4 min	1x
Denaturation	94 °C	45 s	10x
Annealing	52 °C	45 s	
Elongation	72 °C	90 s	
Denaturation	94 °C	45 s	15x
Annealing	56 °C	45 s	
Elongation	72 °C	90 s	
Final elongation	72 °C	5 min	1x
Cooling	15 °C	1 min	1x

C)

Cycle	Temperature	Time	Repeat
Initial denaturation	98 °C	30 s	1x
Denaturation	98 °C	30 s	30x
Annealing	55-60 °C	45 s	
Elongation	72 °C	15 s to 2 min	
Final elongation	72 °C	5 min	1x
Cooling	4 °C	1 min	1x

2.5.6 Cloning of DNA fragments

Vector DNA or PCR amplified DNA fragments were cut by respective restriction endonucleases according to the manufacturer's manual. Restriction digestion mixes of 20 or 50 µL volume were incubated overnight at 37 °C. The DNA was extracted by addition of 200 µL TE-buffer and 300 µL of PCI followed by centrifugation (10 min at 13,200*g*) and precipitation with 600 µL isopropanol. After centrifugation, the DNA pellet was washed with 300 µL of 70% ethanol and dissolved in 15 µL TE-buffer. The cut and purified DNA fragments were ligated overnight at 4 °C. The ligation mix was prepared by mixing in order 3 µL of dH_2O, 1 µL of 10x ligase buffer (Roche, Burgess Hill, UK), 5 µL digested DNA, and 1 µL T4 DNA ligase (Roche, Burgess Hill, UK). The ligated DNA was directly used for transformation of electrocompetent *E. coli* cells.

2.5.7 PAC library construction

Genomic DNA from *S. antibioticus* Tü6040 was used for construction of a genomic library by Bio S&T Inc. (Montreal, Canada) as described previously (Jones *et al.*, 2013). Isolated genomic DNA was cloned into pESAC13, an integrative phage P1-derived artificial chromosome (PAC) vector derived from pPAC-S1, by replacing the carbenicillin resistance gene (*Bam*HI fragment). Selection for pESAC13 was achieved by kanamycin resistance in *E. coli* and thiostrepton resistance in *Streptomyces*. The genomic library consisted of 2,688 individual PAC clones with an average insert size of 110 kb and more than 20-fold genome coverage. PAC clones containing the complete *sim* gene cluster were identified by PCR screening of the genomic library using two primer pairs (**Table 2.4**) for amplification of fragments flanking the *sim* gene cluster on both sides. Six double-positive clones with amplification of both PCR products were identified. Sequencing of one double-positive clone (PAC-12I; **Table 2.4**) by Illumina sequencing (University of Cambridge, Cambridge, UK) revealed a total insert size of 95 kb, containing the minimal 72-kb *sim* gene cluster flanked by genomic

regions of ~19 kb and 4 kb on either side. Thus, PAC-12I was chosen for heterologous expression studies.

2.5.8 Deletion of genes via PCR targeting and complementation

Genes were deleted by PCR targeting as described previously (Gust *et al.*, 2003; Gust *et al.*, 2004), leaving an 81-bp in-frame scar. PCR targeting has been developed to improve genetic manipulation with linear DNA and deletion of DNA fragments from the genome of *Streptomyces* species. The strategy is based on λ-RED recombination with a disruption cassette mediated by λ-RED genes (*exo*, *bet*, *gam*), which encode a 5'-exonuclease (Exo), a single stranded DNA binding protein (Bet) and a polypeptide (Gam) that confers protection against nuclease attack of linear double-stranded DNA (Muniyappa and Radding, 1986). The introduced disruption cassette is flanked by FLP recognition target (FRT) sites, which are recognised and removed by the FLP-recombinase to leave an 81-bp "scar" sequence in the preferred reading frame.

The strategy of PCR targeting was adapted in this study for mutagenesis of the heterologously expressed simocyclinone gene cluster as described recently (Jones *et al.*, 2013). The *sim* gene cluster was isolated on a phage artificial chromosome (PAC) vector from a genomic library of *S. antibioticus* Tü6040 (see **chapter 2.5.7**). The presence of an origin of transfer (*oriT*) on the PAC vector backbone allowed conjugational transfer into *Streptomyces* strains, which is not only more efficient than protoplast transformation but also facilitates mutagenesis of the cluster in *E. coli* before conjugating respective scar mutants into *Streptomyces* strains (**Figure 2.1**). The introduction of large DNA constructs such as PACs into bacteria is particularly challenging and highly sensitive to disruptions. Therefore, *E. coli* strains were transformed by electroporation with purified and concentrated PAC clones and conjugational transfer was carried out overnight with careful handling of the conjugation plate.

For deletion of DNA fragments, an apramycin resistance cassette flanked by FRT sites was PCR amplified from pIJ773_del_oriT using extended primers with 39 nucleotide homology (5'-ends) to the flanking sites of the DNA region to be deleted (**Figure 2.1**). The PAC clone was transformed into *E. coli* BW25113 carrying the temperature-sensitive λ-RED recombination plasmid pIJ790 (*gam*, *bet*, *exo*). Recombination was carried out by induction of λ-RED genes with 10 mM L-arabinose for 3-4 hours, followed by transformation of the respective extended PCR construct and a temperature shift from 30 °C to 37 °C for 16-20 hours. Successful PCR targeting was confirmed by PCR amplification and restriction digestion with *Eco*RI from Apramycin resistant colonies. Mutated DNA constructs were transformed twice into *E. coli* DH5α or DH10B and confirmed again as described to avoid transfer of mixed DNA constructs, which cannot be detected after removal of the resistance marker. The apramycin resistance cassette was removed from the DNA construct by transformation into *E. coli* BT340 carrying the temperature sensitive FLP-recombination plasmid pCP20, followed by overnight growth after a temperature shift from 30 to 43 °C (**Figure 2.1**). Excision of the apramycin resistance cassette left an 81-bp in frame scar, which was analysed by PCR amplification and restriction digestion with *Eco*RI from apramycin sensitive colonies.

For mobilisation of PAC clones, the DNA constructs were conjugated from *E. coli* DH5α into the methylation-deficient *E. coli* strain ET12567 by triparental mating, using *E. coli* TOP10 cells carrying the self-transmissible driver plasmid pR9604 (Jones *et al.*, 2013). PAC clones were conjugated from *E. coli* ET12567/pR9406 into *Streptomyces* strains as previously described (Kieser *et al.*, 2000). After 16-20 hours of incubation on R2 medium without sucrose to reduce the osmotic pressure for the impaired *E. coli* ET12567, plates were overlaid with thiostrepton and nalidixic acid. Exconjugants were streaked on SFM agar with thiostrepton and nalidixic acid and confirmed by PCR amplification and restriction digestion with *Eco*RI. For spore preparation, strains were grown on Instant Mashed Potato agar (2 g/L Smash® instant mashed potato and 2 g/L LabM agar, made up with tap water).

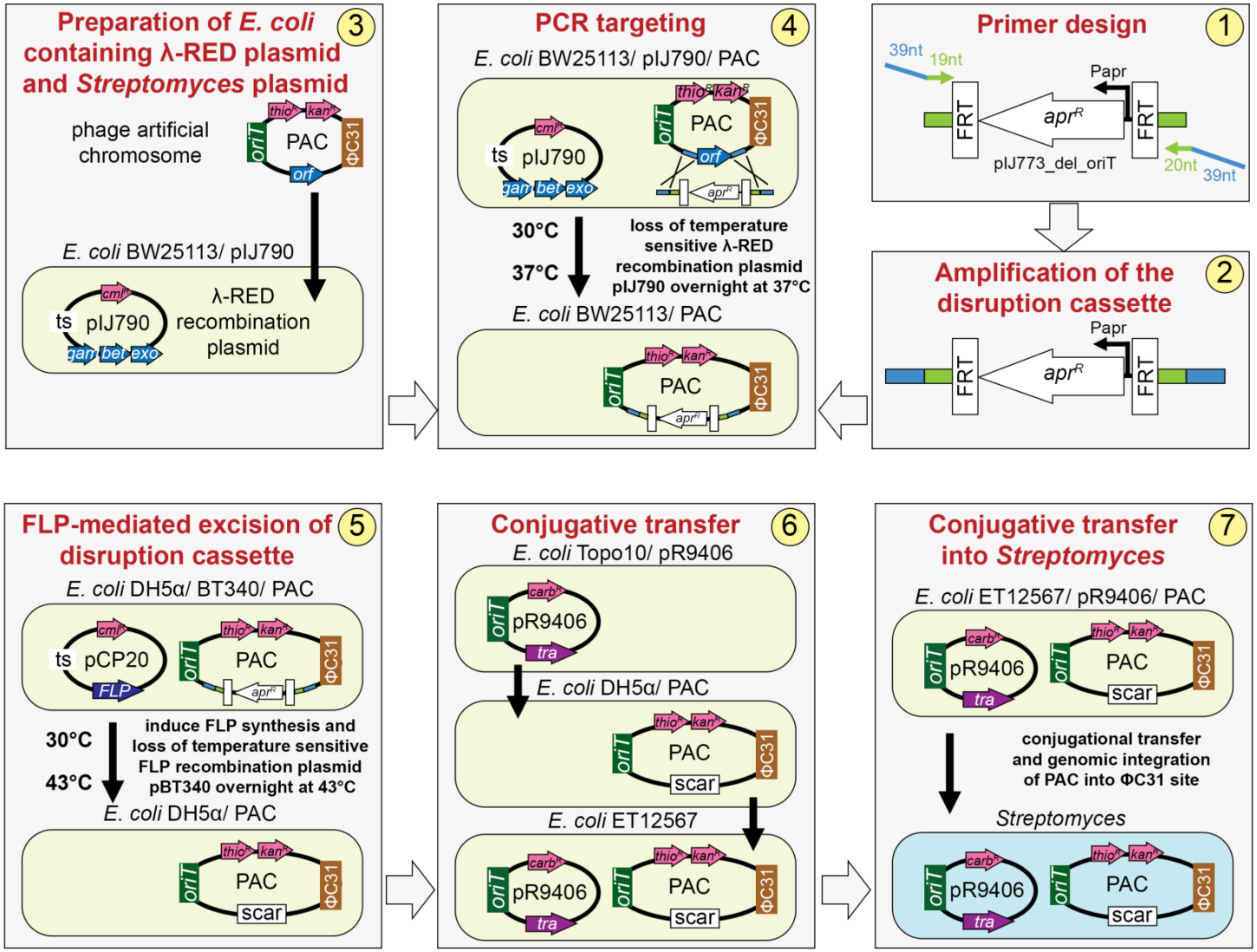


Figure 2.1

Modified strategy for PCR targeting.

The strategy was developed for mutagenesis of the simocyclinone (*sim*) biosynthetic gene cluster cloned into a phage artificial chromosome (PAC) vector. Modified from Gust *et al.* (2003). .

For complementation of SimC7, the respective gene was cloned into pGM1190 (*ptipA::simC7*) to generate pGM1190-NB-C7. Alternatively, *simC7* was fused to the *simB7* promoter (*pB7::simC7*) to generate pGM1190-pB7-NB-C7. Successful complementation was confirmed by PCR and product analysis was confirmed by liquid chromatography-mass spectrometry (LC-MS).

2.5.9 Site-directed mutagenesis of SimC7

The PCR-based *Q5*® site-directed mutagenesis kit (New England Biolabs, Ipswich, UK) was used to generate *simC7* point mutants, using plasmid pET15b-NB-C7 as the template. Pairs of oligonucleotides were designed to amplify linearized plasmid DNA with one oligonucleotide containing the desired mutation (**Table 2.5**). The template plasmid was degraded with *Dpn*I and the linear PCR product was self-ligated in presence of kinase and DNA ligase before transformation into *E. coli* DH5α. Mutations were verified by sequencing. Protein purifications were carried out as described for native SimC7.

2.5.10 Recombineering with single-stranded DNA oligonucleotides

Recombineering is homologous recombination-mediated genetic engineering. Single-stranded linear DNA in form of ~70 base oligonucleotides are sufficient to introduce point mutations or deletions up to 10 kb (Sawitzke *et al.*, 2013). The recombineering efficiency can be visualised by using a control oligonucleotide that introduces a mutation in *galK.* The mutation corrects a *TAG* stop codon to a *TAT* tyrosine codon ($galK_{TYR145UAG}$). Successful recombineering gives red colonies on MacConkey galactose indicator agar (Costantino and Court, 2003).

Electrocompetent, recombineering-proficient cells of *E. coli* HME68 were prepared by adding 0.1% inoculum of an overnight culture to 50 mL of LB medium with appropriate selection. Cultures were grown at 30 °C while

shaking (250 rpm) till OD_{600} of 0.4-0.5. To induce the genes needed for recombineering, cultures were grown for 15 min at 42 °C and then cooled on ice for 5-10 min. Cells were washed twice in 10 mL of ice-cold 10% (v/v) glycerol (5 min at 1,825*g* and 4 °C) and were resuspended in the remaining liquid. Freshly prepared electrocompetent, recombineering-proficient cells (200 µL) were mixed with the desired oligonucleotides for single amino acid substitutions and a control oligo (~2,310 ng per oligo; 1 µL of 100 µM oligo). Cells were electroporated as described (see **chapter 2.4.2**). Cultures were grown shaking for 2 hours at 30 °C. Dilutions of 10^{-4} to 10^{-5} were plated on MacConkey agar with respective antibiotics and kept at 30 °C for 2 days. Single colonies were plated on selective LB plates and used for PCR screening with various primers and restriction digests.

2.5.11 Clone library from genomic DNA

A clone library was made from genomic DNA of *S. antibioticus* Tü6040. Genomic DNA (~10 µg) was partially digested for 1 hour at 37 °C with 0.05 U of *Sau*3AI and separated by agarose gel electrophoresis. Fragments of 10-30 kb were purified and ligated into the *Bam*HI-linearized and dephosphorylated plasmid pSET152. The ligation mix was transformed into library-efficient electrocompetent *E. coli* DH10B (Invitrogen) and the cells were grown overnight at 37 °C on 10 selective LB agar medium. Clones from each plate were pooled and used for triparental mating with *E. coli* DH5α/ pUB307 and *S. lividans* TK24 as described (see **chapter 2.4.3**). Exconjugants were replica-plated on SFM agar and spores were harvested by pooling spores from each plate. The spores were kept at -20 °C.

2.5.12 Sequencing

Genomic regions of interest were PCR-amplified using a proof-reading DNA polymerase (*Pfu*) under conditions described in **chapter 2.5.5**. Amplified DNA fragments were purified via PCR purification kit (Qiagen, Hilden,

Germany) and were sent for sequencing (MWG Eurofins, Ebersberg, Germany).

For genome sequencing, genomic DNA was isolated from *S. antibioticus* Tü6040 grown in 50 mL DNB for 1 week shaking at 30 °C and sample quality was analysed by pulsed-field gel electrophoresis (PFGE) and a NanoDrop spectrophotometer (NanoDrop products, Wilmington, USA). The genome was sequenced and assembled at the Earlham Institute (Norwich, UK) using Pacific Biosciences (PacBio, Menlo Park, Canada) sequencing.

2.6 Expression and purification of recombinant proteins

2.6.1 Protein expression and purification (His-tagged)

The recombinant SimC7 protein had 20 additional amino acids at the N-terminus compared to the native sequence (MGSSHHHHHHSSGLVPRGSH) and a molecular weight of 32,235 Da.

For overexpression of His-tagged recombinant proteins, two 400 mL LB cultures were grown shaking (250 rpm) at 28 °C with 1% inoculum from an overnight culture. The cultures were induced at early-growth phase (OD_{600} ~0.3) with 0.5 mM IPTG and were grown for about 16-18 hours at 18 °C. The cells were harvested by centrifugation (F12 rotor with 3,032*g* for 10 min at 4 °C) using one centrifugation tube per 400 mL culture. The cell pellets were re-dissolved in 20 mL of lysis buffer I and lysed for 5-10 min at room temperature and for 30 min on ice. The solution was sonicated twice for 40 seconds to separate DNA from proteins and the cell debris was separated by centrifugation (SS34 rotor with 38,724*g* for 30 min at 4 °C). The protein solution was mixed with 0.5 mL of Ni-NTA-Agarose beads (Qiagen) and was loaded on a 5 mL polypropylene column (Qiagen). His-tagged proteins were separated from the remaining lysate by applying 700 mL of washing buffer I and the tagged proteins were eluted from the Ni-NTA-Agarose beads by addition of 1 mL protein elution buffer I. The solution was dialysed in a dialysis tube overnight at 4 °C in 2-3 L of storage buffer I and

stored at -80 °C. Isolated proteins were quantified via Bradford protein assay and purity was analysed by SDS-PAGE (**Figure 2.2**).

For protein crystallography and mutant proteins, I established from a number of expression and purification trials that the presence of sodium chloride had a tendency to induce aggregation of protein derived from this construct. For this reason, sodium chloride was either excluded or kept at low concentrations in the buffers used thereafter. 10 mL overnight culture was used to inoculate 1 L of LB medium without sodium chloride containing 40 mg carbenicillin and 10 mg chloramphenicol. The culture was grown at 37 °C to OD_{600} ~0.3, cooled to 18 °C, and protein expression was induced by addition of IPTG to a final concentration of 0.5 mM. The culture was incubated for 16 hours shaking at 18 °C. Harvested cells were resuspended in lysis buffer IV, and lysed in an Avestin EmulsiFlex®-B15 high pressure homogeniser (three cycles with 16,000 psi pressure) (Avestin, Mannheim, Germany). Cell debris was removed by centrifugation (SS34 rotor with 38,724*g* for 30 min at 4 °C). His-tagged protein was purified by Ni-affinity chromatography using 700 mL of washing buffer III and 15 mL of elution buffer III. Isolated proteins were analysed as described above (**Figure 2.2**), buffer exchanged against storage buffer X and concentrated to 4 mg/mL using VivaSpin 15 columns (Vivaproducts, Littleton, USA), and frozen at -80 °C. The CD spectra of all the mutant SimC7 proteins were collected and found to be unchanged relative to that of the wild-type enzyme.

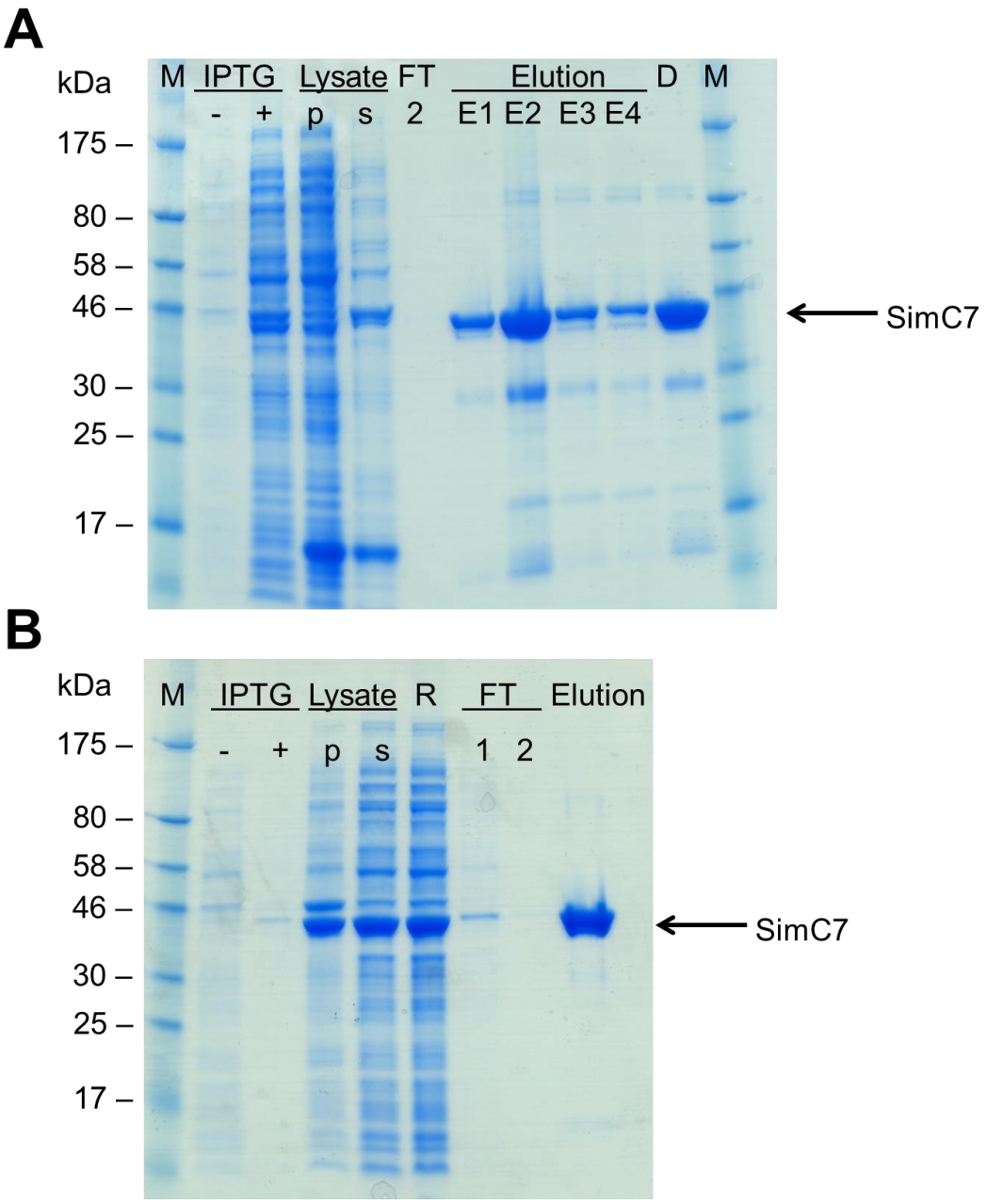


Figure 2.2

Purification of recombinant SimC7 protein from *E. coli*.

SimC7 was overexpressed and purified for enzyme activity assays (A) and for protein crystallography (B) as described in this chapter. Lanes: M, pre-stained broad-range protein marker (NEB); IPTG, culture supernatant before (-) and after (+) induction with 0.5 mM IPTG; Lysate, lysate pellet (p) and supernatant (s); R, remaining supernatant after removal of protein-bound Ni-NTA agarose; FT, flow-through before (1) and after (2) addition of washing buffer; Elution, elution fractions; D, dialysed SimC7 protein.

2.6.2 Protein quantification via Bradford assay

The Bradford assay is a colorimetric method for quantification of proteins. The assay is based on the absorbance shift of Coomassie Brilliant Blue G-250 from a red colour under acidic conditions to a blue dye when bound to cationic or non-polar, hydrophobic amino acid residues. Consequently, the amino acid composition determines the intensity for staining (Bradford, 1976). Bovine serum albumin (BSA) was used for calibration in 1-20 µg/mL final concentration. Aliquots (2 µL) of protein samples or dilutions of BSA in 800 µL water were mixed with 200 µL dye reagent concentrate (Bio-Rad) and were incubated for 5 min in the dark. The absorption at 595 nm wavelength was determined against a protein-free control (blank). Protein concentrations were calculated using the linear equation from BSA standards, taking respective dilutions into account (**Equation 1**).

Equation 1
Calculation of protein concentrations via Bradford assay.

$$(A\ 595\ nm) = 0.1592\ \times \left(Protein\ \left[\frac{\mu g}{mL}\right]\right) + 0.0152$$

$$Protein\ \left[\frac{\mu g}{mL}\right] = \left(\frac{A595\ nm - 0.0152}{0.1592}\right)$$

2.6.3 Discontinuous polyacrylamide gene electrophoresis (SDS-PAGE)

Proteins were separated under denaturing conditions with sodium dodecyl sulphate using discontinuous polyacrylamide gel electrophoresis (SDS-PAGE). Polyacrylamide gels were composed of a lower 12% separating gel (pH 8.8) and an upper 5% stacking gel (pH 6.8) as listed in **Table 2.9**. Protein samples (1-5 µL) were mixed with protein loading buffer (1-2-fold final concentration) and heat-denatured (95 °C for 5-10 min) before loading

them into the gel. The electrophoresis was performed at 80-150 Volts in SDS-PAGE running buffer for about 1-1.5 hours using prestained broad range protein marker (7-175 kDa) (New England Biolabs, Ipswich, UK) for identification of respective protein bands.

Table 2.9

Composition of stacking gel (5%) and separating gel (12%) for SDS-PAGE.

Compound	Stacking gel	Separating gel
Tris	125 mM	375 mM
Acrylamide/Bisacrylamide (30:0.75, v/v)	5 %	12 %
SDS (w/v)	0.1 %	0.1 %
Ammonium persulphate (w/v)	0.1 %	0.1 %
TEMED (v/v)	0.1 %	0.04%
adjusted with HCl to	pH 6.8	pH 8.8

2.6.4 Staining and immobilisation of proteins in SDS-PAGE gels

Most commonly Coomassie Brilliant Blue (CBB) is used that reversibly binds to cationic and non-polar, hydrophobic amino acid residues. SDS-PAGE gels were stained in a Coomassie Brilliant Blue-based staining solution (Instant Blue, Expedeon, Harston, UK) overnight at room temperature under continuous agitation. The background staining was reduced by incubating the gel for 30 min in distilled water.

2.7 Protein crystallography

2.7.1 Protein crystallisation and cryoprotection

All crystallisation experiments were performed with N-terminally His-tagged SimC7 at a concentration of approximately 4 mg/mL and at a temperature of 20 °C. Screening was conducted by sitting-drop vapour diffusion in MRC 96-well crystallisation plates (Molecular Dimensions) with a mixture of 0.3 µL well solution from both commercial and in-house screens (Qiagen: PEG suite, $AmSO_4$ suite; Molecular Dimensions: PACT premier™, JCSG-plus™, structure, Morpheus™, MIDAS™, SG1™; in house: KISS) and 0.3 µL protein

solution with 40 µL of respective mother liquor using an OryxNano crystallisation robot (Doublas Instruments Ltd, Berkshire, UK). Images of screening conditions were taken daily by a Ministrel DT plate imager (Rigaku, Ettlingen, Germany) using visible light and UV light to distinguish between protein and salt crystals. Conditions with small protein crystals required manual optimisation. Promising conditions were optimised manually in 24-well hanging-drop vapour diffusion format using XRL plates (Molecular Dimensions) with drops consisting of 1 µL protein and 1 µL precipitant applied on microscope cover slips and positioned upside down above a well with 1 mL of respective mother liquor sealed with high vacuum grease (Dow Corning, Auburn, USA). Crystal formation was followed manually using a stereoscopic microscope (SMZ800, Nikon, Badhoevedorp, Netherlands) of 10- to 63-fold magnification with an attached digital camera.

Tetragonal crystals of the SimC7 complex with $NADP^+$ were obtained from 16% (w/v) PEG-8000 and 20% (v/v) glycerol (native SimC7) or 1.3 M DL-malic acid (pH 7) supplemented with 1% (v/v) DMSO and 2.5 mM EDTA in the drop (SeMet-labelled SimC7). The native crystals did not require further cryoprotection, while the SeMet crystals were cryoprotected by supplementing the crystallisation solution with 25% (v/v) glycerol. Crystals of *apo*-SimC7 that were isomorphous with those of the complex with $NADP^+$ (*apo* form 1) grew from 25% (w/v) PEG-6000, 20% (v/v) glycerol in 0.1 M MES (pH 6.5), while monoclinic *apo*-SimC7 crystals (*apo* form 2) were obtained from 16% (w/v) PEG-8000, 18% (v/v) glycerol with 10 mM dithiothreitol (DTT) added to the drop; neither form required further cryoprotection. 7-oxo-SD8 was isolated from *S. antibioticus* Tü6040 as described (see **chapter 2.9.1**), and was dissolved in DMSO to a concentration of 50 mM prior to mixing with protein. The cofactor $NADP^+$ (Sigma-Aldrich, Dorset, UK) was dissolved in water to a concentration of 100 mM without adjusting the pH, which was crucial for obtaining crystals. SimC7 was first mixed with $NADP^+$ and then with 7-oxo-SD8, giving a final protein concentration of 0.12 mM (3.8 mg/mL) protein, 1.5 mM of 7-oxo-SD8, and 3.5 mM of $NADP^+$. The final concentration of DMSO did not exceed 3% (v/v). Orthorhombic crystals of the ternary complex of SimC7 bound to

$NADP^+$ and 7-oxo-SD8 grew from 20-25% (w/v) PEG-8000, 20-25% (v/v) glycerol in 0.1 M sodium acetate (pH 4.8), and did not require further cryoprotection.

2.7.2 Selenomethionine labelling of recombinant protein

For *de novo* structure determination, recombinant protein was labelled with selenomethionine (SeMet) by metabolic inhibition (Doublie, 1997). This technique is based on the inhibition of methionine biosynthesis and does not require a methionine-auxotroph strain and therefore no further transformation. Cells were grown in two times 0.5 L M9 minimal medium at 37 °C to OD_{600} ~0.1, amino acids were added (100 mg/L each of threonine, lysine and phenylalanine; 50 mg/L each of leucine, isoleucine and valine; 60 mg/L SeMet), and the cultures were further incubated for 90 min. Protein overexpression was induced with 0.5 mM IPTG overnight at 18 °C. Labelled protein was purified as described above (see **chapter 2.6.1**).

2.7.3 X-ray data collection, structure determination and structure refinement

Crystals were harvested and flash-cooled in liquid nitrogen using LithoLoops (Molecular Dimensions). The mounted crystals were stored in Unipuck cassettes (MiTeGen) prior to transport to the Diamond Light Source (Oxfordshire, UK), where they were transferred robotically to the goniostat on either beamline I03, I04 or I04-1 and maintained at -173 °C with a Cryojet cryocooler (Oxford Instruments). X-ray diffraction data were recorded using a Pilatus 6M hybrid photon counting detector (Dectris), then integrated using XDS (Kabsch, 2010), and scaled and merged using AIMLESS (Evans and Murshudov, 2013) via the XIA2 expert system (Winter, 2010).

Native data from the complex with $NADP^+$ were collected to 1.95 Å and processed in space group $P4_12_12$ with an estimated solvent content of 49.3%, based on one copy of the protein chain in the ASU. A SeMet-labelled

SimC7 crystal was used to collect a single-wavelength anomalous dispersion data set at the selenium K X-ray absorption edge (wavelength 0.9796 Å). The data were processed to 2.05 Å resolution in space group $P4_12_12$ and were isomorphous with the native set. Experimental phases were determined by analysing the SeMet and native data using the SHELX suite (Sheldrick, 2008). SHELXD located 4 selenium sites and, after phasing with SHELXE (which revealed that $P4_12_12$ was the correct enantiomorph), the figure-of-merit was 0.589 to 1.95 Å resolution. These phases were improved by density modification with PARROT (Cowtan, 2010) and used as input to automated building with BUCCANEER (Cowtan, 2006), which was able to fit 233 residues (82% of the native sequence). After combining phases calculated from this model with the original experimental phases using SIGMAA (Read, 1986), PARROT and BUCCANEER were re-run. This time 277 residues (98% of the native sequence) were fitted giving R_{work} and R_{free} values of 0.263 and 0.298, respectively, at 1.95 Å resolution. The model of the complex with $NADP^+$ was completed by several iterations of manual rebuilding in COOT (Emsley and Cowtan, 2004) and restrained refinement in REFMAC5 (Murshudov *et al.*, 1997) using isotropic thermal parameters and TLS group definitions obtained from the TLSMD server (Painter and Merritt, 2006).

A dataset was collected to 1.60 Å resolution from *apo*-SimC7 crystallised in the same form as in complex with $NADP^+$. This was used to generate the model of *apo* form 1 starting from the model of the complex with $NADP^+$ from which the coordinates of the cofactor had been removed. Model building and refinement were completed as above. Crystals of *apo* form 2 yielded a dataset to 1.90 Å resolution after processing in space group C2, with an estimated solvent content of 48.3% based on two copies of the protein in the ASU. The structure was solved by molecular replacement with PHASER (McCoy *et al.*, 2007) using the protein component of $NADP^+$ complex structure as the template, and then refined as above.

Data for the ternary complex of SimC7 with $NADP^+$ and 7-oxo-SD8 were processed to 1.20 Å resolution in space group $P2_12_12_1$, giving an estimated solvent content of 45.3% based on one copy of the protein in the ASU. The

latter was located using PHASER with the same template as before. For this structure it was possible to use anisotropic thermal parameter refinement in REFMAC5.

Model geometries were validated with MOLPROBITY (Davis *et al.*, 2007) before submission to the Protein Data Bank.

A simulated annealing omit electron density map was calculated for the substrate in the ternary complex. For this, the 7-oxo-SD8 was deleted from the coordinates of the final model. The resultant PDB file was used as input to simulated annealing refinement with PHENIX (Adams *et al.*, 2010) from a starting temperature of 5000 K after applying small random shifts to the model ('shake' term set to 0.3).

2.8 Enzyme assays *in vitro*

2.8.1 DNA gyrase inhibition assays

Assays were performed as described previously (Flatman *et al.*, 2005; Hearnshaw *et al.*, 2014). DNA gyrase from *E. coli* (A_2B_2 complex, Inspiralis Ltd., Norwich, UK) was diluted in gyrase dilution buffer (Inspiralis Ltd., Norwich, UK) as needed. Purified simocyclinones were dissolved in DMSO. Ciprofloxacin (Sigma-Aldrich, Dorset, UK) was dissolved in water. MGD8N2A was made by Axel Zeeck (Edwards *et al.*, 2009).

DNA supercoiling assays were carried out in 30 µL reaction volumes, containing DNA gyrase (22 nM), relaxed pBR322 DNA (0.5 µg; 6 nM) in presence of ATP (1 mM) using gyrase supercoiling buffer (Inspiralis Ltd., Norwich, UK). Reactions were initiated by the addition of DNA gyrase, followed by incubation at 37 °C for 30 min. The reactions were stopped by the addition of 30 µL chloroform-isoamylalcohol (24:1) and 15 µL of STEB solution, followed by brief vortexing and centrifugation (16,100*g* for 5 min). The upper phase was analysed on 1% (w/v) agarose gels.

DNA cleavage assays were carried out in presence of 1 µM ciprofloxacin under the conditions described above in relaxation buffer (Inspiralis Ltd., Norwich, UK). The reaction was either carried out in forward direction (0.5 µg; 6 nM relaxed pBR322 DNA in presence of 1 mM ATP) or reverse reaction (0.3 µg; 3.6 nM supercoiled pBR322 DNA in absence of ATP). After initial incubation for 30 min at 37 °C, the protein-DNA complex was separated by addition of proteinase K (0.2 µg/mL) and SDS (0.2%, v/v). The mixture was incubated for further 30 min at 37 °C before the reaction was stopped.

2.8.2 Surface plasmon resonance (SPR)

The binding of simocyclinones to DNA gyrase was analysed by surface plasmon resonance (SPR) at 25 °C on a BiacoreT100 instrument (GE Healthcare, Amersham, UK). Purified GyrA55 protein was immobilised on a CM5 sensor chip. The surface carboxyl groups on the chip were activated by addition of 35 µL of a mixture of N-hydroxysuccinimide (50 mM) and 1-ethyl-3-(3-diaminopropyl)carbodiimide (0.2 M) from the Biacore amine coupling kit. The protein was diluted in HBS-EP+ buffer (10 mM Na-HEPES; 150 mM NaCl; 3 mM EDTA; 0.05% (v/v) surfactant P20; pH 7.4) and immobilised on the chip (10 µL/min) until a response of 3000 resonance units was reached. The remaining ester groups on the chip were deactivated by addition of 35 µL of ethanolamine. The system was equilibrated with HBS-EP+ running buffer with 1% (v/v) DMSO and a flow rate of 30 µL/min. Defined concentrations of compounds were added, previously diluted in HBS-EP+ running buffer with a final DMSO concentration of 1% (v/v). Protein-ligand interactions were analysed with a binding time of 2 min followed by a dissociation time of 2 min at a flow rate of 5 µL/min. The protein was regenerated by washing the chip twice with 1 M NaCl for 1 min at a flow rate of 30 µL/min. The maximum binding response (R_{max}) was calculated from the molecular weight ratio of compound (MW analyte) and protein (MW ligand) multiplied by the response of protein bound to the chip and the binding

stoichiometry (S) (**Equation 2**). Data were analysed using single kinetics with a 1:1 binding stoichiometry.

Equation 2

Calculation of the maximum binding response (R_{max}).

$$R_{max} = \left(MW_{analyte} / MW_{ligand}\right) \times R_{ligand} \times S$$

2.8.3 Circular dichroism (CD) spectroscopy

Circular dichroism (CD) spectroscopy experiments were performed using a Chirascan-Plus CD spectrophotometer (Applied Photophysics, Surrey, UK). Concentrated proteins in storage buffer X were buffer exchanged against 20 mM potassium phosphate buffer (pH 7.2) with 5% (v/v) glycerol and diluted to 0.2 mg/mL (6.25 µM). CD analysis was carried out at 20 °C using a quartz glass cuvette with a 0.5 mm path length. CD spectra were averaged from four scans collected between 190 nm and 260 nm wavelength, using a bandwidth of 2.0 nm, a step size of 0.5 nm, and one time point per second.

2.8.4 SimC7 oligomeric state determination by size exclusion chromatography (SEC)

Size exclusion chromatography (SEC) was used to determine the oligomeric state of SimC7. Concentrated protein samples (4 mg/mL) were prepared by exchanging storage buffer X with running buffer III using VivaSpin 500 columns (Vivaproducts, Littleton, USA). A Superdex 75 10/300 GL column attached to an AKTA protein crystallography machine (GE Healthcare, Amersham, UK) was equilibrated with 2 column volumes (60 mL) of running buffer III before analysis of protein samples (100 µL) using a constant flow

rate of 0.5 mL/min. Absorption was recorded at 280 nm wavelength for 1.25 column volumes.

2.8.5 Ketoreductase activity assays

Ketoreductase assays were performed in reaction mixtures consisting of the substrate 7-oxo-SD8 (25-200 µM), cofactor NAD(P)H (0.3 mM) or NAD^+ (300 mM), HEPES (50 mM, pH 7.2), and enzyme (24.18 µg/mL, 500 nM). Reactions were initiated with SimC7 and quenched at specific time points with equal volumes of methanol. Stopped reactions were incubated for 10 min at 100 °C and centrifuged (13,000*g* for 10 min at 4 °C) to remove precipitated protein. SD8 formation was quantified by analytical reversed-phase HPLC using a linear gradient of 25-95% methanol over 20 min (see **chapter 2.9.4**). SD8 concentrations were calculated by comparison with a standard curve (7.5-2,100 pmol, **Figure 2.3a** and **Figure 2.4**), which was calculated from a serial dilution of SD8 prepared in methanol (0.25-70 µM) under the conditions described above. Samples were analysed in duplicates and SD8 standards were analysed in triplicates. Product formation was confirmed by SD8 and 7-oxo-SD8 standards as well as LC-MS.

SimC7 variants were assayed as described above using 16.12 µg/mL (500 nM) enzyme, 0.3 mM NADH, 0.2 mM 7-oxo-SD8, and HEPES (50 mM, pH 7.2). Each mutant protein was analysed in three independent experiments. Data were averaged from two technical replicates for each SimC7 variant using a shortened linear gradient of 70-95% methanol over 10 min (see **chapter 2.9.4**). Substrate conversion was determined by quantification of the peak area for the reaction product SD8 based on a standard curve (7.5-9,000 pmol, **Figure 2.3b** and **Figure 2.4**) calculated from a serial dilution of SD8 (0.25-300 µM) prepared in methanol.

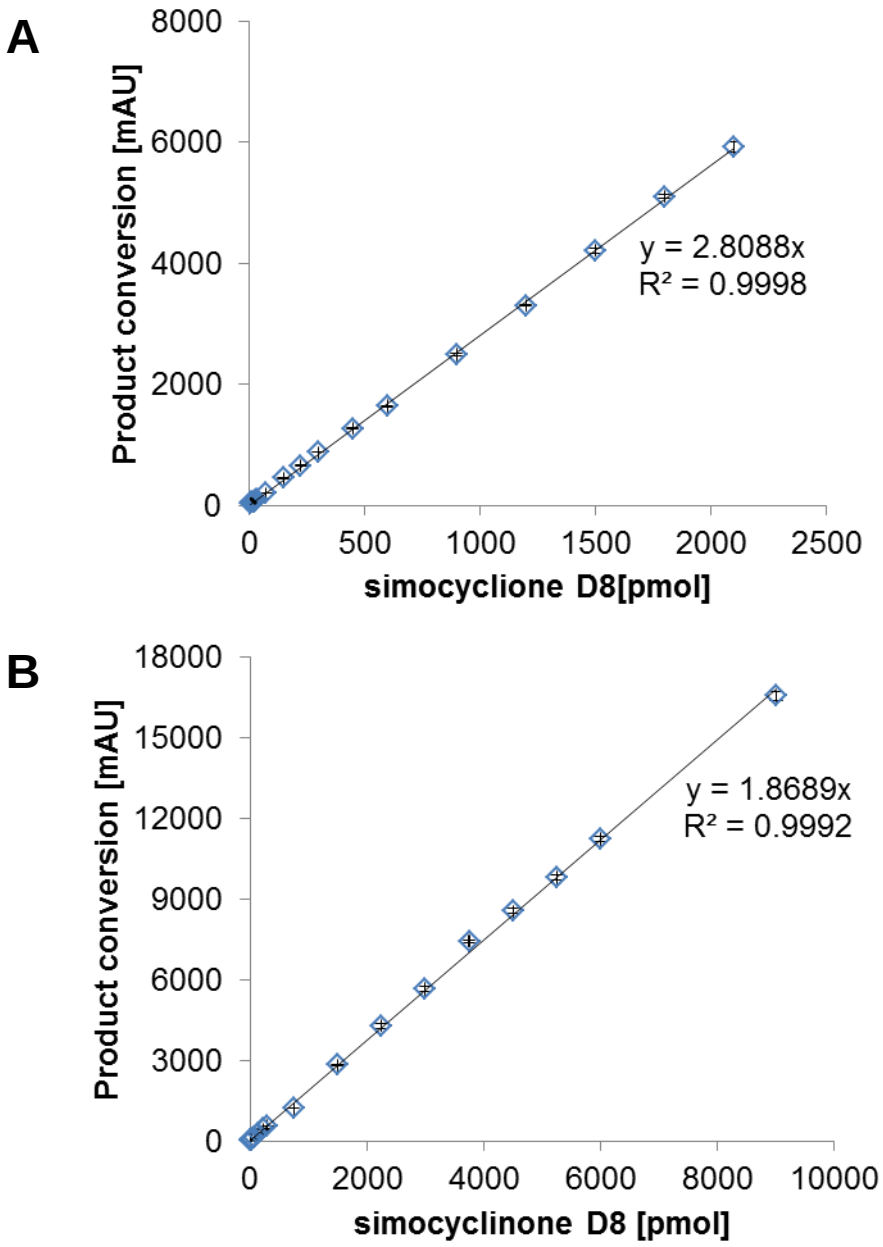


Figure 2.3

SD8 standard curve for determination of product conversion by (A) SimC7 and (B) SimC7 variants.

Standards (A: 7.5-2,100 pmol; B: 7.5-9,000 pmol) were analysed by HPLC in triplicate and the average peak area was determined. Error bars indicate the standard deviation between results of three sampling repeats. The coefficient of determination (R^2) was (A) 0.9998 and (B) 0.9992.

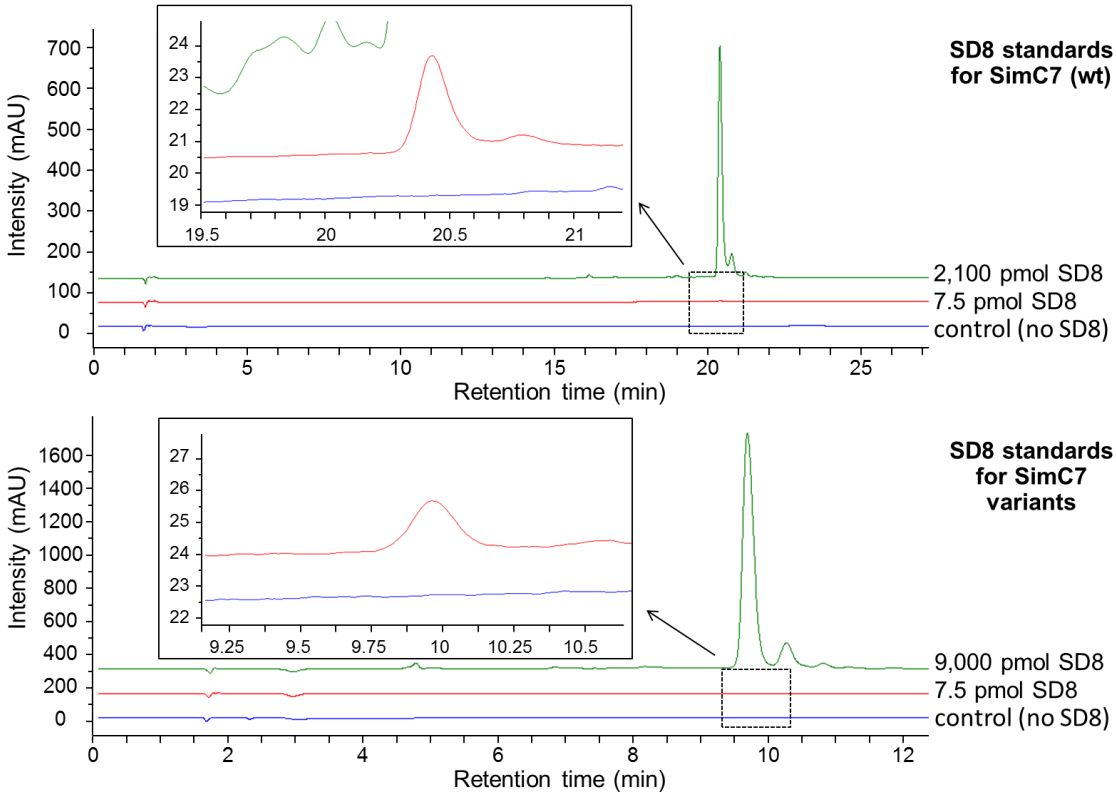


Figure 2.4

Detection limit for determination of product conversion by (A) SimC7 and (B) SimC7 variants.

HPLC chromatograms at 360 nm wavelength are show for the most dilute (7.5 pmol) and most concentrated (2,100 pmol for SimC7 wildtype and 9,000 pmol for SimC7 variants) SD8 standards used for calibration (see **Figure 2.3**).

2.9 Extraction and analysis of simocyclinones

2.9.1 Extraction of simocyclinones

For simocyclinone extraction, *Streptomyces* strains were grown for 6 days shaking at 30 °C in 50 mL Tu2 medium. For large-scale extraction of 7oxoSD8, four 1 L cultures of *S. coelicolor* M1152ex1::12IΔC7 were grown in Tu2 medium. Cultures were adjusted to pH 4.0 with 10% HCl solution and simocyclinones were extracted twice using ethyl acetate (1 volume for small cultures and 0.2 volumes for large cultures) while shanking for about 30 min at room temperature. The phases were separated by centrifugation (1,026*g* for 10 min) and the organic phase was vacuum-concentrated and kept at -20 or -80 °C.

Alternatively, the adsorptive hydrophobic resin Diaion® HP20 (Sigma-Aldrich, Dorset, UK) was used to absorb simocyclinones from the production medium. The resin was activated overnight in 3-4 volumes of methanol. The methanol was removed by washing the resin 2-3 times in dH_2O using a vacuum filter and was then kept in a semi-dry state before adding 50 g/L (5% w/v) to the culture medium. For simocyclinone extraction, cultures were centrifuged and methanol was added to the pellet (resin and mycelium) to dissolve simocyclinones from the resin. The liquid was vacuum-concentrated and samples were kept at -20 °C.

2.9.2 ^{13}C-enrichment of simocyclinones

For stable isotope labelling with carbon 13 (^{13}C), cultures were grown as described (see **chapter 2.9.1**). One 1 L culture of *S. coelicolor* M1152ex1::12IΔC7 was grown in Tu2 medium shaking at 30 °C and spiked with a total of 10 mL of 1 M sterile-filtered [1,2-$^{13}C_2$] acetate over 6 days. Aliquots of 1.5 mL were added on days 0, 1, and 3, and two times 0.75 mL on days 4, 5, and 6 with a final concentration of 0.75-1.5 mM of [1,2-$^{13}C_2$] acetate at each time of addition.

2.9.3 Liquid chromatography-mass spectrometry (LC-MS)

Simocyclinone extracts were analysed via liquid chromatography-mass spectrometry (LC-MS). Simocyclinone extracts were dissolved in methanol and samples (10 µL) were injected onto an HPLC column (Phenomenex Luna-C_{18} 3µ C18(2) 100A, 100 mm × 2 mm) with a Phenomenex SecurityGuard Prep Cardridge Holder Kit (2 mm) attached to a Thermo ion-trap LC-MS (LCQ DECA XP^{plus}) equipped with a Surveyor HPLC system (Thermo Fisher, Loughborough, UK). Samples were separated using a linear gradient with 0.1% formic acid as mobile phase A and methanol as mobile phase B at a flow rate of 1 mL/min. Mobile phase B was kept at 25% for 1 min followed by a gradient of 25-95% over 20 min, then 95% for 4 min, and 25% for 6 min. Samples were analysed by positive electrospray ionisation (ESI) MS and UV-vis chromatography. The PDA collected spectra from 200 nm to 600 nm at 1 Hz, additionally fixed wavelengths were collected at 210, 340, and 365 nm (9 nm bandwidth) at 10 Hz. Data were collected from fragments with a mass-to-charge (*m/z*) ratio of 150-2000 and data-dependent tandem mass spectrometry (MS2) was carried out for the most abundant ions above *m/z* 400 at an isolation width of *m/z* 4.0 and 30% collision energy. For quantification of SD8, an external calibration was used based on UV absorbance at 360±5 nm obtained from LC-MS. The spray chamber conditions were 350 °C capillary temperature, 50 units sheath gas, 5 units aux gas, and a source voltage of 3.8 kV using a steel needle kit. The column was maintained at 30 °C and the autosampler maintained the samples at 10 °C.

Later samples were analysed on a Shimadzu liquid chromatography mass spectrometry ion trap time-of-flight (LC-MS IT-TOF) (Shimadzu Scientific Instruments, Columbia, USA) using a shortened gradient and a flow rate of 0.5 mL/min. Mobile phase B was kept at 25% for 1 min followed by a gradient of 25-95% over 7 min, then 95% for 6 min, and 25% for 2 min. Samples were analysed by positive electrospray MS and UV-vis chromatography. The PDA collected spectra from 200 nm to 600 nm at 4.167 Hz with a time constant of 0.16 s. Data were collected from fragments

with a mass-to-charge (*m/z*) ratio of 180-1500 with a maximum ion accumulation time of 20 ms and automatic sensitivity control set to a target of 70% optimum base peak intensity. Tandem mass spectrometry (MS2) was carried out for the most abundant precursor ions with isolation width of *m/z* 3.0, 50% collision energy and 50% collision gas. The spray chamber conditions were 250 °C curved desorbation line, 300 °C heat block, 1.5 L/min nebulizer gas, and drying gas “on”. The instrument was calibrated with sodium trifluoroacetate cluster ions according to the manufacturer’s instructions.

2.9.4 High-pressure liquid chromatography (HPLC)

For purification of simocyclinones, dried culture extracts were dissolved in a minimal amount of methanol or a 10:1 methanol-acetone mixture and were separated by high-pressure liquid chromatography (HPLC).

For preparative HPLC, samples (350 µL) were injected onto an HPLC column (Phenomenex Gemini-NX 5µ C18 110A AXIA Packed, 150 mm × 21.20 mm) with a Phenomenex SecurityGuard Prep Cartridge Holder Kit (21.20 mm) attached to a Dionex UltiMate 3000 HPLC machine. Separation was performed using a linear gradient with 0.1%formic acid as mobile phase A and methanol as mobile phase B. The gradient was 75-95% mobile phase B over 20 min followed by 95% mobile phase B for 5 min at a flow rate of 21 mL/min, and absorbance was recorded at 360 nm.

For analytical HPLC, samples (30 µL) were separated on an HPLC column (Phenomenex Gemini-NX 3µ C18 110A, 150 mm × 4.6 mm) attached to an Agilent 1200 series HPLC machine using a linear gradient with 0.1% formic acid as mobile phase A and methanol as mobile phase B. The gradient was either 25-95% or 75-95% solvent B over 20 min at a flow rate of 1 mL/min. Absorbance was recorded with a diode array detector at 210 nm, 230 nm, and 360 nm for simocyclinones and 535 nm as reference. The gradient was shortened for enzyme activity assays with SimC7 variants to 70-95%

methanol against 0.1% formic acid over 10 min, followed by 95% methanol for 5 min (see **chapter 2.8.5**).

2.9.5 Normal phase liquid chromatography (LC)

Crude extracts of simocyclinones were separated by normal phase liquid chromatography (LC) on a Biotage chromatography machine (Biotage, Uppsala, Sweden). Samples were purified on a pre-packed silica column (Biotage® SNAP cartridge KP-SIL 50 g, 40- to 65-µm particle size, 39 mm × 81 mm; Biotage, Uppsala, Sweden) using a gradient elution from CH_2Cl_2-MeOH (95:5) to CH_2Cl_2-MeOH (70:30) over 4 hours at a flow rate of 50 mL/h (Holzenkampfer *et al.*, 2002). A methanol gradient (5-30% methanol) against dichloromethane was performed at a flow rate of 10 mL/min for 2 hours. Fractions of 10 mL were collected and the presence of simocyclinones was verified by thin layer chromatography, HPLC, and LC-MS.

2.9.6 Thin layer chromatography (TLC)

Thin layer chromatography (TLC) is a fast and simple way to identify chemical compounds present in a mixture and to determine their purity based on their distinct chemical properties. Samples were spotted (1 µL) on normal phase (silica coated aluminium) TLC plates (GE Healthcare, Amersham, UK), dried, and separated by placing in a glass chamber filled with ~1 cm of solvent. The solvent mixture was optimized so that the retention factor value (Rf value), which is the distance travelled by the component relative to the distance travelled by the solvent, was ~0.3. The presence of simocyclinones was confirmed under UV-light.

2.9.7 Small molecule size exclusion chromatography (SEC)

Size exclusion chromatography (SEC) separates molecules in solution according to their size. Simocyclinone samples were dissolved in methanol

and loaded on a column packed with 30 g of Sephadex LH-20 (18- to 111-µm particle size) (Sigma-Aldrich, Dorset, UK) using a mobile phase of 10% acetone, 80% methanol, and 10% water. Fractions of 10-15 mL were collected and presence of simocyclinones was confirmed by HPLC and LC-MS. Respective fractions were combined and dried *in vacuo*.

2.9.8 Nuclear magnetic resonance (NMR) spectroscopy

The structure of simocyclinones were analysed by nuclear magnetic resonance (NMR) spectroscopy. HPLC-purified compounds were dissolved in DMSO-d_6 and then analysed on Bruker AVANCE-DRX 700 and Avance III 400 spectrometers at 298 K. Data were processed and analysed using TopSpin v3.2 software (Bruker). Proton (^{1}H) coupling constants were determined using experiments from the Avance III 400 spectrometer. All other assignments were made using experiments from the DRX 700 spectrometer.

For proton (^{1}H) NMR spectroscopy, a high number of scans was chosen to amplify the signal for simocyclinones (160 scans). The proton (^{1}H) NMR experiment was repeated after addition of 5 µL deuterated water (D_2O) to confirm the signals of hydroxyl groups. For carbon (^{13}C) NMR spectroscopy, an attached proton test (APT) experiment was used to facilitate the assignment of protons attached to each carbon. Whereas carbons in methane (CH) and methyl (CH_3) groups yield positive signals, carbons in quaternary (C) and methylene (CH_2) groups yield negative signals.

Beside proton (^{1}H) and carbon (^{13}C) NMR spectroscopy, several two-dimensional NMR techniques were applied to confirm the structure of 7-oxo-SD8. Correlation spectroscopy (COSY) and total correlation spectroscopy (TOCSY) were used to determine neighbouring hydrogen atoms, heteronuclear single quantum coherence spectroscopy (HSQC) was used to assign hydrogens to respective carbon atoms, and heteronuclear multiple-bond correlation spectroscopy (HMBC) was used to detect

heteronuclear correlations between hydrogens and carbons over longer ranges of about 2-4 bonds.

2.10 Software for *in silico* analysis

Software programs used for *in silico* analyses are listed in **Table 2.10**.

Table 2.10

List of software used for bioinformatics.

Software	Purpose	Reference
antiSMASH (v2.0-3.0) (http://antismash.secondarymetabolites.org/)	Genome mining of biosynthetic gene clusters	(Weber *et al.*, 2015)
BioEdit (v7.1.11) (http://www.mbio.ncsu.edu/bioedit/bioedit.html)	Alignment of DNA and protein sequences	(Hall, 1999)
BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi)	Homology searches of DNA and protein sequences	(Altschul *et al.*, 1990; Camacho *et al.*, 2009)
Buccaneer (ccp4i)	Automated structure building	(Cowtan, 2006)
CCP4i (v6.5.019) (http://www.ccp4.ac.uk/ccp4i_main.php)	Macromolecular crystallography suite	(Potterton *et al.*, 2003)
ChemDraw (v13) (http://www.cambridgesoft.com)	Property prediction of chemicals	Cambridgesoft (Cambridge, USA)
DNAPlotter (v1.11) (http://www.sanger.ac.uk/science/tools/dnaplotter)	Genome visualisation	(Carver *et al.*, 2009)
EBI-PISA (http://www.ebi.ac.uk/pdbe/pisa/)	Protein, interfaces, structures, and assemblies server	(Krissinel and Henrick, 2007)
ESPript (v3.0) (http://espript.ibcp.fr/ESPript/ESPript/)	Protein sequence alignment and secondary structure information server	(Robert and Gouet, 2014)
Gepard (v1.30) (http://mips.gsf.de/services/analysis/gepard)	Dotplot calculation of DNA sequences	(Krumsiek *et al.*, 2007)
Mega (v5.05) (http://www.megasoftware.net/)	Alignment of DNA and protein sequences	(Tamura *et al.*, 2011)
MolProbity (http://molprobity.biochem.duke.edu/)	Protein X-ray structure validation server	(Davis *et al.*, 2007; Chen *et al.*, 2010)
Parrot (ccp4i)	Density modification	(Cowtan, 2010)

Software	Purpose	Reference
PerlPrimer (v1.1.21) (http://perlprimer.sourceforge.net)	Oligonucleotide design for PCR	(Marshall, 2004)
Phaser (ccp4i)	Molecular replacement	(McCoy *et al.*, 2007)
Phenix (https://www.phenix-online.org/)	Automated structure determination; simulated annealing refinement	(Adams *et al.*, 2010)
Phyre2 (v2.0) (http://www.sbg.bio.ic.ac.uk/phyre2)	Protein folding recognition server	(Kelley and Sternberg, 2009; Kelley *et al.*, 2015)
Pfam (v26.0; Nov. 2011) (http://pfam.sanger.ac.uk/)	Annotation of protein domains	(Finn *et al.*, 2010)
Prodigal (v1.20) (http://prodigal.ornl.gov/)	ORF prediction for *Streptomyces* DNA	(Hyatt *et al.*, 2010)
PyMOL (v1.7.1) (http://sourceforge.net/projects/pymol/)	Visualisation of protein structures	The PyMOL Molecular Graphics System, Schrödinger, LLC.
RAST (v2.0) (http://rast.nmpdr.org/)	Prokaryotic genome annotation server	(Aziz *et al.*, 2008)
Refmac5	Restrained refinement	(Murshudov *et al.*, 1997)
ShelX suite (http://shelx.uni-ac.gwdg.de/SHELX/)	Experimental phasing of macromolecules	(Sheldrick, 2008)
SigmaA (ccp4i)	Combining of model and experimental phases	(Read, 1986)
Smart (v6.0 and v7.0) (http://smart.embl-heidelberg.de/)	Protein domain annotation server	(Schultz *et al.*, 1998; Letunic *et al.*, 2012)
WinCoot (v0.7), **Coot** (v0.8.2) (http://www.ysbl.york.ac.uk/~emsley/coot)	Structural alignment of proteins	(Emsley and Cowtan, 2004; Emsley *et al.*, 2010)
wwPDB (http://wwpdb-validation.wwpdb.org/validservice/)	Protein X-ray structure validation server	(Berman *et al.*, 2003)

Chapter 3
Heterologous expression of simocyclinones (Schäfer et al., 2015)

3.1 Introduction 111

3.2 Isolation of the simocyclinone (*sim*) biosynthetic gene cluster .. 112

3.3 Optimisation of the heterologous system 114

3.4 Discussion..... 118

3.5 Acknowledgements 119

3.1 Introduction

Understanding the complex biosynthesis of secondary metabolites such as the simocyclinones requires the genetic manipulation of the corresponding biosynthetic gene clusters. However, the natural producer of the simocyclinones, *S. antibioticus* Tü6040, grows very slowly and is particularly challenging to manipulate genetically. Although there are a few reports of successful directed mutations of the *sim* gene cluster in *S. antibioticus* Tü6040 (Galm *et al.*, 2002; Trefzer *et al.*, 2002; Galm, 2003), the procedures used had low and inconsistent success rates. Therefore, heterologous expression and manipulation of the *sim* cluster was desirable.

The *sim* gene cluster was initially isolated on four overlapping cosmids (Trefzer *et al.*, 2002). This was sufficient for sequencing of the ~72-kb *sim* cluster but did not permit straightforward reconstruction of the cluster as a single contiguous clone. A newly developed approach allows cloning and manipulation of genomic fragments of up to 150 kb in phage artificial chromosomes (PACs) (Jones *et al.*, 2013), enabling heterologous expression of large gene clusters such as the *sim* gene cluster.

Here I show that the *sim* gene cluster can be cloned on a single PAC, which facilitates genetic manipulations and product analysis by using a heterologous system. I show that *S. coelicolor* M1152 is able to produce simocyclinones heterologously and that overexpression of the simocyclinone-specific efflux pump SimEx1 improves growth and survival of M1152, most likely by increased export of simocyclinones.

3.2 Isolation of the simocyclinone (*sim*) biosynthetic gene cluster

The entire *sim* gene cluster was isolated on a single phage artificial chromosome (PAC) clone (PAC-12I) after PCR screening of a genomic library generated in *E. coli* according to a recently developed protocol for working with large gene clusters (Jones *et al.*, 2013). The genomic library was generated and screened by Bio S&T (Montreal, Canada). Subsequent sequencing of the 95-kb insert revealed the presence of the whole 72-kb minimal *sim* gene cluster consisting of 49 genes, flanked by genomic regions of ~19 kb and 4 kb on either side (**Figure 1.6** and **Figure 3.1**).

Comparison of the sequencing results obtained from PAC-12I with previously published sequences of the *sim* cluster (Galm *et al.*, 2002; Trefzer *et al.*, 2002) revealed no significant differences (i.e. no frameshifts or premature stop codons in biosynthetic genes).

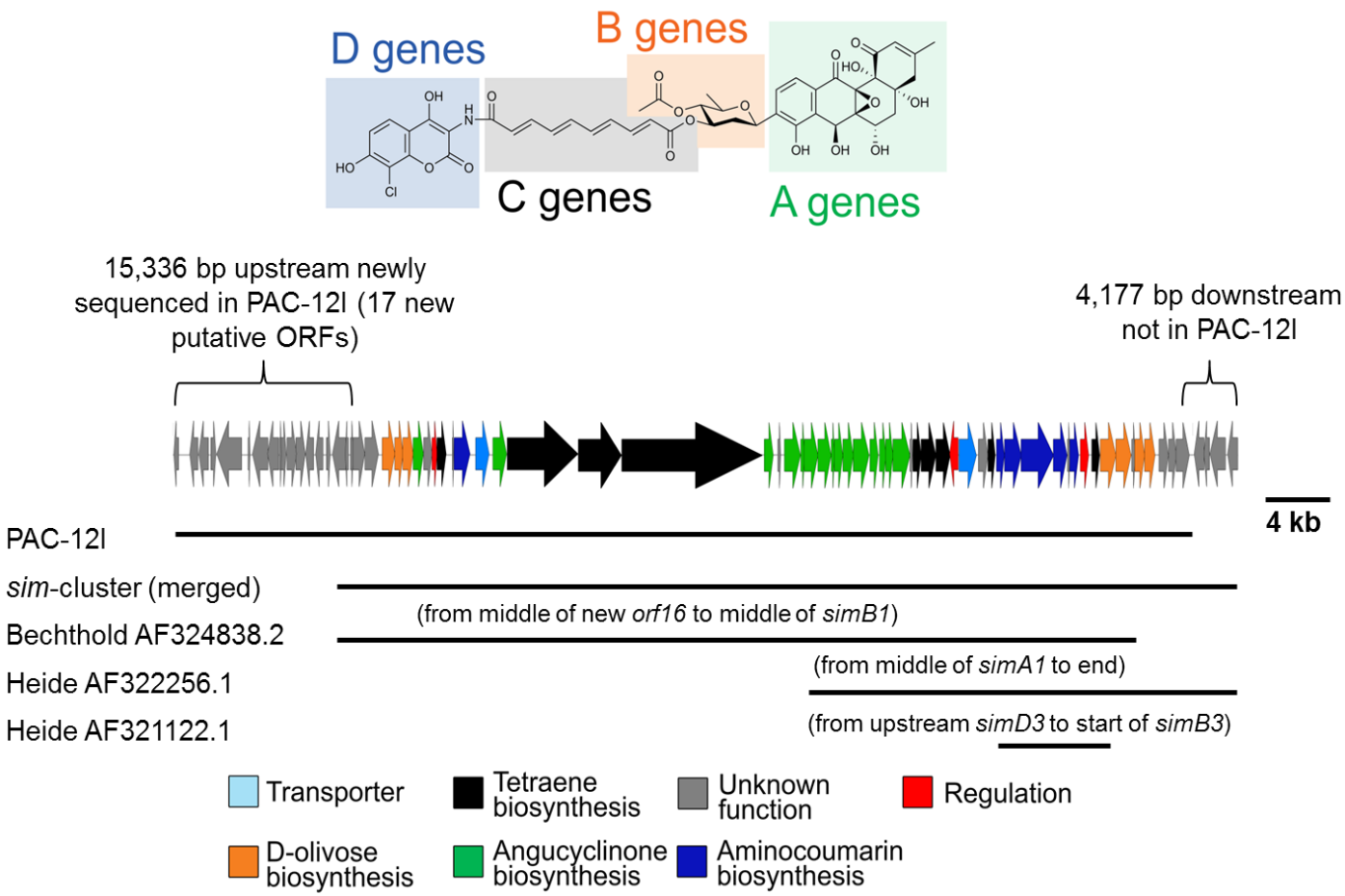


Figure 3.1

Comparison of sequenced simocyclinone (*sim*) gene cluster fragments.

The *sim* gene cluster has been sequenced independently by two research groups (Galm *et al.*, 2002; Trefzer *et al.*, 2002). The newly sequenced genomic DNA fragment in PAC-12I covers the entire minimal *sim* gene cluster as well as ~19 kb and 4 kb sequence on either side (grey open reading frames). For gene annotations see **Figure 1.6**.

3.3 Optimisation of the heterologous system

Clone PAC-12I was conjugated into the heterologous host *S. coelicolor* M1152 using the methylation-deficient *E. coli* strain ET12567 carrying the self-transmissible helper plasmid pR9406. *S. coelicolor* M1152 was chosen as heterologous host because it is a fast growing strain specifically engineered for improved heterologous production and analysis of secondary metabolites. The strain has no antimicrobial activity and a simplified extracellular metabolite profile for high-performance liquid chromatography (HPLC) and mass spectrometry analyses due to the deletion of four endogenous secondary metabolite gene clusters (Gomez-Escribano and Bibb, 2011). Beside the deletion of its endogenous gene clusters, a point mutation in the *rpoB* gene encoding for the RNA polymerase β-subunit further contributes to a pleiotropic increase in expression levels of secondary metabolite gene clusters (Gomez-Escribano and Bibb, 2011).

Unexpectedly, introduction of PAC-12I dramatically reduced growth and sporulation of *S. coelicolor* M1152. Several attempts were made to improve the viability of the heterologous host and overcome the toxicity of simocyclinones (see **chapter 6**). The best results were achieved after overexpression of the simocyclinone export protein SimEx1. Overexpression of the simocyclinone efflux pump SimEx1 has been shown previously to confer resistance against externally added SD8. Resistance of *S. lividans* against SD8 increased from an MIC of 2 µg/mL to 65 µg/mL after introduction of an integrative plasmid (pIJ10480) carrying *simEx1* expressed from the strong constitutive promoter *ermEp** (*ermEp*::simEx1*) (Le *et al.*, 2009). Consequently, pIJ10480 was introduced into M1152, resulting in strain *S. coelicolor* M1152ex1, which improved resistance against externally added SD8 ~100-fold from 1 µg/mL to 100 µg/mL. Indeed, the modified heterologous host M1152ex1 showed improved growth and sporulation in the presence of the *sim* gene cluster (i.e. introduction of PAC-12I) compared to M1152 (**Figure 3.2**). Normal growth was observed for M1152ex1 and M1152ex1 with a control construct lacking the *sim* gene cluster (PAC-12I∅sim) (**Figure 3.2**). Nevertheless, growth of M1152ex1 expressing the

sim gene cluster remained impaired, raising further questions about the basis of resistance in the native producer *S. antibioticus* Tü6040.

Strain M1152ex1 with PAC-12I was grown in defined medium (Tu2 medium) optimised for SD8 production with L-glutamine as nitrogen source and glycerol as carbon source (Schimana *et al.*, 2001). Production of SD8 was generally two- to five-fold higher in the heterologous host M1152ex1 (24-56 mg/mL) compared to *S. antibioticus* Tü6040 (9-14 mg/L). As expected, no simocyclinones were detectable from M1152ex1 with a control construct lacking the *sim* gene cluster (PAC-12I∅sim). All cultures carrying the control construct PAC-12I∅sim were unimpaired in growth, showing that the *sim* gene cluster is responsible for the observed toxicity.

Heterologous production of simocyclinones in M1152ex1 was monitored over 14 days to determine the time range of maximum production. Optimum yield was achieved after 6 days of growth in production medium and correlated with the growth of the heterologous host (**Figure 3.2** and **Figure 3.3**). In comparison, *S. antibioticus* Tü6040 produced simocyclinones much earlier and reached maximum production levels after only 3 days of growth, as shown previously (Schimana *et al.*, 2000). Since strain M1152ex1 gave best results for heterologous expression of simocyclinones, this strain was used for all further experiments in this study.

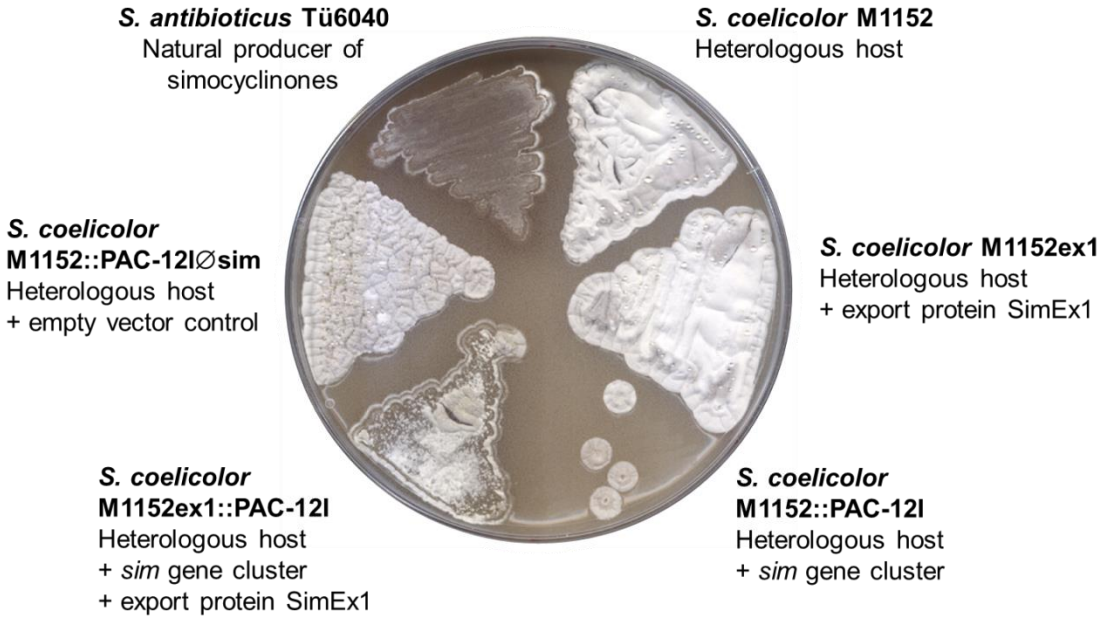


Figure 3.2

Growth of *Streptomyces* strains.

Growth of the natural simocyclinone producer *S. antibioticus* Tü6040 in comparison with the heterologous host *S. coelicolor* M1152. Growth and sporulation of M1152 was strongly impaired in presence of the *sim* gene cluster, which could be partially restored by overexpression of the simocyclinone efflux pump SimEx1. Strains were grown on SFM agar for 9 days at 30 °C.

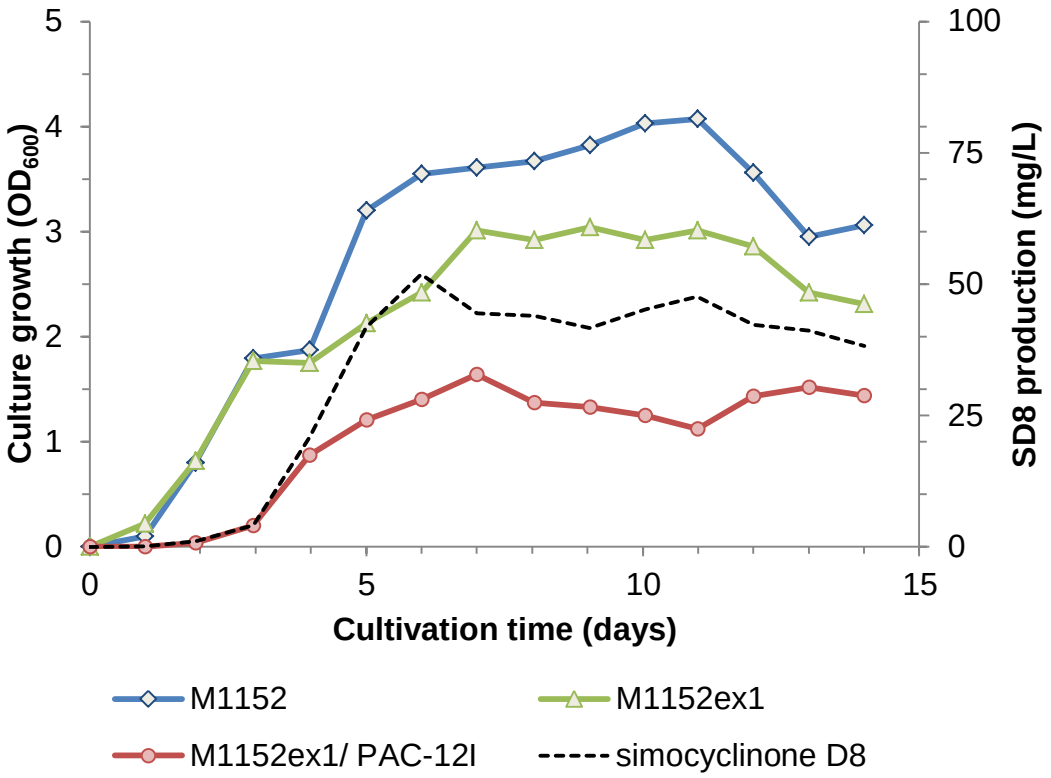


Figure 3.3

Growth and SD8 production of the heterologous host.

Growth of *S. coelicolor* M1152, M1152ex1 and M1152ex1 carrying the *sim* gene cluster (PAC-12I) was monitored over 14 days. Biosynthesis of SD8 (black line) by strain M1152ex1/ PAC-12I was quantified by LC/MS. Simocyclinones were isolated as described and SD8 was quantified by measuring the peak area from the UV-vis chromatogram from LC-MS analysis. Cultures were inoculated with 10^7 spores (M1152 and M1152ex1) or 10^8 spores (M1152ex1/ PAC-12I).

3.4 Discussion

Heterologous expression is a convenient tool for structure–function analysis of secondary metabolites, and improved derivatives of *S. coelicolor* and *Streptomyces avermitilis* have recently been developed for this purpose (Komatsu *et al.*, 2010; Gomez-Escribano and Bibb, 2011; Komatsu *et al.*, 2013). Here, I demonstrate the utility of this approach for the manipulation and functional analysis of simocyclinone, an antibiotic naturally made by *S. antibioticus* Tü6040, a strain that is particularly challenging for genetic manipulation. Despite the large size of the *sim* gene cluster (49 genes covering ~72 kb), it was straightforward to isolate a PAC clone carrying the entire cluster and to introduce it by conjugation into an engineered *S. coelicolor* host, where the clone integrated irreversibly and replicated stably as part of the chromosome. The same approach was recently used for the successful heterologous expression in *S. coelicolor* of the 83.5-kb FK506 gene cluster from *Streptomyces tsukubaensis* (Jones *et al.*, 2013). The engineered *S. coelicolor* host often gives a greater yield than the natural producer (Gomez-Escribano and Bibb, 2011; Jones *et al.*, 2013), and this was also true for simocyclinone. A further advantage of this approach is that genetic modification of the target cluster can be carried out in *E. coli*, giving full access to the advanced recombineering tools available in this host. The explosion of whole genome sequencing has revealed tens of thousands of new secondary metabolic gene clusters (Medema *et al.*, 2011; Cimermancic *et al.*, 2014), many in rare and difficult-to-culture actinomycetes for which no genetic tools have been developed. As a consequence, the heterologous expression approach exemplified here is an attractive option for characterizing many of these clusters and is likely to be of increasing importance in the future.

3.5 Acknowledgements

I would like to thank Bio S&T (Montreal, Canada) for generating and screening a genomic library using genomic DNA isolated from *S. antibioticus* Tü6040 to isolate the simocyclinone biosynthetic gene cluster. I would also like to thank the staff at the University of Cambridge for sequencing PAC clone PAC-12I carrying the entire *sim* gene cluster.

Chapter 4

SimC7 is a novel NAD(P)H-dependent ketoreductase essential for the antibiotic activity of the DNA gyrase inhibitor simocyclinone (Schäfer et al., 2015)

4.1 Introduction 121

4.2 Structural analysis of the novel simocyclinone 7-oxo-SD8 122

4.3 Complementation of *simC7* mutant 146

4.4 SimC7 is an NAD(P)H-dependent ketoreductase that acts on the angucyclic polyketide 148

4.5 The C-7 hydroxyl group is required for the antibiotic activity of simocyclinone 150

4.6 Discussion 154

4.7 Acknowledgements 158

4.1 Introduction

The 49-gene simocyclinone (*sim*) biosynthetic cluster of *S. antibioticus* Tü6040 has been sequenced and a hypothetical biosynthetic pathway has been proposed (Galm *et al.*, 2002; Trefzer *et al.*, 2002). Although two *sim* cluster transcriptional regulators have been studied (Le *et al.*, 2009; Le *et al.*, 2011a; Le *et al.*, 2011b; Horbal *et al.*, 2012), to date, only one biosynthetic enzyme has been characterised due to the genetic inaccessibility of the native producer and until recently limitations in available techniques to analyse large biosynthetic gene clusters heterologously. SimD5 (SimL) catalyses the presumed last step in the pathway, functioning as an amide-bond-forming ligase that attaches the aminocoumarin to the tetraene linker (Luft *et al.*, 2005; Pacholec *et al.*, 2005; Anderle *et al.*, 2007).

Trefzer et al (2002) proposed that the tetraene linker in SD8 would be the product of a modular type I polyketide synthase, SimC1ABC, working *in trans* with two monofunctional enzymes, SimC6 and SimC7. One of these monofunctional enzymes, SimC7, a member of the short-chain dehydrogenase/reductase (SDR) superfamily, was proposed to supply the dehydratase activity missing from two modules of the polyketide synthase (Trefzer *et al.*, 2002).

Here I determine the true function of SimC7 experimentally and show that it is in fact an NAD(P)H-dependent ketoreductase that acts on the C-7 carbonyl group of the angucyclinone during simocyclinone biosynthesis. I show that deletion of *simC7* from the *sim* gene cluster results in production of 7-oxo-SD8, a novel simocyclinone, with a non-reduced carbonyl group at position 7 of the angucyclinone of simocyclinones. I manually optimise the purification of 7-oxo-SD8 and verify its structure using nuclear magnetic resonance (NMR) spectroscopy and high-resolution mass spectrometry (MS). Furthermore, I show that the C-7 hydroxyl group generated by SimC7 is essential for the antibiotic activity of simocyclinone and I propose that the inactivity of 7-oxo-SD8 arises from an intramolecular hydrogen bond between the non-reduced 7-carbonyl group and the neighbouring 8-hydroxyl group.

4.2 Structural analysis of the novel simocyclinone 7-oxo-SD8

In order to determine the role of SimC7 in simocyclinone biosynthesis, I created an in-frame deletion in *simC7* in PAC-12I using λ-Red-mediated PCR targeting, and introduced the resulting PAC clone (PAC-12IΔC7) into the optimised heterologous host *S. coelicolor* M1152ex1. Fermentation extracts of the resulting *simC7* mutant strain were analysed by HPLC and the major biosynthetic product had a later retention time than that of SD8 (**Figure 4.1**). The UV-vis spectrum of the new compound was similar to that of SD8 but with reduced absorbance at 275 nm and a slight shift in the absorbance maxima to 245 nm and 361 nm (**Figure 4.2**). Unexpectedly, the major biosynthetic product of the *simC7* mutant had a mass 2 Da lighter than SD8. All of the A-, B- and C-group simocyclinone intermediates made by the *simC7* mutant were also 2 Da lighter than the equivalent molecules made by the wild-type *sim* cluster, suggesting that the 2 Da difference lay in the angucyclic polyketide, and not the tetraene linker.

This suggested that one of the five hydroxyl groups present in the angucyclic polyketide moiety of SD8 remained as a carbonyl group in this compound, and that SimC7 acts as a ketoreductase. To test this hypothesis, I purified the new simocyclinone intermediate from 4 L culture and determined its structure. I manually optimised the purification of the new compound by using organic extraction with acidified ethyl acetate (pH 4), several pre-purification steps (filtering through glass wool, normal phase liquid chromatography, dissolving in acetone, size exclusion chromatography) followed by optimised preparative reverse-phase HPLC. For purification, the dried extract from 4 L culture medium (~1.5 g with ~15% compound) was dissolved in methanol, filtered through glass wool and purified by normal-phase liquid chromatography as described (Holzenkampfer *et al.*, 2002). The presence of the compound was verified by thin layer chromatography and LC-MS and selected fractions were pooled, dried *in vacuo*, and dissolved in acetone. The acetone-soluble fraction was dried and dissolved in methanol for size exclusion chromatography. Following elution using methanol, fractions were collected and the presence of the compound was confirmed

by LC-MS. Fractions containing the compound were combined, dried *in vacuo*, and the concentrate was dissolved in a methanol–acetone mix (1:1) and purified further by preparative HPLC, which I optimised for purification of the new compound using analytical HPLC. After freeze drying, I obtained ~63 mg of the new simocyclinone derivative 7-oxo-SD8 in a fine yellow powder (yield ~27%; purity >98%).

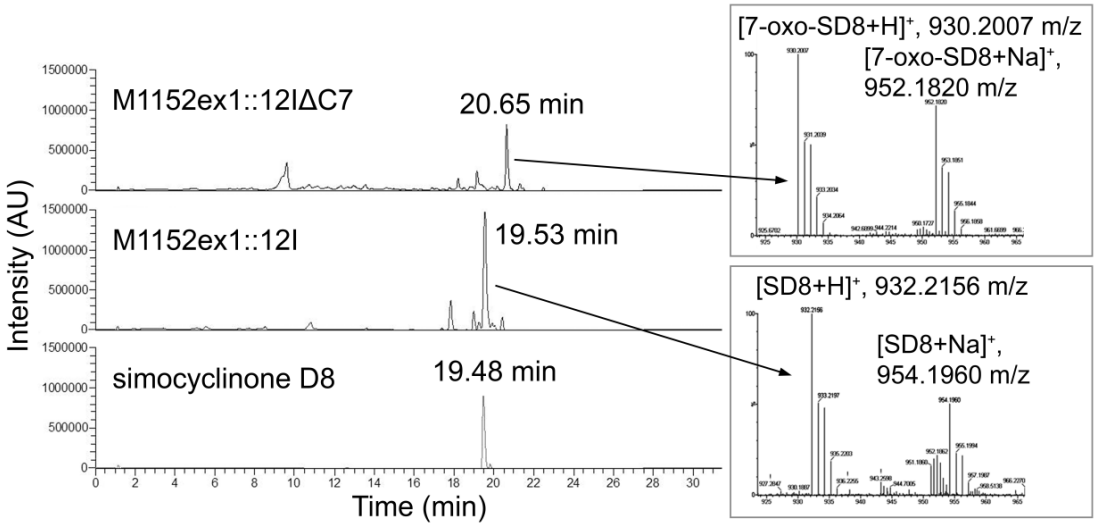


Figure 4.1

Analysis of simocyclinones produced by the *simC7* deletion mutant.

LC-MS-UV analysis of simocyclinones was carried out using culture extracts from the heterologous host *S. coelicolor* M1152ex1 expressing the *sim* gene cluster. Chromatograms (200-700 nm) of culture extracts from clones carrying the *simC7* mutant cluster on PAC-12IΔsimC7 (top) and the complete *sim* cluster on PAC-12I (middle), alongside an isolated standard of SD8 (bottom), are shown on the left side. Molecular weights were confirmed by high-resolution mass spectrometry (parent peaks are indicated). Cultures were grown for 6 days at 30 °C in production medium. A close-up of the total ion chromatogram (TIC) for 7-oxo-SD8 (top) and SD8 (bottom) is shown on the right side.

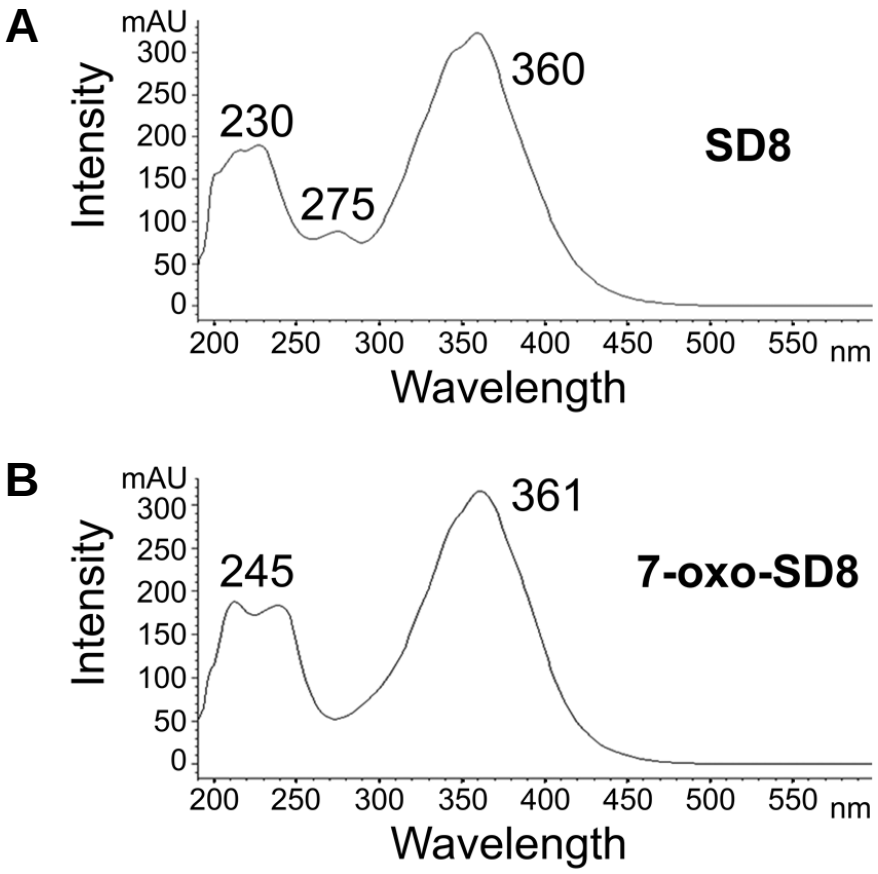


Figure 4.2

Absorbance spectra of simocyclinones.

The absorbance spectra of (A) SD8 and (B) 7-oxo-SD8 are shown with indicated wavelength of absorbance maxima. Compounds were dissolved in methanol (50 µg/mL) and were analysed by HPLC-UV at pH 7 as described (see **chapter 2.9.4**). The absorbance spectra of both compounds were extracted and compared at respective retention times. The extinction coefficient (ε) at 340 nm was 52,400 M-1 cm-1 for SD8 and 56,600 M-1 cm-1 for 7-oxo-SD8 using an Agilent 1200 series instrument with a cell path length of 10 mm at 50 µM sample concentration. In comparison, the extinction coefficient of NADH and NADPH is 6,220 M-1 cm-1 at 340 nm.

The molecular formula of the new compound was determined using high resolution mass spectrometry which confirmed it was missing two hydrogen atoms (*m/z* (ESI) for $C_{46}H_{40}ClNO_{18}$: calculated 930.2006, observed 930.2009, Δ = 0.3 ppm $[M+H]^+$) compared to SD8 (*m/z* (ESI) for $C_{46}H_{42}ClNO_{18}$: calculated 932.2163, observed 932.2156, Δ = -0.8 ppm $[M+H]^+$) (**Figure 4.1**). MS/MS analysis revealed similar fragmentation patterns for the new compound and SD8, with all ion fragments containing the angucyclic polyketide moiety having *m/z* ratios reduced by 2 Da compared to SD8. Notably, the ion fragment for the angucyclinone itself, which was only detectable at low abundance in the SD8 spectra, was one of the most abundant peaks for the new compound (**Figure 4.3**).

I analysed the structure of the new compound by one-dimensional (^{1}H-NMR and ^{13}C-NMR) and two-dimensional (COSY, TOCSY, HSQC and HMBC) NMR spectroscopy techniques. Signals from hydroxyl groups were determined by comparing the proton NMR spectra before and after addition of 5 µL deuterated water (D_2O) to the sample, which sequesters the signals of exchangeable hydrogens. Full NMR spectra can be found at the end of this chapter (**Figure 4.9-4.14**). Examination of the NMR spectra suggested the new molecule was structurally similar to SD8. However, the signal for the H-6 proton was shifted downfield 0.4 ppm and no signal could be observed for the H-7 proton consistent with a structural change in the angucyclic polyketide (**Figure 4.4**, **Figure 4.9** and **Table 4.1**). Further analysis of proton-proton and proton-carbon correlations from two-dimensional NMR data (COSY, TOCSY, HSQC and HMBC) enabled me to determine the position of the structural change in the new molecule (**Figure 4.5** and **Figure 4.6** and **Figure 4.10-4.13**). I confirmed appropriate proton-proton correlations for H-5 and H-6 which were supported by relevant proton-carbon correlations. In addition, I identified a signal shifted downfield for C-7 in the ^{13}C spectrum from 63.5 ppm to 194.9 ppm typical for a carbonyl group (**Figure 4.14** and **Table 4.2**) and cross-peaks to a C-7 resonance could be identified in the HMBC experiments (**Figure 4.5** and **Figure 4.12**).

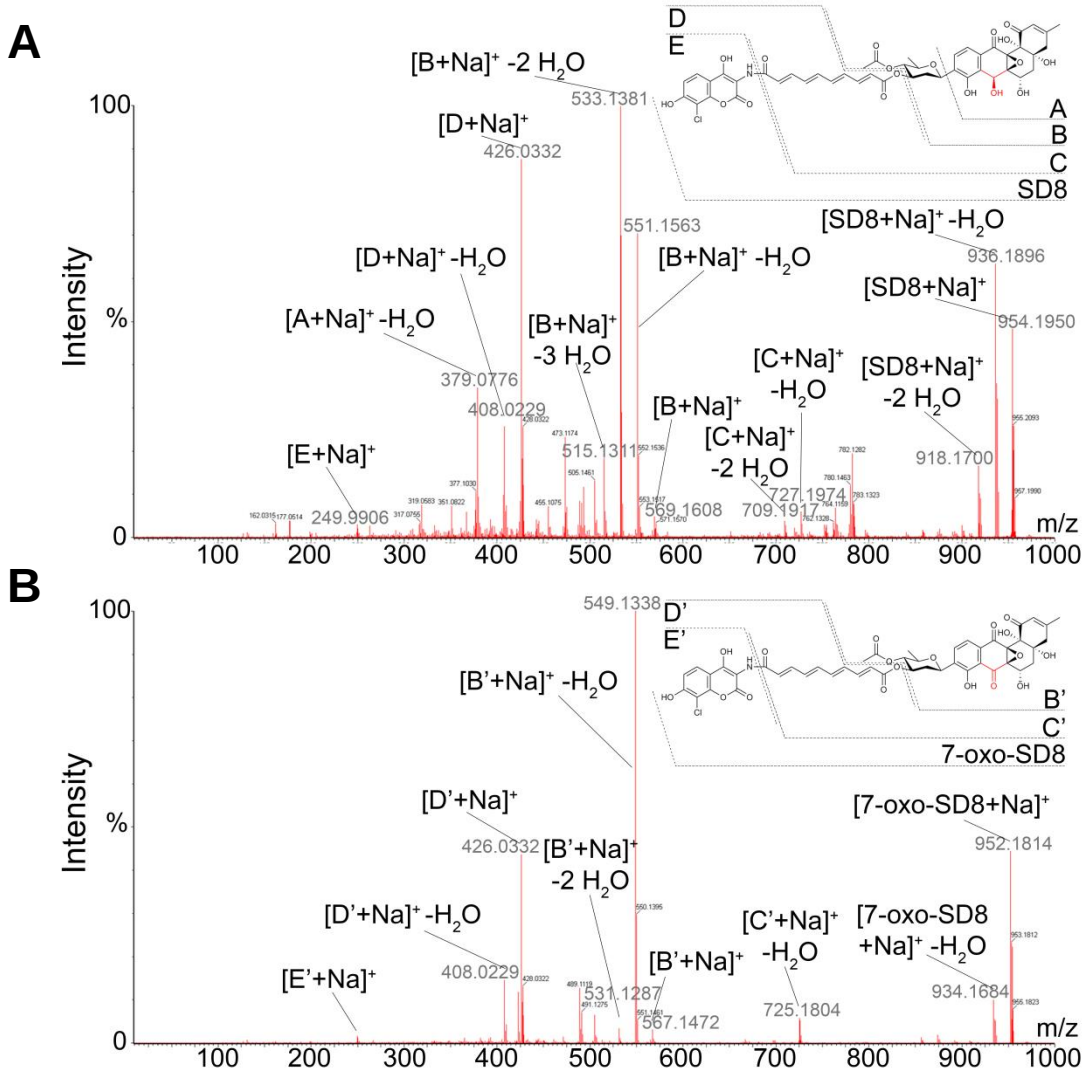


Figure 4.3

High resolution tandem mass spectra for SD8 and 7-oxo-SD8.

Fragmentation patterns are shown for the sodium adducts of (A) SD8 and (B) 7-oxo-SD8. Selected fragments are highlighted, consistent with the mass difference of 2 Da being located in the angucyclinone.

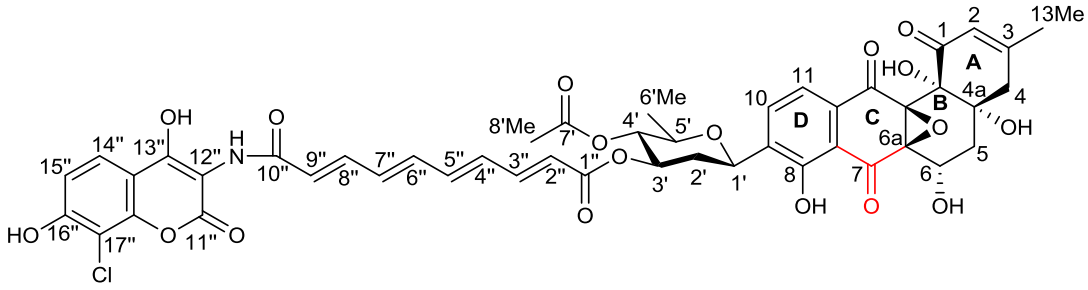


Figure 4.4

Numbering of carbon atoms in 7-oxo-SD8 for NMR spectroscopy.

Numbering of atoms was adopted from structural determination of SD8 for comparison (see **Table 4.1** and **Table 4.2**). The carbonyl group at position 7 in the angucyclinone is highlighted in red (Holzenkampfer *et al.*, 2002).

Table 4.1

^{1}H NMR data (700 MHz) for 7-oxo-SD8 in *d_6*-DMSO.

Data for SD8 were taken from (Holzenkampfer *et al.*, 2002) for comparison.

^{1}H NMR	SD8		7-oxo-SD8	
Proton(s)	δH (ppm)	δH (J in Hz)	δH (ppm)	δH (J in Hz)
6'-CH_3	1.16	d, 6.0	1.19	d, 6.02
2'-CH_2 (a)	1.55	ddd, 11.5, 11.5, 11.5 (ax)	1.59	dd, 11.73, 11.11
13-CH_3	1.81	s	1.86	s
5-H (a)	1.91	ddd, 13.5, 7.5, 1.0	1.92	dd, 14.59, 5.09
5-H (b)	1.98	m	2.08	m
8'-CH_3	2.01	s	2.03	s
4-CH_2 (a)	2.23	d, 19.0	2.38	d, 19.32
2'-CH_2 (b)	2.42	ddd, 1.0, 4.5, 11.5 (eq)	2.47	d, 5.65 (1)
4-CH_2 (b)	2.70	d, 19.0	2.64	d, 19.32
5'-H	3.78	dq, 6.0, 9.0 (ax)	3.87	dq, 6.09, 15.50
6-H	4.42	dd, 7.5, 7.5	4.82	m (2)
4'-H	4.74	dd, 9.0, 9.0 (ax)	4.77	dd, 9.41, 9.41
OH (a)	4.82	---	---	---
1'-H	4.96	dd, <1, 11.5 (ax)	5.05	dd, 0.94, 11.48
3'-H	5.18	ddd, 4.5, 9.0, 11.5 (ax)	5.23	m
12b-OH	5.27	n/a	5.88	s
6-OH	5.58	n/a	5.43	d, 7.78
7-H	5.75	s	---	---
2-H	5.90	s	5.93	q, 1.16 ($_2$J)
2''-H	6.00	d, 15.0	6.03	d, 15.06
9''-H	6.59	d, 15.0	6.58	d, 15.18
4''-H	6.61-6.73	m	6.60-6.77	m
7''-H	6.61-6.73	m	6.60-6.77	m
5''-H	6.79-6.88	m	6.80-6.92	m
6''-H	6.79-6.88	m	6.80-6.92	m
8''-H	7.23	m/ dd, 15.0, 11.0	7.23	m
14''-H	7.23	m	7.23	m
15''-H	7.23	m	7.23	m
3''-H	7.26	dd, 15.0, 11.0	7.23	m
11-H	7.33	d, 8.0	7.58	d, 7.91
10-H	7.45	d, 8.0	7.93	d, 8.03
OH (b)	9.58	n/a	9.81	s
OH (c)	9.85	n/a	10.31	s
NH	11.3	n/a	11.98	s
13''-OH	12.6	n/a	12.61	s

(1) Due to solvent overlap and poor resolution only a doublet observed.

(2) Due to overlap with 4'-H signal could not extract J value.

Table 4.2

^{13}C NMR data (700 MHz) for 7-oxo-SD8 in d_6-DMSO.

Data for SD8 were taken from (Holzenkampfer *et al.*, 2002) for comparison.

^{13}C NMR		SD8	7-oxo-SD8	^{13}C-labelled 7-oxo-SD8				
position	carbon	δC (ppm)	δC (ppm)	δC (ppm)	δC (J in Hz)	^{13}C-enrichment	peak	coupling *
6’	CH_3	17.70	18.21	17.73	---	---	-	
8’	CH_3	20.50	21.07	20.60	29.75	yes	-	7'
13	CH_3	23.10	23.78	23.31	16.81	yes	-	3
2’	CH_2	36.50	36.74	36.28	---	---	+	
5	CH_2	38.70	40.49	40.10	n/a	---	-	n/a (2)
4	CH_2	42.70	44.48	44.00	16.26	yes	+	4a
6	CH	62.60	61.66	61.19	15.40	yes	-	n/a (2)
7	CH or C	63.50	194.87	194.40	27.10	yes	+	6a
6a	C	65.40	64.20	63.72	27.18	yes	+	7
12a	C	65.70 (1)	65.40	64.95	26.87	yes	+	12
1’	CH	70.60	70.50	71.38	---	---	-	
3’	CH	71.60	71.85	71.38	---	---	-	
4a	C	72.00	72.73	72.25	16.30	yes	+	4
5’	CH	73.10	73.64	73.17	---	---	-	
4’	CH	74.10	74.43	73.95	---	---	-	
12b	C	75.10	76.31	75.84	18.86	yes	+	1
12’’	C	105.30	105.47	104.67	n/a	---	+	
17’’	C	114.40	113.77	114.96	---	---	+	
13a’’	C	115.30	115.85	115.43	---	---	+	

^{13}C NMR		SD8	7-oxo-SD8	^{13}C-labelled 7-oxo-SD8				
position	carbon	δC (ppm)	δC (ppm)	δC (ppm)	δC (J in Hz)	^{13}C-enrichment	peak	coupling *
15''	CH	115.80	116.44	115.92	---	---	-	
11	CH	118.10	117.14	119.37	30.05	yes	-	11a
14''	CH	119.60	119.85	n/a	---	---	-	
2''	CH	121.50	120.01	121.40	38.43	yes	-	1''
2	CH	121.80	120.62	121.80	29.22	yes (3)	-	--- (3)
9''	CH	124.60	122.69	121.78	32.52	yes	-	10''
7a	C	124.80	125.69	125.39	32.23	yes	-	8
10	CH	125.80	130.44	129.96	29.22	yes	+	9
11a	C	127.60	132.50	132.03	30.02	yes	-	11
4''	CH	133.10	133.58	133.07	26.00	yes	-	3''
9	C	134.20	134.54	136.96	29.22	yes	+	10
7''	CH	134.40	135.10	134.67	26.74	yes	-	8''
6''	CH	139.00	137.43	138.93	25.64	yes	+	5''
5''	CH	140.20	139.51	140.83	25.64	yes	-	6''
8''	CH	141.10	140.96	140.48	26.74	yes	-	7''
3''	CH	144.40	141.40	144.69	26.00	yes	-	4''
17a''	C	145.60	145.14	145.84	---	---	+	
16''	C	151.00	146.22	150.85	---	---	+	
8	C	152.60	151.41	156.73	32.23	yes	+	7a
3	C	157.70	157.74	158.68	16.97	yes	+	14
11''	CO	158.70	159.15	159.05	---	---	+	
13''	C	185.70	159.45	159.05	---	---	+	
1''	CO	165.20	165.80	165.33	38.40	yes	+	2''

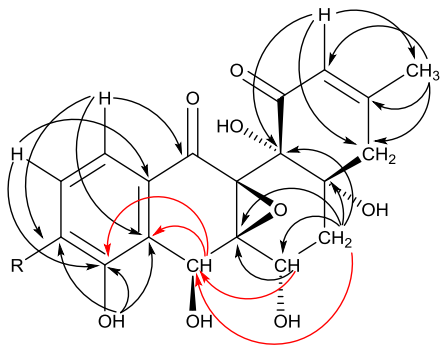
^{13}C NMR		**SD8**	**7-oxo-SD8**	**^{13}C-labelled 7-oxo-SD8**				
position	**carbon**	**δC (ppm)**	**δC (ppm)**	**δC (ppm)**	**δC (J in Hz)**	**^{13}C-enrichment**	**peak**	**coupling ***
10''	CO	166.10	166.55	165.97	32.52	yes	+	9''
7'	CO	169.60	170.27	169.81	29.75	yes	+	8'
12	CO	190.00	188.47	188.00	26.87	yes	+	12a
1	CO	196.00	195.52	195.06	18.86	yes	+	12b

* Coupling carbon atom.

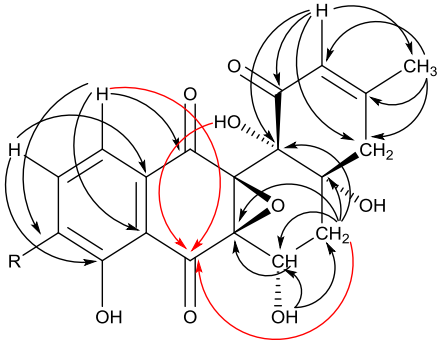
(1) Only visible by feeding of [1,3-$^{13}C_2$] malonic acid.

(2) Overlap of solvent signal with signal for C-5.

(3) ^{13}C-enrichment but no coupling partner.



SD8



7-oxo-SD8

Figure 4.5

Selected HMBC correlations for the angucyclic polyketide moiety of SD8 and of the new simocyclinone intermediate 7-oxo-SD8.

Key correlations for the C-7 position are depicted in red. R includes the deoxysugar, tetraene linker and the aminocoumarin.

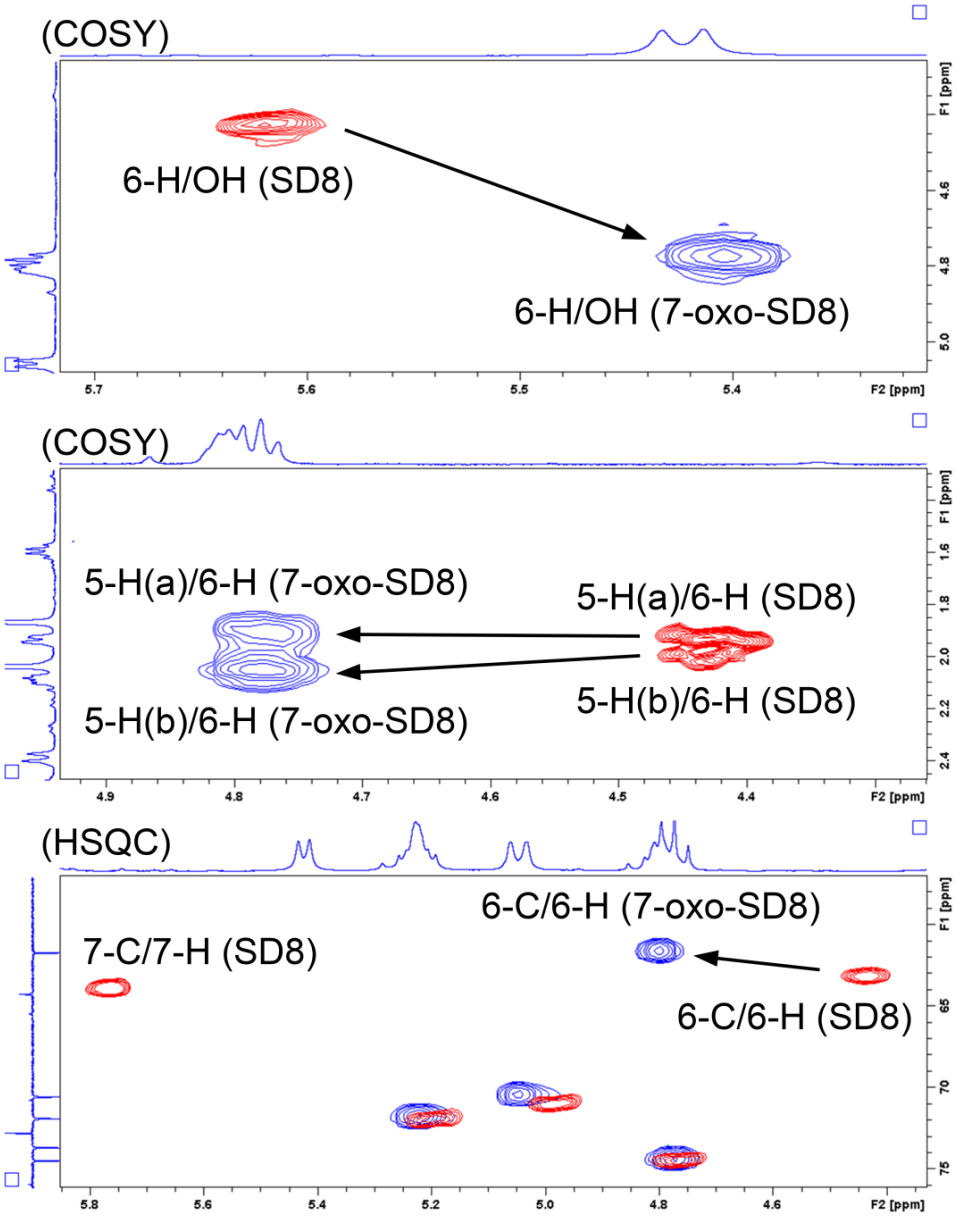


Figure 4.6

Differences in 2D NMR spectra of SD8 and 7-oxo-SD8.

Shifted peaks and cross-peaks are shown of SD8 (red) and 7-oxo-SD8 (blue) from 2D NMR experiments (correlated spectroscopy and heteronuclear single quantum coherence). Projections of proton and carbon signals of 7-oxo-SD8 are indicated for each map. These are entirely consistent with the introduction of a carbonyl group at the C-7 position.

Because signals for ternary carbons without direct hydrogen contacts are generally very weak, I decided to confirm the C-7 carbonyl carbon by stable isotope labelling with carbon 13 (^{13}C). Therefore, I grew the *simC7* mutant strain in presence of [1,2-$^{13}C_2$] acetate for 6 days and isolated the new compound by extraction with ethyl acetate (~285 mg with ~12% compound), dissolving the dried extract in acetone, and purification of the acetone-soluble fraction by preparative HPLC. After freeze-drying, I obtained 26 mg of the ^{13}C-labelled 7-oxo-SD8 (yield ~78%; purity >99%). According to the proposed biosynthetic pathway and previous labelling experiments I expected incorporation of labelled acetate into 31 carbons of the angucyclinone, the tetraene, and the acetyl group attached to the deoxysugar (Holzenkampfer and Zeeck, 2002). The remaining 15 carbons of the aminocoumarin and the deoxysugar would not be labelled because their carbon skeletons originate from amino acid and sugar metabolism, respectively. Indeed, the feeding experiment resulted in ~30% incorporation of ^{13}C-labelled acetate based on an observed average mass shift of ~10 m/z (detectable mass shift of up to 18-19 m/z) from LC-MS analysis (**Figure 4.7**). As expected, NMR analysis revealed incorporation of ^{13}C-labelled acetate into the angucyclinone, the tetraene, and the acetyl group attached to the deoxysugar but no incorporation into the deoxysugar itself or the aminocoumarin (**Figure 4.8a** and **Table 4.2**). Importantly, the NMR data confirmed the presence of a carbonyl group at position 7 in the angucyclinone. Identical carbon-carbon coupling constants were observed for the shifted signal peak of C-7 and the signal peak for C-6a (**Figure 4.8b**, **Figure 4.15** and **Table 4.2**).

In conclusion, I have shown that the product of the *simC7* mutant had a carbonyl group instead of a hydroxyl group at position C-7 in the angucyclinone. The new simocyclinone intermediate was named 7-oxo-simocyclinone D8 (7-oxo-SD8) according to the structural change compared to SD8 (**Figure 4.4** and **Figure 4.5**).

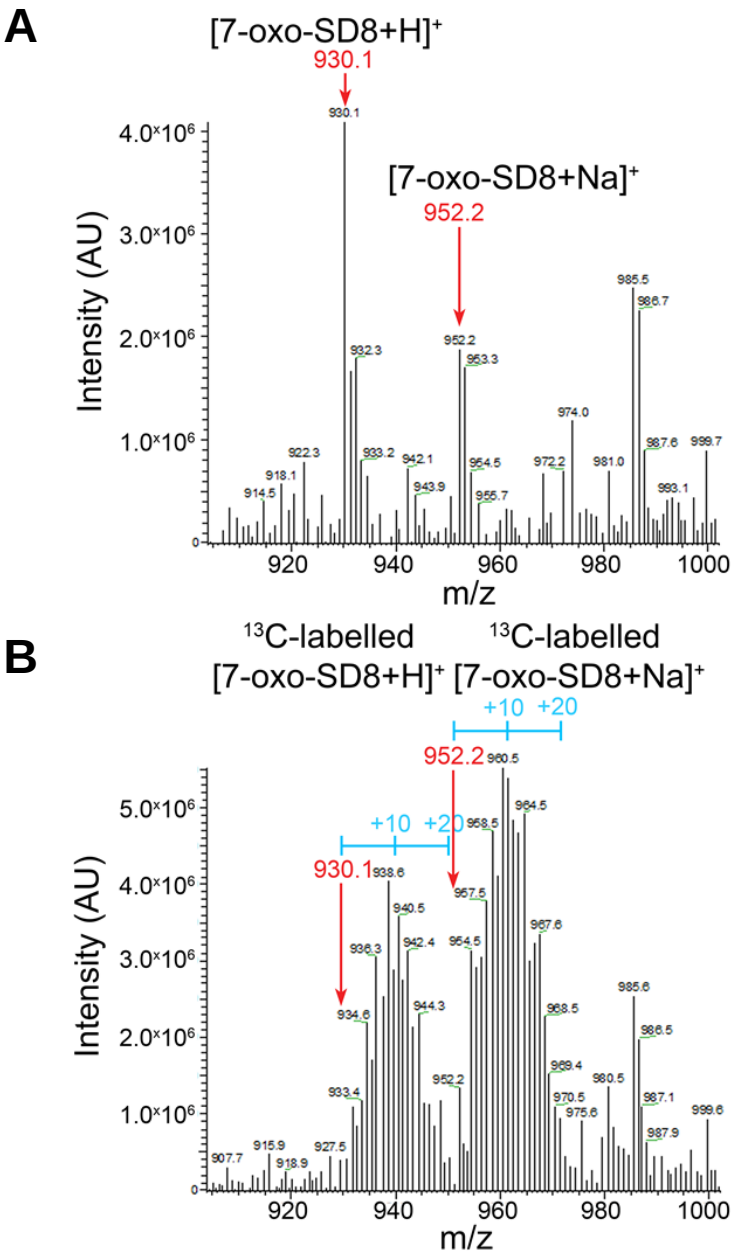


Figure 4.7

Isotope patters of (A) purified 7-oxo-SD8 and (B) ^{13}C-enriched 7-oxo-SD8.

Data were extracted from the total ion chromatogram at 20 min retention time after LC-MS analysis as described. The parent ion of proton and sodium adducts are indicated in red for unlabelled 7-oxo-SD8.

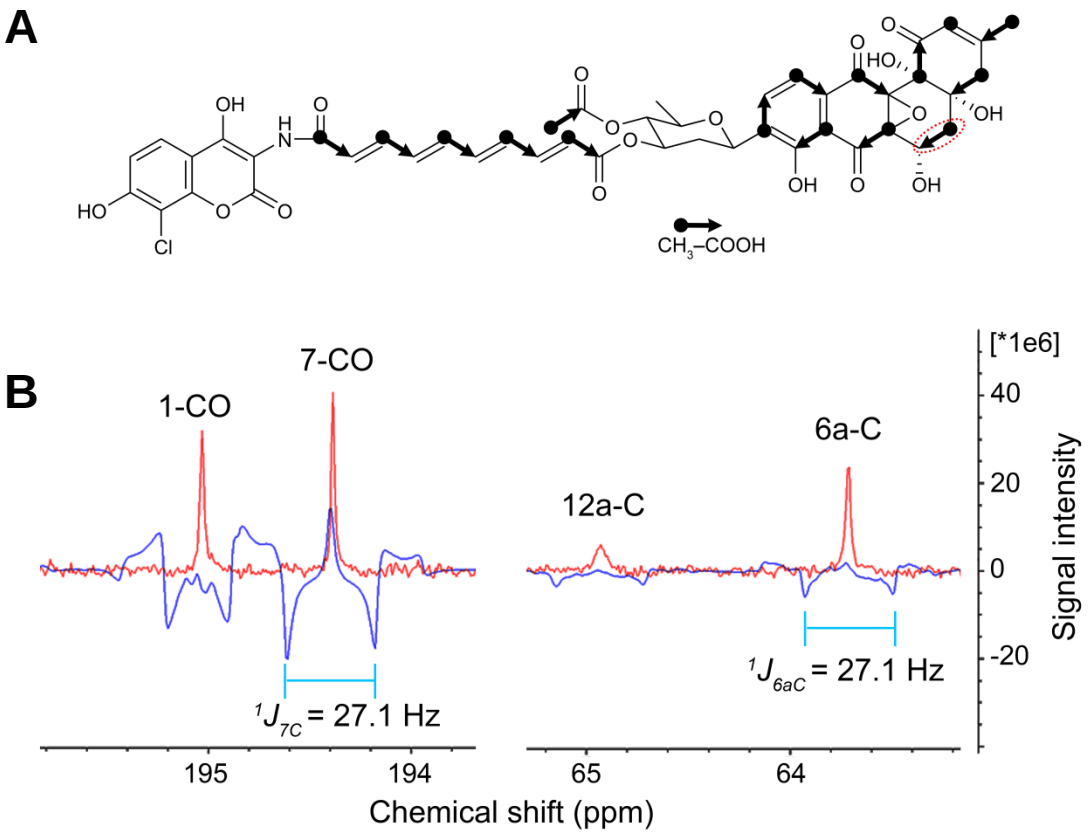


Figure 4.8

Structural validation of ^{13}C-labelled 7-oxo-SD8.

(A) Incorporation pattern of [1,2-$^{13}C_2$] acetate (arrows) into 7-oxo-SD8. The coupling constant for the carbon at position 5 could not be identified due to signal overlap with the solvent DMSO-d_6 (red circle). No coupling partner could be identified for the carbon at position 2. (B) ^{13}C-NMR spectra of unlabelled (red) and labelled (blue) 7-oxo-SD8. The carbon signals of unlabelled and labelled 7-oxo-SD8 have identical chemical shifts. Incorporation of [1,2-$^{13}C_2$]-acetate as shown in (A) causes coupling of intact acetate units, resulting in identical coupling constants for the respective two neighbouring carbons. Identical coupling constants were observed for the carbonyl group at position 7 and the ternary carbon at position 6a (see **Table 4.2**).

It is worth mentioning that the absence of a coupling partner for carbon C-2 in the angucyclic polyketide supports a decarboxylation event at this position, which is most likely carried out by the predicted decarboxylase SimA12 (**Figure 4.8a**, **Figure 4.15** and **Table 4.2**). Although incorporation of ^{13}C-labelled precursors has been done for SD8, the use of [1,3-$^{13}C_2$] malonic acid and [2-^{13}C] malonic acid did not allow any conclusion about the decarboxylation side due to missing information about carbon-carbon coupling and the direct use of acetate was not possible as it inhibited simocyclinone production in *S. antibioticus* Tü6040 (Holzenkampfer *et al.*, 2002). The incorporation pattern of ^{13}C-labelled acetate into the angucyclinone of simocyclinones was identical to the incorporation pattern observed for landomycin A (Weber *et al.*, 1994) and urdamycin A (Rohr *et al.*, 1989) suggesting a common biosynthetic pathway for the formation of the benz[*α*]anthracene backbone.

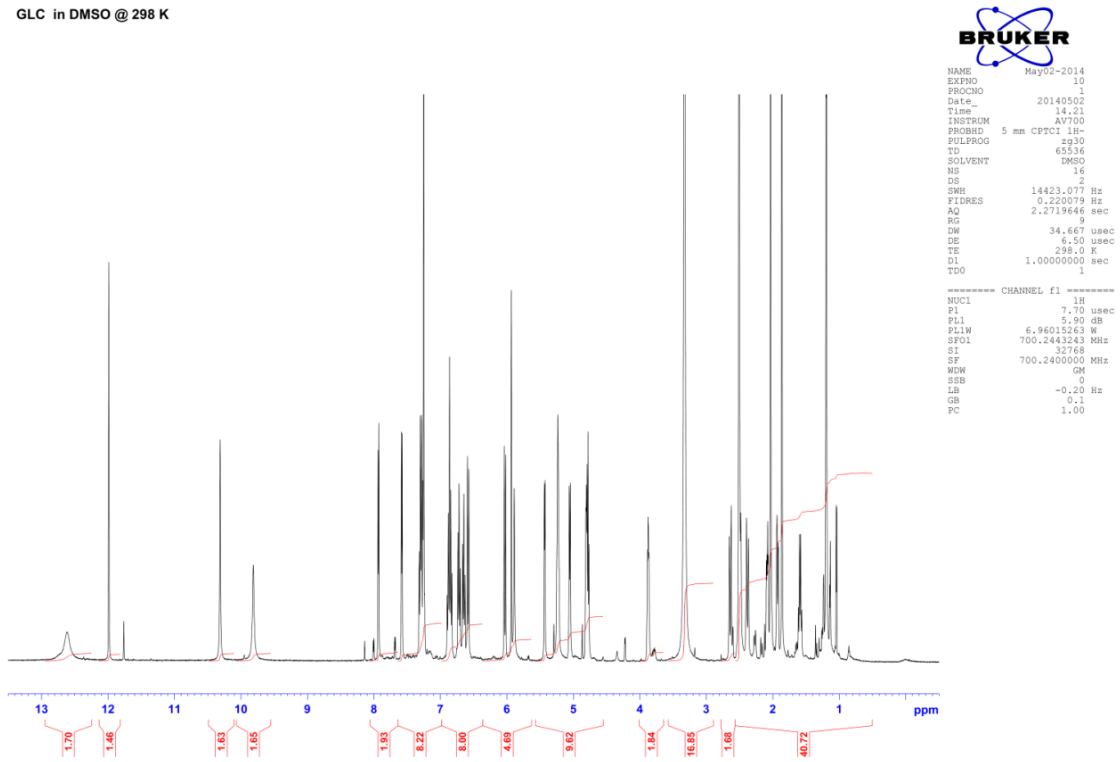


Figure 4.9

^{1}H NMR spectrum of 7-oxo-SD8.

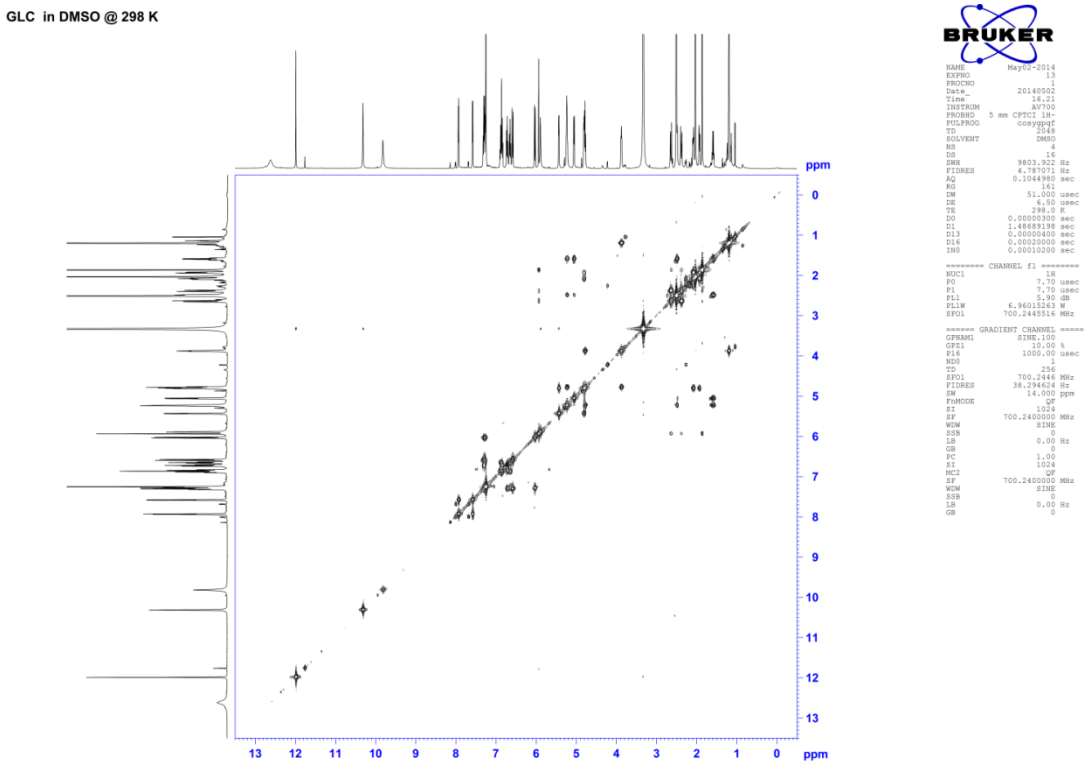


Figure 4.10

Correlated spectroscopy (COSY) spectrum of 7-oxo-SD8.

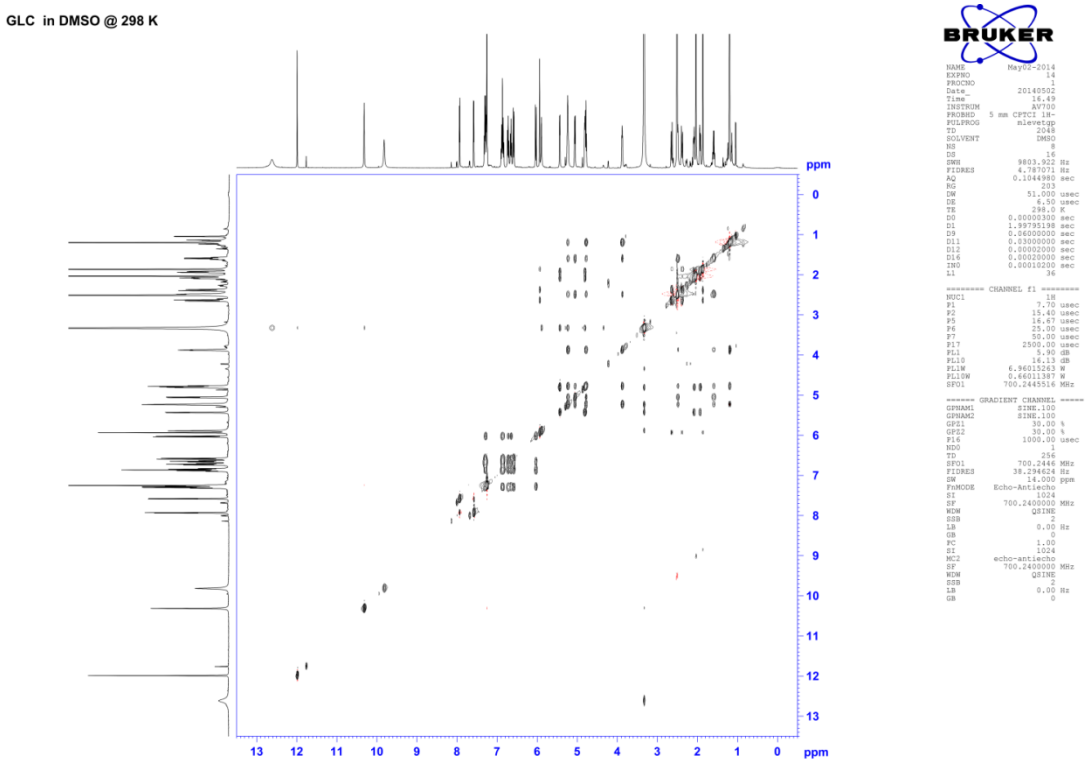


Figure 4.11

Total correlated spectroscopy (TOCSY) spectrum of 7-oxo-SD8.

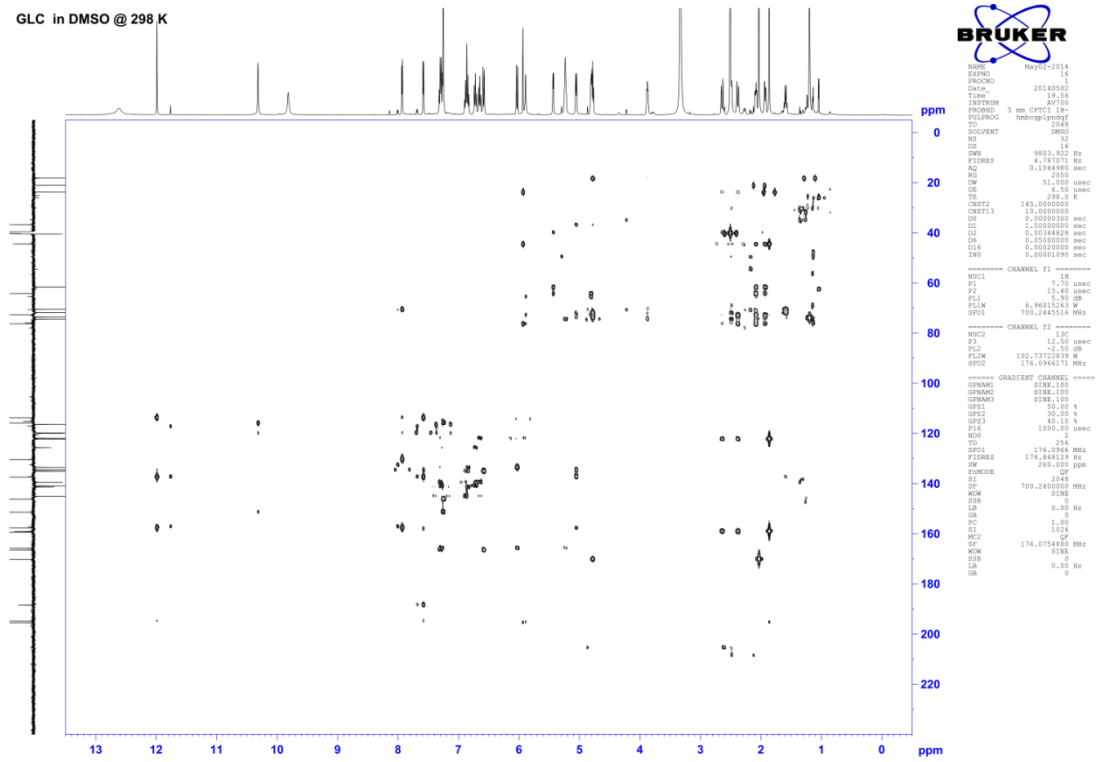


Figure 4.12

Heteronuclear multiple bond correlation (HMBC) spectrum of 7-oxo-SD8.

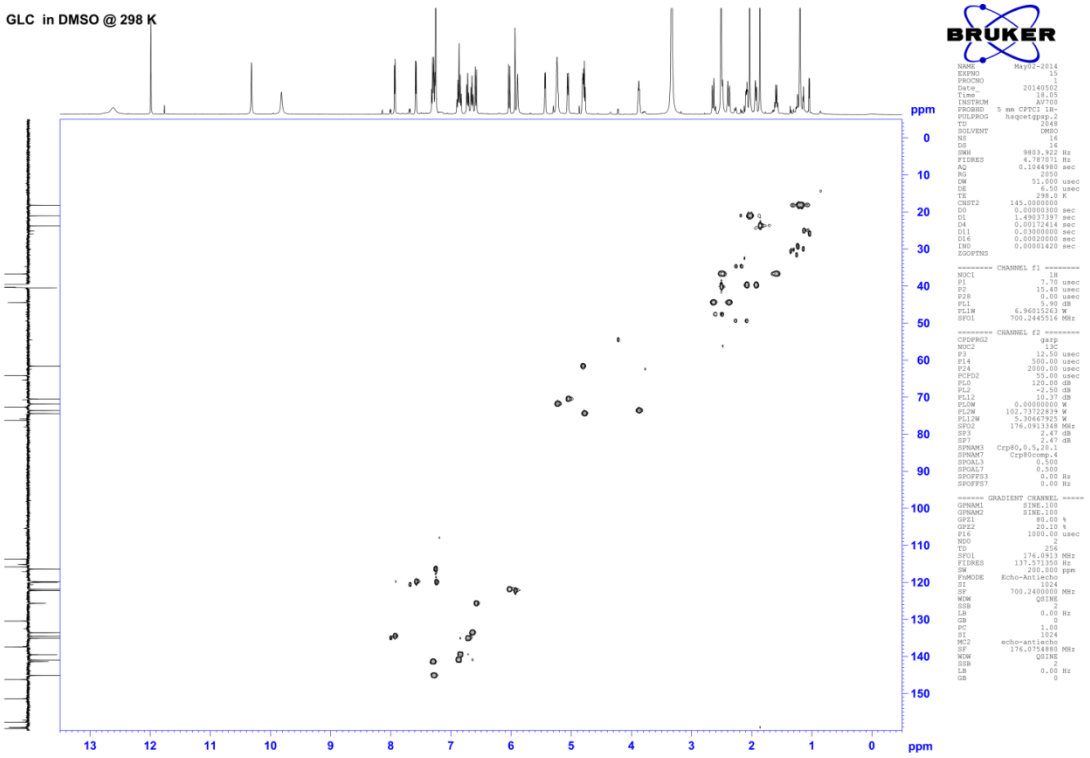


Figure 4.13

Heteronuclear single quantum coherence (HSQC) spectrum of 7-oxo-SD8.

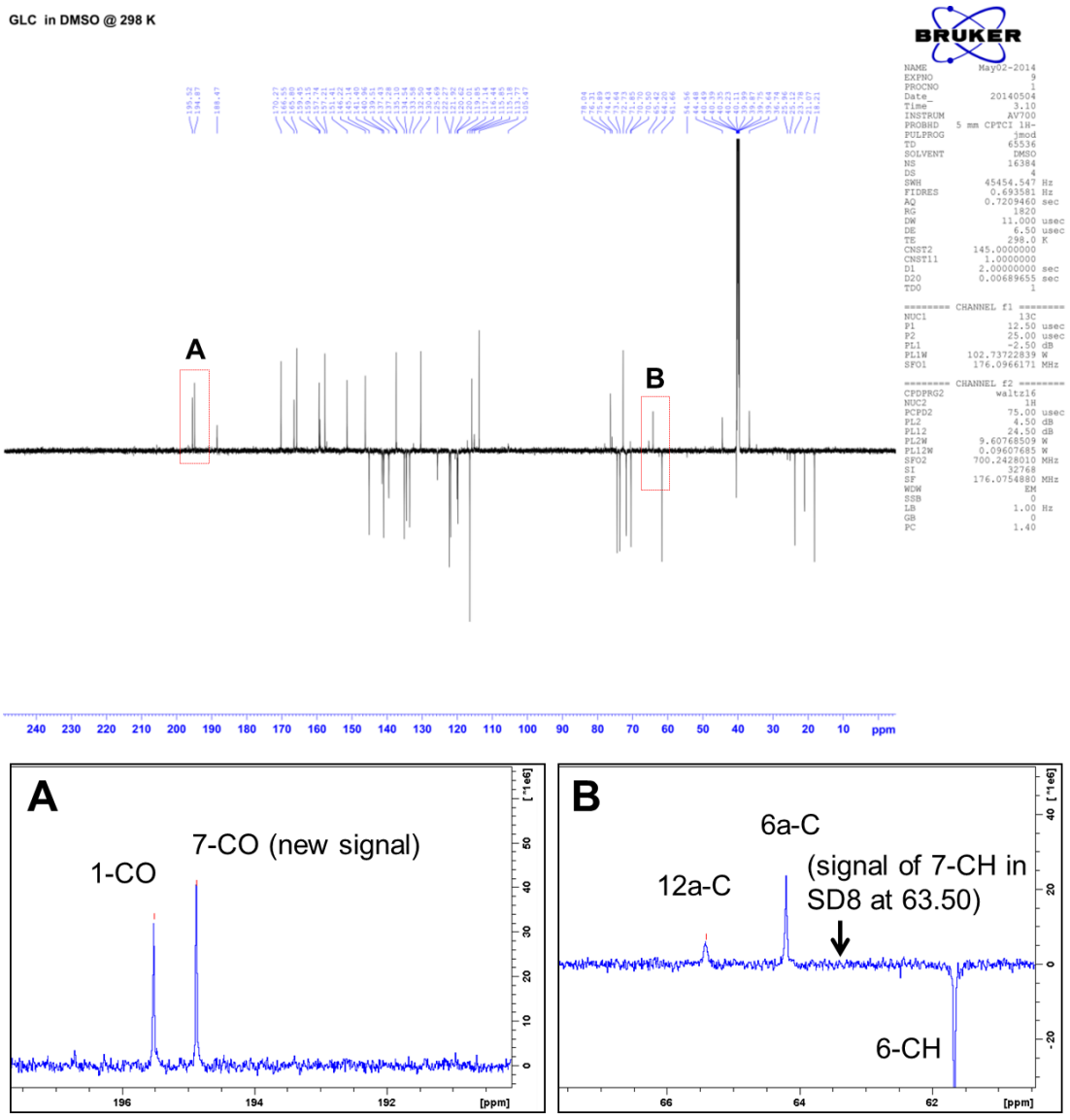


Figure 4.14

Attached proton test (APT) ^{13}C-NMR spectrum of 7-oxo-SD8.

The boxes below the NMR spectrum show a close-up of (A) the region for the 7-carbonyl group in 7-oxo-SD8 and (B) the absence of a signal for the 7-CH group present in SD8.

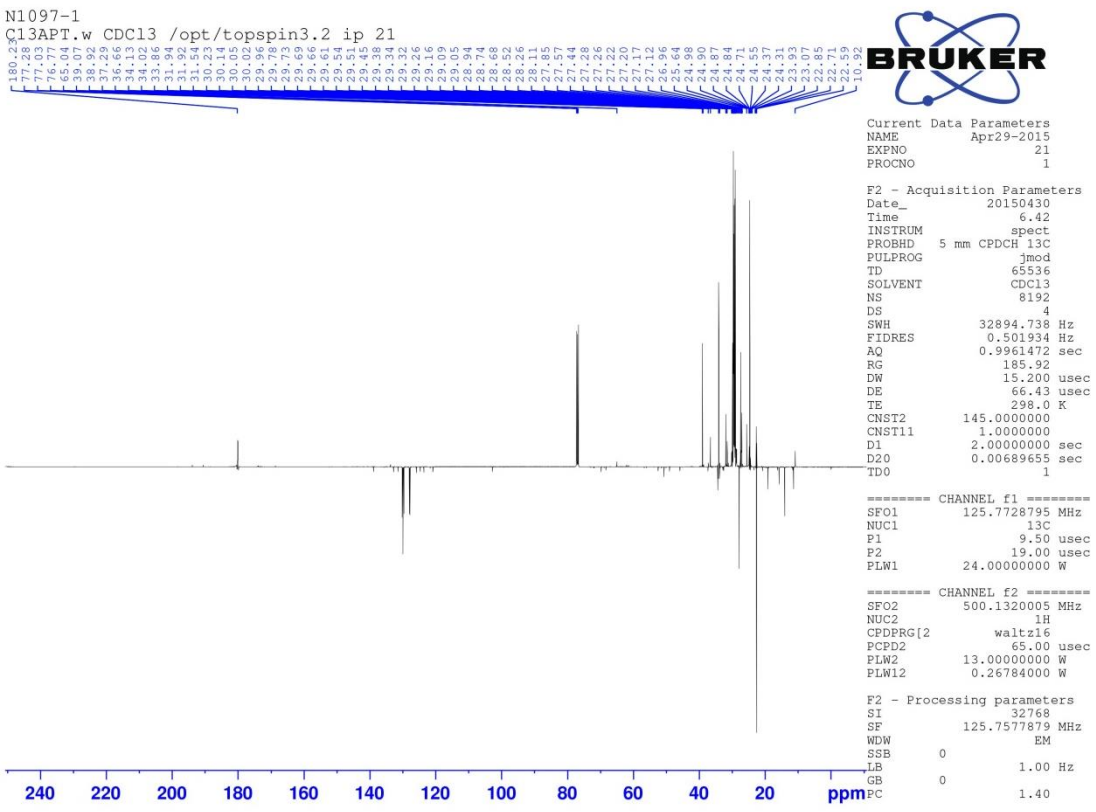


Figure 4.15

Attached proton test (APT) ^{13}C-NMR spectrum of 7-oxo-SD8 after ^{13}C-enrichment with [1,2-$^{13}C_2$]-acetate.

4.3 Complementation of *simC7* mutant

To determine if the *simC7* phenotype was caused solely by loss of SimC7 function, I complemented the mutant with an *in trans* copy of the gene. I cloned the *simC7* gene into pGM1190 and introduced the construct into the *simC7* mutant (*S. coelicolor* M1152ex1::12IΔC7). Although the complemented strain was able to produce SD8, the yield was very low (**Figure 4.16**), probably because the *simC7* gene lacked a promoter to drive its expression. *simC7* is the last gene in a putative seven-gene operon ranging from *simB7* to *simC7* suggests that the putative native promoter of *simC7* is upstream to *simB7* (**Figure 3.1**). Therefore, I created a second construct and drove expression of *simC7* from its putative native promoter by fusing the *simC7* gene directly to an ~400-bp fragment spanning the intergenic region upstream of *simB7*. I cloned this fusion into the vector pGM1190 and introduced this construct into the *simC7* mutant, which restored SD8 production to the *simC7* mutant (**Figure 4.16**).

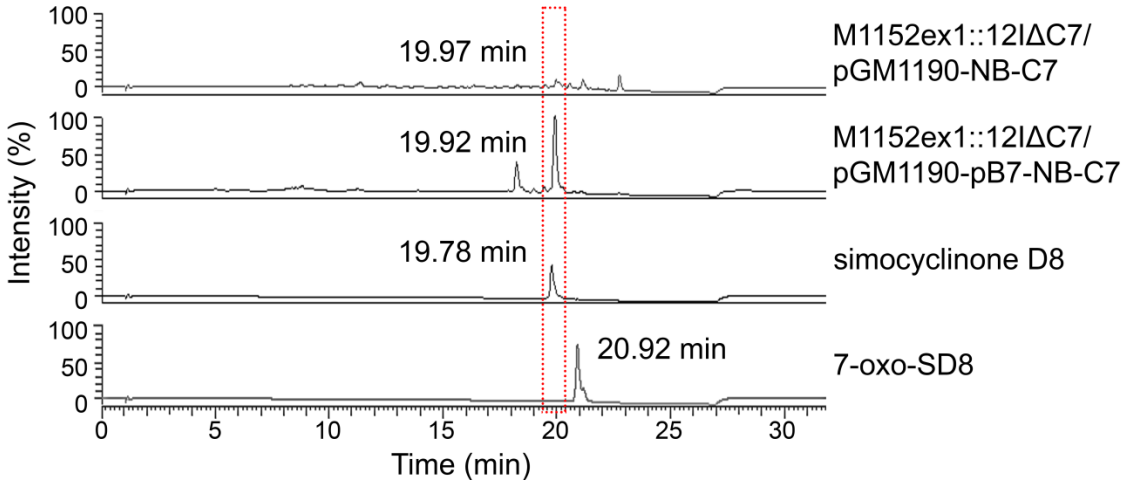


Figure 4.16

Complementation of *simC7* deletion mutant.

LC-MS-UV analysis of simocyclinones confirmed the complementation of *simC7* in the heterologous host *S. coelicolor* M1152ex1 carrying the mutated *sim* cluster without *simC7*. Chromatograms (200-700 nm) are shown of culture extracts from complemented strains with *simC7* under control of the *tipA* promoter (top) or its putative natural promoter (middle) compared to standards of SD8 and 7-oxo-SD8 (bottom). Molecular weights were confirmed by high-resolution mass spectrometry (not shown). Cultures were grown for 6 days at 30 °C in production medium.

4.4 SimC7 is an NAD(P)H-dependent ketoreductase that acts on the angucyclic polyketide

The fact that the *simC7* mutant produced 7-oxo-SD8 suggested that SimC7 is a ketoreductase involved in biosynthesis of the angucyclic polyketide and is not a dehydratase that acts during formation of the tetraene linker as previously proposed (Trefzer *et al.*, 2002). To test this hypothesis, I overexpressed an N-terminally His-tagged version of SimC7 in *E. coli* and assayed the purified recombinant protein for its ability to use 7-oxo-SD8 as a substrate. Both the substrate and the product of the reaction have strong chromophores with only minor differences in their absorbance spectra, and because their spectra also overlap those of $NAD(P)^+$/NAD(P)H, it was not possible to assay SimC7 using spectrophotometric methods. Instead, I followed the reaction by HPLC-UV (with SD8 as a standard) and found that SimC7 readily converted 7-oxo-SD8 into SD8 (**Figure 4.17b**). SimC7 was able to use NADH or NADPH for the reduction (**Figure 4.17b**) but showed a preference for NADPH (data not shown). Although I tried to measure the substrate conversion in presence of NADPH, the reaction was almost completed after 30 seconds, which did not allow me to collect a sufficient amount of samples for kinetic analysis. Reduced concentrations of SimC7 during the assays did not improve the measurements and the reaction was too slow below an enzyme concentration of 10 nM. It was also possible to assay the reaction in the reverse direction, following the oxidation of SD8 to 7-oxo-SD8 by SimC7 (**Figure 4.17c**). This reaction was much slower and required a very high concentration of the NAD^+ cofactor (300 mM). In all cases, the identity of the reaction product was confirmed by tandem mass spectrometry.

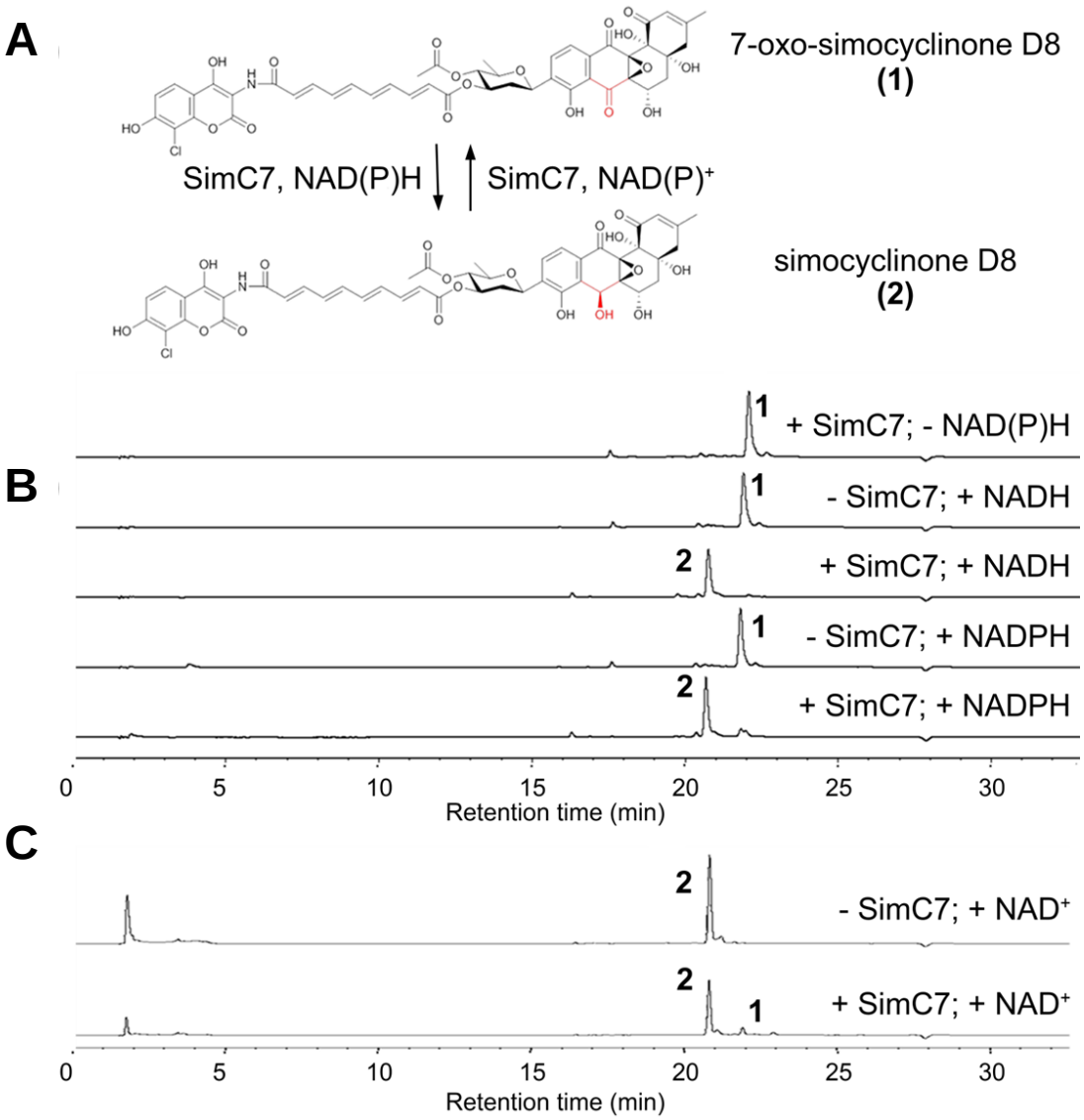


Figure 4.17

Ketoreductase activity of SimC7.

(A) SimC7 mediates conversion of the novel simocyclinone intermediate 7-oxo-SD8 (1) into SD8 (2). (B) NAD(P)H-dependent conversion of 7-oxo-SD8 into SD8 and (C) $NAD(P)^+$-dependent conversion of SD8 into 7-oxo-SD8, monitored by reversed-phase HPLC. Samples were incubated for 1 h at room temperature using the substrate 7-oxo-SD8 (100 µM), the cofactor NADH (0.3 mM) or NAD^+ (300 mM) and the enzyme SimC7 (500 nM) in HEPES buffer (50 mM, pH 7.2). Reactions (250 µL) were stopped by quenching in methanol (1:1) and heat denaturation for 10 min at 100 °C.

4.5 The C-7 hydroxyl group is required for the antibiotic activity of simocyclinone

The activity of 7-oxo-SD8 was tested *in vivo* and *in vitro*. Wild-type *E. coli* and other Gram-negative bacteria are resistant to simocyclinones because the compounds cannot penetrate the outer membrane (Schimana *et al.*, 2000). Therefore, I used an *E. coli* strain (NR698), which is sensitive to simocyclinones due to an in-frame deletion in the *imp* (increased membrane permeability) gene (Ruiz *et al.*, 2005). The MIC for SD8 was 0.3 µM. In contrast, strain NR698 grew in presence of 7-oxo-SD8 at concentrations up to 17.5 µM, an ~60-fold increase in MIC. In comparison, a wild type *E. coli* strain (ATCC 25922™), which is naturally resistant against simocyclinones, had an increased MIC for SD8 of ~87.5 µM and grew unimpaired in the presence of high levels of 7-oxo-SD8 (MIC >100 µM). The contrasting antibiotic activities of SD8 and 7-oxo-SD8 against whole cells suggested that the oxidation state of the oxygen at the C-7 position might be important for its ability to block DNA supercoiling by DNA gyrase. To answer this question, I analysed the activity of SD8 and 7-oxo-SD8 as DNA gyrase inhibitors *in vitro*. In line with previous studies (Edwards *et al.*, 2009; Oppegard *et al.*, 2009; Alt *et al.*, 2011), I found that SD8 inhibited supercoiling by DNA gyrase with IC_{50} = 0.1-0.5 µM (**Figure 4.18a**). In contrast, 7-oxo-SD8 was almost 3 orders of magnitude less active (IC_{50} = 50-100 µM) (**Figure 4.18a**). The fluoroquinolone ciprofloxacin links DNA gyrase to its substrate by stabilizing the DNA-protein cleavage complex, leading to a characteristic “cleavage band” on gels (**Figure 4.18b**) (Collin *et al.*, 2011). The presence of SD8 prevents the enzyme from binding to its substrate and thereby blocks DNA cleavage by gyrase (Edwards *et al.*, 2009). I tested the ability of 7-oxo-SD8 to abrogate the ciprofloxacin-stimulated cleavage of DNA by gyrase, using SD8 as a control (**Figure 4.18b**) (Flatman *et al.*, 2005). As shown previously (Edwards *et al.*, 2009), SD8 prevented DNA cleavage at low concentrations (e.g., 0.5 µM). In contrast, much higher concentrations of 7-oxo-SD8 (>10 µM) were required for inhibition of DNA cleavage by gyrase (**Figure 4.18b**).

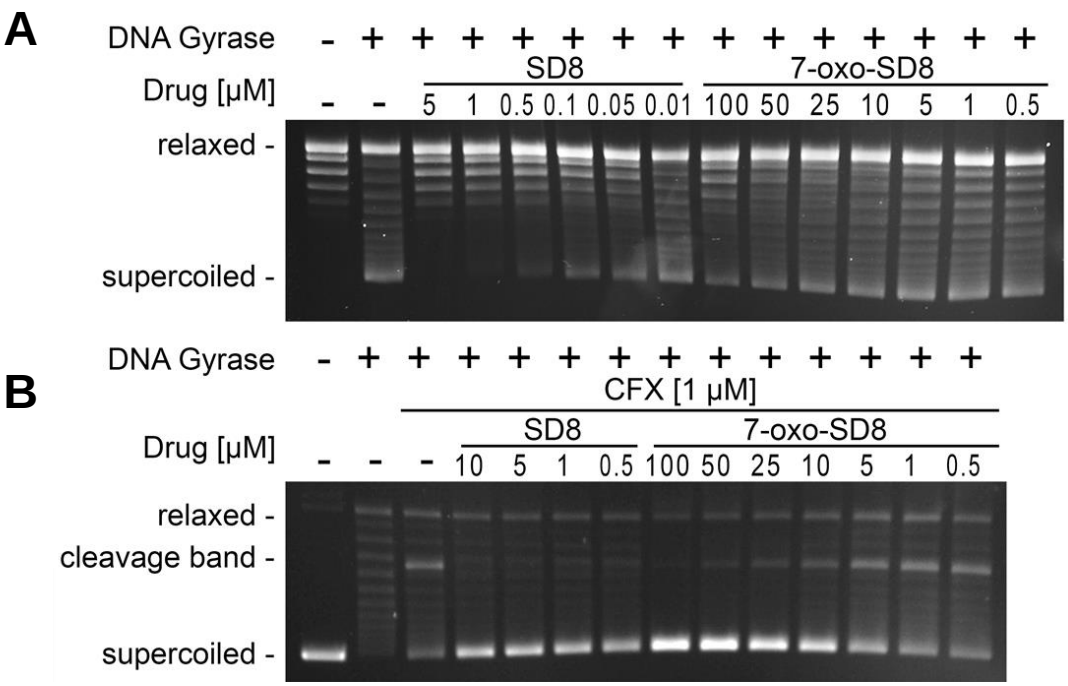


Figure 4.18

Inhibition of DNA gyrase by 7-oxo-SD8.

(A) Supercoiling and (B) cleavage relaxation assays with DNA gyrase in the presence of 7-oxo-SD8 or SD8. The reaction mixtures contained *E. coli* DNA gyrase and varying concentrations of either SD8 or 7-oxo-SD8. For relaxation assays, the mixtures were incubated without ATP and with ciprofloxacin (1 µM).

To investigate the binding of 7-oxo-SD8 to GyrA, I used surface plasmon resonance as described previously (Edwards *et al.*, 2009). Therefore, the GyrA protein was immobilised on a CM5 sensor chip by amine coupling via reactive esters, which makes use of the N-terminus and ε-amino groups of lysine residues of the protein. I found that SD8 bound to the GyrA N-terminal domain with a similar affinity to that reported previously, but that only weak, non-specific binding could be seen for the 7-oxo analog (data not shown). Taken together, these data suggest that 7-oxo-SD8 binds gyrase 2-3 orders of magnitude more weakly than SD8. MGD8N2A is a simocyclinone analog (generated by chemical hydrolysis of SD8) that lacks the angucyclic polyketide and the deoxysugar (**Figure 4.19**). This analog has been tested previously and found to have greatly reduced activity against DNA gyrase (IC_{50} = 50 μM) in comparison to SD8 (IC50 = 0.1–0.6 μM) (Edwards *et al.*, 2009). I repeated these experiments with MGD8N2A and obtained a similar value (IC_{50} = 25 μM). Thus, comparing the IC_{50} values for 7-oxo-SD8 and MGD8N2A, it is clear that the presence of a carbonyl group at the C-7 position has a similarly negative effect on the activity of simocyclinone as does complete loss of the angucyclinone.

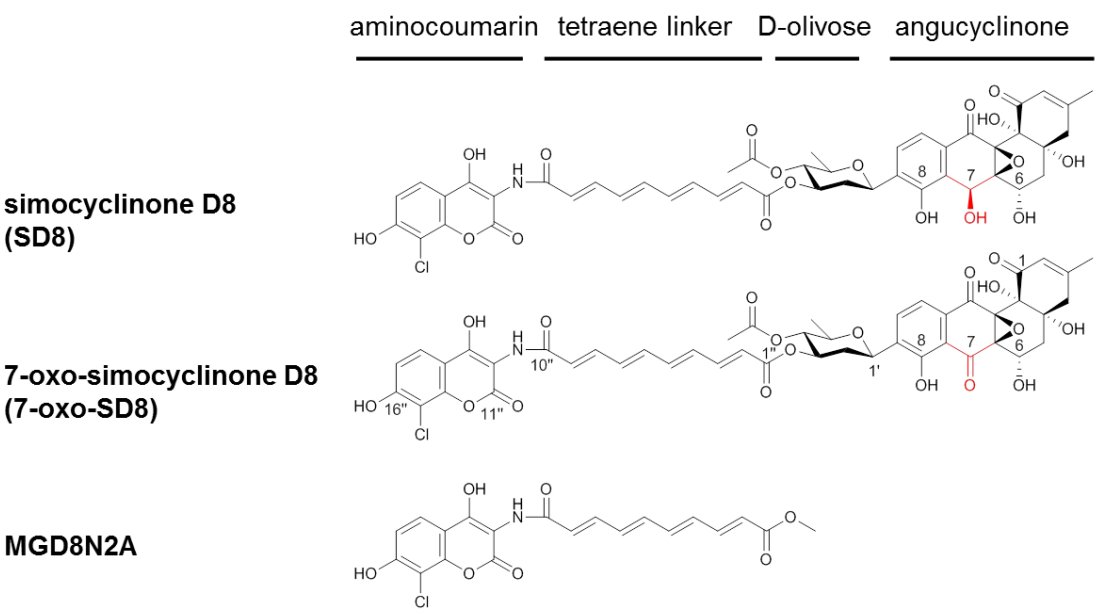


Figure 4.19

Chemical structures of SD8, 7-oxo-SD8 and MGD8N2A.

The structural difference between SD8 and 7-oxo-SD8 at position C-7 is highlighted in red.

4.6 Discussion

To examine the prediction that SimC7 is a dehydratase involved in the biosynthesis of the tetraene linker, the ~72-kb *sim* gene cluster was isolated on a single PAC clone and simocyclinones were expressed heterologously in an *S. coelicolor* strain engineered for improved antibiotic production. Deletion of *simC7* from the PAC clone resulted in the production of a novel simocyclinone, 7-oxo-SD8, which unexpectedly carried a normal tetraene linker but was altered in the angucyclinone moiety. I went on to demonstrate that SimC7 is an NAD(P)H-dependent ketoreductase that converts 7-oxo-SD8 into SD8 and that reduction of the keto group, catalysed by SimC7, is essential for the antibiotic activity of simocyclinone.

The biochemical function of SimC7

SimC7 was proposed to be a dehydratase involved in the biosynthesis of the tetraene linker (Trefzer *et al.*, 2002), but I have demonstrated that SimC7 is actually an NAD(P)H-dependent ketoreductase that reduces a carbonyl group at the C-7 position of the angucyclinone moiety. The fermentation product of the simC7 mutant, 7-oxo-SD8, was almost inactive as a DNA gyrase inhibitor, showing that reduction of the 7-oxo functional group by the ketoreductase activity of SimC7 is essential for the biological function of simocyclinones. Although SimC7 readily converts 7-oxo-SD8 into SD8, synthesis of the angucyclic polyketide is normally completed before it is linked to the D-olivose sugar and the other moieties of simocyclinone (Schimana *et al.*, 2001). This means that the natural substrate of SimC7 would be an A-group intermediate (i.e., one having only the angucyclic polyketide moiety) carrying a C-7 carbonyl. This species was detected neither in the native producer *S. antibioticus* (Schimana *et al.*, 2001; Galm *et al.*, 2002; Trefzer *et al.*, 2002) nor in the heterologous host carrying the wild-type *sim* cluster, but as expected, a species with the appropriate mass was detected among the intermediates made by the *simC7* mutant. Having

determined the true function of SimC7, it remains unclear how the tetraene linker of simocyclinone is assembled and how dehydration takes place.

SimC7 (30 kDa; 284 amino acids) is a member of the SDR family, one of the largest protein superfamilies, having more than 120,000 representatives in the databases. SDR proteins are diverse, with low overall amino acid sequence identity (typically 20-30% in pairwise comparisons), and they are principally characterized by the presence of a predicted pyridine nucleotide-binding Rossmann fold, comprising a parallel β-sheet flanked by three helices on each side. SDR proteins also have diverse biochemical activities, including acting as dehydratases, reductases, dehydrogenases, decarboxylases, and epimerases (Kallberg *et al.*, 2010; Persson and Kallberg, 2013). Despite its function as an angucyclinone ketoreductase, based on sequence identity, SimC7 seems to be most similar to various SDR sugar epimerases, and SimC7 secondary and tertiary structure predictions using the Phyre2 server (Kelley and Sternberg, 2009; Kelley *et al.*, 2015) also show highest similarity to sugar-modifying enzymes, such as epimerases, reductases, and dehydratases. No close homologs in the databases are known to function as ketoreductases. Although none is closely related to SimC7, several SDR ketoreductases have been identified that act on polyketides. These include SDR proteins that reduce the carbonyl group at position C-9 in the biosynthetic pathways for jadomycin (JadE), actinorhodin (ActKR), and hedamycin (HedKR) (Kulowski *et al.*, 1999; Korman *et al.*, 2008; Javidpour *et al.*, 2011a) and SDR proteins that reduce the carbonyl group at position C-6 in the biosynthetic pathways for landomycins (LanV) and urdamycins ($UrdM_{red}$) (Paananen *et al.*, 2013; Patrikainen *et al.*, 2014). Close homologs of these enzymes are encoded in the *sim* gene cluster, raising the possibility that SimA6 and SimA9 function to reduce the angucyclinone C-10 and C-6 carbonyl groups, respectively, during the biosynthesis of simocyclinone.

SimC7 function is vital for producing simocyclinones with antibiotic activity

Why does such a small structural change render 7-oxo-SD8 effectively inactive as a DNA gyrase inhibitor? The crystal structure of the GyrA–SD8 complex revealed how DNA gyrase binds the angucyclinone (Hearnshaw *et al.*, 2014) (**Figure 1.8** and **Figure 4.20**). The C-6 hydroxyl group makes a hydrogen bond to Met120, the C-7 hydroxyl group makes hydrogen bonds to Pro79 and Arg121 (via a water molecule), and the C-8 hydroxyl group makes hydrogen bonds to His80 and Arg121. One speculative possibility is that the presence of a C-7 carbonyl group in 7-oxo-SD8 leads to the formation of an intramolecular hydrogen bond with the neighbouring C-8 hydroxyl group, thus simultaneously breaking contacts with His80 and a highly coordinated water molecule held in place by Pro79 and Arg121 (**Figure 4.20**). His80 in particular seems to play a crucial role in binding simocyclinone, as the SD8 IC_{50} is increased 230-fold when this residue is mutated to alanine (Edwards *et al.*, 2009). In addition, the presence of a carbonyl group at C-7 will alter the conformation of the angucyclic polyketide, which may affect other bonding interactions with GyrA. Thus, the introduction of a carbonyl group at position C-7 is likely to break several hydrogen bonds that secure the angucyclic polyketide in its binding pocket in addition to the loss of the above mentioned hydrogen bond between the C-7 hydroxyl group and Pro79 and Arg121 via water molecule.

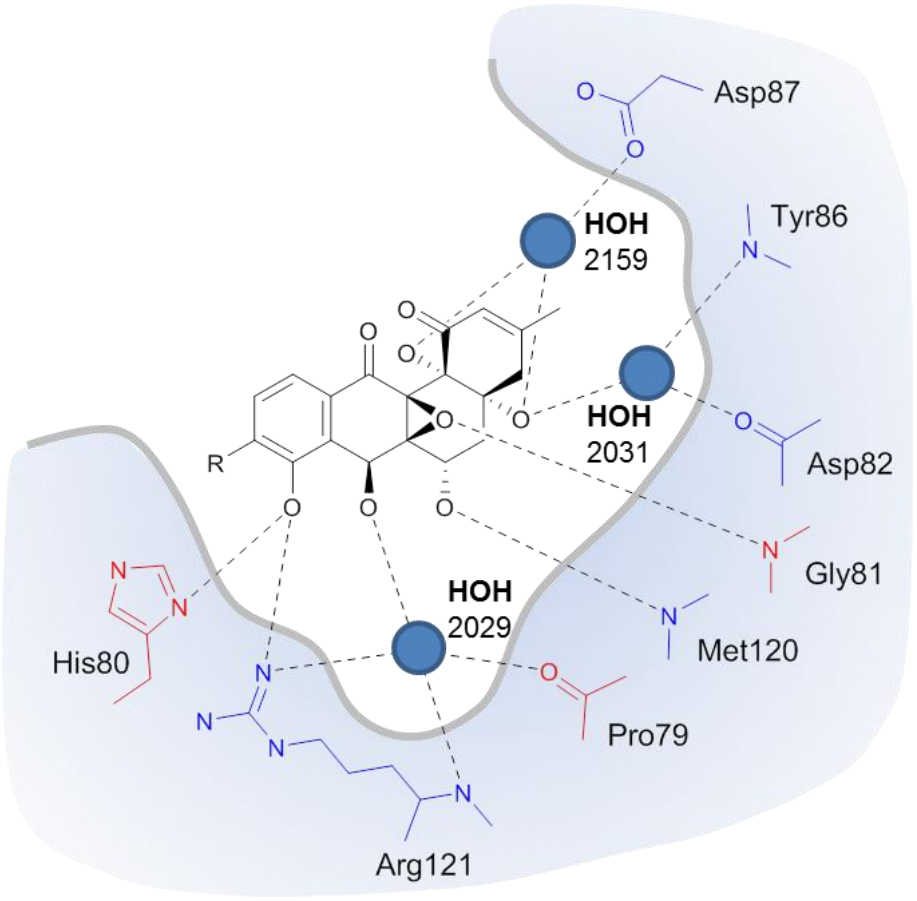


Figure 4.20

The angucyclinone binding pocket in GyrA.

SD8 is contacted by residues from both GyrA subunits; those from one subunit are shown in red and those from the other subunit are shown in blue. For clarity, all hydrogens have been omitted. Interactions were identified from the crystal structure of the GyrA–SD8 complex (PDB accession number 4CKL) (Hearnshaw *et al.*, 2014).

4.7 Acknowledgements

I would like to thank Professor Gregory L. Challis and Ivan Prokes (experimental officer in NMR) at the University of Warwick for collaboration and recording NMR spectra of my samples on a Bruker AVANCE-DRX 700 instrument as well as Sergey Nepogodiev in NMR at the John Innes Centre for helping me with NMR analysis of my samples on a Bruker Avance III 400 instrument. Further, I would like to thank Clare E.M. Stevenson and Stephen Hearnshaw for providing material and running SPR experiments. I would also like to thank the Maxwell group for providing material and in particular Stephen Hearnshaw for helping me with gyrase inhibition assays. I would like to thank Professor Barrie Wilkinson for helping me with HPLC and NMR experiments and Lionel Hill for his support with LC-MS experiments. I would like to thank Professor Hans-Peter Fiedler for providing purified simocyclinone D8.

Chapter 5
Substrate-Assisted Catalysis in Polyketide Reduction: Structural Analysis of the Ketoreductase SimC7 from Simocyclinone Biosynthesis (Schäfer et al., 2016)

5.1 Introduction 160

5.2 Optimisation of crystal conditions for SimC7 163

5.3 Overall structure of SimC7 164

5.4 Oligomeric state of SimC7 173

5.5 Structural homologs of SimC7 176

5.6 Cofactor binding 180

5.7 Substrate binding 181

5.8 A novel catalytic mechanism for an angucyclinone ketoreductase 185

5.9 Mutagenesis of the SimC7 active site 189

5.10 Discussion 192

5.11 Acknowledgements 193

5.1 Introduction

Angucyclin(on)es form the largest group of polycyclic aromatic polyketides, many with anticancer and antibacterial activities (Kharel *et al.*, 2012). They share a polyketide-derived tetracyclic benz[*a*]anthracene carbon skeleton, but numerous structures are generated by a range of tailoring reactions, including O- or C-linked deoxysugar glycosylation to form angucyclines (Rohr and Thiericke, 1992). In general, these tailoring enzymes are not well understood. An important challenge, therefore, is to define the step catalysed by each enzyme, and to determine their reaction mechanisms. This knowledge is particularly relevant to the rational engineering of angucyclin(on)e biosynthetic pathways for novel therapeutics.

The angucyclinone moiety of SD8 is synthesised by a type II polyketide synthase (PKS) (SimA1-3) and multiple tailoring enzymes (SimA4-13, SimC7) that catalyse cyclisation, aromatisation, oxidation and reduction reactions. In recent years, several ketoreductases of the short-chain dehydrogenase/reductase (SDR) family have been characterised that act on angucyclinones or related polyketides. The reduction of carbonyl groups at the C-6 and C-9 positions of polyketides have been functionally characterised and the crystal structures of the corresponding SDR ketoreductases (KRs) have revealed their catalytic reaction mechanisms (**Figure 5.1**) and the factors that determine their stereoselectivity. The ketoreductases LanV and $UrdM_{red}$ act on the C-6 carbonyl group of angucyclic polyketides from the landomycin and urdamycin pathways (Paananen *et al.*, 2013; Patrikainen *et al.*, 2014). In contrast, the ketoreductases ActKR and HedKR act on the C-9 carbonyl group of early intermediates in the actinorhodin and hedamycin polyketide pathways (Korman *et al.*, 2004; Korman *et al.*, 2008; Javidpour *et al.*, 2011a; Javidpour *et al.*, 2011b; Javidpour *et al.*, 2013). The LanV, $UrdM_{red}$, ActKR and HedKR structures revealed the catalytic Ser-Tyr-Lys triad characteristic of SDR enzymes, where the latter two residues form a YxxxK motif. In these classical SDR proteins the conserved active site tyrosine serves as central acid-base catalyst that donates a proton to the substrate (**Figure 5.1**). The

adjacent lysine residue lowers the pKa of the tyrosine hydroxyl group and often contributes directly to a proton relay mechanism, and the hydroxyl group of the serine stabilises and polarises the carbonyl group of the substrate (Kavanagh *et al.*, 2008).

At the sequence level, SimC7 shares little similarity with any characterised ketoreductase, even with functionally analogous polyketide ketoreductases. The striking differences between the amino acid sequence of SimC7 and those of the polyketide ketoreductases HedKR, ActKR, LanV and $UrdM_{red}$ suggested that SimC7 might have a novel catalytic mechanism. To investigate this possibility, I solved several crystal structures of SimC7.

Here I present crystal structures of SimC7 alone (*apo*), in complex with $NADP^+$, and in a ternary complex with both $NADP^+$ and 7-oxo-SD8, determined to 1.6 Å, 1.95 Å and 1.2 Å resolution, respectively. My results show that SimC7 is structurally distinct from previously characterised polyketide ketoreductases and, importantly, lacks the canonical SDR Ser-Tyr-Lys catalytic triad (Oppermann *et al.*, 2003; Kavanagh *et al.*, 2008; Kallberg *et al.*, 2010; Persson and Kallberg, 2013). Based on the atomic-resolution structure of the substrate-complex and active site mutations I propose that SimC7 catalyses a substrate-assisted, two-step reaction for reduction of the C-7 carbonyl group involving an unusual phenolate intermediate.

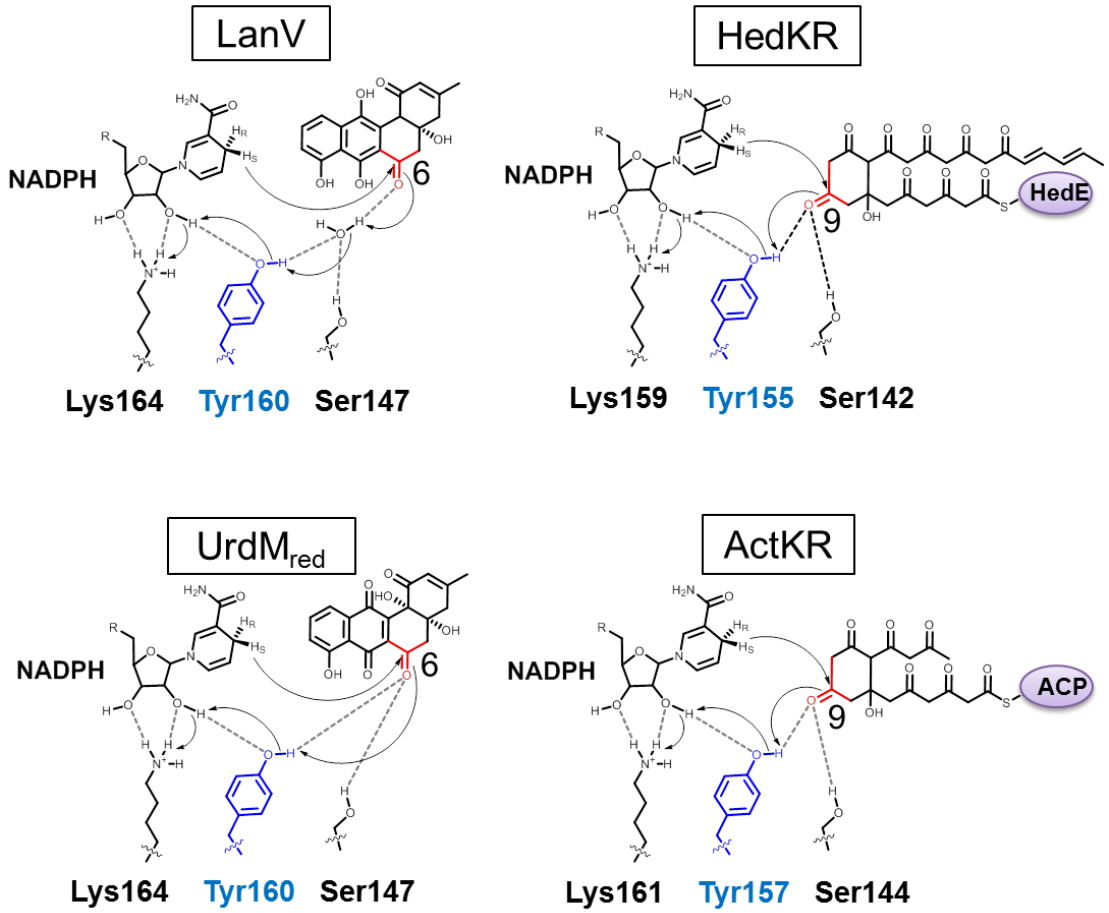


Figure 5.1

Comparison of proposed proton relay mechanisms for classic SDR ketoreductases.

The reduction of carbonyl groups by SDR ketoreductases follows a common mechanism that involves a characteristic catalytic triad (Ser-Tyr-Lys). The active site tyrosine (blue) acts as acid-base catalyst that transfers a proton to the substrate. The carbonyl group is reduced via direct hydride transfer from the cofactor NADPH followed by a proton relay system as indicated (black arrows). Hydrogen bonds are dashed lines. The proton of lysine is replenished by adjacent water molecules.

5.2 Optimisation of crystal conditions for SimC7

Recombinant SimC7 crystallises preferentially in the presence of organic acids at a pH range below (pH 3.8-4.8) or above (pH 6.5-9) its predicted isoelectric point (pI = 5.96) (see Materials and Methods, **chapter 2.7.1**). Therefore, I tested the effect on crystal growth of various monobasic acids (formic acid, acetic acid, propionic acid, caproic acid), dibasic acids (oxalic acid, malic acid), and tribasic acids (citric acid). Best results were achieved in the presence of 0.1 M acetic acid (pH 4.8), followed by formic acid and propionic acid under equivalent conditions. Diffraction data for the trimeric complex of protein, cofactor and simocyclinone were of reproducibly high resolution in the presence of acetic acid (1.2-1.5 Å). Saturated solutions of neutralised citric acid (1.1-1.5 M, pH 6.5) or DL-malic acid (1.3-1.6 M, pH 7.0), which are structurally very similar, resulted in large crystals of bipyramid shape. Unfortunately, the diffraction data from crystals grown in the presence of citric or malic acid showed strongly increased mosaicity and co-crystallisations with SD8 and 7-oxo-SD8 were unsuccessful. Evaluation of the structural data from respective conditions revealed that the larger organic acids occupied the enzyme's active site as indicated in several unrefined structures (data not shown), which most likely stabilized the enzyme and promoted crystal growth. Similar results were initially obtained for saturated ammonium sulphate (1.4-1.8 M) which was not considered for further optimisations due to the presence of salt molecules in the active site and the very irregular appearance of the crystals.

In summary, low concentrations of short monobasic acids promote crystal growth and result in reproducibly high-resolution protein crystal structures (1.2-1.5 Å), whereas longer acids or high concentrations of di- or tribasic acids have adverse effects and occupy the substrate binding pocket of SimC7.

5.3 Overall structure of SimC7

Several crystal structures of SimC7 were solved, including SimC7 alone (1.6 Å resolution), in complex with $NADP^+$ (1.95 Å), and in a ternary complex with both $NADP^+$ and 7-oxo-SD8 (1.2 Å) (**Table 5.1** and **Table 5.2**). SimC7 is made up of two domains, the larger of these is the nucleotide binding domain that adopts a classic Rossmann fold (Rossmann *et al.*, 1975), consisting of a seven-stranded parallel β-sheet (β1-6 and β9), which is flanked by three parallel α-helices on both sides (α1, α6, α10 and α2, α3, α4, respectively) (**Figure 5.2a-b** and **Figure 5.3**). The smaller substrate-binding domain, characteristic of the so-called 'extended' SDR subfamily (Kavanagh *et al.*, 2008), is mainly α-helical and is largely formed by two insertions in the nucleotide binding domain (between β6 and α6, and between β9 and α10). Notably, the latter insertion contains a 'lid' motif consisting of two anti-parallel α-helices (α8 and α9) that folds over the active site (**Figure 5.2a-b** and **Figure 5.3**). The substrate-binding domain is completed by a short helical segment at the C-terminus of the polypeptide chain.

Overall all SimC7 structures are very similar (**Table 5.3**), with the notable exception of the lid motif, which displays a number of different conformations (**Figure 5.2c** and **Table 5.3**). Although the changes are not large (maximum Cα-Cα shift 5.35 Å; **Table 5.3**), there is a clear closure of the lid over the bound substrate (**Figure 5.2c**), suggesting a role in gating access to the active site and/or substrate capture. Moreover, in the ternary complex the underside of the lid contributes to the tight, highly hydrophobic substrate binding pocket (**Figure 5.4** and **Figure 5.5**) that provides the necessary environment for catalysis (see below).

Table 5.1

X-ray data collection and processing.

Data set	Binary SeMet	Binary native	Apo form 1	Apo form 2	Ternary complex
Beamline	I03	I03	I04-1	I04-1	I04
Wavelength (Å)	0.9796	1.0052	0.9173	0.9173	0.9795
Detector	Pilatus 6M	Pilatus 6M	Pilatus 6M	Pilatus 6M	Pilatus 6M
Resolution range (Å)[a]	51.48 – 2.05 (2.10 – 2.05)	50.74 – 1.95 (2.00 – 1.95)	53.40 – 1.60 (1.64 – 1.60)	54.88 – 1.90 (1.95 – 1.90)	47.52 – 1.20 (1.23 – 1.20)
Space Group	$P4_12_12$	$P4_12_12$	$P4_12_12$	C2	$P2_12_12_1$
a, b, c (Å)	51.5, 51.5, 213.2	52.2, 52.2, 214.0	52.4, 52.4, 213.6	107.0, 64.8, 91.5	51.9, 53.7, 102.3
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 105.5, 90.0	90.0, 90.0, 90.0
Total observations[a]	457751 (25343)	563166 (40131)	435288 (32092)	336053 (23051)	1159483 (81348)
Unique reflections[a]	19033 (1335)	22650 (1625)	40514 (2936)	47271 (3433)	90009 (6590)
Multiplicity[a]	24.1 (19.0)	24.9 (24.7)	10.7 (10.9)	7.1 (6.7)	12.9 (12.3)
Mean $I/\sigma(I)$[a]	16.5 (1.3)	18.4 (2.2)	20.0 (1.5)	13.6 (1.0)	17.5 (2.1)
Completeness (%)[a]	99.8 (99.7)	99.9 (99.8)	99.9 (99.2)	98.5 (96.1)	100.0 (99.5)
R_{merge}[a,b]	0.122 (2.745)	0.118 (1.845)	0.066 (1.814)	0.087 (1.990)	0.070 (1.208)
R_{meas}[a,c]	0.125 (2.820)	0.120 (1.883)	0.069 (1.903)	0.094 (2.157)	0.073 (1.249)
$CC_{½}$[a,d]	0.999 (0.673)	0.999 (0.819)	1.000 (0.701)	0.999 (0.543)	0.999 (0.760)
Wilson *B* value ($Å^2$)	45.3	34.0	21.6	34.2	11.9

[a] Values for the outer resolution shell are given in parentheses.

[b] $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$.

[c] $R_{meas} = \sum_{hkl} [N/(N-1)]^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl, $\langle I(hkl)\rangle$ is the weighted average intensity for all observations i of reflection hkl and N is the number of observations of reflection hkl.

[d] $CC_{½}$ is the correlation coefficient between symmetry-related intensities taken from random halves of the dataset.

Table 5.2

Refinement of X-ray structures.

Data set	Binary native	Apo form 1	Apo form 2	Ternary complex
Resolution range (Å)[a]	50.74 – 1.95 (2.00 – 1.95)	53.40 – 1.60 (1.64 – 1.60)	54.88 – 1.90 (1.95 – 1.90)	47.52 – 1.20 (1.23 – 1.20)
Reflections: working/free[b]	21479/1171	38504/2010	44913/2357	85532/ 4476
Final R_{work}[a,c]	0.182 (0.282)	0.189 (0.297)	0.195 (0.342)	0.126 (0.218)
Final R_{free}[a,c]	0.217 (0.316)	0.218 (0.303)	0.229 (0.348)	0.149 (0.230)
Cruickshank DPI (Å)[d]	0.139	0.087	0.150	0.032
R.m.s. bond deviations (Å)	0.010	0.010	0.009	0.010
R.m.s. angle deviations (°)	1.40	1.38	1.29	1.52
No. of protein residues (ranges)	280 (1 to 280)	280 (1 to 280)	A chain:278 (2-279); B chain 281 (-1 to 279)[*]	278 (2-279)
No. of heterogen residues: cofactor/ 7-oxo/ water/ other	1/0/104/0	0/0/167/0	0/0/202/2	1/1/324/1
Mean *B*-factors: protein/ cofactor/ 7-oxo/ water/ other/ overall (Å^2)	47/46[#]/-/46/-/47	35/-/-/39/-/35	50/-/-/49/58/50	17/11/21/32/28/19
Ramachandran plot: favoured/ allowed/ disallowed (%)[e]	98.6/1.4/0.0	99.3/0.7/0.0	98.4/1.2/0.4	99.1/0.9/0.0
PDB code	5L3Z	5L40	5L45	5L4L

[a] Values for the outer resolution shell are given in parentheses.

[b] The data set was split into "working" and "free" sets consisting of 95 and 5% of the data, respectively. The free set was not used for refinement.

[c] The R-factors R_{work} and R_{free} are calculated as follows: $R = \sum(|F_{obs} - F_{calc}|)/\sum|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively.

[d] Diffraction precision indicator based on R_{free} based as calculated by *REFMAC5* (Murshudov *et al.*, 1997)

[e] As calculated using *MOLPROBITY* (Davis *et al.*, 2007)

[*] Two residues were visible for the N-terminal His-tag in the B-chain. Since the numbering scheme was based on the wild-type sequence, these residues were labelled as “-1” and “0”.

[#] Cofactor was refined with occupancy of 0.7.

Table 5.3

Comparison of SimC7 structures.

Pairwise superpositions of all protein structures determined herein; Root-mean-square deviations (r.m.s.d.) were determined by LSQKAB (Kabsch, 1976). For Apo form 2, the two molecules in the ASU were treated as separate models.

Overall r.m.s.d. (Å) [max. Cα shift (Å)/ corresp. res.]	**Apo form 1**	**Apo form 2: A-chain**	**Apo form 2: B-chain**	**Binary native**	**Ternary complex**
Apo form 1	0.00	1.13 [4.18/ P37][*] [3.82/ A231]	0.58 [1.71/ P230]	0.18[#] [0.40/ R222]	0.84 [3.67/ G227]
Apo form 2: A-chain		0.00	0.88 [3.15/ E232]	1.10 [3.92/A231]	1.43 [5.35/ G227]
Apo form 2: B-chain			0.00	0.58 [1.65/ P230]	0.98 [4.20/ G227]
Binary native				0.00	0.83 [3.67/ G227]
Ternary complex					0.00

* With the exception of this comparison, all the others showed the biggest shift in the lid region (residues 216-246 inclusive). Here the biggest shift was due to a *trans/cis* isomerisation of Pro37 (see main text for further explanation). The second largest shift (in the lid) is also shown.

[#] N. B. Apo form 1 and Binary native SimC7 are isomorphous.

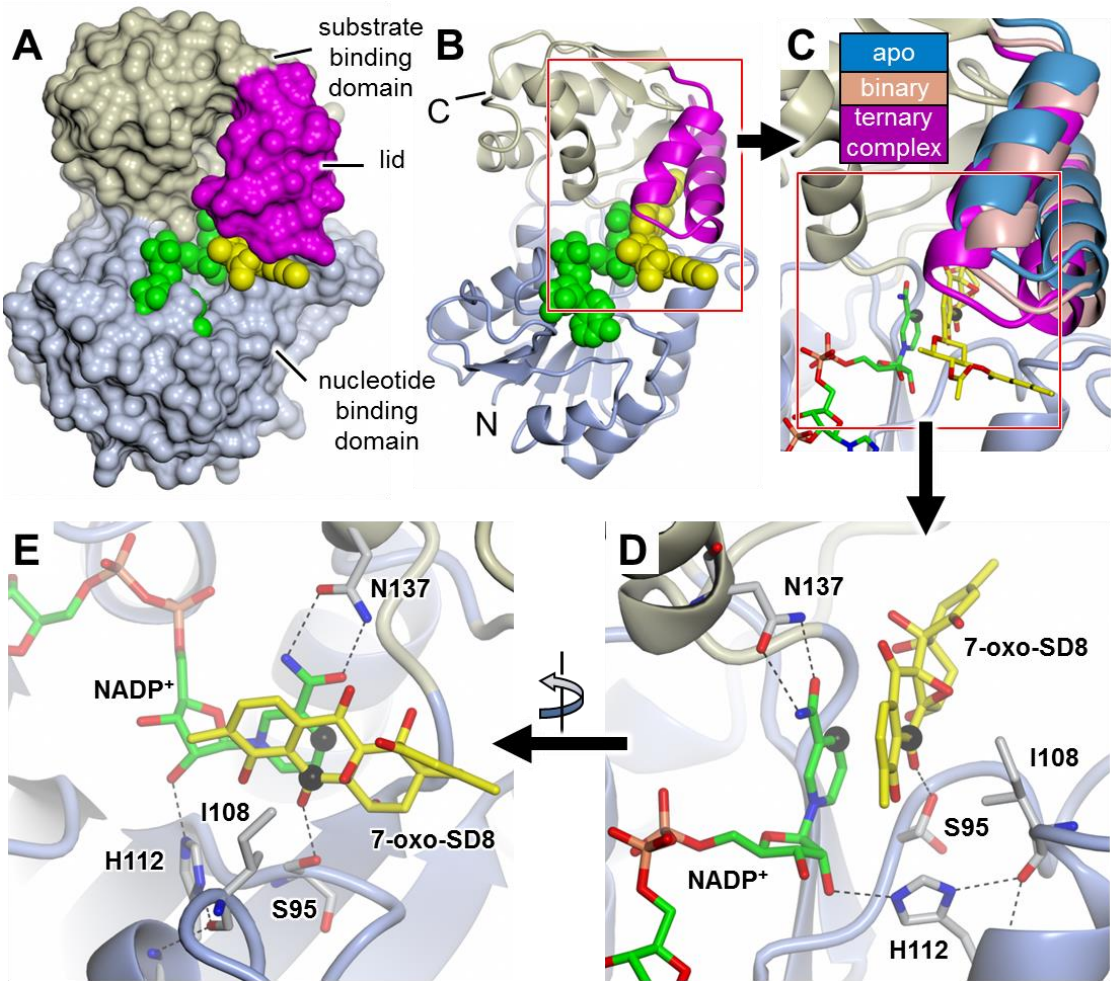


Figure 5.2

Crystal structure of SimC7.

S. antibioticus SimC7 displayed as (A) a molecular surface and (B) in cartoon representation where pale blue, beige and magenta colouration indicate the nucleotide binding domain, the substrate binding domain and the lid motif, respectively. Ligands are shown as van der Waals spheres, where green indicates the cofactor and yellow the 7-oxo-SD8 substrate. (C) Close-up of the region highlighted in part (B) showing conformational changes in the lid between the *apo* (chain A, form 2; blue), the binary complex (salmon) and the ternary complex (magenta); only the core protein structure and ligands (stick representation) for the latter are shown. C-4 of the cofactor and C-7 of the substrate are highlighted by black spheres, showing C-7 of the substrate is positioned 3 Å from C-4 of the nicotinamide ring, poised for direct hydride transfer. (D) and (E) are orthogonal close-ups of the region

highlighted in part (C) showing the detail of the active site of the ternary complex including the Ser95-Ile108-His112 ‘catalytic triad’ residues, and Asn137, which is important in maintaining the *syn*-conformation of the cofactor. For clarity, the lid motif has been omitted and only the angucyclic polyketide moiety of the substrate is shown. Part (E) is also reproduced as a stereo image in **Figure 5.8** (see **chapter 5.7**), which also shows a difference electron density map calculated from the final model after simulated annealing refinement with the substrate omitted.

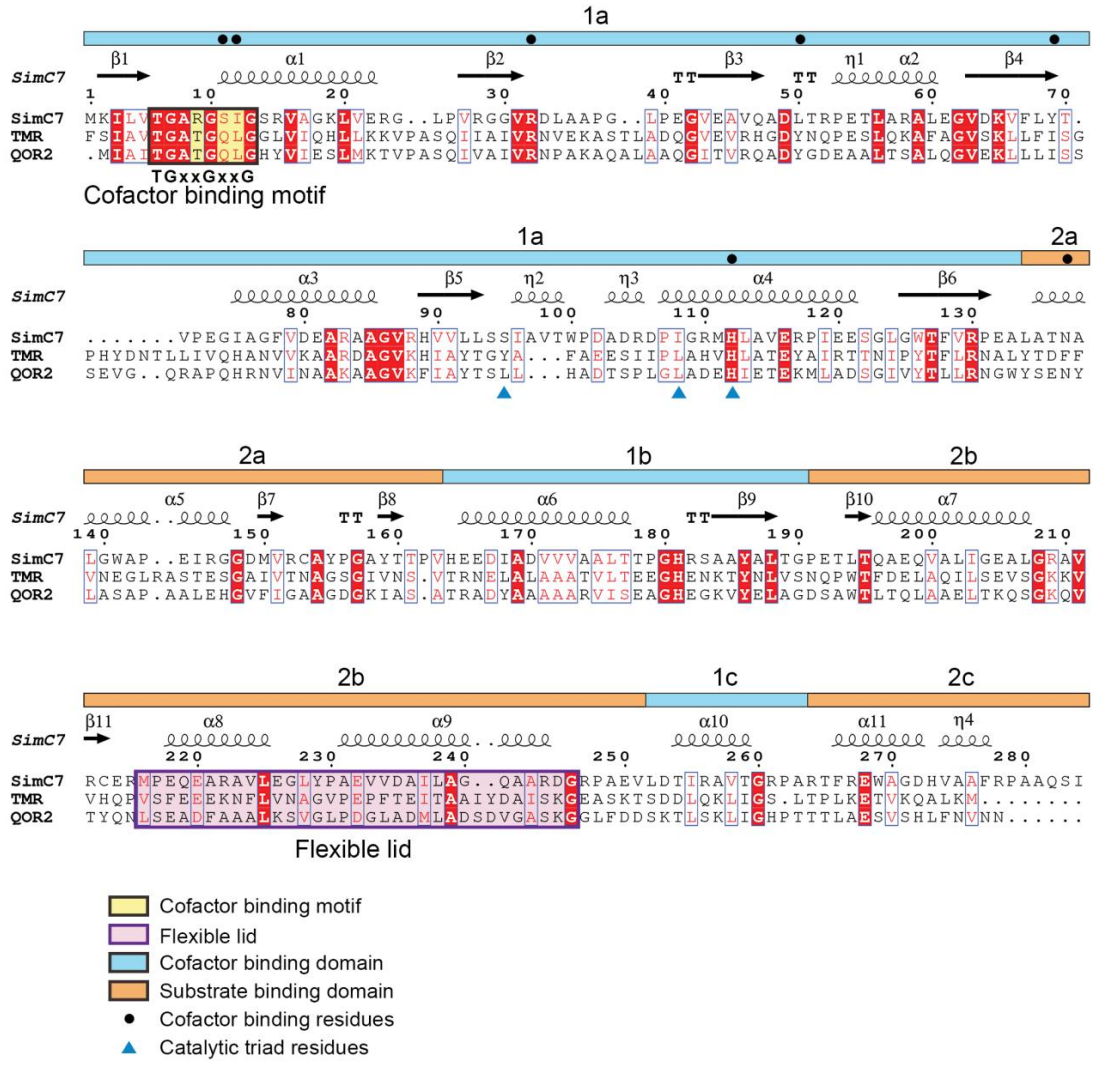


Figure 5.3

Structure-based sequence alignment of SimC7 with triphenylmethane reductase (TMR) from *Citrobacter* sp. and quinone oxidoreductase (QOR2) from *E. coli*.

Conserved residues are shown in white on a red background and similar residues are shown in red font. Secondary structure for SimC7 is shown above the alignment (α for α-helix and β for β-strand, TT for β-turn and η for 3_{10} helix). The bars above the alignment indicate the cofactor binding domain (1a-c in blue) and the substrate binding domain (2a-c in orange) with black dots indicating residues that interact with the $NADP^+$ cofactor. The conserved N-terminal cofactor-binding motif (TGxxGxxG, yellow with black frame) and the flexible lid (Met216 to Gly246, pink frame) are highlighted. Blue triangles indicate the positions equivalent to the active site triad residues in canonical SDR proteins. The active site residues of TMR and QOR2 are unknown.

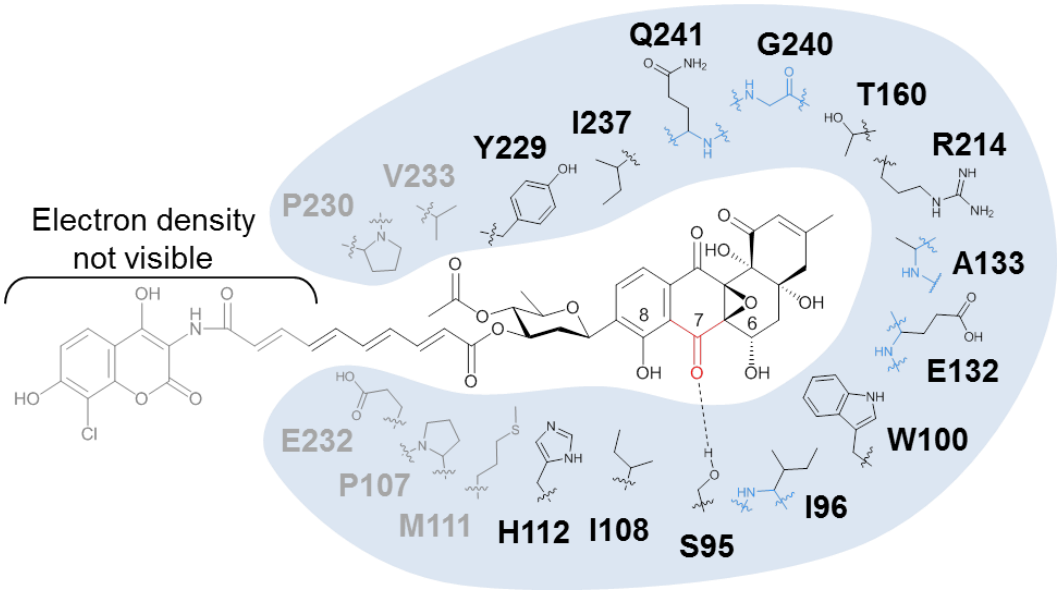


Figure 5.4

Schematic representation of the highly hydrophobic substrate binding pocket of SimC7.

The substrate is bound by only one direct hydrogen bond between the C-7 carbonyl group in the angucyclinone moiety and the side-chain hydroxyl of Ser95. Although Ser95 is positioned within hydrogen bonding distance to the C-7 carbonyl group (2.7 Å distance from oxygen to oxygen atom; ~1.7 Å estimated distance for the respective hydrogen bond), Ser95 is not required for the reaction (see **chapter 5.9**). This interaction may assist in positioning the substrate and to facilitate the reaction. Interacting residues are shown in black (side-chain interactions) or blue (backbone interactions). The hydrophobic residues shown in grey line the entrance to the substrate pocket but do not interact directly with the bound angucyclinone. One face of the pocket is formed by the cofactor itself (not shown). The electron density for the aminocoumarin and the adjacent part of the tetraene linker (grey), which protrude from the active site pocket, was not visible in the complex structure (see **Figure 5.8a-b**).

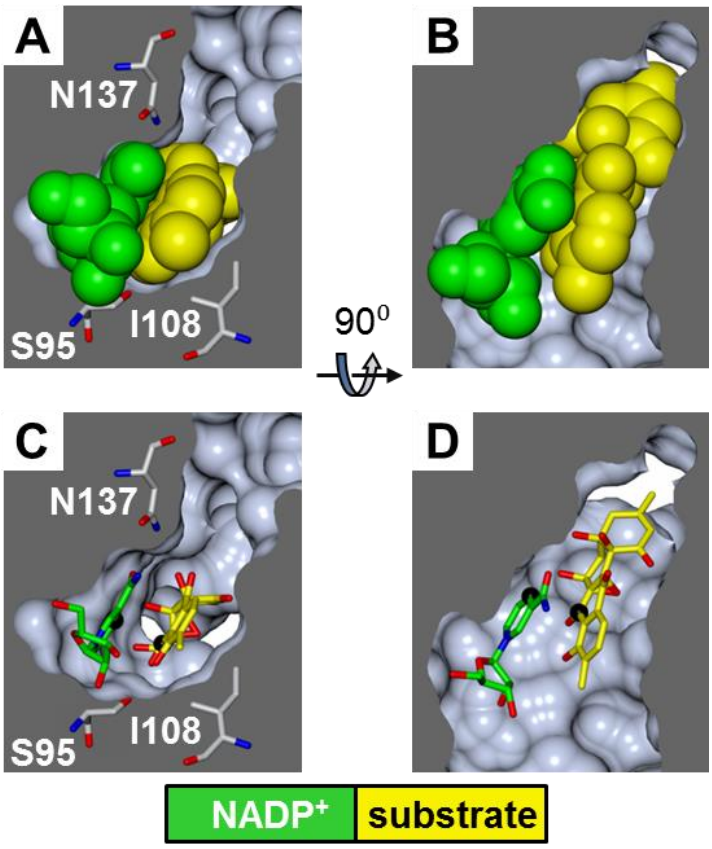


Figure 5.5

SimC7 has a very constricted active site.

Orthogonal cross-sections through a molecular surface centred on the active site revealing how tightly the cofactor (green) and the substrate (yellow) are bound in the active site pocket. For clarity, only the nicotinamide ribosyl moiety of the cofactor and the polyketide moiety of the substrate are shown. The surface of the protein is shown in pale blue and the interior of the protein is shown in grey. In (A) and (C) the view corresponds roughly to that shown in **Figure 5.2d**, whereas (B) and (D) show the view from above relative to **Figure 5.2d**. The ligands are depicted as van der Waals spheres in (A) and (B), and in stick representation in (C) and (D). In the lower panels, C-4 of the cofactor and C-7 of the substrate are highlighted by black spheres, showing C-7 of the substrate is exactly positioned 3 Å from C-4 of the nicotinamide ring, poised for direct hydride transfer. Active site residues that are out of the plane of view are not shown.

5.4 Oligomeric state of SimC7

Three different crystal forms were obtained for SimC7. SimC7 alone (*apo*) crystallised in two crystal forms, form 1 [space group $P4_12_12$; one molecule per asymmetric unit, (ASU)] and form 2 (space group C2; two molecules per ASU). The complex with $NADP^+$ was isomorphous with *apo* form 1 (**Table 5.1**). Crystal contacts within crystal form 1 and form 2 gave rise to two distinctly different two-fold symmetric putative dimers, both having substantial interfacial areas (1046 and 1246 $Å^2$, respectively; **Table 5.4**). By contrast, SimC7 in complex with $NADP^+$ and 7-oxo-SD8 produced a new crystal form (space group $P2_12_12_1$) in which no substantial interfaces were present, the largest covering only 496 $Å^2$, and this did not generate a two-fold symmetric dimer. In parallel, size exclusion chromatography was used to ascertain the oligomeric states of the *apo* form and the binary complex in solution. In both cases, the protein was exclusively monomeric (**Figure 5.6**).

The chromatographic and crystallographic results first appeared to be in conflict. Therefore, the characteristics of the potential SimC7 dimer interfaces were compared with the closest structural homolog of SimC7 where both crystallographic and experimental data supported the existence of a homodimer, namely triphenylmethane reductase (TMR) from *Citrobacter* sp. (PDB accession number 2VRB) (Kim *et al.*, 2008c). Each structure was analysed using the PISA server (Krissinel, 2015), the ProFace server (Saha *et al.*, 2006) and the Shape Complementarity program (Lawrence and Colman, 1993). From the various outputs, the most striking results were the hydrophobic P-values determined by PISA and the Residue Propensity Scores reported by ProFace (see legend to **Table 5.4** for definitions). For the SimC7 structures, these results diminished the significance of both types of dimer, whereas for TMR, they were fully supportive of a dimer being biologically relevant, despite having a somewhat smaller dimer interface area than the values obtained for both interface types in SimC7. Indeed, assessing the biologically relevant assembly from interfacial area alone can lead to erroneous conclusions (Ponstingl *et al.*, 2000; Bahadur *et al.*, 2004; Krissinel, 2010).

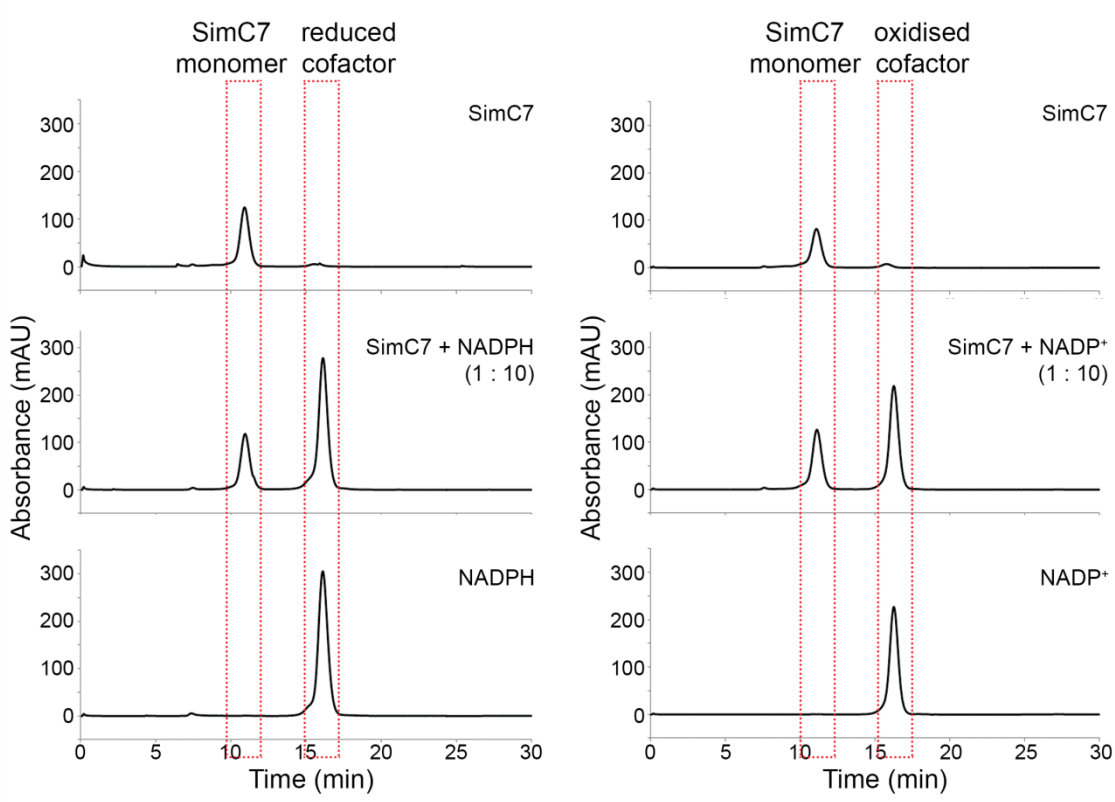


Figure 5.6

Size exclusion chromatography of SimC7 in the presence and absence of cofactor.

The oligomeric state of SimC7 was analysed by size exclusion chromatography in the absence (top) or presence (middle) of cofactor. SimC7 eluted as monomer under all conditions tested with a retention time of 11 min. The second peak seen at 16 min was caused by the absorbance of the cofactor (bottom). Conditions were equivalent to crystallisation screens with concentrated protein (4 mg/mL) and ten-fold molar excess of reduced (NADPH) or oxidised ($NADP^+$) cofactor using a mobile phase of 20 mM potassium phosphate (pH 7.2) with 5% (v/v) glycerol. For analysis with cofactor, SimC7 was pre-incubated on ice for 4 hours.

Table 5.4

Interfaces of crystallographically observed dimers.

Structure	PDB code	Space group	Operator	Interface area (Å^2)[a]	$\Delta^i G$ (kcal/mol)[a.b]	Hydrophobic P-value[a,c]	H-bonds/ salt bridges[a]	Residue Propensity Score[d]	Shape comple-mentarity[e]
SimC7 Apo form 1	5L40	$P4_12_12$	y+1,x-1,-z	1064	-9.0	0.266	10/4	-2.74	0.66
SimC7 Apo form 2	5L45	C2	2-fold NCS[‡]	1246	-11.4	0.422	5/3	-1.66	0.58
SimC7 Binary	5L3Z	$P4_12_12$	y+1,x-1,-z	1032	-11.2	0.219	8/2	-3.54	0.70
TMR Binary	2JL1[*]	$C222_1$	-x,y,-z+1/2	892	-13.9	0.041	6/8	1.40	0.75

[a]Values from the PISA server (http://www.ebi.ac.uk/pdbe/pisa/) (Krissinel, 2015).

[b] Indicates the solvation free energy gain upon formation of the interface, in kcal mol^{-1}. The value is calculated as the difference in total solvation energies of isolated and interfacing structures. Negative $\Delta^i G$ corresponds to hydrophobic interfaces, or positive protein affinity.

[c] The probability of finding a patch of the same area on the protein surface that would be more hydrophobic than the interface. Thus, lower values indicate specific hydrophobic patches, which are likely to be preferred in protein-protein interactions, while larger values (≥ 0.5) indicate interfaces that are non-specific.

[d] As calculated by the Proface server (http://resources.boseinst.ernet.in/resources/bioinfo/interface/) (Saha *et al.*, 2006), where on average, the interfaces of functional homodimers (and protein-protein complexes in general) have positive values, and large crystal-packing interfaces have negative values.

[e] Values calculated by the program SC (Lawrence and Colman, 1993), being a measure of the extent to which, on average, the normal vectors between closest-neighbour opposing points within the molecular interface are antiparallel. A value of 1.0 implies that the surfaces fit exactly because all such vectors are perfectly antiparallel, whereas values below 0.65 indicate relatively poor shape complementarity.

[*] TMR also crystallises in another form (PDB accession number 2VRC) where the ASU contains four molecules arranged as two homodimers, both being equivalent to the dimer generated by the crystallographic symmetry in this crystal form. This dimer does not resemble either of the two putative SimC7 dimers.

[‡] NCS = non-crystallographic symmetry.

Therefore, an oligomeric state inferred from a crystal structure should be treated as a hypothesis unless confirmed by experiment. Moreover, the fact that I observed both distinctly different dimers when the protein is in the *apo* form raises concerns about their relevance. Consequently, both of the dimeric assemblies observed for SimC7 are most likely crystallographic artefacts.

5.5 Structural homologs of SimC7

Structures annotated as SDR proteins (PFAM family PF00106) are prevalent in the Protein Data Bank with >600 entries (Finn *et al.*, 2010; Berman *et al.*, 2014). To look for structural homologs of SimC7, I carried out a structure-based similarity search using the DALI server (Holm and Sander, 1995) (**Table 5.5**). Strikingly, characterised angucyclinone ketoreductases ranked very low in the search, the closest match being LanV (**Figure 5.7**), which was the 166th ranked hit after filtering for sequence redundancy (**Table 5.5**). Instead, the two most structurally similar proteins to SimC7 were quinone oxidoreductase (QOR2) from *Escherichia coli* (PDB accession number 2ZCV) (Kim *et al.*, 2008b) (**Figure 5.7**) and triphenylmethane reductase (TMR) from *Citrobacter* sp. KCTC 18061P (PDB accession number 2VRB) (Kim *et al.*, 2008c). QOR2 and TMR share with SimC7 the ability to reduce substrates with extensively conjugated pi-systems but have roles in the detoxification of xenobiotics rather than the biosynthesis of natural products. The majority of the closest structural homologs of known function are involved in sugar biosynthesis, many of them epimerases such as UDP-hexose 4-epimerase from *Thermatoga maritima* (Shin *et al.*, 2015). In addition, a handful of entries are annotated as non-catalytic regulatory proteins, such as human HSCARG (PDB accession number 2EXX) (Zheng *et al.*, 2007), which is proposed to control gene expression in response to fluctuations in the intracellular $NADP^+$:NADPH ratio. Also scoring highly were a number of reductases from plant sources, including cinnamoyl-CoA reductase (PDB accession number 4R1S) (Pan *et al.*, 2014), a key enzyme in lignin biosynthesis.

Table 5.5

Selected structural homologues of SimC7.

Protein	Source	Biological unit[a]	PDB code[b]	Ligand bound[c]	Resol. (Å)	DALI output					"Catalytic" triad	Reference
						Rank[d]	Z-score	R.m.s.d. (Å)	aligned residues	Identity (%)		
SimC7	*Streptomyces antibioticus* Tü6040	Monomer	5L4L	NADP$^+$ & 7oxo-SD8	1.20	-	-	0.00	278	100	Ser95, Ile108, His112	This work
QOR2	*Escherichia coli*	Monomer	2ZCV	NADP$^+$	1.60	1	27.4	3.1	266	23	Leu104, Leu114, His118	(Kim *et al.*, 2008b)
TMR	*Citrobacter* sp. KCTC 18061P	Dimer	2JL1	NADP$^+$	1.96	2	27.0	3.2	268	22	Tyr107, Leu117, His121	(Kim *et al.*, 2008c)
HSCARG	*Homo sapiens*	Monomer/ Dimer	2EXX	NADP$^+$	2.40	4	24.8	3.1	263	22	Leu114, His129, Lys133	(Zheng *et al.*, 2007)
UDP-hexose 4-epimerase	*Thermotoga maritima*	Dimer	4ZRM	NADP$^+$ & UDP-Glc	2.00	6	22.9	2.5	247	20	Thr117, Tyr143, Lys147	(Shin *et al.*, 2015)
Cinnamoyl-CoA reductase	*Petunia hybrid*	(Monomer)	4R1S	NADP$^+$	1.60	8	22.4	3.0	246	22	Ser123, Tyr157, Lys161	(Pan *et al.*, 2014)
LanV	*Streptomyces cyanogenus*	Dimer	4KWI	NADP$^+$ & Lando	2.00	166	14.3	3.0	173	22	Ser147, Tyr160, Lys164	(Paananen *et al.*, 2013)
$UrdM_{red}$	*Streptomyces fradiae*	(Tetramer)	4OSP	NADP$^+$ & rabelomycin	2.25	167	14.3	3.2	177	20	Ser147, Tyr160, Lys164	(Patrikainen *et al.*, 2014)
HedKR	*Streptomyces griseoruber*	(Tetramer)	3SJU	NADP$^+$	2.40	N/A	15.5	4.0	185	19	Ser142, Tyr155, Lys159	(Javidpour *et al.*, 2011a)
ActKR	*Streptomyces coelicolor*	Tetramer	2RHC	NADP$^+$ & Emodin	2.10	N/A	15.1	3.9	185	18	Ser144, Tyr157, Lys161	(Korman *et al.*, 2008)

[a] Experimentally determined (e.g. by size exclusion chromatography), unless shown in brackets, in which case it was inferred from the crystal structure alone. For HSCARG, the oligomeric state (as judged by dynamic light scattering) was dependent on the $NADP^+$ concentration.

[b] Results of a DALI search (http://ekhidna.biocenter.helsinki.fi/dali_server) (Holm and Sander, 1995) using the SimC7 ternary complex structure as the template. The hits were filtered for redundancy; where a relevant ligand bound structure exists for a particular enzyme, this is the one that is shown.

[c] Lando = 11-deoxy-6-oxylandomycinone. Where the cofactor is underlined, this indicates that the nicotinamide ring adopts the *syn* configuration like SimC7; otherwise the configuration is *anti*.

[d] Ranking of the DALI hit (after redundancy filtering at 90% sequence identity). For some classes of enzymes, e.g. the sugar epimerases, only the top hit is shown. The lower ranking hits LanV and $UrdM_{red}$ were also chosen as they were relevant to the discussion. For the same reason, ActKR and HedKR were included, despite not being picked up by a blind DALI search.

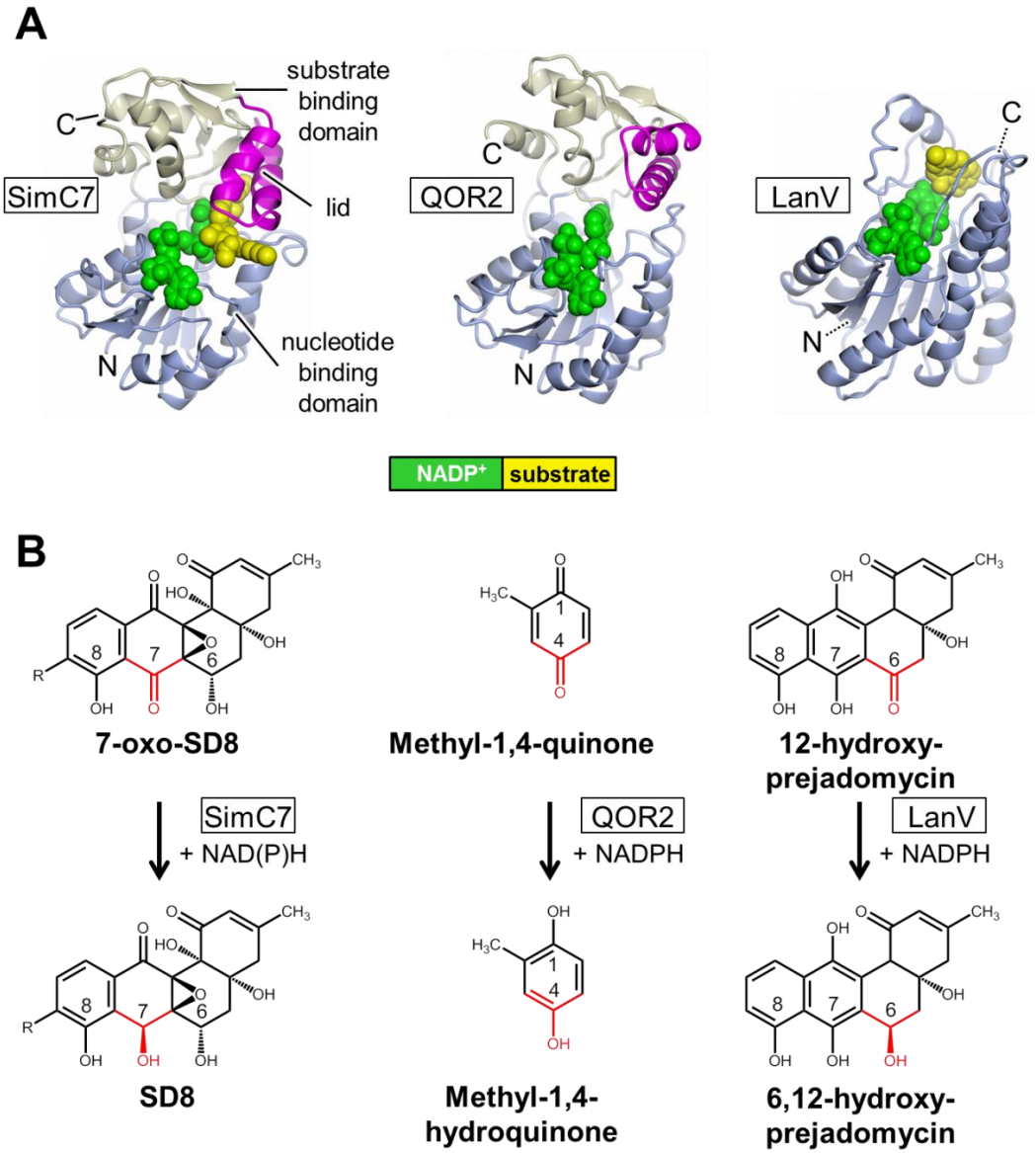


Figure 5.7

Comparison of the ternary complexes of SimC7 and LanV and the binary complex of QOR2 and respective reaction mechanisms.

(A) Structures are depicted in cartoon representation, where pale blue, beige and magenta colouration indicate the nucleotide binding domain, the substrate binding domain and the lid motif, respectively (N.B there is no distinct substrate binding domain or lid motif in LanV). The ligands are shown as van der Waals spheres, with the $NADP^+$ cofactor shown in green and the substrate (where present) shown in yellow. See **Table 5.5** for more extensive comparison of SimC7 structural homologues. (B) Reaction mechanism for SimC7 (only the angucyclinone is shown for 7-oxo-SD8 and SD8), QOR2 and LanV.

In terms of their overall structures, the closest DALI hits, together with SimC7, fall into the extended SDR subfamily of proteins, which are characterised by having two distinct domains that together form a partially occluded active site pocket at their junction (**Figure 5.7a-b**). In contrast, the more distantly related angucyclinone ketoreductases belong to the classical SDR subfamily (Kavanagh *et al.*, 2008), where three insertions within the core Rossmann fold motif delineate a more accessible active site cavity, but do not constitute a well-defined substrate binding domain (**Figure 5.7c**).

5.6 Cofactor binding

The $NADP^+$ cofactor in the binary SimC7 complex is bound in an extended configuration on top of the Rossmann fold motif and held by a network of hydrogen bonds such that the nicotinamide moiety adopts the *syn*-conformation relative to the adjacent ribose sugar, with the *Si*-face of the ring directed toward the substrate binding pocket (**Figure 5.2d-e**, **Figure 5.5** and **Figure 5.8**). This conformation is coordinated by Asn137 and His112, which make hydrogen bonds with the carboxamide of the nicotinamide ring and the C-2′ hydroxyl group of the ribose sugar, respectively.

SimC7 accepts NADH and NADPH as cofactor (see **chapter 4.4**). Nevertheless, SimC7 has a strong preference for NADPH, which is likely due to the additional favourable contacts with the 2′ phosphate, specifically salt bridges with the guanidinium groups of Arg9 and Arg32. Further hydrogen bonds are made by the side-chain of Asp49 and the backbone amide of Leu50 with the adenine moiety, and by the side-chain of Tyr69 and the backbone amides of Ser11 and Ile12 with the diphosphate group of the cofactor, where the latter two amino acids correspond to the second ‘xx’ dipeptide in the TGxxGxxG cofactor binding motif characteristic of the extended SDR subfamily (Kavanagh *et al.*, 2008) (**Figure 5.3**). The functional significance of the first and second Gly residues of the motif is also apparent in this protein-cofactor complex, since side-chains at these

positions would result in steric clashes with the cofactor. In contrast, the previously identified polyketide ketoreductases LanV, $UrdM_{red}$, HedKR and ActKR all belong to the classical SDR subfamily and therefore contain a different cofactor binding motif (TGxxxGxG) and the classic active site motif YxxxK.

5.7 Substrate binding

In the ternary complex with substrate, determined at 1.2 Å resolution, the angucyclic ring system of 7-oxo-SD8 binds adjacent and parallel to the nicotinamide ring of the cofactor (**Figure 5.2d-e** and **Figure 5.8**), where it adopts a relatively planar conformation differing from the conformations seen in the DNA gyrase-SD8 and SimR-SD8 complexes, in which the A-ring of the angucyclinone in SD8 is tilted upwards towards the epoxide (Le *et al.*, 2011b; Hearnshaw *et al.*, 2014) (**Figure 5.9**). This planar conformation is most likely enforced by the shape of the very constricted and highly hydrophobic substrate pocket created by residues Ile96, Trp100, Ile108, His112, Glu132, Ala133, Thr160, Tyr229, Ile237, Gln241, and Arg244 (**Figure 5.4** and **Figure 5.5**). Within this hydrophobic pocket, 7-oxo-SD8 is bound only by a single direct hydrogen bond between the side-chain of Ser95 and the C-7 carbonyl oxygen of the angucyclinone moiety, which may help to position the latter exactly above the C-4 position of the nicotinamide ring, where it is poised for direct hydride transfer (highlighted by black spheres in **Figure 5.2c-e**, **Figure 5.5** and **Figure 5.8**). As mentioned above, the natural substrate for SimC7 is likely to be a 7-oxo angucyclinone intermediate lacking the deoxysugar, the tetraene linker and the aminocoumarin. Consistent with this, only the angucyclinone moiety is buried in the active site of SimC7 in the ternary complex. Roughly half of the tetraene linker is visible in the electron density, projecting away from the protein surface, and the aminocoumarin ring is not resolved at all (**Figure 5.8**).

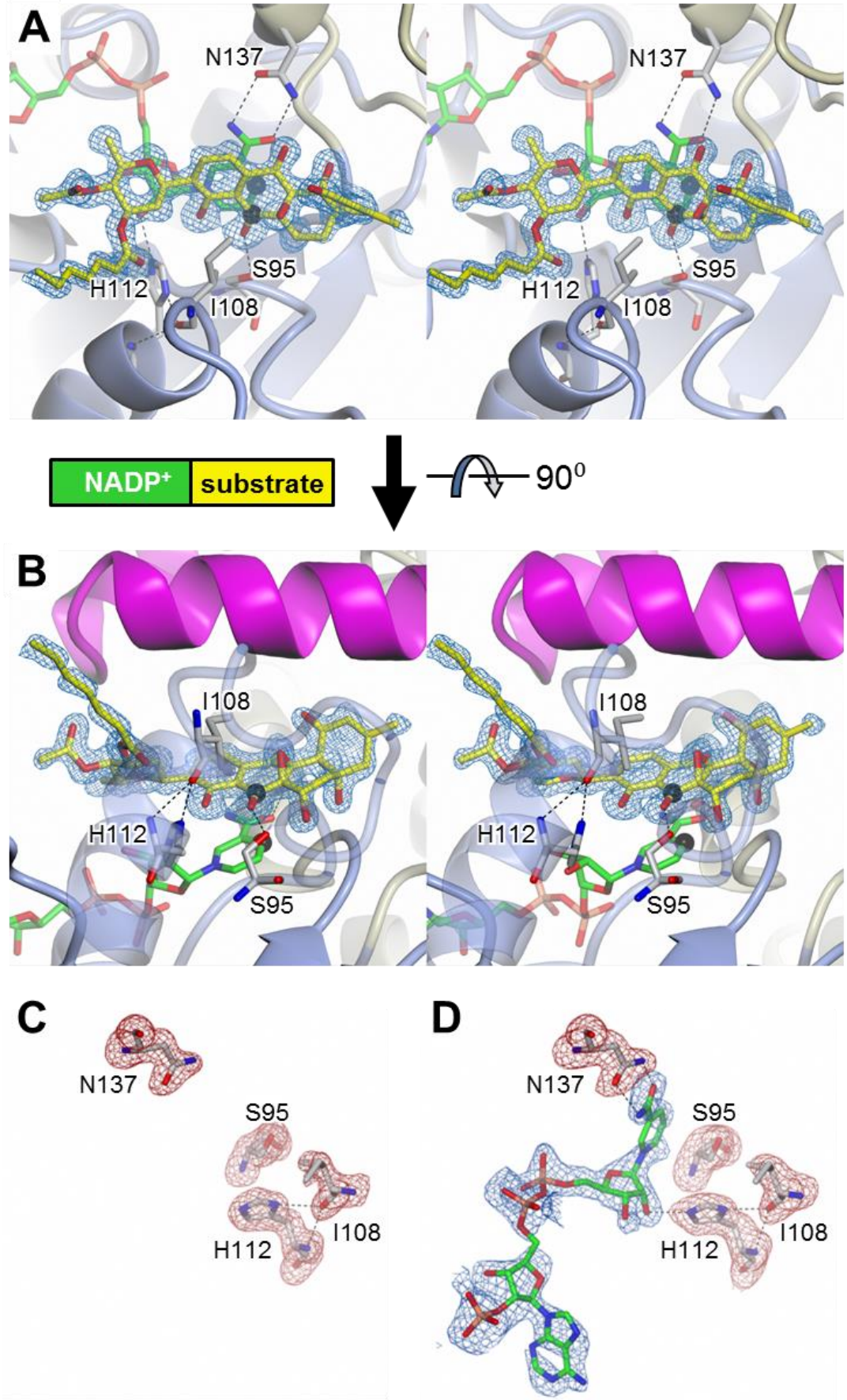

A
N137
S95
H112
I108
N137
S95
H112
I108
NADP+
substrate
90⁰
B
I108
H112
S95
I108
H112
S95
C
N137
S95
H112
I108
D
N137
S95
H112
I108

Figure 5.8

Simulated annealing omit electron density maps.

Parts (A) and (B) show orthogonal stereo views of the active site of the SimC7 ternary complex. Ligands are depicted in stick representation, where green indicates carbon atoms of the cofactor and yellow, carbons of the substrate; C-4 of the cofactor and C-7 of the substrate are specifically highlighted by small black spheres. Superposed upon the substrate in blue is a simulated annealing omit map calculated at 1.2 Å resolution (see Materials and Methods, **chapter 2.7.3**). Also shown are the catalytic triad residues, as well as Asn137, which is important in maintaining the *syn* conformation of the cofactor. The view in part (A) is equivalent to that in **Figure 5.2e** (again the lid motif has been removed for clarity). In part (B), the protein backbone encompassing the β5-α4 loop and bearing the catalytic triad has been shown in transparent mode as it would otherwise obscure the detail of the active site; the lid is in magenta. Parts (C) and (D) show omit maps (red density) for just the active site residues in Apo form 1 (1.6 Å resolution) and the binary complex (1.95 Å resolution), respectively. In the latter panel a separate omit map is also shown (blue density), which was calculated from the final model lacking only the cofactor. The view in parts (C) and (D) is roughly equivalent to that used in **Figure 5.2d**. All omit maps were contoured at ~3.0 σ.

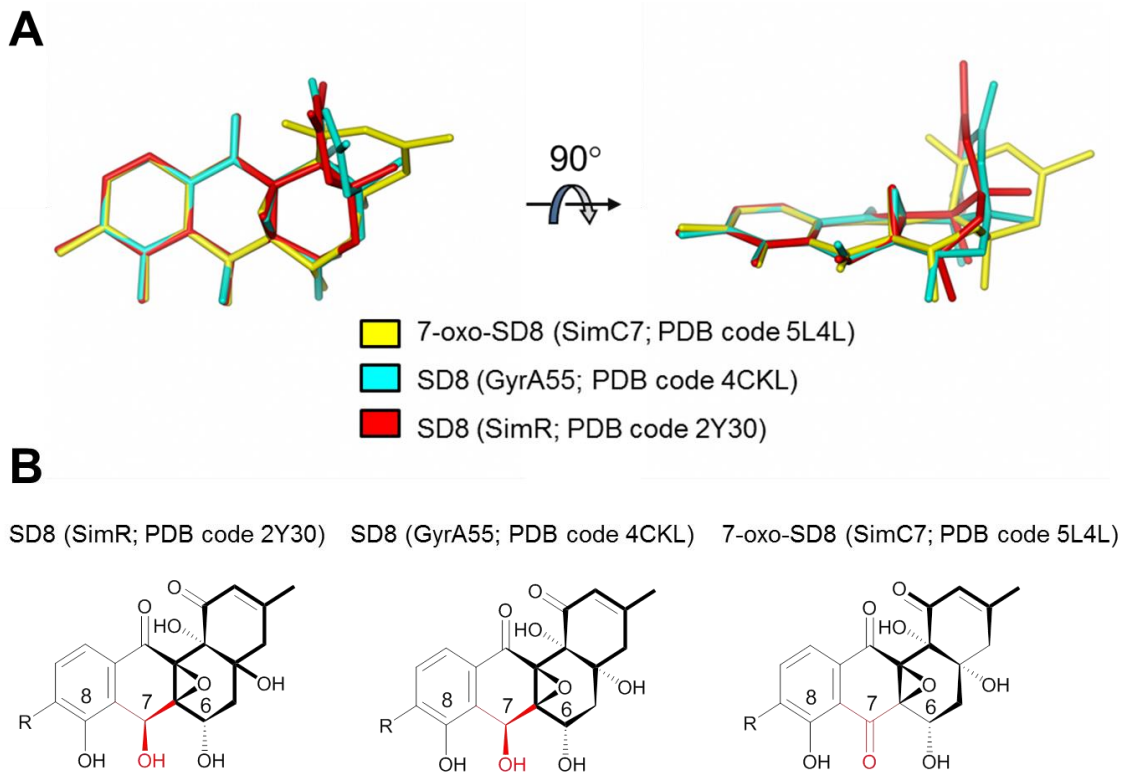


Figure 5.9

Comparison of the 7-oxo-SD8 conformation in the ternary complex of SimC7 with SD8 conformations observed previously in other complexes.

(A) Orthogonal views showing a least-squares superposition of the angucyclic polyketide moiety of 7-oxo-SD8 (based on the 6 carbon atoms of ring D) upon SD8 from the complex with GyrA55 and SD8 from the complex with SimR. (B) Chemical drawings of the angucyclinone moiety of SD8 and 7-oxo-SD8 to visualise conformational differences between the three crystal structures. The C-7 position is highlighted in red.

5.8 A novel catalytic mechanism for an angucyclinone ketoreductase

SimC7 lacks the canonical Ser-Tyr-Lys SDR catalytic triad (**Table 5.5**). While the serine is conserved (Ser95), the other two residues (i.e. the YxxxK motif) are not, being instead replaced by Ile108 and His112, respectively. This Ser-Ile-His triad is unlike any described for the five subfamilies of SDRs defined by Kavanagh (Kavanagh *et al.*, 2008). Particularly surprising is the absence of the tyrosine residue that acts as the acid/base catalyst in the classical SDR mechanism. Inspection of the structure of the ternary complex shows that none of the five tyrosine residues in SimC7 is sufficiently close to C-7 of the angucyclinone ring system of the substrate to play a direct role in catalysis. Further, the structure also shows that there is no alternative residue that could act as an acid/base catalyst. Consequently, SimC7 must perform ketoreduction of 7-oxo-SD8 via a novel mechanism.

Based on the structure of the ternary complex of the enzyme with $NADP^+$ and 7-oxo-SD8, I propose a simple two-step mechanism for SimC7 that does not depend on catalytic residues in the protein, but rather takes advantage of the specific properties of the substrate itself, and is thus a novel example of substrate-assisted catalysis (Dall'Acqua and Carter, 2000). In the first step, the hydrophobic environment of the substrate binding pocket and the juxtaposition of the quinone-like C-ring and the phenyl-like D-ring of the angucyclinone favour the formation of an intramolecular hydrogen bond between the proton on the C-8 hydroxyl group and the oxygen of the neighbouring C-7 carbonyl group (**Figure 5.10**). This enhances the polarisation of the latter such that the electrophilicity of C-7 is increased, making it a good acceptor for a direct hydride transfer from the 4-*pro*-S position of the nicotinamide ring. Crucially, the hydride donor and acceptor carbon atoms are only 3.0 Å apart in the crystal structure. The C-7 hydroxyl group is then formed by internal proton transfer from the neighbouring C-8 hydroxyl group, generating a phenolate intermediate in which the negative charge on the C-8 oxygen atom is stabilised by the aromatic D-ring. In the second step of the reaction, the phenolate intermediate leaves the substrate

binding pocket and the proton required to reinstate the C-8 hydroxyl group is recovered by abstraction from bulk water, which is not possible within the confines of the active site (**Figure 5.10**). Exchange of the negatively charged reaction intermediate is most likely accelerated by repulsion from the hydrophobic active site cavity. Finally, the direct hydride attack from below the angucyclic polyketide unambiguously explains the 7-*S*-stereochemistry of simocyclinones. In support of this proposed mechanism, molecular modelling (Chem3D Pro v.13) predicts the existence of the key intramolecular hydrogen bond between the C-8 hydroxyl group and the C-7 carbonyl group of 7-oxo-SD8, and an equivalent intramolecular hydrogen bond is observed in the small molecule crystal structure of panglimycin, a closely related polyketide (Fotso *et al.*, 2008). Furthermore, molecular modelling (Chem3D Pro v.13) also predicts that the C-8 hydroxyl group will have the most acidic and exchangeable proton in the angucyclic polyketide, with an estimated pKa of 6.9-7.7, and, therefore, could readily transfer to the neighbouring C-7 oxygen at the end of step 1. Indeed, there is no other possible proton donor (neither protein nor water derived) sufficiently close to O-7 to perform this role.

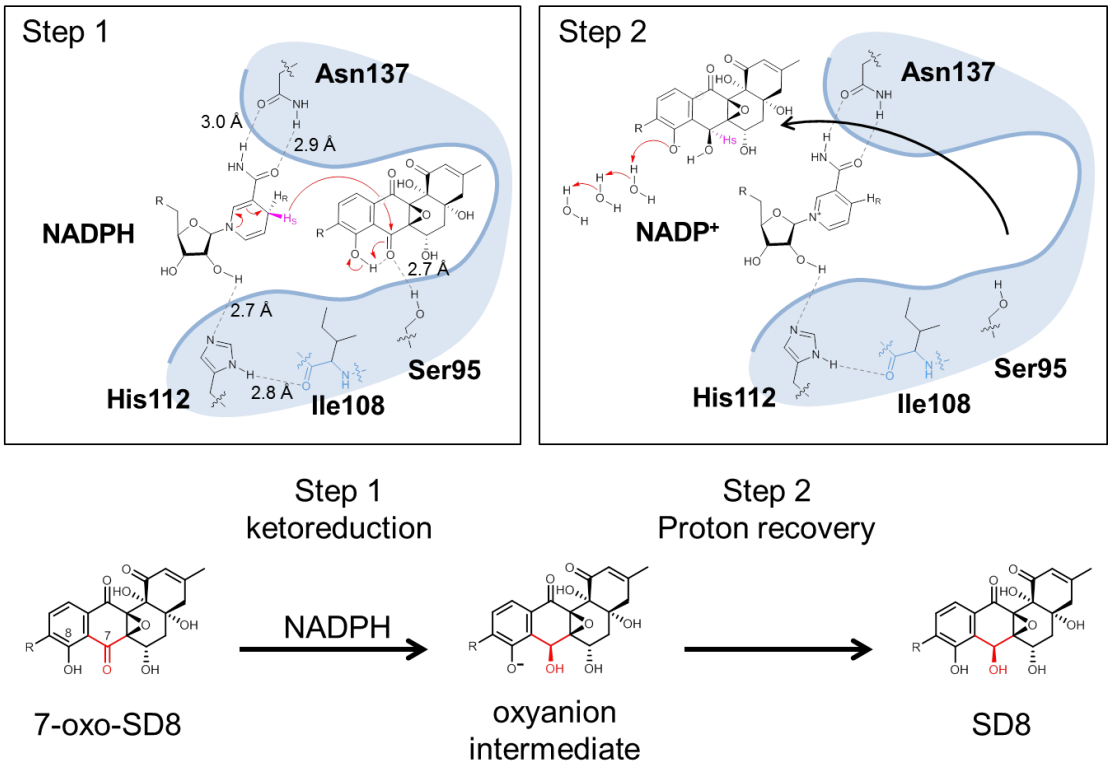


Figure 5.10

Proposed SimC7 reaction mechanism.

SimC7 has an atypical catalytic triad comprised of Ser95, Ile108, and His112. In the first step of the proposed mechanism, the C-7 carbonyl group of the substrate (7-oxo-SD8) is reduced by transfer of the 4-*pro*-*S* hydride of the cofactor onto the C-7 carbon of the substrate. This transfer from below the C-ring results in the characteristic 7-*S*-stereochemistry of SD8. Ketoreduction at position C-7 is completed by an intramolecular proton transfer from the neighbouring C-8 hydroxyl group of the angucyclinone; the resultant negative charge on the latter is stabilised by the adjacent aromatic ring system (ring D). In the second step, the C-8 oxyanion intermediate regains a proton from bulk water after leaving the substrate binding pocket. Note that there are no water molecules in the active site pocket that could contribute to the reaction mechanism. In the ternary complex, the nearest water to O-7 of the angucyclic polyketide is ~5.5 Å away, and the nearest water to O-8 is ~4.9 Å away. Due to steric constraints within the pocket, neither could approach the substrate oxygen atoms without a repositioning of the substrate or a conformational change in the protein.

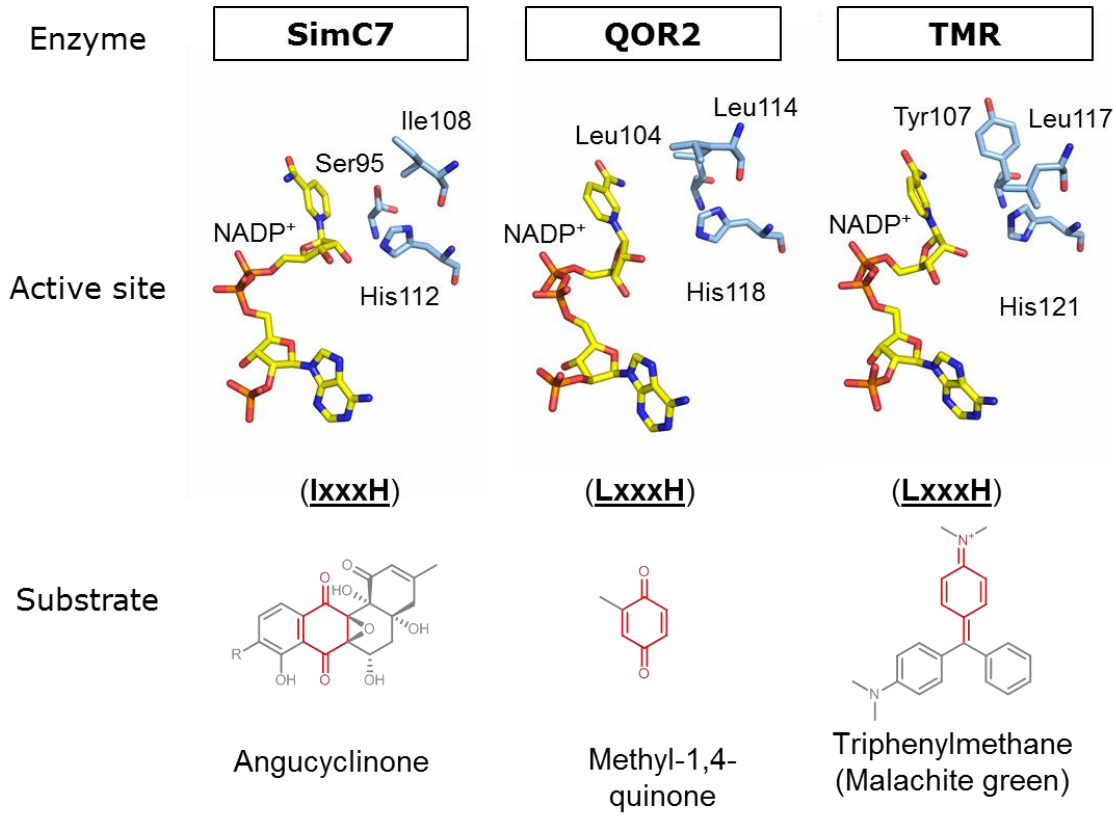


Figure 5.11

Variations in the active site triad and substrates between SimC7, QOR2 and TMR.

Arrangement of the cofactor (yellow) and the protein residues (blue) that constitute the active site triad in SimC7, quinone oxidoreductase (QOR2, PDB accession number 2ZCV) from *E. coli*, and triphenylmethane reductase (TMR, PDB accession number 2JL1) from *Citrobacter* sp. The substrate of each enzyme is shown below. The natural substrate of TMR is not known.

5.9 Mutagenesis of the SimC7 active site

To investigate the potential roles of the SimC7 'catalytic triad' residues in the proposed reaction mechanism, I mutagenized Ser95, Ile108, and His112 (**Table 5.6**). In the wild-type enzyme, the hydroxyl group of Ser95 could aid catalysis by helping to bind and correctly orient the substrate, and by providing additional polarisation to the C-7 carbonyl group via a hydrogen bond, the latter role being consistent with the function proposed for the structurally equivalent Ser/Thr residues in the classical SDR mechanism (Oppermann *et al.*, 2003; Kavanagh *et al.*, 2008; Korman *et al.*, 2008; Kallberg *et al.*, 2010; Persson and Kallberg, 2013). Despite the serine being within hydrogen bonding distance of the C-7 carbonyl group, the serine residue retains activity (86% substrate conversion relative to the wildtype enzyme), suggesting that this is not the proton donor and Ser95 is not crucial for activity (**Table 5.6**). Therefore, SimC7 must be able to orient 7-oxo-SD8 for catalysis without the necessity for the hydrogen bond from Ser95, and this likely arises because the active site cavity of the complex with $NADP^+$ closely matches the shape of the substrate (**Figure 5.5**).

Ile108 contributes to the hydrophobic surface that forms one side of the active site cavity and appears to have a rather non-specific role in helping to trap the angucyclinone group against the nicotinamide ring of the cofactor (**Figure 5.5**). Unsurprisingly, an I108A substitution had almost no effect on the enzyme, whereas changing it to aspartate (I108D) abolished activity (**Table 5.6**). In the latter case, the introduction of a negative charge would interfere with the hydrophobic environment and could disrupt the intramolecular hydrogen bond in the substrate, both being necessary for catalysis.

Finally, mutation of His112 (H112A, H112N, and H112Q) strongly reduced or abolished enzyme activity (**Table 5.6**), consistent with its key role in binding and positioning the cofactor via a hydrogen bond to the 2′-hydroxyl of the nicotinamide ribosyl moiety (**Figure 5.2d-e** and **Figure 5.10**).

To exclude that the reduced substrate conversion by certain SimC7 mutants (i.e. mutations I108D, H112A, H112N, and H112Q) was due to incorrect or incomplete protein folding, I recorded the far UV circular dichroism (CD) spectra of respective mutant enzymes (**Figure 5.12**). The CD spectral signatures of all tested SimC7 mutants were almost identical to the wild type protein, confirming that the above mentioned changes in substrate conversion were entirely caused by respective introduced point mutations.

Table 5.6

Enzymatic activities of SimC7 active site mutants.

Substrate conversion was analysed using HPLC by quantifying the reaction product, SD8, against purified SD8 standards. Results and standard deviations are shown from three independent experiments. Specific activity is given as enzyme activity per milligram of protein (i.e. nanomoles of product formed per minute per milligram of protein).

Mutant	Conversion (%)	Specific activity (nmol min^{-1} mg^{-1})
Wildtype	100.0 ± 7.3	13.9 ± 1.0
S95A	86.1 ± 7.1	11.9 ± 1.0
I108A	91.5 ± 13.8	12.7 ± 1.9
I108D	1.2 ± 0.8	0.2 ± 0.1
H112A	12.5 ± 3.6	1.7 ± 0.5
H112N	5.8 ± 2.5	0.8 ± 0.3
H112Q	1.0 ± 1.4	0.1 ± 0.2

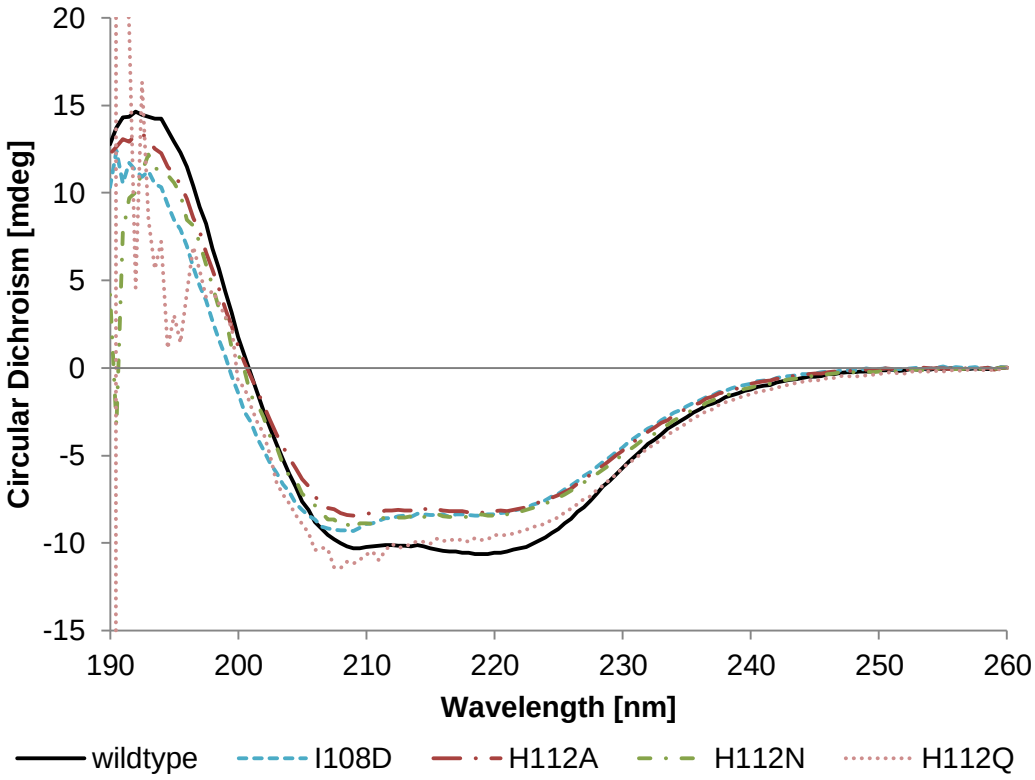


Figure 5.12

Far UV circular dichroism (CD) spectra of inactive SimC7 mutants.

The CD spectra of all inactive SimC7 mutants were nearly identical to the CD spectrum recorded for the wildtype enzyme. SimC7 has a secondary structure composition of 44% α-helix, 20% β-sheet, 10% β-turn, and 25% other.

5.10 Discussion

Substrate-assisted catalysis refers to a functional group in a substrate that contributes to enzyme catalysis. This has been widely demonstrated for hydrolytic enzymes (serine proteases, GTPases, type II restriction endonucleases, lysozymes and chitinases) but there are few examples from other enzyme groups like the oxidoreductases. Substrate-assisted catalysis can occur by different mechanisms including general base catalysis and stabilisation of intermediates (Dall'Acqua and Carter, 2000).

The sequence of SimC7 is distinct from previously characterised polyketide ketoreductases and the structural data reported in this study suggest that SimC7 catalyses a novel substrate-assisted, two-step reaction for the reduction of the C-7 carbonyl group. This mechanism involves the intramolecular transfer of a substrate-derived proton to generate a phenolate intermediate, obviating the need for proton transfer for a canonical SDR active site tyrosine. Like SimC7 (Ser-Ile-His), the two closest structural homologs, TMR (Tyr-Leu-His) and QOR2 (Leu-Leu-His), also have unusual active site triads (**Figure 5.11** and **Table 5.6**). Thus, SimC7, TMR and QOR2 share an I/LxxxH motif and have substrates with extensively conjugated pi-systems (**Figure 5.11**). No enzyme-substrate complex crystal structures have been described for TMR or QOR2 and no firm proposals exist for their catalytic mechanisms, but their substrate structures (**Figure 5.11**) and the data provided here suggest they are also likely to employ non-canonical mechanisms. The data from this study therefore point to members of the extended SDR subfamily having I/LxxxH active site motifs as a source of new biochemistry.

5.11 Acknowledgements

I would like to thank Dr David M. Lawson and Clare E.M. Stevenson for providing material for crystallisation screens, for harvesting crystals for X-ray crystallography and for data collection and refinement. Protein structures were generated by me but final refinements before submission as well as Figures 5.2, 5.5, 5.7, 5.8 and 5.9 and Tables 5.1-5.5 were generated by Dr. David Lawson. I would also like to thank the staff at the Diamond Lightsource (Oxford, UK) for access to beamlines I03, I04, and I04-1.

Chapter 6

Alternative approaches to manipulate simocyclinone biosynthesis for optimised heterologous production

6.1 Introduction 195

6.2 Addition of the adsorptive hydrophobic resin HP20..... 195

6.3 Spontaneous resistant mutants of *S. coelicolor* M1152 198

6.4 Alternative heterologous hosts..... 202

6.5 Mutational analysis of the *sim* gene cluster in *S. antibioticus* Tü6040..... 203

6.6 Evidence that simocyclinone intermediates are the cause of the toxicity to *S. coelicolor* M1152, and that their target is not DNA gyrase 204

6.7 Genome sequencing reveals *S. antibioticus* Tü6040 to be a 'talented producer' of secondary metabolites..... 205

6.8 Acknowledgements..... 218

6.1 Introduction

The natural producer of simocyclinones, *S. antibioticus* Tü6040, is highly resistant to SD8 (>100 µg/mL) (Schimana *et al.*, 2000). The two membrane transport proteins SimEx1 and SimEx2 from the *sim* biosynthetic gene cluster were assumed to confer resistance to simocyclinones (Galm *et al.*, 2002; Trefzer *et al.*, 2002). Unexpectedly, introduction of the *sim* cluster into *S. coelicolor* M1152 strongly impaired growth and sporulation of the strain. The reason for this toxicity to the heterologous host remained unclear.

Here I describe several alternative approaches I took in an attempt to improve manipulation and heterologous production of simocyclinones. I have evidence that production of simocyclinone intermediates is toxic to the heterologous hosts. My experiments indicate that early simocyclinone intermediates lacking the tetraene linker and the aminocoumarin may have inhibitory activity against unknown intracellular targets. Sequencing of the genome of *S. antibioticus* Tü6040 revealed a vast range of secondary metabolite gene clusters but no further obvious resistance mechanisms against simocyclinones could be identified.

6.2 Addition of the adsorptive hydrophobic resin HP20

Adsorptive polymeric resins have been shown to improve production of specific secondary metabolites (Nielsen *et al.*, 2010; Gonzalez-Menendez *et al.*, 2014). The application of polymeric resins dates back to the 1970s and they have since been used during fermentation processes for bacteria and fungi to significantly increase the yield of secondary metabolites, prevent their degradation or decrease their toxic side effects (Phillips *et al.*, 2013). Therefore, I added an adsorptive hydrophobic resin (5% w/v of Diaion® HP20) to the production medium and grew the improved heterologous host expressing the mutant *sim* cluster (*S. coelicolor* M1152ex1::12IΔC7) as described. Extracts from culture supernatants and the hydrophobic resin were analysed by LC-MS-UV and compared to cultures grown in parallel from the same pre-cultures without hydrophobic resin.

Addition of the adsorptive resin HP20 strongly repressed mycelial growth. Furthermore, cultures produced less of a dark pigment, which was generally observed for M1152 and M1152ex1 expressing the *sim* cluster with *simC7* deleted. Unfortunately, the total yield of 7-oxo-SD8 extracted from cultures grown in the presence of the HP20 resin was ~4-8-fold reduced (0.3-0.5 mg/mL) compared to cultures grown without resin (2.0-2.6 mg/mL) (**Figure 6.1**). In addition, concentrated crude extracts from cultures grown in the presence of HP20 appeared green in colour in contrast to the typical yellow-brown colour of concentrated simocyclinone extracts (**Figure 6.1**). The reason for this colour change remained unclear but was not caused by changes in pH or temperature.

In summary, addition of an adsorptive hydrophobic resin did not improve heterologous production of simocyclinones. Instead, the presence of HP20 resin decreased culture growth and production of simocyclinones, which could have been caused by mechanical stress in addition to the strongly impaired growth phenotype of *S. coelicolor* M1152 and M1152ex1.

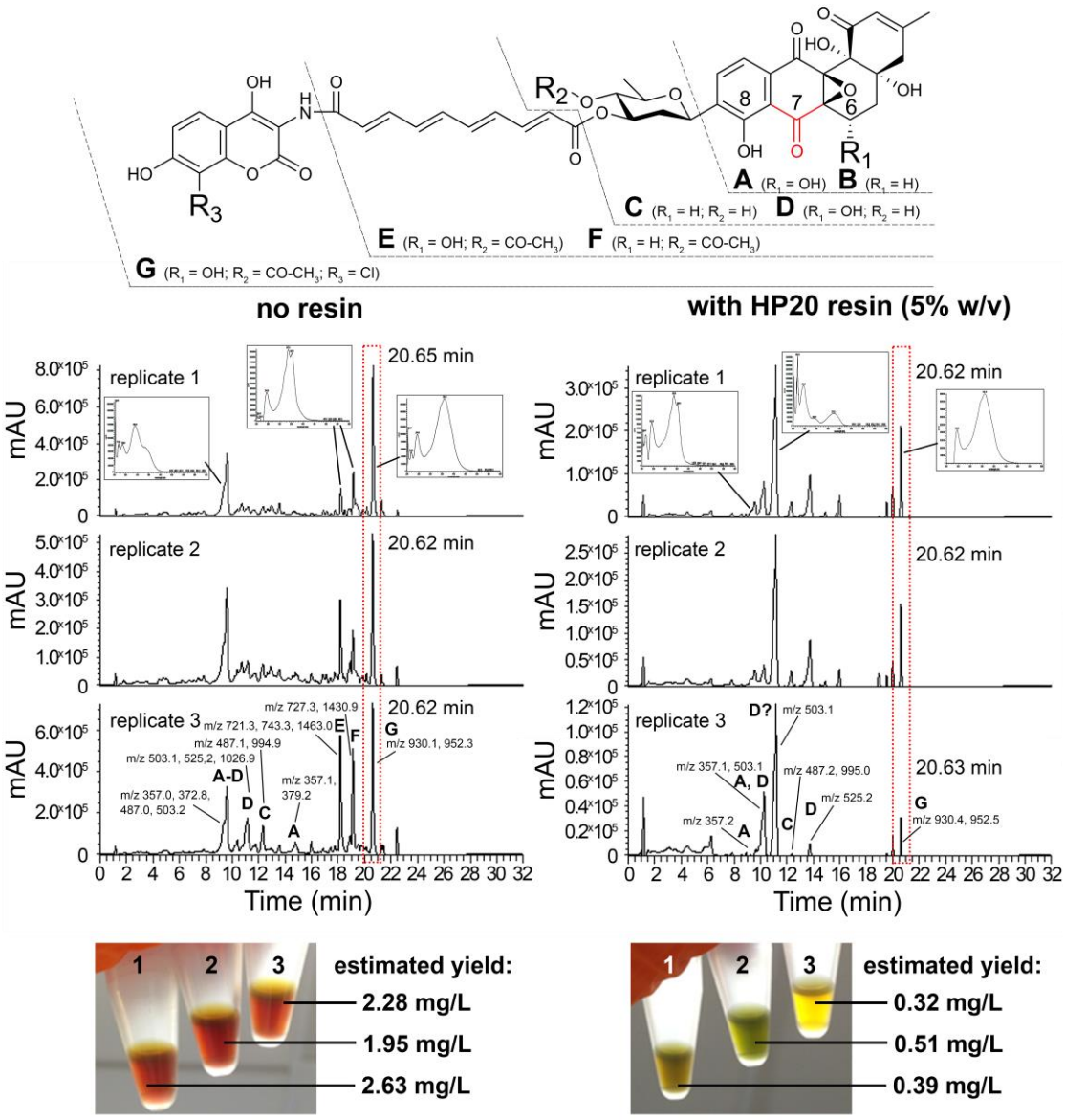


Figure 6.1

Production of 7-oxo-SD8 in absence and presence of an absorptive hydrophobic resin.

Cultures of M1152ex1::12IΔC7 were grown in the absence (left) and presence (right) of HP20 resin (Diaion®; 5% w/v). Cultures were grown in triplicate for 6 days at 30 °C and concentrated extracts with a mixture of simocyclinone intermediates were analysed by LC-MS-UV. Crude extracts (2.5-fold concentrated in methanol) from cultures grown with HP20 resin had a green colour (bottom right) compared to normal extracts (bottom left). Estimated yields of 7-oxo-SD8 are indicated for each culture. Masses (*m/z*) and putative simocyclinone intermediates (A-G) are indicated. The origin of

the additional absorbance peak at 11 minutes of cultures grown with HP20 resin shows a different absorbance spectrum compared to early angucyclinone intermediates. The most abundant ion (*m/z* 503.1) however implies that it could be simocyclinone A2.

6.3 Spontaneous resistant mutants of *S. coelicolor* M1152

Spontaneous resistant mutants often arise around threshold levels of antimicrobial agents as a consequence of genomic mutations (Edwards *et al.*, 2009; Toprak *et al.*, 2012). Once a bacterial strain has acquired resistance, it is able to grow in the presence of the antimicrobial agent and can gradually increase its resistance threshold by acquiring additional mutations, which can result in a highly resistant strain.

Previous results have shown that the outer membrane-permeable *E. coli* strain NR698, which is sensitive to SD8, spontaneously develops resistance to SD8 when exposed to concentrations slightly above its MIC threshold. Sequencing revealed that this resistance is mediated by point mutations in the SD8 binding pocket of DNA gyrase, which was confirmed by assaying the mutant DNA gyrases in the presence of SD8 *in vitro*, and by structural studies (Edwards *et al.*, 2009; Hearnshaw *et al.*, 2014). Since these mutants were simple to isolate in *E. coli*, a similar approach was used to select for resistant mutants of *S. coelicolor* M1152. A confluent lawn of M1152 was grown for 6 days at 30 °C (~10^5 spores) in the presence of low amounts of SD8 applied to paper discs, which were placed directly on the plates (0/ 0.25/ 0.5/ 1/ 5/ 10 µg/mL). Resistant colonies, which grew within the zone of inhibition, were picked and cultured separately. The increased resistance of these mutants was confirmed against higher amounts of SD8 (30 µg/mL and 60 µg/mL) using paper disc assays on SMMS agar (**Figure 6.2**). All the SD8-resistant *S. coelicolor* M1152 colonies were analysed for mutations in the N-terminal domain of GyrA by PCR amplification and sequencing of the relevant part of the *gyrA* gene. However, none of the selected colonies contained mutations in *gyrA*, suggesting they arose from decreased

membrane permeability or increased efflux of SD8. However, since the SD8-resistant mutants could still be advantageous as an optimised heterologous system, the *sim* cluster (i.e. PAC-12I) was introduced twice into all 24 tested strains (**Table 6.1**). Although each mutant strain showed increased resistance against externally applied SD8, expression of the *sim* gene cluster still impaired the growth and sporulation of all the mutant strains to approximately the same degree as for the parent M1152 strain. Thus, the acquired mutations did not improve intracellular resistance to simocyclinone intermediates, consistent with the hypothesis that the mutant strains had decreased membrane permeability to SD8. Interestingly, several mutated strains with strongly increased resistance to SD8 also showed a delay in development or were affected in sporulation before introduction of the *sim* gene cluster (**Table 6.1**). However, there was no correlation between strain viability before and after introduction of the *sim* gene cluster.

In conclusion, *S. coelicolor* M1152 is able to develop spontaneous resistance to SD8 after exposure to low amounts of SD8, but the mutations responsible did not occur in *gyrA* as previously observed for SD8-resistant *E. coli* mutants. Further, the mutations did not confer resistance against intercellular simocyclinones, suggesting they cause changes in membrane permeability.

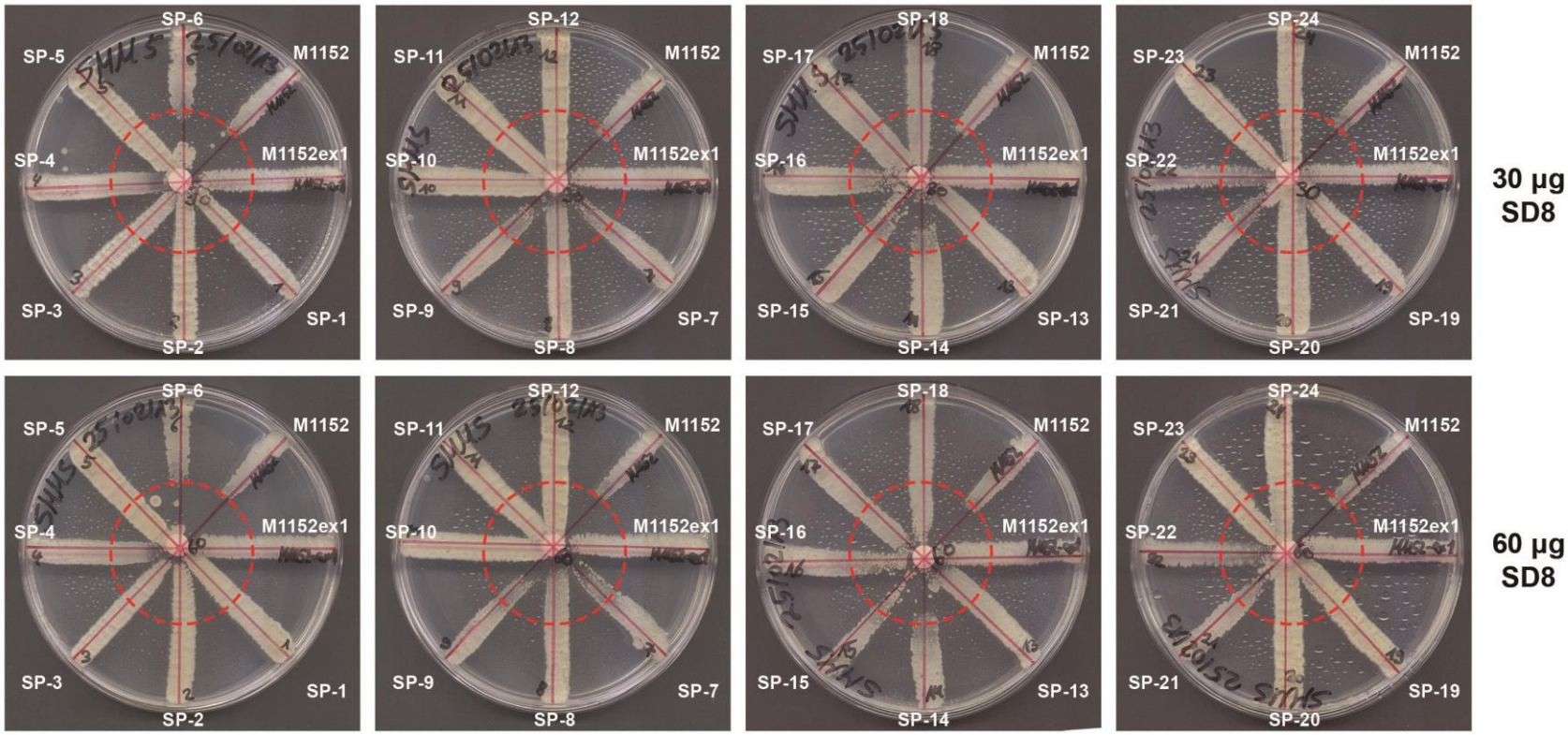


Figure 6.2

Spontaneous mutants of *S. coelicolor* M1152 with increased resistance to SD8.

Strains (SP-1 to SP-24) were grown on SMMS agar for 6 days at 30 °C in presence of 30 µg or 60 µg of SD8 on a paper disc before evaluation of the resistance levels compared to M1152 (red circles). Controls strains M1152 (sensitive) and M1152ex1 (resistant) were added on each plate.

Table 6.1

Spontaneous mutant strains of *S. coelicolor* M1152 with increased SD8 resistance.

Spore development and sporulation are indicated before introduction of the *sim* gene cluster. Sporulation was strongly impaired for all mutant strains (SP-1 to SP-24) after introduction of the *sim* gene cluster.

M1152 mutant	SD8 resistance	Spore development	Sporulation	Exconjugants (1^{st}/ 2^{nd} test)
SP-1	high	3 days	good	151 / 40
SP-2	high	3 days	low/ no spores	55 / 2
SP-3	high	3 days	low/ no spores	7 / 0
SP-4	medium	7 days	low/ no spores	n/a / 0
SP-5	high	3 days	good	121 / 30
SP-6	very low	3 days	good	50 / 0
SP-7	low	3 days	good	16 / 1
SP-8	medium	4 days	low/ no spores	208 / 11
SP-9	low	3 days	low/ no spores	7 / 0
SP-10	high	3 days	low/ no spores	62 / 1
SP-11	high	3 days	good	202 / 2
SP-12	high	4 days	low/ no spores	7 / 0
SP-13	medium	4 days	low/ no spores	>>300 / >>300
SP-14	very low	3 days	good	10 / 1
SP-15	low	3 days	good	60 / 9
SP-16	low	3 days	low/ no spores	31 / 1
SP-17	high	4 days	low/ no spores	38 / 2
SP-18	medium	---	no spores	n/a / n/a
SP-19	high	4 days	low/ no spores	12 / 60
SP-20	high	3 days	good	103 / 461
SP-21	medium	7 days	low/ no spores	n/a / 7
SP-22	medium	---	no spores	n/a / n/a
SP-23	high	3 days	good	8 / 310
SP-24	high	3 days	good	3 / 141

6.4 Alternative heterologous hosts

The strategy of using heterologous hosts has led to the discovery of multiple natural products and vastly facilitates the analysis and manipulation of biosynthetic pathways. The engineered heterologous host *S. coelicolor* M1152 is generally an excellent choice for this kind of work. Deletion of the endogenous antibiotic gene clusters in M1152 usually causes increased heterologous metabolite production because the endogenous antibiotics otherwise serve as major sinks for carbon. In addition, M1152 has no anti-microbial activity and has a much simplified extracellular metabolite profile for HPLC and LC-MS analyses (Gomez-Escribano and Bibb, 2011). Nevertheless, the increased production of introduced secondary metabolite gene clusters can have adverse effects on M1152 (J. P. Gomez-Escribano, personal communication). As a consequence, I also investigated heterologous hosts other than *S. coelicolor* M1152.

Commonly used heterologous hosts include *Streptomyces albus* (Sanchez *et al.*, 2002; Gullon *et al.*, 2006; Yan *et al.*, 2012), *Streptomyces lividans* (Motamedi and Hutchinson, 1987; Lacalle *et al.*, 1992; Fischer *et al.*, 2003), and *Streptomyces venezuelae* (Park et al., 2009). In particular, the angucyclinone and antitumor agent oviedomycin from *Streptomyces antibioticus* ATCC 11891™, which is structurally related to the angucyclinone moiety of SD8, has been characterised using the heterologous host *S. albus* (Mendez *et al.*, 2002). Three further structurally related polyketides, namely tetrangulol, tetrangomycin and the tricyclic fridamycin from *Streptomyces rimosus* NRRL 3016, have been analysed heterologously in *S. lividans* (Hong *et al.*, 1997).

Therefore, I attempted to introduce PAC-12I carrying the *sim* gene cluster into the potential heterologous hosts *Streptomyces venezuelae* ATCC 10712™, *Streptomyces albus* J1074, and *Streptomyces lividans* 1326. Conjugations were performed using the methylation-deficient *E. coli* strain ET12567 carrying the self-transmissible helper plasmid pR9406. No exconjugants were obtained for *S. venezuelae* and *S. albus*. Several potential exconjugants of *S. lividans* were later shown to be false positives

that did not contain the *sim* gene cluster. In fact, only *S. coelicolor* derived strains yielded exconjugants reliably.

In conclusion, none of the tested alternative strains was suitable for heterologous expression of simocyclinones. This might be due to phylogenetic differences of these strains or the fact that wildtype *S. coelicolor* produces more secondary metabolites, which could have facilitated survival for M1152 in comparison to the three tested alternative hosts.

6.5 Mutational analysis of the *sim* gene cluster in *S. antibioticus* Tü6040

Although the natural producer is difficult to manipulate, I tested if the desired mutations could be made in *S. antibioticus* Tü6040. This was feasible in principle given that prior to my arrival in the lab, Tung Le had succeeded in making a null mutation in the gene encoding the transcriptional regulator SimReg3, which is directly next to *simC7*. The *simReg3* gene was replaced with an apramycin resistance cassette using double homologous recombination with a *sim* cluster-containing cosmid (1K3-neo; T. Le., unpublished). Thus, the same approach was used to delete *simC7*, which is present on the same cosmid. The mutant construct was conjugated into *S. antibioticus* Tü6040 from *E. coli* ET12567/pR9406, however no exconjugants were obtained, despite varying numerous parameters individually (i.e. concentration of spores, different spore stocks, ratio between *E. coli* and *S. antibioticus*, time to dry conjugation plates, time of antibiotic overlay, amount of antibiotics used).

As an alternative approach, a 5.2-kb DNA fragment containing a central apramycin resistance cassette, which replaced *simC7* (*simC7::apr*), was cloned into the *Hin*dIII site of plasmid pGM1190-Kan-oriT (N. Holmes, unpublished). This plasmid is a derivative of pGM1190, a temperature-sensitive and non-integrative plasmid with intermediate copy number, which was successfully conjugated into *S. antibioticus* Tü6040. Unfortunately, exconjugant colonies did not grow at temperatures above 32 °C, which was

required for loss of the temperature-sensitive plasmid to select for recombinant colonies. Further growth tests with *S. antibioticus* Tü6040 at different temperatures revealed a narrow temperature range for growth around 25-32 °C, which might reflect the adaption of this strain to its natural environment in Argentina, South America (Schimana *et al.*, 2000).

In summary, mutational analysis of *S. antibioticus* Tü6040 was not successful, even after multiple attempts to delete *simC7* from the genome. These attempts confirm the genetic intractability of strain Tü6040.

6.6 Evidence that simocyclinone intermediates are the cause of the toxicity to *S. coelicolor* M1152, and that their target is not DNA gyrase

Several test conjugations were carried out during optimisation of the heterologous system. Among others, a mutant *sim* cluster was introduced into M1152 and M1152ex1 carrying an in-frame deletion in three genes (PAC-12IΔC1ABC) encoding a type I PKS predicted to be responsible for biosynthesis of the tetraene linker. Production of C- and D-group simocyclinones was completely abolished due to the absence of the PKS enzymes involved in tetraene linker biosynthesis. Simocyclinones of the A-group (i.e. aglycones) and B-group (i.e. angucyclinone with deoxysugar) were detectable by LC-MS analysis, although none of these intermediates accumulated in the culture supernatant. Unexpectedly, the absence of C- and D-group simocyclinones did relieve the toxicity to the heterologous host, which remained impaired in growth and sporulation, similar to M1152 expressing the wild-type *sim* cluster. *E. coli* DNA gyrase is only very weakly inhibited by simocyclinone intermediates lacking the aminocoumarin end of the molecule (Edwards *et al.*, 2009). These results provide evidence that one or more intermediates of the A- and/or B-group are highly toxic to *S. coelicolor,* and that their target is unlikely to be DNA gyrase.

Interestingly, the number of exconjugants obtained with PAC-12IΔC1ABC was consistently more than 100-fold higher than with PAC-12I carrying the

wildtype *sim* cluster. Given this is not caused by the relief of toxicity, it is possible that strongly improved conjugation efficiency is simply a consequence of the reduced size of this construct (~24 kb removed).

6.7 Genome sequencing reveals *S. antibioticus* Tü6040 to be a 'talented producer' of secondary metabolites

The genome of *S. antibioticus* Tü6040 was sequenced, in part to look for further potential simocyclinone resistance mechanisms, such as *gyrA* genes encoding simocyclinone-insensitive GyrA subunits. Therefore, genomic DNA was isolated and the quality and concentration was determined via pulsed-field gel electrophoresis (PFGE) and NanoDrop spectrophotometer, respectively. The sample was of high purity and good quality as reflected by a ratio of absorbance at 260 nm and 280 nm (260/280) of 1.85 (expected: 1.80) and pulse-field or normal gel electrophoresis (**Figure 6.3**). The sample was pre-diluted 10-fold (~200 ng/µL) before Pacific Biosciences (PacBio)-based sequencing using three SMRT cells. The sequencing statistics are summarised in **Table 6.2**.

One single assembled sequence covered most of the draft genome (contig 0, 9,731,730 bp, **Table 6.2**) whereas the remaining six assembled sequences were of comparably short lengths (2.5-180 kb) and either aligned internally to the longest sequence (contigs 1 and 5) or did not match a continuous alignment (contigs 3-6, and 9). This has been observed previously and was identified as systematic error for this sequencing method (Gomez-Escribano *et al.*, 2015). Therefore, the remaining short sequences were not considered for the draft genome. The open reading frames of the genome sequence (contig_0) were identified and annotated using Prodigal and RAST (Aziz *et al.*, 2008; Hyatt *et al.*, 2010). The draft genome contained a total of 8,872 predicted genes including 93 RNA-encoding genes (21 rRNAs and 72 tRNAs). It is likely that the ~10 Mb draft genome sequence covers most of the genome of *S. antibioticus* Tü6040, given a genome size range of 8.7-11.9 Mb for the *Streptomyces* genus (Zhou *et al.*, 2012). In order to identify

the ends of the genome, I analysed all sequences for terminal inverted repeats (TIRs) and helicase genes, which are typically found at the end of the linear genomes of this genus (Bey *et al.*, 2000). The absence of both indicated that the ends of the *S. antibioticus* Tü6040 genome were not sequenced.

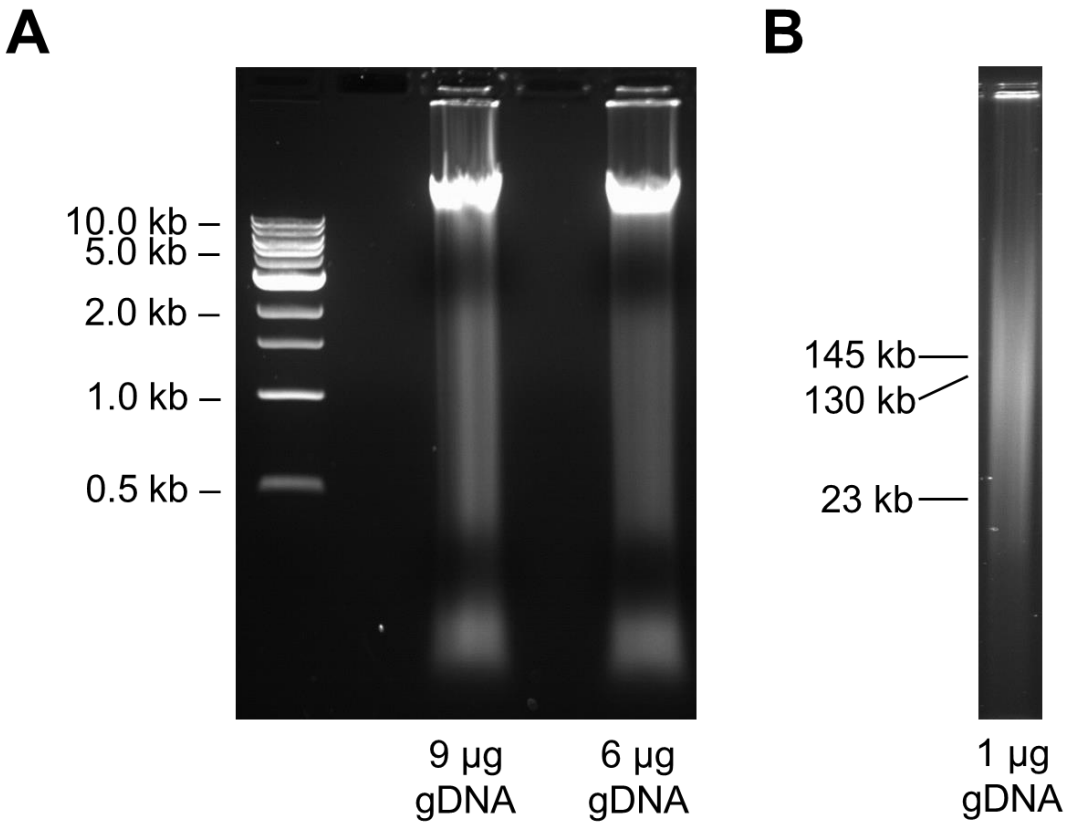


Figure 6.3

Purity and quality of isolated genomic DNA from *S. antibioticus* Tü6040. Genomic DNA (gDNA) was isolated as described and dissolved in TE-buffer overnight. DNA was separated by agarose gel electrophoresis (A) and pulse-field gel electrophoresis (B).

Table 6.2

Summary of sequencing results for genomic DNA from *S. antibioticus* Tü6040.

Metric	Value
Total sequences (raw data)	225,083
Sequence length (raw data)	500-28,119 bp
GC content (raw data)	68%
Polished contigs[a]	7
Contig_0	9,731,730 bp
Contig_1	25,696 bp
Contig_3	101,892 bp
Contig_4	23,779 bp
Contig_5	15,175 bp
Contig_6	180,596 bp
Contig_9	2,521 bp
Number of reads	188,099
Max contig length	9,731,730 bp
N50 contig length[b]	9,731,730 bp
Sum of contig lengths	10,081,389 bp

a) Automated assembly by the Earlham Institute(Norwich, UK).

b) Shortest sequence length at 50% of the genome.

Since SD8 targets DNA gyrase, my first aim was to identify and compare the GyrA and GyrB protein sequences from the natural simocyclinone producer *S. antibioticus* Tü6040 and the heterologous host *S. coelicolor*. I compiled the predicted proteome of *S. antibioticus* Tü6040 in a database and searched for homologs of GyrA (Sco3873) and GyrB (Sco3874) from *S. coelicolor* using a protein-based BLAST search (Altschul *et al.*, 1990; Camacho *et al.*, 2009). In total, four genes were identified, encoding the subunits of DNA gyrase (GyrA and GyrB) and the structurally related enzyme Topo IV (ParC and ParE). As expected, the genes encoding DNA gyrase were located in a single operon (GyrA encoded by *sant4394*: 4,915,060::4,917,654; GyrB encoded by *sant4395:* 4,917,697::4,919,766) and genes encoding for Topo IV were located in proximity to each other in the genome (ParC encoded by *sant2337*: 2,634,569::2,637,028 ; ParE encoded by *sant2357*: 2,655,425::2,657,548) (**Figure 6.4**).

Sequence alignments revealed high similarity between the two GyrA proteins (95% identical, 810/857; 98% positive, 837/857; 0% gaps 0/857) and between the two GyrB proteins (91% identical, 626/686; 95% positive, 654/686; 1% gaps, 5/686), with lower similarities between GyrA and ParC (43% identical, 339/790; 59% positive, 465/790; 6% gaps, 457/790) and between GyrB and ParE (47% identical, 331/710; 61% positive, 430/710; 8% gaps, 58/710). Importantly, there were no amino acid sequence differences in the aminocoumarin- and angucyclinone-binding pockets between GyrA from *S. antibioticus* and *S. coelicolor* (**Figure 6.5**). Furthermore, no additional obvious resistance genes such as membrane transport proteins could be identified in the entire draft genome or the remaining short contigs.

The draft genome was analysed for secondary metabolite gene clusters using antiSMASH v3.0 (Weber *et al.*, 2015), which identified 74 potential gene clusters (**Figure 6.4** and **Table 6.3**). This makes *S. antibioticus* Tü6040 a metabolically 'talented' producer dedicating ~23% of its genome for secondary metabolism, which is substantially more compared to other genomes. Less than 7% of prokaryotic species devote more than 7.5% of their genome to secondary metabolite biosynthesis including other natural product producing *Streptomyces* species (Cimermancic *et al.*, 2014). Despite

the fact that the ClusterFinder algorithm used by the antiSMASH software detects mostly putative saccharide gene clusters, which may not produce any metabolites with anti-microbial activity, the remaining 39 gene clusters constitute ~15% of the *S. antibioticus* genome, which is still much more compared to other sequenced strains. The *sim* gene cluster was located near the centre of the genome (predicted by antiSMASH: 4,518,198::4,434,712; proposed minimal *sim* cluster: 4,437,587::4,509,853). The sequence of the *sim* gene cluster of the genome draft showed no significant differences compared to the previously sequenced PAC-12I clone and the original published sequences (Galm *et al*., 2002; Trefzer *et al*., 2002). Several further gene clusters shared remarkable sequence similarities with genes of biosynthetic enzymes involved in simocyclinone or aminocoumarin biosynthesis. These similarities include biosynthetic genes for the simocyclinone tetraene linker (gene cluster 13; *simC1ABC*, *simC2*, *simA7*; 46-54% protein sequence identity), the simocyclinone angucyclic polyketide (gene cluster 38; *simA1-6*; 49-61% protein sequence identity), and the aminocoumarin deoxysugar noviose (gene clusters 5 and 24; nov/clo/couW-V-U-T-S-P; 51-63% and 50-77% protein sequence identity).

Remarkably, the longest protein encoded in the draft *S. antibioticus* genome has a length of 7,118 amino acids (751 kDa). This enormous protein is part of an unusual non-ribosomal peptide synthetase trans-acyltransferase polyketide synthase (NRPS-transAT-PKS) cluster (cluster 64). In addition, this large enzyme shares similarity with LnmJ (7,349 amino acids) from the biosynthetic gene cluster of the anticancer agent leinamycin (LNM) (Tang *et al.*, 2004). In fact, cluster 64 shares further sequence similarity with the leinamycin biosynthetic gene cluster, although it can reasonably be assumed that the two products are structurally distinct because there are also substantial differences between the two clusters (**Table 6.3**).

In summary, genome sequencing of the natural simocyclinone producer *S. antibioticus* Tü6040 revealed a large genome (~10 Mb) with more than 70 potential secondary metabolite gene clusters. No obvious further simocyclinone resistance genes lying outside the *sim* cluster could be identified bioinformatically. Intriguingly, the GyrA and GyrB subunits of

S. antibioticus Tü6040 are nearly identical to those of the simocyclinone-sensitive heterologous host *S. coelicolor*. It thus remains unclear why the *sim* cluster is so toxic to *S. coelicolor* but not to *S. antibioticus* Tü6040.

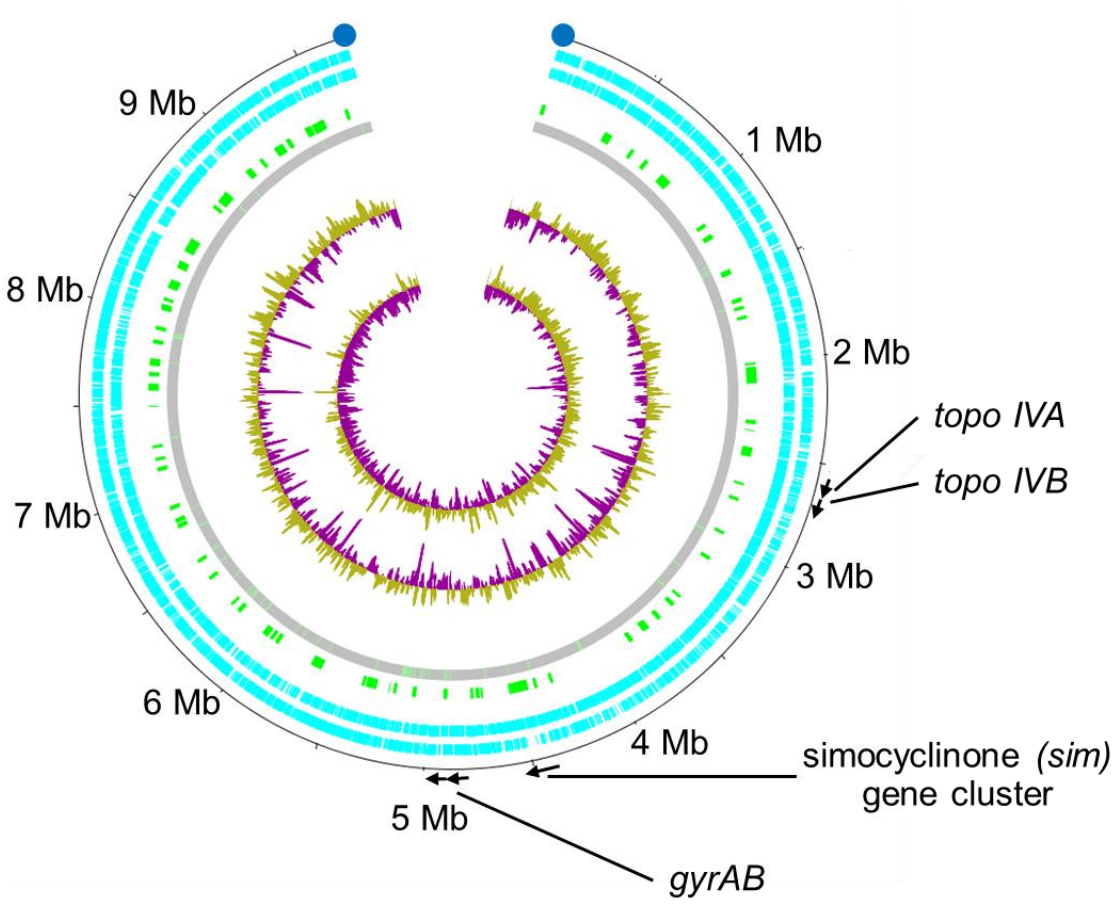


Figure 6.4

Circular representation of the *S. antibioticus* Tü6040 genome.

From outside to inside: open reading frames (cyan), secondary metabolite gene clusters predicted by antiSMASH v3 (green), GC-plot (above average in yellow; below average in purple), GC-skew (above average in yellow; below average in purple). The graph was created with DNAPlotter.

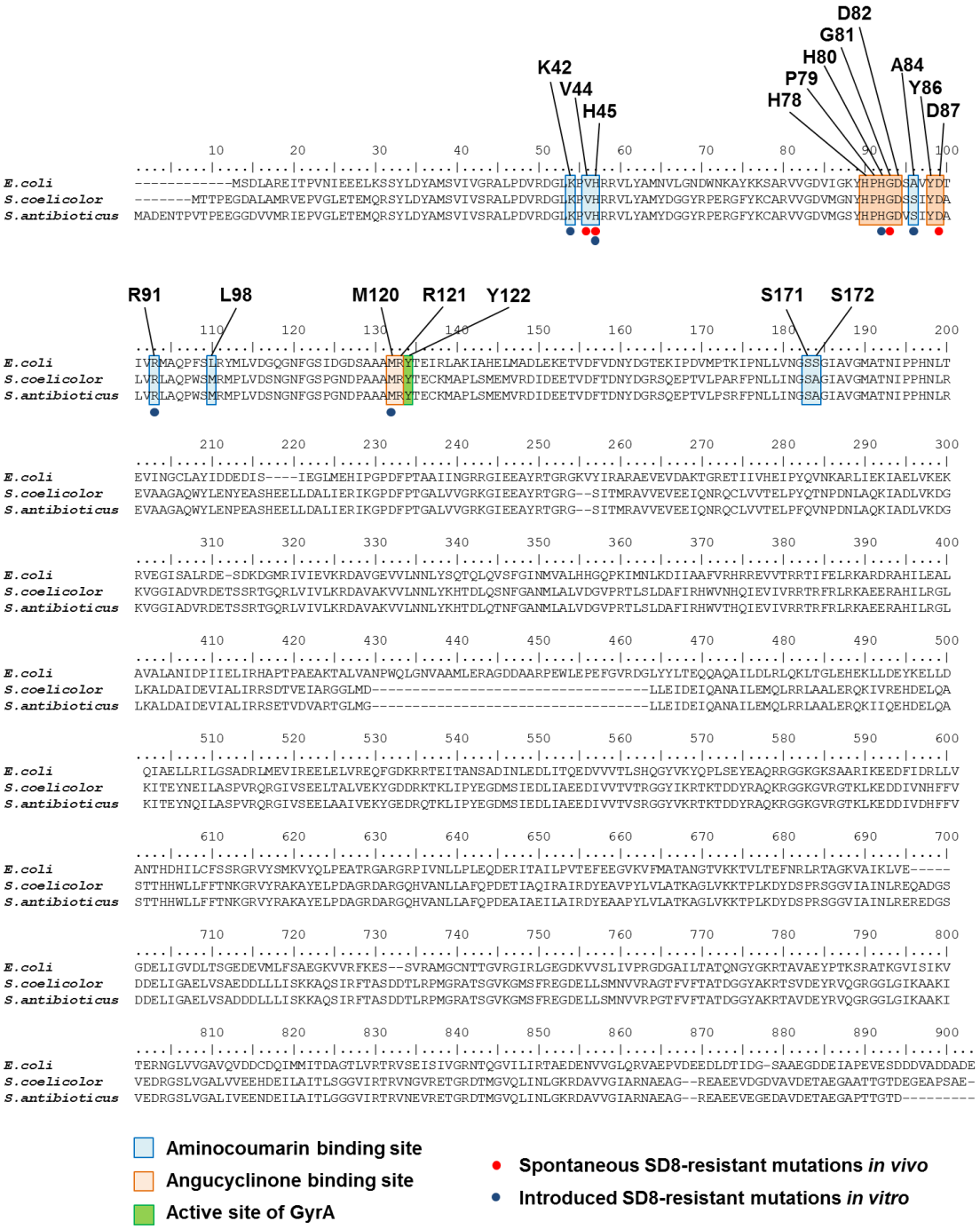


Figure 6.5

Comparison of the SD8 binding site in the DNA gyrase subunit A (GyrA).

The amino acid sequence alignment is shown for the highly similar GyrA proteins from *S. coelicolor* and *S. antibioticus* Tü6040. For comparison, the identified SD8 binding site and respective SD8-resistance mutations are highlighted for GyrA from *E. coli* (compare with **Figure 1.8**).

Table 6.3

Predicted biosynthetic gene clusters from *S. antibioticus* Tü6040.

Gene clusters were predicted by antiSMASH (Weber *et al.*, 2015). (Cf: cluster identified by ClusterFinder; yellow: simocyclinone biosynthetic gene cluster; orange: cluster containing homologs to genes involved in simocyclinone biosynthesis).

Cluster	Type (predicted)	From	To	Length	Most similar known cluster
1	Lantipeptide	8,953	31,736	22,783	Labyrinthopeptin A1-A3 gene cluster
2	T1pks	399,560	447,632	48,072	Laidlomycin biosynthetic gene cluster
3	Terpene	570,719	591,711	20,992	Kirromycin biosynthetic gene cluster
4	Cf_fatty_acid	673,476	694,729	21,253	-
5	Cf_saccharide-lantipeptide-bacteriocin	795,487	860,780	65,293	Informatipeptin biosynthetic gene cluster
6	Cf_saccharide	1,161,628	1,183,112	21,484	Tetronasin biosynthetic gene cluster
7	Cf_fatty_acid	1,231,232	1,258,058	26,826	Pristinamycin biosynthetic gene cluster
8	Terpene-cf_saccharide	1,454,017	1,484,381	30,364	Hopene biosynthetic gene cluster
9	Cf_saccharide	1,626,230	1,647,411	21,181	-
10	Cf_fatty_acid	1,672,687	1,693,673	20,986	-
11	Cf_saccharide	1,726,463	1,748,082	21,619	Kanamycin biosynthetic gene cluster
12	Siderophore	1,997,164	2,010,270	13,106	-
13	Cf_fatty_acid -t1pks	2,024,216	2,119,853	95,637	Rifamycin biosynthetic gene cluster
14	Terpene	2,326,886	2,349,042	22,156	-
15	Bacteriocin	2,383,692	2,395,041	11,349	-
16	T1pks-cf_fatty_acid	2,478,711	2,534,195	55,484	Tetronasin biosynthetic gene cluster
17	Siderophore	2,684,184	2,696,130	11,946	-
18	Microcin	2,772,636	2,792,730	20,094	-

Cluster	Type (predicted)	From	To	Length	Most similar known cluster
19	Cf_saccharide	2,977,683	2,998,603	20,920	-
20	Cf_saccharide	3,142,948	3,163,760	20,812	A-503083 biosynthetic gene cluster
21	Terpene	3,412,973	3,434,160	21,187	Albaflavenone biosynthetic gene cluster
22	Cf_saccharide	3,473,223	3,495,055	21,832	-
23	Microcin	3,581,388	3,601,482	20,094	Avilamycin A biosynthetic gene cluster
24	Cf_saccharide	3,667,891	3,708,105	40,214	Clorobiocin biosynthetic gene cluster
25	Cf_saccharide	3,779,772	3,806,537	26,765	-
26	Microcin	4,287,295	4,307,389	20,094	-
27	Cf_saccharide	4,380,643	4,401,545	20,902	Gilvocarcin biosynthetic gene cluster
28	Aminocoumarin-t2pks-t1pks-cf_fatty_acid	4,434,712	4,532,592	97,880	Simocyclinone biosynthetic gene cluster
29	Siderophore	4,534,375	4,547,729	13,354	-
30	Bacteriocin	4,693,312	4,705,588	12,276	-
31	Cf_saccharide	4,712,140	4,735,109	22,969	-
32	Cf_saccharide	4,746,848	4,767,549	20,701	-
33	Cf_saccharide	4,896,386	4,917,921	21,535	Granaticin biosynthetic gene cluster
34	Cf_saccharide	5,074,533	5,096,863	22,330	-
35	Microcin	5,187,631	5,207,725	20,094	-
36	Butyrolactone	5,226,150	5,237,091	10,941	Methylenomycin gene cluster
37	Cf_saccharide	5,298,875	5,320,278	21,403	-
38	T2pks	5,318,146	5,360,670	42,524	Spore pigment biosynthetic gene cluster
39	Cf_saccharide	5,369,628	5,390,650	21,022	-
40	T1pks-nrps-cf_fatty_acid	5,627,392	5,695,424	68,032	Elaiophylin biosynthetic gene cluster
41	Cf_saccharide	5,912,882	5,933,910	21,028	-

Cluster	Type (predicted)	From	To	Length	Most similar known cluster
42	Microcin	5,944,750	5,964,844	20,094	-
43	Cf_saccharide	5,980,001	6,021,221	41,220	-
44	Cf_saccharide	6,186,009	6,208,063	22,054	-
45	Siderophore	6,246,518	6,258,887	12,369	Desferrioxamine B gene cluster
46	Cf_saccharide	6,434,666	6,456,057	21,391	-
47	Cf_saccharide	6,547,762	6,568,976	21,214	-
48	Terpene	6,772,088	6,795,621	23,533	Kanamycin biosynthetic gene cluster
49	Cf_fatty_acid	6,805,740	6,828,449	22,709	Fredericamycin biosynthetic gene cluster
50	Cf_saccharide	6,884,506	6,905,951	21,445	Herboxidiene biosynthetic gene cluster
51	Cf_saccharide	7,114,485	7,135,663	21,178	-
52	Cf_saccharide	7,172,762	7,193,850	21,088	-
53	Cf_saccharide	7,245,235	7,266,494	21,259	Herboxidiene biosynthetic gene cluster
54	Ectoine	7,485,301	7,496,152	10,851	Ectoine biosynthetic gene cluster
55	Microcin-cf_saccharide	7,572,312	7,611,006	38,694	Amychelin biosynthetic gene cluster
56	Cf_saccharide	7,654,318	7,676,278	21,960	-
57	Other	7,714,241	7,758,101	43,860	Pristinamycin biosynthetic gene cluster
58	Cf_saccharide	7,840,751	7,861,701	20,950	Streptomycin biosynthetic gene cluster
59	Cf_saccharide	7,920,499	7,941,695	21,196	-
60	Cf_saccharide	8,036,666	8,070,375	33,709	-
61	Microcin	8,088,906	8,109,000	20,094	-
62	Hglks-t1pks	8,162,601	8,214,979	52,378	Cinnamycin biosynthetic gene cluster
63	T3pks	8,288,570	8,329,694	41,124	Herboxidiene biosynthetic gene cluster
64	Nrps-transatpks-cf_fatty_acid	8,392,372	8,474,411	82,039	Leinamycin biosynthetic gene cluster
65	Cf_saccharide	8,679,410	8,701,785	22,375	-

Cluster	Type (predicted)	From	To	Length	Most similar known cluster
66	Cf_saccharide-t3pks	8,724,532	8,780,918	56,386	-
67	Cf_fatty_acid	8,777,125	8,798,417	21,292	-
68	Nrps	8,950,075	8,993,329	43,254	Stenothricin biosynthetic gene cluster
69	Cf_saccharide	9,054,511	9,075,668	21,157	-
70	Nrps-t1pks	9,158,635	9,214,423	55,788	-
71	Terpene-melanin	9,237,694	9,264,599	26,905	Melanin biosynthetic gene cluster
72	Other	9,362,573	9,403,283	40,710	Kanamycin biosynthetic gene cluster
73	Nrps-lantipeptide	9,407,925	9,482,796	74,871	Herboxidiene biosynthetic gene cluster
74	Lantipeptide	9,608,484	9,631,393	22,909	-

6.8 Acknowledgements

I would like to thank the Earlham Institute (formerly The Genome Analysis Centre, TGAC) (Norwich, UK) for sequencing the genome of *S. antibioticus* Tü6040. I would like to thank Professor Barrie Wilkinson for providing material and heling me with culturing *Streptomyces* strains in presence of the adsorptive hydrophobic resin HP20.

Chapter 7
General discussion

7.1 The role of SimC7 during angucyclinone biosynthesis 220

7.2 Tetraene biosynthesis 226

7.3 Ketoreductases involved in simocyclinone biosynthesis 230

7.4 Natural resistance mechanism against simocyclinones 231

7.1 The role of SimC7 during angucyclinone biosynthesis

Since the simocyclinone biosynthetic pathway remains largely hypothetical, it is unknown in which order the angucyclinone is assembled and modified. Consequently, it is possible that reduction of the C-7 carbonyl group by SimC7 might not be the last step in angucyclinone biosynthesis. So, as a speculative example, it might be that in *S. antibioticus* epoxide formation occurs after the C-7 carbonyl group is reduced, but that the angucyclinone carrying the epoxide is still accepted as substrate by SimC7 *in vitro* (**Figure 7.1**). Acid-induced hydrolysis of the epoxide group in the angucyclinone was observed in early crystal structures after prolonged incubation of protein crystals under acidic conditions. Hydrolysis of the epoxide group was accompanied by increased flexibility of the A- and B-ring as indicated by a reduced electron density, whereas the position of the C- and D-ring remained unchanged (data not shown). Mass spectrometry analysis of SD8 and 7-oxo-SD8 incubated under equivalent acidic conditions revealed that the angucyclic polyketide skeleton remained intact but hydrolysis occurred of the epoxide group as well as the amide bond and the two carboxylic acid esters.

Thus, although acid-induced hydrolysis proceeded very slowly and almost certainly happened after crystal formation, it shows that the enzyme locks the C- and D-ring of the substrate in place for the reaction, whereas the remaining part of the substrate may adopt different conformations.

Biosynthesis of angucyclin(on)es is generally very complex and involves multiple tailoring enzymes that catalyse cyclisation, aromatisation, oxidation and reduction reactions. The predicted tailoring enzymes involved in biosynthesis of the angucyclinone in simocyclinone include two cyclases (SimA4 and SimA5) to generate the characteristic angucyclic ring system, three oxygenases (SimA7, SimA8, and SimA13) that introduce two hydroxyl groups and the epoxide, as well as four ketoreductases (SimC7, SimA6, SimA9, and SimA10) to reduce various carbonyl groups (Galm *et al.*, 2002; Trefzer *et al.*, 2002) (**Figure 7.1**).

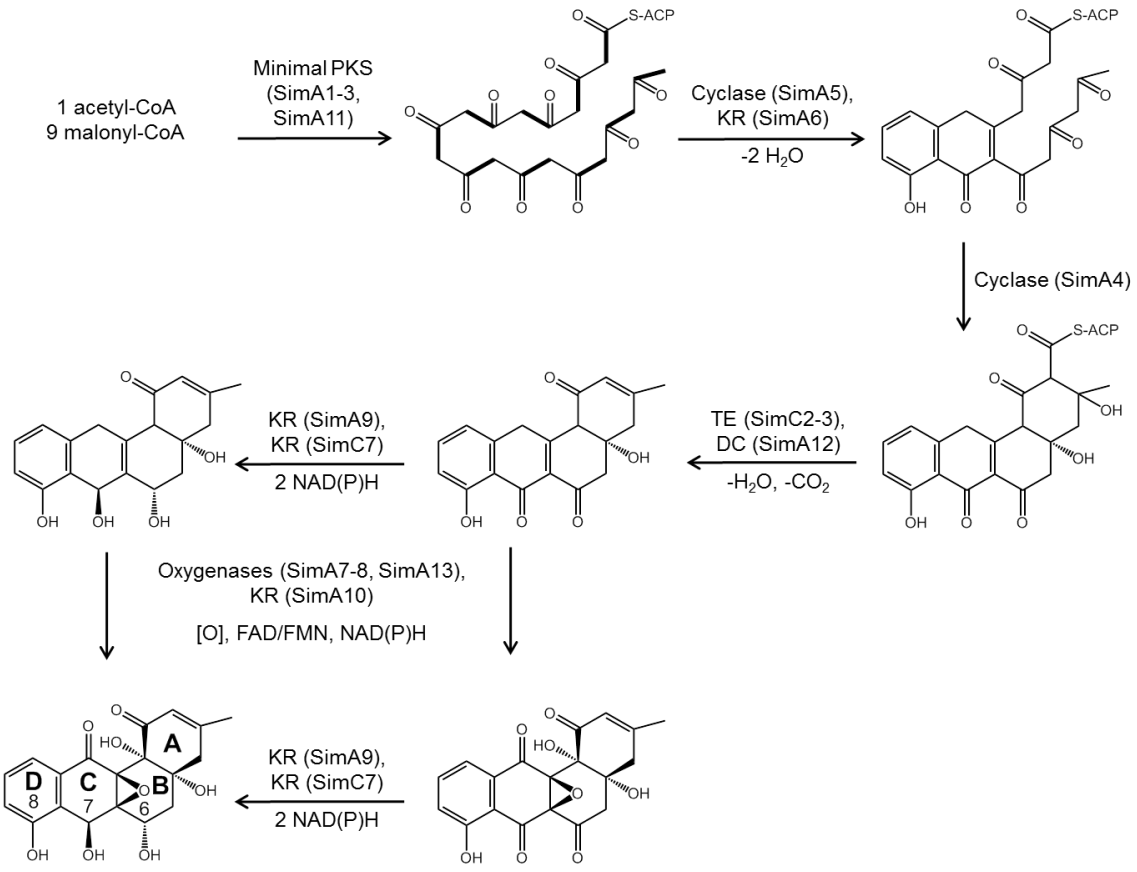


Figure 7.1

Putative biosynthetic pathway for the angucyclinone moiety of simocyclinone.

Biosynthesis of the angucyclinone begins with a polyketide skeleton, which is made by a type II PKS (KS, CLF, ACP) and subsequently cyclised to an angucyclic intermediate with four hexameric rings. The angucyclic intermediate is subsequently decarboxylated, oxygenated and reduced in multiple steps. The pathway is based on studies on close homologs and the results of this study. The order of oxygenation and ketoreduction is unknown and could be different than shown. The incorporated acetate units are highlighted in bold for the initial polyketide chain (compare with **Figure 4.8**). (KR, ketoreductase; TE, thioesterase; DC, dicarboxylase)

The D-ring of angucyclinones is cyclised first, followed by the C-ring, so that the core structure of the SimC7 substrate, an aromatic 8-phenyl adjacent to a 7-carbonyl, is generated very early in the pathway (Zhou *et al.*, 2010). The product of the minimal type II PKS, a linear poly-β-ketone, is extremely reactive and cyclises spontaneously. Consequently, the first regio-selective intramolecular aldol condensation has to follow immediately after completion of the polyketide chain as it determines the structure of the final molecule. The first cyclisation step and the corresponding cyclases have been well characterised and the most common patterns for aromatic polyketides are C7-C12 and C9-C14 cyclisations (Zhou *et al.*, 2010). In jadomycin biosynthesis for example, the cyclase JadD initiates cyclisation of the D-ring of the final angucyclinone by catalysing an aldol condensation between positions C-7 and C-12 of the polyketide chain. After cyclisation of the D-ring, JadD cyclises the C-ring and passes the intermediate on to a second cyclase named JadI, which subsequently cyclises the remaining two rings (Kulowski *et al.*, 1999). Homologs of JadD and JadI are present in the *sim* gene cluster (SimA5 and SimA4, respectively), suggesting a similar mechanism during simocyclinone biosynthesis (**Figure 7.1**).

Based on their structures (**Figure 7.2**), there are four other angucyclinones where a SimC7-like mechanism might generate a C-7 hydroxyl group – panglimycin A from *Streptomyces* sp. ICBB8230, its related stereoisomer elmycin A from *Streptomyces* sp. DSM4201 and DSM4202 (Fotso *et al.*, 2008), and the recently identified angucyclinones kiamycin and grisemycin from the marine isolate *Streptomyces griseus* sp. M268 (Xie *et al.*, 2012; Xie *et al.*, 2016) (**Figure 7.2**). However, the biosynthetic gene clusters for these four compounds have yet to be reported, and so it is not known whether a SimC7-like enzyme is involved in their synthesis.

Interestingly, most other angucyclin(on)es have a carbonyl group at position C-7. The urdamycins (Rohr *et al.*, 1989), the landomycins (Weber *et al.*, 1994) and the saquayamycins (Uchida *et al.*, 1985) are quinones, but intermediates in their biosynthetic pathways have a hydroxyl group at position C-7. However, these C-7 hydroxyl groups arise from cyclisation/aromatisation of the C-ring, not from ketoreduction (**Figure 7.2**).

The cyclases from several gene clusters have been characterised, including JadI (see above), TcmN, and PgaF (Kulowski *et al.*, 1999; Metsa-Ketela *et al.*, 2003; Thompson *et al.*, 2004). For example, the 7-hydroxyl group of prejadomycin (a shared intermediate in the biosynthesis of the jadomycins, landomycins and oviedomycin) originates from cyclisation of the C-ring and subsequent aromatisation via enol tautomerisation (Lombo *et al.*, 2009; Kharel and Rohr, 2012). This is also the case for the angucyclinone azicemicin A (Ogasawara and Liu, 2009), as well as the angucyclin(on)e-related tetracenomycins, chromomycin A_3 and mithramycins (Menendez *et al.*, 2004; Thompson *et al.*, 2004). Similarly, the 7-hydroxyl group in urdamycin C and D is a consequence of lactone formation at position C-12 and subsequent enol tautomerisation (Rohr *et al.*, 1988).

Among the vast range of known angucyclin(on)es are several compounds with structures similar to the known SimC7 substrate 7-oxo-SD8. These include angucyclinones such as oviedomycin, rabelomycin, gaudimycin, and chemomicin A as well as angucyclines such as landomycins, urdamycins, aquayamycin, and moromycins (**Figure 7.2**). In addition, various other natural product classes including flavonoids, anthraquinones and tetracyclines share a structurally-related carbon skeleton similar to angucyclinones with a carbonyl group adjacent to and in plane with a phenyl ring system (**Figure 7.3**). Depending on the substrate specificity of SimC7, these structurally related compounds could serve as potential SimC7 substrates, which would allow engineering of novel natural products. This could result in a range of novel angucyclin(on)e derivatives with a 7-hydroxyl group, which could be screened for improved biological activities. In case SimC7 accepts more diverse substrates such as for example the plant-derived flavonoids (**Figure 7.3**), these novel derivatives could also be screened for their biological activities. The substrate specificity of SimC7 could be tested by introducing the *simC7* complementation construct (pGM1190-pB7-NB-C7) into producer strains. Alternatively, substrate conversion could be analysed in a purified *in vitro* system as described in this study, where purified substrates are available.

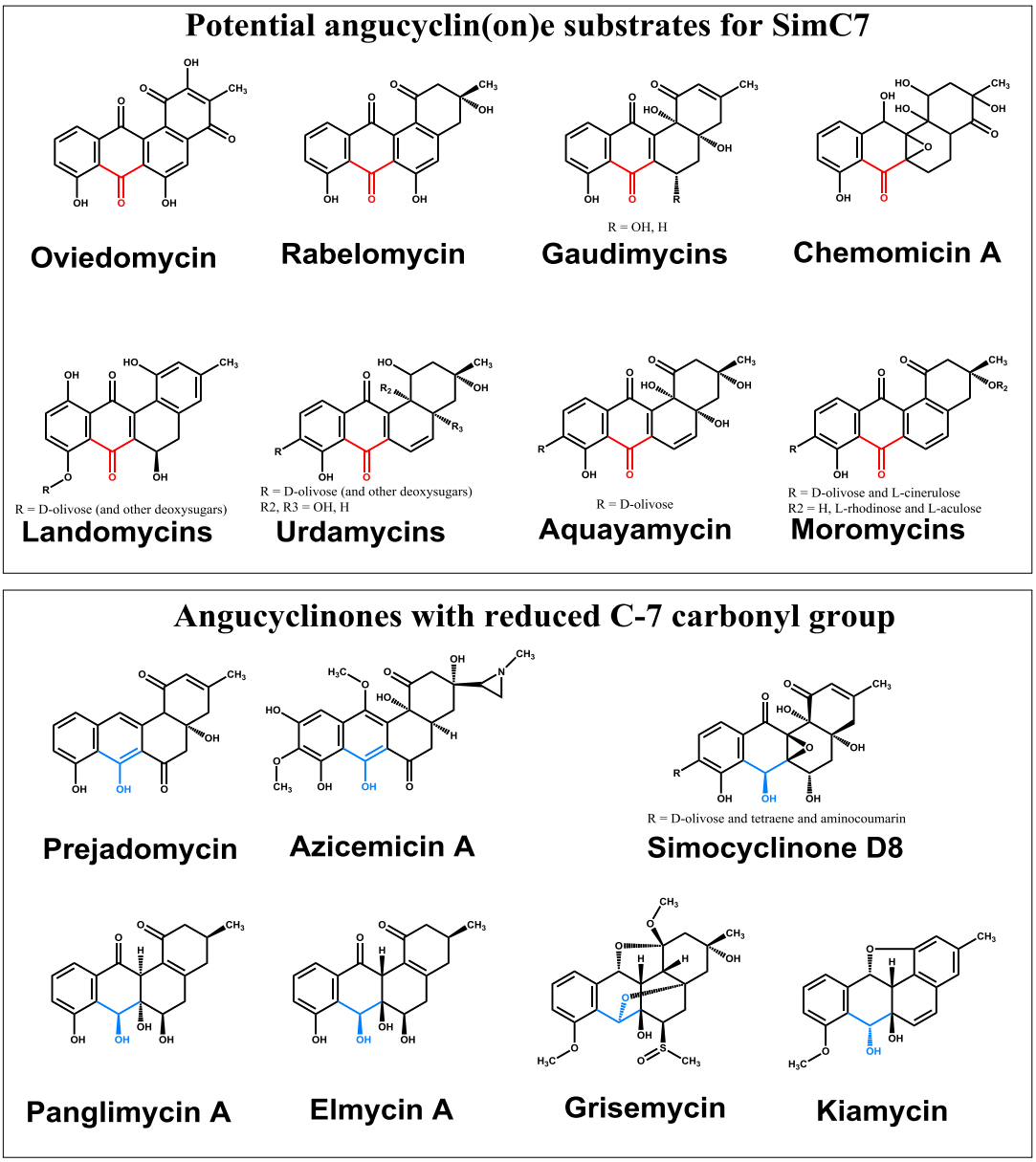


Figure 7.2

Potential angucyclin(on)e substrates for SimC7 and examples of angucyclin(on)es with 7-hydroxyl group.

Angucyclin(on)es with a 7-carbonyl group (red) are commonly found in bacteria and there are few examples of compounds with a 7-hydroxyl group (blue) as shown in the upper and lower panel, respectively. Reduction of the 7-carbonyl group arises generally from C-ring cyclisation/aromatisation (e.g. prejadomycin and azicemicin A), not from ketoreduction. Angucyclinones with a SimC7-like mechanism (i.e. a non-aromatic C-ring) include panglimycin A, elmycin A, grisemycin, and kiamycin.

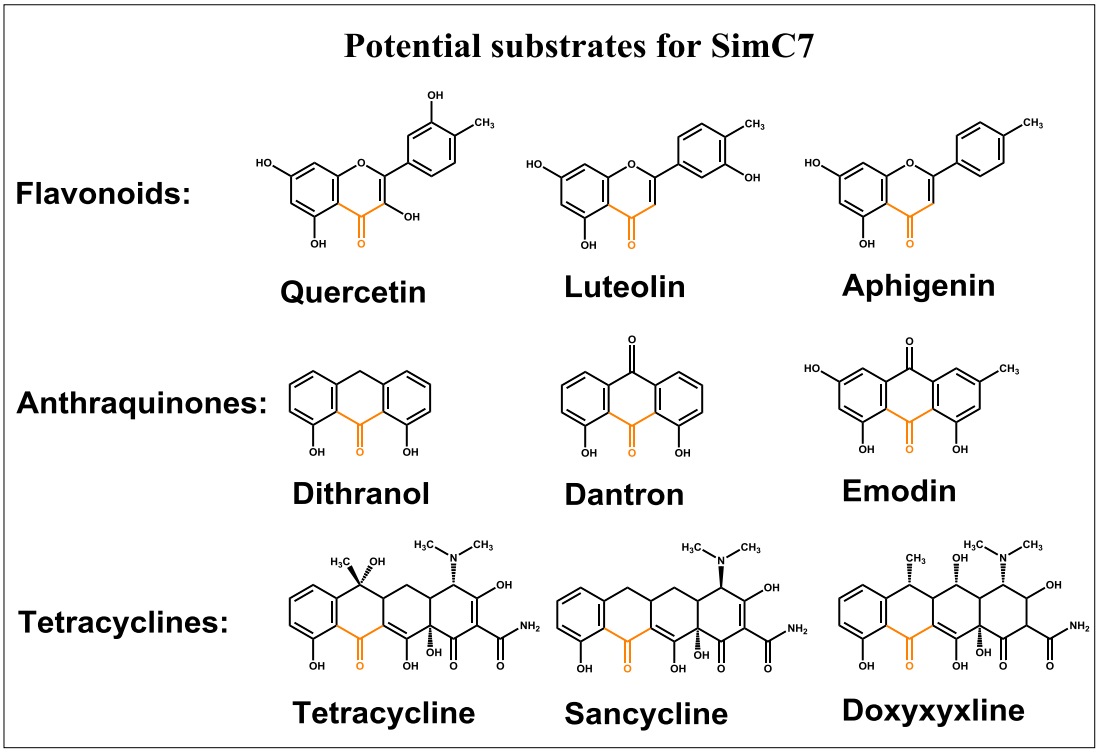


Figure 7.3

Further potential SimC7 substrates with related structural motifs.

Several flavonoids, anthraquinones and tetracyclines have an angucyclinone-related carbon skeleton with a carbonyl group (orange) adjacent to a phenyl ring system. Depending on the substrate-specificity of SimC7, respective compounds could be reduced by SimC7.

7.2 Tetraene biosynthesis

The fact that SimC7 is not involved in the biosynthesis of the tetraene linker of simocyclinone leaves unanswered the question as to how tetraene biosynthesis proceeds. One speculative possibility is that the type I PKS SimC1ABC contains at least one iterative module to produce a longer polyketide intermediate (**Figure 7.4**).

One piece of evidence consistent with this speculation is the predicted dioxygenase function of SimC4. Dioxygenases incorporate both atoms of molecular oxygen into the reaction product. For example, β-carotene dioxygenase cleaves a central double bond of β-carotene to generate two aldehydes (retinaldehyde, also known as all-trans-retinal) (Bugg, 2003). Consequently, the tetraene intermediate would be at least one acetyl-group longer than the final product. Cleavage of this intermediate would result in a tetraene linker with an aldehyde derived from molecular oxygen on one end and a carboxylic acid derived from acetate on the other end. The acetyl precursor units would remain complete after cleavage, which would result in a labelling pattern identical to the proposed terminally introduced aldehyde and thus cannot be distinguished experimentally in the ^{13}C-labelling data obtained in this and previous studies (Holzenkampfer *et al.*, 2002).

Another indication for a potential iterative PKS module involved in tetraene biosynthesis can be found in fumagillin, a meroterpenoid produced by the pathogenic fungus *Aspergillus fumigatus* (Lin *et al.*, 2013). Fumagillin contains a tetraene (i.e. decatetraenedioic acid), which is structurally identical to the tetraene produced during simocyclinone biosynthesis. In *A. fumigatus*, the tetraene is produced by a single iterative PKS module (fma-PKS) as a C_{12} polyketide product, which is released by an α,β-hydrolase (fma-AT) and subsequently modified into a dicarboxylic acid (Lin *et al.*, 2013). Strikingly, the iterative PKS from *A. fumigatus* has identical domain architecture to SimC1B, including the non-functional enoyl reductase domain. This further supports the prediction of a longer polyketide intermediate in tetraene linker biosynthesis in *S. antibioticus* Tü6040.

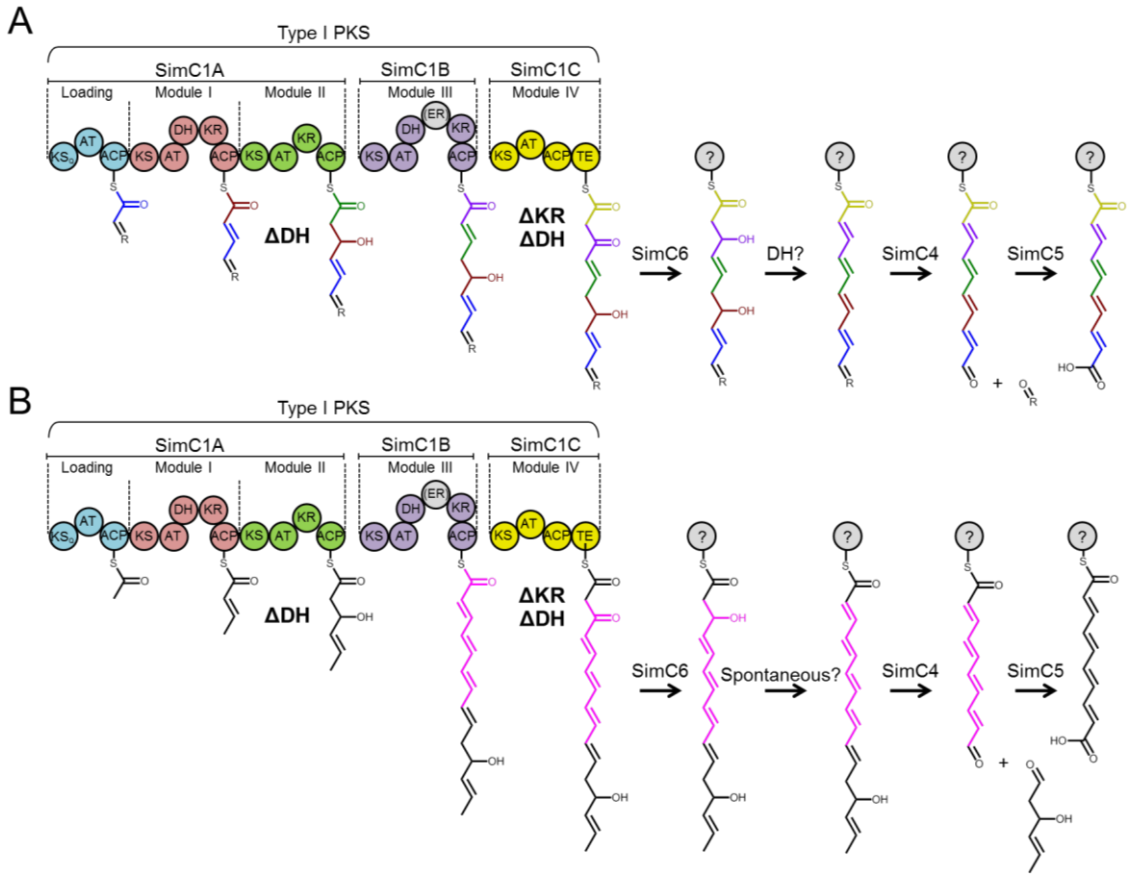


Figure 7.4

Potential biosynthetic pathway for the simocyclinone tetraene linker.

Tetraene linker biosynthesis in *S. antibioticus* Tü6040 involves a modular type I PKS (SimC1A-C) and several additional monofunctional enzymes. It remains unknown if the linker contains a longer starter unit (A) or if the PKS has an iterative module (B, highlighted in pink), which would result in a longer intermediate. This putative intermediate is then cleaved into two aldehydes by the dioxygenase SimC4. Figure adapted from Trefzer *et al.* (2002) (see **Figure 1.5**).

In order to investigate the biosynthesis of the simocyclinone tetraene linker (**Figure 1.5** and **Figure 7.4**) it would be necessary to analyse the assembly of the polyketide intermediate made by the type I PKS enzyme SimC1ABC. This can be done *in vivo* by chemical chain termination, a method specifically developed for determination of polyketide intermediates (Tosin *et al.*, 2010). Chemical probes such as the non-hydrolysable analogues malonyl (carba)dethia coenzyme A and *N*-acetylcysteamine thioesters (SNACs) compete with acyl carrier protein bound substrates (β-keto ACP thioesters) for chain elongation and cause premature release of truncated polyketide intermediates. Because these intermediates are not reloaded onto the polyketide, they accumulate and can be characterised by LC-MS analysis (Tosin *et al.*, 2010). Feeding cultures with synthetic chain terminators is a convenient method that is widely used in current mutasynthetic approaches. For example, this method led to the identification of intermediates from a non-iterative type I PKS involved in biosynthesis of erythromycin A (Tosin *et al.*, 2011), a partially reducing iterative type I PKS involved in biosynthesis of micacocidin (Kage *et al.*, 2015), and an iterative type I PKS involved in biosynthesis of 6-methylsalicylic acid (Parascandolo *et al.*, 2016).

There are two challenges to overcome to investigate tetraene linker biosynthesis via this approach. First, simocyclinone biosynthesis involves two PKS-based pathways and second, inactivation of the type II PKS for the angucyclinone completely abolishes simocyclinone production and most likely production of the tetraene intermediate (Trefzer *et al.*, 2002). These two problems could be overcome by comparing the polyketide intermediates generated by the wild-type *sim* cluster with those from a mutant lacking the biosynthetic genes for the tetraene linker (*ΔsimC1ABC*). In addition, biosynthesis of the simocyclinone tetraene could be further analysed by inactivating or removing domains or modules from the type I PKS enzyme SimC1ABC. Alternatively, the *simC1ABC* could be expressed separately in a heterologous system, which would simplify the metabolite profile and facilitate mutagenesis. Cloning and engineering of the large PKS genes (~22 kb) could be facilitated by using the recently developed strategy of

transformation-associated recombination (TAR) cloning (Yamanaka *et al.*, 2014).

Since I have shown that SimC7 is a ketoreductase involved in angucyclinone biosynthesis, the tetraene dehydratase (initially predicted to be SimC7) remains unidentified. One explanation could be that dehydration occurs spontaneously. Aldols (also named β-hydroxy ketones) can spontaneously convert into α,β-unsaturated carbonyls. The driving force for this spontaneous dehydration is the formation of a conjugated system with the adjacent carbonyl group. Dehydration proceeds even more rapidly under acidic conditions or at increased temperatures. Indeed, non-enzymatic dehydration of polyketide natural products has been observed previously (Kitao *et al.*, 1980; Fu *et al.*, 1994; Kulowski *et al.*, 1999) or was predicted to occur spontaneously based on the absence of dehydratase enzymes in the corresponding gene clusters (Nowak-Thompson *et al.*, 1997; Chiang *et al.*, 2011). If non-enzymatic dehydration occurs spontaneously during simocyclinone biosynthesis, it could be identified by the above mentioned chain termination approach, which should reveal the predicted tetraene linker intermediates.

A recent alternative hypothesis is that the tetraene linker is made by an iterative minimal PKS rather than a modular type I PKS (Bilyk *et al.*, 2016). Bilyk *et al.* (2016) identified *Kitasatospora* sp. (*smc* gene cluster) and *Streptomyces* sp. NRRL B-24484 (*sml* gene cluster) as new simocyclinone producers. Strikingly, however, they found there were no homologs of *simC1ABC* in the *smc* and *sml* biosynthetic gene clusters, or anywhere else on the chromosome. Therefore, they proposed that the tetraene linker in these two strains is made by a fatty acid related PKS enzyme, comprised of an acyl carrier protein (ACP, encoded by *smcP* in *Kitasatospora* sp.) and three 3-oxoacyl-ACP-synthases (I-III: *smcKSI*, *smcKSII*, and *smcX5* in *Kitasatospora* sp.) (Bilyk *et al.*, 2016). Homologs of these genes are also present at one end of the *sim* gene cluster in *S. antibioticus* Tü6040 (*orf1-3* and *simX5*, **Figure 1.6**) but it remains to be tested experimentally whether these fatty acid-related PKS genes are involved in tetraene linker biosynthesis in *S. antibioticus* Tü6040. Deletion of the *simC1ABC* genes in

S. antibioticus Tü6040 abolished production of simocyclinones that contained the tetraene linker (C- and D-group simocyclinones) but the deletion had no effect on earlier simocyclinone intermediates (A- and B-group simocyclinones), suggesting the type I PKS encoded by these genes is essential for tetraene linker biosynthesis (data not shown). Nevertheless, it would be interesting to test the involvement of *orf1-3* and *simX5* in tetraene biosynthesis in *S. antibioticus* Tü6040 by deleting them. The alternative model for tetraene linker formation in *Kitasatorspora* sp. would require dehydration of the entire polyketide chain and it was speculated that one or more dehydratases from fatty acid biosynthesis might be recruited for this purpose (Bilyk *et al.*, 2016).

7.3 Ketoreductases involved in simocyclinone biosynthesis

In addition to the novel angucyclinone ketoreductase SimC7 described here, there are three further SDR ketoreductases predicted to be involved in simocyclinone biosynthesis – SimA6, SimA9, and SimA10 (Galm *et al.*, 2002; Trefzer *et al.*, 2002). Close homologs of SimA6 (i.e. ActKR and HedKR) and SimA9 (i.e. LanV and $UrdM_{red}$) have been characterised that act on the C-9 and C-6 positions, respectively (see **chapter 5**). In contrast, the function of SimA10, an NAD(P)H-dependent flavin mononucleotide (FMN) ketoreductase, remains unknown. SimA10 has homologs in other angucycline gene clusters, for example UrdO and SaqO from urdamycin and saquayamycin biosynthesis, respectively (Faust *et al.*, 2000; Erb *et al.*, 2009), but their biosynthetic roles remain unknown. SimA10 could be involved in recycling the cofactor for another biosynthetic enzyme. For example, the related FMN reductase JadY from jadomycin biosynthesis is involved in recycling oxidised flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) back to their reduced forms ($FMNH_2$ and $FADH_2$). The reduced cofactor is required for JadG, an oxygenase that catalyses a unique B-ring opening C-C bond cleavage reaction needed for formation of an *N*-heterocycle (Fan *et al.*, 2012). Because SimA10 is related

to JadY (30% identity; 40% similarity) it is possible that SimA10 assists another enzyme during simocyclinone biosynthesis (**Figure 7.1**).

It would be interesting to investigate the potential involvement of SimA10 in simocyclinone biosynthesis since there is nothing known about its function. This could be done by mutagenesis of *simA10*, provided that the heterologous host carrying the mutagenized PAC is viable (see below).

7.4 Natural resistance mechanism against simocyclinones

One of the most important challenges that needs to be addressed is the toxicity of simocyclinones to heterologous hosts. Since the natural producer of simocyclinone is difficult to manipulate, an efficient heterologous system is crucial for successful studies on simocyclinone biosynthesis. This raises the obvious question: What is the resistance mechanism of the natural simocyclinone producer *S. antibioticus* Tü6040? So far, only constitutive overexpression of the simocyclinone efflux pump SimEx1 has been shown to confer significant levels of resistance to SD8 (Le *et al.*, 2009). Perhaps high-level transcription of *simEx1* requires additional transcriptional regulators that are not encoded on the PAC-12I clone, but are encoded elsewhere on the *S. antibioticus* genome. Regulatory crosstalk may even exist between the gene clusters for simocyclinone biosynthesis and other secondary metabolites, as observed for example in *Aspergillus nidulans*, where overexpression of a positive regulator from the silent *inp* gene cluster on chromosome II induces expression of a second gene cluster for asperfuranone biosynthesis on chromosome VIII (Bergmann *et al.*, 2010). Alternatively, there may well be an additional resistance mechanism in the natural simocyclinone producer that is absent from *S. coelicolor*. Evidence for this idea comes from the impaired development of the *S. coelicolor* heterologous host expressing a tetraene linker deletion construct (PAC-12IΔ*C1ABC*). The absence of the type I PKS genes (*simC1ABC*) makes this mutant unable to produce C- and D-group simocyclinones, yet the strain is still very sick. This seems to imply that the angucyclinone moiety of

simocyclinone inhibits a cellular target other than DNA gyrase in *S. coelicolor*.

The biological activities of angucyclines and angucyclinones are varied, covering cytotoxicity (including antitumor activity), enzyme inhibition, inhibition of platelet aggregation (thrombocytes), antibacterial activities, and antiviral activities (Rohr and Thiericke, 1992; Kharel *et al.*, 2012). Most interesting are the enzyme inhibitory activities of structurally related angucyclin(on)es (**Figure 7.2**). For example, aquayamycin is a non-competitive inhibitor of tyrosine hydroxylase (noradrenaline biosynthesis), dopamine β-hydroxylase (adrenaline biosynthesis), and tryptophan 5-monooxygenase. Vineomycin A_1 is a non-competitive inhibitor of prolyl hydroxylase, an essential enzyme for collagen biosynthesis (Rohr and Thiericke, 1992). Vineomycin C and saquayamycin A_1 weakly inhibit the inducible nitric oxide synthase (iNOS). Several glycol-modified urdamycins were found to have potent activity against xanthine oxidase at sub-micromolar concentrations. Gephyromycin, which contains an ether-bridged structure between C-3 and C-12a, is a strong glutaminergic agonist, causing influx of Ca^{2+} ions into neurons and the eventual release of glutamate and aspartate into the extracellular space (Kharel *et al.*, 2012). Furthermore, structural homologs of the angucyclinone moiety of simocyclinone have antitumor activity (landomycins, oviedomycin, saquayamycins), antibacterial activity (Gram positive: saquayamycins; Gram-positive and Gram-negative: urdamycins, rabelomycins, 8-O-methyl-rabelomycin), and weak antifungal activity (elmycins, emycin A, and sakyomicins) (Kharel *et al.*, 2012).

Additional factors that contribute to simocyclinone resistance could be identified by screening a genomic library for resistant colonies. This could be done using the genomic PAC library created by Bio S&T (Montreal, Canada) or by generating a new cosmid library, which has the advantage of screening much shorter DNA fragments than the ~110-kb inserts of the PAC-library. The screening could be carried out in *S. coelicolor* or *S. lividans* against purified SD8 or crude simocyclinone extracts. In addition, the screening could be repeated with (semi-)purified early simocyclinone intermediates to

test if toxicity originates from additional cellular targets of the angucyclinone moiety.

Several spontaneous *gyrA* mutations are known to confer simocyclinone resistance in *E. coli* (Edwards *et al.*, 2009). Equivalent mutations could be introduced into *gyrA* in *S. coelicolor*, and the high sequence similarity between DNA gyrase from *E. coli* and *S. coelicolor* implies that they should improve resistance to SD8. However, the suggestion that the angucyclinone moiety by itself is toxic to *S. coelicolor* and that its cellular target is not DNA gyrase means this approach by itself may not lead to a robust and vigorous heterologous expression system for simocyclinones.

References

Adams, P. D., P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallogr. Sect. D **66**(Pt 2): 213-221.

Albert, A. C., F. Spirito, N. Figueroa-Bossi, L. Bossi and A. R. Rahmouni (1996). "Hyper-negative template DNA supercoiling during transcription of the tetracycline-resistance gene in topA mutants is largely constrained in vivo." Nucleic Acids Res **24**(15): 3093-3099.

Alt, S., L. A. Mitchenall, A. Maxwell and L. Heide (2011). "Inhibition of DNA gyrase and DNA topoisomerase IV of Staphylococcus aureus and Escherichia coli by aminocoumarin antibiotics." J Antimicrob Chemother **66**(9): 2061-2069.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.

Anderle, C., S. Hennig, B. Kammerer, S. M. Li, L. Wessjohann, B. Gust and L. Heide (2007). "Improved mutasynthetic approaches for the production of modified aminocoumarin antibiotics." Chem Biol **14**(8): 955-967.

Aziz, R. K., D. Bartels, A. A. Best, M. DeJongh, T. Disz, R. A. Edwards, K. Formsma, S. Gerdes, E. M. Glass, M. Kubal, F. Meyer, G. J. Olsen, R. Olson, A. L. Osterman, R. A. Overbeek, L. K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G. D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke and O. Zagnitko (2008). "The RAST Server: rapid annotations using subsystems technology." BMC Genomics **9**: 75.

Bahadur, R. P., P. Chakrabarti, F. Rodier and J. Janin (2004). "A dissection of specific and non-specific protein-protein interfaces." J Mol Biol **336**(4): 943-955.

Bahassi, E. M., M. H. O'Dea, N. Allali, J. Messens, M. Gellert and M. Couturier (1999). "Interactions of CcdB with DNA gyrase - Inactivation of GyrA, poisoning of the gyrase-DNA complex, and

the antidote action of CcdA." Journal of Biological Chemistry **274**(16): 10936-10944.

Bates, A. D. and A. Maxwell (2007). "Energy coupling in type II topoisomerases: why do they hydrolyze ATP?" Biochemistry **46**(27): 7929-7941.

Bax, B. D., P. F. Chan, D. S. Eggleston, A. Fosberry, D. R. Gentry, F. Gorrec, I. Giordano, M. M. Hann, A. Hennessy, M. Hibbs, J. Z. Huang, E. Jones, J. Jones, K. K. Brown, C. J. Lewis, E. W. May, M. R. Saunders, O. Singh, C. E. Spitzfaden, C. Shen, A. Shillings, A. J. Theobald, A. Wohlkonig, N. D. Pearson and M. N. Gwynn (2010). "Type IIA topoisomerase inhibition by a new class of antibacterial agents." Nature **466**(7309): 935-U951.

Bennett, P. M., J. Grinsted and M. H. Richmond (1977). "Transposition of TnA does not generate deletions." Mol Gen Genet **154**(2): 205-211.

Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill and D. A. Hopwood (2002). "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)." Nature **417**(6885): 141-147.

Bergmann, S., A. N. Funk, K. Scherlach, V. Schroeckh, E. Shelest, U. Horn, C. Hertweck and A. A. Brakhage (2010). "Activation of a Silent Fungal Polyketide Biosynthesis Pathway through Regulatory Cross Talk with a Cryptic Nonribosomal Peptide Synthetase Gene Cluster." Applied and Environmental Microbiology **76**(24): 8143-8149.

Berman, H., K. Henrick and H. Nakamura (2003). "Announcing the worldwide Protein Data Bank." Nat Struct Biol **10**(12): 980.

Berman, H. M., G. J. Kleywegt, H. Nakamura and J. L. Markley (2014). "The Protein Data Bank archive as an open data resource." J Comput Aided Mol Des **28**(10): 1009-1014.

Bey, S. J., M. F. Tsou, C. H. Huang, C. C. Yang and C. W. Chen (2000). "The homologous terminal sequence of the Streptomyces lividans chromosome and SLP2 plasmid." Microbiology-Uk **146**: 911-922.

Bierman, M., R. Logan, K. O'Brien, E. T. Seno, R. N. Rao and B. E. Schoner (1992). "Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp." Gene **116**(1): 43-49.

Bilyk, O., E. Brotz, B. Tokovenko, A. Bechthold, T. Paululat and A. Luzhetskyy (2016). "New Simocyclinones: Surprising Evolutionary and Biosynthetic Insights." Acs Chemical Biology **11**(1): 241-250.

Bisang, C., P. F. Long, J. Cortes, J. Westcott, J. Crosby, A. L. Matharu, R. J. Cox, T. J. Simpson, J. Staunton and P. F. Leadlay (1999). "A chain initiation factor common to both modular and aromatic polyketide synthases." Nature **401**(6752): 502-505.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem **72**: 248-254.

Bugg, T. D. H. (2003). "Dioxygenase enzymes: catalytic mechanisms and chemical models." Tetrahedron **59**(36): 7075-7101.

Bunton, C. A., G. W. Kenner, M. J. Robinson and B. R. Webster (1963). "Experiments Related to Biosynthesis of Novobiocin and Other Coumarins." Tetrahedron **19**(6): 1001-&.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer and T. L. Madden (2009). "BLAST+: architecture and applications." BMC Bioinformatics **10**: 421.

Carver, T., N. Thomson, A. Bleasby, M. Berriman and J. Parkhill (2009). "DNAPlotter: circular and linear interactive genome visualization." Bioinformatics **25**(1): 119-120.

Chain, E., H. W. Florey, M. B. Adelaide, A. D. Gardner, N. G. Heatley, M. A. Jennings, J. Orr-Ewing, A. G. Sanders and L. F. Peltier (2005). "The Classic - Penicillin as a chemotherapeutic agent (Reprinted from Lancet, vol 24, pg 226-231, 1940)." Clinical Orthopaedics and Related Research(439): 23-26.

Chambers, H. F. and F. R. Deleo (2009). "Waves of resistance: Staphylococcus aureus in the antibiotic era." Nat Rev Microbiol **7**(9): 629-641.

Champoux, J. J. (2001). "DNA topoisomerases: Structure, function, and mechanism." Annual Review of Biochemistry **70**: 369-413.

Chater, K. F. and L. C. Wilde (1976). "Restriction of a Bacteriophage of Streptomyces-Albus G Involving Endonuclease Sali." Journal of Bacteriology **128**(2): 644-650.

Chen, H. and C. T. Walsh (2001). "Coumarin formation in novobiocin biosynthesis: beta-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI." Chem Biol **8**(4): 301-312.

Chen, V. B., W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson (2010). "MolProbity: all-atom structure validation for macromolecular crystallography." Acta Crystallogr D Biol Crystallogr **66**(Pt 1): 12-21.

Cherepanov, P. P. and W. Wackernagel (1995). "Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant." Gene **158**(1): 9-14.

Chiang, Y. M., K. M. Meyer, M. Praseuth, S. E. Baker, K. S. Bruno and C. C. C. Wang (2011). "Characterization of a polyketide synthase in Aspergillus niger whose product is a precursor for both dihydroxynaphthalene (DHN) melanin and naphtho-gamma-pyrone." Fungal Genetics and Biology **48**(4): 430-437.

Chiriac, A. I., F. Kloss, J. Kramer, C. Vuong, C. Hertweck and H. G. Sahl (2015). "Mode of action of closthioamide: the first member of the polythioamide class of bacterial DNA gyrase inhibitors." Journal of Antimicrobial Chemotherapy **70**(9): 2576-2588.

Cimermancic, P., M. H. Medema, J. Claesen, K. Kurita, L. C. W. Brown, K. Mavrommatis, A. Pati, P. A. Godfrey, M. Koehrsen, J. Clardy, B. W. Birren, E. Takano, A. Sali, R. G. Linington and M. A. Fischbach (2014). "Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters." Cell **158**(2): 412-421.

Collin, F., S. Karkare and A. Maxwell (2011). "Exploiting bacterial DNA gyrase as a drug target: current state and perspectives." Appl Microbiol Biotechnol **92**(3): 479-497.

Costantino, N. and D. L. Court (2003). "Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants." Proc Natl Acad Sci U S A **100**(26): 15748-15753.

Cowtan, K. (2006). "The Buccaneer software for automated model building. 1. Tracing protein chains." Acta Crystallogr. Sect. D **62**(Pt 9): 1002-1011.

Cowtan, K. (2010). "Recent developments in classical density modification." Acta Crystallogr. Sect. D **66**(Pt 4): 470-478.

Dall'Acqua, W. and P. Carter (2000). "Substrate-assisted catalysis: Molecular basis and biological significance." Protein Science **9**(1): 1-9.

Datsenko, K. A. and B. L. Wanner (2000). "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products." Proc Natl Acad Sci U S A **97**(12): 6640-6645.

Davis, I. W., A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, 3rd, J. Snoeyink, J. S. Richardson and D. C. Richardson (2007). "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids." Nucleic Acids Res. **35**(Web Server issue): W375-383.

DiMasi, J. A., R. W. Hansen and H. G. Grabowski (2003). "The price of innovation: new estimates of drug development costs." J Health Econ **22**(2): 151-185.

Doublie, S. (1997). "Preparation of selenomethionyl proteins for phase determination." Methods Enzymol **276**: 523-530.

Edwards, M. J., R. H. Flatman, L. A. Mitchenall, C. E. Stevenson, T. B. Le, T. A. Clarke, A. R. McKay, H. P. Fiedler, M. J. Buttner, D. M. Lawson and A. Maxwell (2009). "A crystal structure of the bifunctional antibiotic simocyclinone D8, bound to DNA gyrase." Science **326**(5958): 1415-1418.

Edwards, M. J., M. A. Williams, A. Maxwell and A. R. McKay (2011). "Mass spectrometry reveals that the antibiotic simocyclinone D8 binds to DNA gyrase in a "bent-over" conformation: evidence of positive cooperativity in binding." Biochemistry **50**(17): 3432-3440.

Ehmann, D. E. and S. D. Lahiri (2014). "Novel compounds targeting bacterial DNA topoisomerase/DNA gyrase." Current Opinion in Pharmacology **18**: 76-83.

Emmerson, A. M. and A. M. Jones (2003). "The quinolones: decades of development and use." Journal of Antimicrobial Chemotherapy **51**: 13-20.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr. D Biol. Crystallogr. **60**(Pt 12): 2126-2132.

Emsley, P., B. Lohkamp, W. G. Scott and K. Cowtan (2010). "Features and development of Coot." Acta Crystallogr D Biol Crystallogr **66**(Pt 4): 486-501.

Erb, A., A. Luzhetskyy, U. Hardter and A. Bechthold (2009). "Cloning and Sequencing of the Biosynthetic Gene Cluster for Saquayamycin Z and Galtamycin B and the Elucidation of the Assembly of Their Saccharide Chains." Chembiochem **10**(8): 1392-1401.

Eustaquio, A. S., B. Gust, U. Galm, S. M. Li, K. F. Chater and L. Heide (2005). "Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters." Appl Environ Microbiol **71**(5): 2452-2459.

Evans, P. R. and G. N. Murshudov (2013). "How good are my data and what is the resolution?" Acta Crystallogr D Biol Crystallogr **69**(Pt 7): 1204-1214.

Fan, K. Q., G. H. Pan, X. J. Peng, J. T. Zheng, W. B. Gao, J. Wang, W. S. Wang, Y. Li and K. Q. Yang (2012). "Identification of JadG as the B Ring Opening Oxygenase Catalyzing the Oxidative C-C Bond Cleavage Reaction in Jadomycin Biosynthesis." Chemistry & Biology **19**(11): 1381-1390.

Faust, B., D. Hoffmeister, G. Weitnauer, L. Westrich, S. Haag, P. Schneider, H. Decker, E. Kunzel, J. Rohr and A. Bechthold (2000). "Two new tailoring enzymes, a glycosyltransferase and an oxygenase, involved in biosynthesis of the angucycline antibiotic urdamycin A in Streptomyces fradiae Tu2717." Microbiology **146 (Pt 1)**: 147-154.

Finn, R. D., J. Mistry, J. Tate, P. Coggill, A. Heger, J. E. Pollington, O. L. Gavin, P. Gunasekaran, G. Ceric, K. Forslund, L. Holm, E. L. Sonnhammer, S. R. Eddy and A. Bateman (2010). "The Pfam protein families database." Nucleic Acids Res **38**(Database issue): D211-222.

Fischer, C., F. Lipata and J. Rohr (2003). "The complete gene cluster of the antitumor agent gilvocarcin V and its implication for the biosynthesis of the gilvocarcins." Journal of the American Chemical Society **125**(26): 7818-7819.

Flatman, R. H., A. J. Howells, L. Heide, H. P. Fiedler and A. Maxwell (2005). "Simocyclinone D8, an inhibitor of DNA gyrase with a novel mode of action." Antimicrob Agents Chemother **49**(3): 1093-1100.

Fleming, A. (1929). "On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ." British journal of experimental pathology **10**(3): 226-236.

Foss, M. H., K. A. Hurley, N. A. Sorto, L. L. Lackner, K. M. Thornton, J. T. Shaw and D. B. Weibel (2011). "N-Benzyl-3-sulfonamidopyrrolidines Are a New Class of Bacterial DNA Gyrase Inhibitors." Acs Medicinal Chemistry Letters **2**(4): 289-292.

Fotso, S., T. Mahmud, T. M. Zabriskie, D. A. Santosa, Sulastri and P. J. Proteau (2008). "Angucyclinones from an Indonesian Streptomyces sp." Journal of Natural Products **71**(1): 61-65.

Fu, H., R. Mcdaniel, D. A. Hopwood and C. Khosla (1994). "Engineered Biosynthesis of Novel Polyketides - Stereochemical Course of 2 Reactions Catalyzed by a Polyketide Synthase." Biochemistry **33**(31): 9321-9326.

367 Galm, U. (2003). Molekularbiologische und biochemische Untersuchungen zur Biosynthese der Aminocoumarinantibiotika Simocyclinon D8, Novobiocin, Clorobiocin und Coumermycin A1. Tübingen, Faculty for chemistry and pharmacy, Eberhard-Karls-University**:** 1-175.

Galm, U., M. A. Dessoy, J. Schmidt, L. A. Wessjohann and L. Heide (2004a). "In vitro and in vivo production of new aminocoumarins by a combined biochemical, genetic, and synthetic approach." Chem Biol **11**(2): 173-183.

Galm, U., S. Heller, S. Shapiro, M. Page, S. M. Li and L. Heide (2004b). "Antimicrobial and DNA gyrase-inhibitory activities of novel clorobiocin derivatives produced by mutasynthesis." Antimicrob Agents Chemother **48**(4): 1307-1312.

Galm, U., J. Schimana, H. P. Fiedler, J. Schmidt, S. M. Li and L. Heide (2002). "Cloning and analysis of the simocyclinone biosynthetic gene cluster of Streptomyces antibioticus Tu 6040." Arch Microbiol **178**(2): 102-114.

Gellert, M., K. Mizuuchi, M. H. Odea and H. A. Nash (1976). "DNA Gyrase - Enzyme That Introduces Superhelical Turns into DNA."

Proceedings of the National Academy of Sciences of the United States of America **73**(11): 3872-3876.

Goetschi, E., P. Angehrn, H. Gmuender, P. Hebeisen, H. Link, R. Masciadri and J. Nielsen (1993). "Cyclothialidine and Its Congeners - a New Class of DNA Gyrase Inhibitors." Pharmacology & Therapeutics **60**(2): 367-380.

Gomez-Escribano, J. P. and M. J. Bibb (2011). "Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters." Microb Biotechnol **4**(2): 207-215.

Gomez-Escribano, J. P., J. F. Castro, V. Razmilic, G. Chandra, B. Andrews, J. A. Asenjo and M. J. Bibb (2015). "The Streptomyces leeuwenhoekii genome: de novo sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2." BMC Genomics **16**: 485.

Gonzalez-Menendez, V., F. Asensio, C. Moreno, N. de Pedro, M. C. Monteiro, M. de la Cruz, F. Vicente, G. F. Bills, F. Reyes, O. Genilloud and J. R. Tormo (2014). "Assessing the effects of adsorptive polymeric resin additions on fungal secondary metabolite chemical diversity." Mycology **5**(3): 179-191.

Grant, S. G., J. Jessee, F. R. Bloom and D. Hanahan (1990). "Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants." Proc Natl Acad Sci U S A **87**(12): 4645-4649.

Gullon, S., C. Olano, M. S. Abdelfattah, A. F. Brana, J. Rohr, C. Mendez and J. A. Salas (2006). "Isolation, characterization, and heterologous expression of the biosynthesis gene cluster for the antitumor anthracycline steffimycin." Applied and Environmental Microbiology **72**(6): 4172-4183.

Gust, B., G. L. Challis, K. Fowler, T. Kieser and K. F. Chater (2003). "PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin." Proc Natl Acad Sci U S A **100**(4): 1541-1546.

Gust, B., G. Chandra, D. Jakimowicz, T. Yuqing, C. J. Bruton and K. F. Chater (2004). "Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces." Adv Appl Microbiol **54**: 107-128.

132 Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41:** 95-98.

Hanahan, D. (1983). "Studies on transformation of Escherichia coli with plasmids." J Mol Biol **166**(4): 557-580.

Hearnshaw, S. J., M. J. Edwards, C. E. Stevenson, D. M. Lawson and A. Maxwell (2014). "A New Crystal Structure of the Bifunctional Antibiotic Simocyclinone D8 Bound to DNA Gyrase Gives Fresh Insight into the Mechanism of Inhibition." J Mol Biol **426**(10): 2023-2033.

Heide, L. (2009). "The aminocoumarins: biosynthesis and biology." Nat Prod Rep **26**(10): 1241-1250.

Heide, L. (2013). "New aminocoumarin antibiotics as gyrase inhibitors." Int J Med Microbiol **304**(1): 31-36.

Holdgate, G. A., A. Tunnicliffe, W. H. Ward, S. A. Weston, G. Rosenbrock, P. T. Barth, I. W. Taylor, R. A. Pauptit and D. Timms (1997). "The entropic penalty of ordered water accounts for weaker binding of the antibiotic novobiocin to a resistant mutant of DNA gyrase: a thermodynamic and crystallographic study." Biochemistry **36**(32): 9663-9673.

Holm, L. and C. Sander (1995). "DALI: a network tool for protein structure comparison." Trends. Biochem. Sci. **20**(11): 478-480.

Holzenkampfer, M., M. Walker, A. Zeeck, J. Schimana and H. P. Fiedler (2002). "Simocyclinones, novel cytostatic angucyclinone antibiotics produced by Streptomyces antibioticus Tu 6040 II. Structure elucidation and biosynthesis." J Antibiot (Tokyo) **55**(3): 301-307.

Holzenkampfer, M. and A. Zeeck (2002). "Biosynthesis of simocyclinone D8 in an 18O2-rich atmosphere." J Antibiot (Tokyo) **55**(3): 341-342.

Hong, H. J., M. I. Hutchings, L. M. Hill and M. J. Buttner (2005). "The role of the novel Fem protein VanK in vancomycin resistance in Streptomyces coelicolor." J Biol Chem **280**(13): 13055-13061.

Hong, S. T., J. R. Carney and S. J. Gould (1997). "Cloning and heterologous expression of the entire gene clusters for PD 116740 from Streptomyces strain WP 4669 and tetrangulol and tetrangomycin from Streptomyces rimosus NRRL 3016." Journal of Bacteriology **179**(2): 470-476.

Hopwood, D. A., T. Kieser, H. M. Wright and M. J. Bibb (1983). "Plasmids, Recombination and Chromosome Mapping in Streptomyces-Lividans-66." Journal of General Microbiology **129**(Jul): 2257-2269.

Hopwood, D. A. and H. M. Wright (1978). "Bacterial protoplast fusion: recombination in fused protoplasts of Streptomyces coelicolor." Mol Gen Genet **162**(3): 307-317.

Horbal, L., Y. Rebets, M. Rabyk, R. Makitrynskyy, A. Luzhetskyy, V. Fedorenko and A. Bechthold (2012). "SimReg1 is a master switch for biosynthesis and export of simocyclinone D8 and its precursors." AMB Express **2**(1): 1.

Hyatt, D., G. L. Chen, P. F. Locascio, M. L. Land, F. W. Larimer and L. J. Hauser (2010). "Prodigal: prokaryotic gene recognition and translation initiation site identification." BMC Bioinformatics **11**: 119.

Javidpour, P., J. Bruegger, S. Srithahan, T. P. Korman, M. P. Crump, J. Crosby, M. D. Burkart and S. C. Tsai (2013). "The Determinants of Activity and Specificity in Actinorhodin Type II Polyketide Ketoreductase." Chemistry & Biology **20**(10): 1225-1234.

Javidpour, P., A. Das, C. Khosla and S. C. Tsai (2011a). "Structural and biochemical studies of the hedamycin type II polyketide ketoreductase (HedKR): molecular basis of stereo- and regiospecificities." Biochemistry **50**(34): 7426-7439.

Javidpour, P., T. P. Korman, G. Shakya and S. C. Tsai (2011b). "Structural and biochemical analyses of regio- and stereospecificities observed in a type II polyketide ketoreductase." Biochemistry **50**(21): 4638-4649.

Jones, A. C., B. Gust, A. Kulik, L. Heide, M. J. Buttner and M. J. Bibb (2013). "Phage p1-derived artificial chromosomes facilitate heterologous expression of the FK506 gene cluster." PLoS One **8**(7): e69319.

Kabsch, W. (1976). "A solution for the best rotation to relate two sets of vectors." Acta Crystallogr. Sect. D **32**: 922-923.

Kabsch, W. (2010). "XDS." Acta Crystallogr. D Biol. Crystallogr. **66**(Pt 2): 125-132.

Kage, H., E. Riva, J. S. Parascandolo, M. F. Kreutzer, M. Tosin and M. Nett (2015). "Chemical chain termination resolves the timing of ketoreduction in a partially reducing iterative type I polyketide

synthase." Organic & Biomolecular Chemistry **13**(47): 11414-11417.

Kallberg, Y., U. Oppermann and B. Persson (2010). "Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models." Febs Journal **277**(10): 2375-2386.

Kavanagh, K., H. Jornvall, B. Persson and U. Oppermann (2008). "The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes." Cellular and Molecular Life Sciences **65**(24): 3895-3906.

Kelley, L. A., S. Mezulis, C. M. Yates, M. N. Wass and M. J. Sternberg (2015). "The Phyre2 web portal for protein modeling, prediction and analysis." Nat Protoc **10**(6): 845-858.

Kelley, L. A. and M. J. E. Sternberg (2009). Protein structure prediction on the web: a case study using the Phyre server, Nature Protocols. **4:** 363-371.

Kessler, A., W. Dittrich, M. Betzler and H. Schrempf (1989). "Cloning and analysis of a deletable tetracycline-resistance determinant of Streptomyces lividans 1326." Mol Microbiol **3**(8): 1103-1109.

Kharel, M. K., P. Pahari, M. D. Shepherd, N. Tibrewal, S. E. Nybo, K. A. Shaaban and J. Rohr (2012). "Angucyclines: Biosynthesis, mode-of-action, new natural products, and synthesis." Nat Prod Rep **29**(2): 264-325.

Kharel, M. K. and J. Rohr (2012). "Delineation of gilvocarcin, jadomycin, and landomycin pathways through combinatorial biosynthetic enzymology." Curr Opin Chem Biol **16**(1-2): 150-161.

Kieser, T., M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood (2000). Practical Streptomyces genetics.

Kim, C. G., J. Lamichhane, K. I. Song, V. D. Nguyen, D. H. Kim, T. S. Jeong, S. H. Kang, K. W. Kim, J. Maharjan, Y. S. Hong, J. S. Kang, J. C. Yoo, J. J. Lee, T. J. Oh, K. Liou and J. K. Sohng (2008a). "Biosynthesis of rubradirin as an ansamycin antibiotic from Streptomyces achromogenes var. rubradiris NRRL3061." Arch Microbiol **189**(5): 463-473.

Kim, I. K., H. S. Yim, M. K. Kim, D. W. Kim, Y. M. Kim, S. S. Cha and S. O. Kangl (2008b). "Crystal structure of a new type of NADPH-dependent quinone oxidoreductase (QOR2) from Escherichia coli." Journal of Molecular Biology **379**(2): 372-384.

Kim, M. H., Y. Kim, H. J. Park, J. S. Lee, S. N. Kwak, W. H. Jung, S. G. Lee, D. Kim, Y. C. Lee and T. K. Oh (2008c). "Structural Insight into Bioremediation of Triphenylmethane Dyes by Citrobacter sp Triphenylmethane Reductase." Journal of Biological Chemistry **283**(46): 31981-31990.

Kitao, C., H. Tanaka, S. Minami and S. Omura (1980). "Bioconversion and biosynthesis of nanaomycins using cerulenin, a specific inhibitor of fatty acid and polyketide biosyntheses." J Antibiot (Tokyo) **33**(7): 711-716.

Komatsu, M., K. Komatsu, H. Koiwai, Y. Yamada, I. Kozone, M. Izumikawa, J. Hashimoto, M. Takagi, S. Omura, K. Shin-ya, D. E. Cane and H. Ikeda (2013). "Engineered Streptomyces avermitilis host for heterologous expression of biosynthetic gene cluster for secondary metabolites." ACS Synth Biol **2**(7): 384-396.

Komatsu, M., T. Uchiyama, S. Omura, D. E. Cane and H. Ikeda (2010). "Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism." Proc Natl Acad Sci U S A **107**(6): 2646-2651.

Korman, T. P., J. A. Hill, T. N. Vu and S. C. Tsai (2004). "Structural analysis of actinorhodin polyketide ketoreductase: Cofactor binding and substrate specificity." Biochemistry **43**(46): 14529-14538.

Korman, T. P., Y. H. Tan, J. Wong, R. Luo and S. C. Tsai (2008). "Inhibition kinetics and emodin cocrystal structure of a type II polyketide ketoreductase." Biochemistry **47**(7): 1837-1847.

Krissinel, E. (2010). "Crystal contacts as nature's docking solutions." J Comput Chem **31**(1): 133-143.

Krissinel, E. (2015). "Stock-based detection of protein oligomeric states in jsPISA." Nucleic Acids Research **43**(W1): W314-W319.

Krissinel, E. and K. Henrick (2007). "Inference of macromolecular assemblies from crystalline state." J Mol Biol **372**(3): 774-797.

Krumsiek, J., R. Arnold and T. Rattei (2007). "Gepard: a rapid and sensitive tool for creating dotplots on genome scale." Bioinformatics **23**(8): 1026-1028.

Kulowski, K., E. Wendt-Pienkowski, L. Han, K. Q. Yang, L. C. Vining and C. R. Hutchinson (1999). "Functional characterization of the jadI gene as a cyclase forming angucyclinones." Journal of the American Chemical Society **121**(9): 1786-1794.

Lacalle, R. A., J. A. Tercero and A. Jimenez (1992). "Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts." EMBO J **11**(2): 785-792.

Laponogov, I., M. K. Sohi, D. A. Veselkov, X. S. Pan, R. Sawhney, A. W. Thompson, K. E. McAuley, L. M. Fisher and M. R. Sanderson (2009). "Structural insight into the quinolone-DNA cleavage complex of type IIA topoisomerases." Nature Structural & Molecular Biology **16**(6): 667-669.

Lawrence, M. C. and P. M. Colman (1993). "Shape complementarity at protein/protein interfaces." J. Mol. Biol. **234**(4): 946-950.

Le, T. B., H. P. Fiedler, C. D. den Hengst, S. K. Ahn, A. Maxwell and M. J. Buttner (2009). "Coupling of the biosynthesis and export of the DNA gyrase inhibitor simocyclinone in Streptomyces antibioticus." Mol Microbiol **72**(6): 1462-1474.

Le, T. B., M. A. Schumacher, D. M. Lawson, R. G. Brennan and M. J. Buttner (2011a). "The crystal structure of the TetR family transcriptional repressor SimR bound to DNA and the role of a flexible N-terminal extension in minor groove binding." Nucleic Acids Res **39**(21): 9433-9447.

Le, T. B., C. E. Stevenson, H. P. Fiedler, A. Maxwell, D. M. Lawson and M. J. Buttner (2011b). "Structures of the TetR-like simocyclinone efflux pump repressor, SimR, and the mechanism of ligand-mediated derepression." J Mol Biol **408**(1): 40-56.

Letunic, I., T. Doerks and P. Bork (2012). "SMART 7: recent updates to the protein domain annotation resource." Nucleic Acids Res **40**(Database issue): D302-305.

Lin, H. C., Y. H. Chooi, S. Dhingra, W. Xu, A. M. Calvo and Y. Tang (2013). "The Fumagillin Biosynthetic Gene Cluster in Aspergillus fumigatus Encodes a Cryptic Terpene Cyclase Involved in the Formation of beta-trans-Bergamotene." J Am Chem Soc.

Liu, L. F., C. C. Liu and B. M. Alberts (1980). "Type II DNA topoisomerases: enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break." Cell **19**(3): 697-707.

Lombo, F., M. S. Abdelfattah, A. F. Brana, J. A. Salas, J. Rohr and C. Mendez (2009). "Elucidation of oxygenation steps during oviedomycin biosynthesis and generation of derivatives with increased antitumor activity." Chembiochem **10**(2): 296-303.

Lu, J., S. Patel, N. Sharma, S. M. Soisson, R. Kishii, M. Takei, Y. Fulkuda, K. J. Lumb and S. B. Singh (2014). "Structures of Kibdelomycin Bound to Staphylococcus aureus GyrB and ParE Showed a Novel U-Shaped Binding Mode." Acs Chemical Biology **9**(9): 2023-2031.

Luft, T., S. M. Li, H. Scheible, B. Kammerer and L. Heide (2005). "Overexpression, purification and characterization of SimL, an amide synthetase involved in simocyclinone biosynthesis." Arch Microbiol **183**(4): 277-285.

MacNeil, D. J., K. M. Gewain, C. L. Ruby, G. Dezeny, P. H. Gibbons and T. MacNeil (1992). "Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector." Gene **111**(1): 61-68.

Marshall, O. J. (2004). "PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR." Bioinformatics **20**(15): 2471-2472.

Maxwell, A. (1997). "DNA gyrase as a drug target." Trends Microbiol **5**(3): 102-109.

McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read (2007). "Phaser crystallographic software." J. Appl. Crystallogr. **40**(Pt 4): 658-674.

Medema, M. H., K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano and R. Breitling (2011). "antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences." Nucleic Acids Res **39**(Web Server issue): W339-346.

Mendez, C., E. Kunzel, F. Lipata, F. Lombo, W. Cotham, M. Walla, D. W. Bearden, A. F. Brana, J. A. Salas and J. Rohr (2002). "Oviedomycin, an unusual angucyclinone encoded by genes of the oleandomycin-producer Streptomyces antibioticus ATCC11891." J Nat Prod **65**(5): 779-782.

Menendez, N., M. Nur-e-Alam, A. F. Brana, J. Rohr, J. A. Salas and C. Mendez (2004). "Biosynthesis of the antitumor chromomycin A3 in Streptomyces griseus: analysis of the gene cluster and rational design of novel chromomycin analogs." Chem Biol **11**(1): 21-32.

Metsa-Ketela, M., K. Palmu, T. Kunnari, K. Ylihonko and P. Mantsala (2003). "Engineering anthracycline biosynthesis toward angucyclines." Antimicrobial Agents and Chemotherapy **47**(4): 1291-1296.

Motamedi, H. and C. R. Hutchinson (1987). "Cloning and heterologous expression of a gene cluster for the biosynthesis of tetracenomycin C, the anthracycline antitumor antibiotic of Streptomyces glaucescens." Proc Natl Acad Sci U S A **84**(13): 4445-4449.

Muniyappa, K. and C. M. Radding (1986). "The Homologous Recombination System of Phage-Lambda - Pairing Activities of Beta-Protein." Journal of Biological Chemistry **261**(16): 7472-7478.

Murshudov, G. N., A. A. Vagin and E. J. Dodson (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallogr. D Biol. Crystallogr. **53**(Pt3): 240-255.

Muth, G., B. Nussbaumer, W. Wohlleben and A. Puhler (1989). "A Vector System with Temperature-Sensitive Replication for Gene Disruption and Mutational Cloning in Streptomycetes." Molecular & General Genetics **219**(3): 341-348.

Nielsen, D. R., G. S. Amarasiriwardena and K. L. J. Prather (2010). "Predicting the adsorption of second generation biofuels by polymeric resins with applications for in situ product recovery (ISPR)." Bioresource Technology **101**(8): 2762-2769.

Nollmann, M., N. J. Crisona and P. B. Arimondo (2007). "Thirty years of Escherichia coli DNA gyrase: from in vivo function to single-molecule mechanism." Biochimie **89**(4): 490-499.

Nowak-Thompson, B., S. J. Gould and J. E. Loper (1997). "Identification and sequence analysis of the genes encoding a polyketide synthase required for pyoluteorin biosynthesis in Pseudomonas fluorescens Pf-5." Gene **204**(1-2): 17-24.

Ogasawara, Y. and H. W. Liu (2009). "Biosynthetic Studies of Aziridine Formation in Azicemicins." Journal of the American Chemical Society **131**(50): 18066-+.

Okanishi, M., K. Suzuki and H. Umezawa (1974). "Formation and reversion of Streptomycete protoplasts: cultural condition and morphological study." J Gen Microbiol **80**(2): 389-400.

Oppegard, L. M., B. L. Hamann, K. R. Streck, K. C. Ellis, H. P. Fiedler, A. B. Khodursky and H. Hiasa (2009). "In vivo and in vitro patterns of the activity of simocyclinone D8, an angucyclinone antibiotic from Streptomyces antibioticus." Antimicrob Agents Chemother **53**(5): 2110-2119.

Oppermann, U., C. Filling, M. Hult, N. Shafqat, X. Wu, M. Lindh, J. Shafqat, E. Nordling, Y. Kallberg, B. Persson and H. Jornvall (2003). "Short-chain dehydrogenases/reductases (SDR): the 2002 update." Chem Biol Interact **143-144**: 247-253.

Paananen, P., P. Patrikainen, P. Kallio, P. Mantsala, J. Niemi, L. Niiranen and M. Metsa-Ketela (2013). "Structural and functional analysis of angucycline C-6 ketoreductase LanV involved in landomycin biosynthesis." Biochemistry **52**(31): 5304-5314.

Pacholec, M., C. L. Freel Meyers, M. Oberthur, D. Kahne and C. T. Walsh (2005). "Characterization of the aminocoumarin ligase SimL from the simocyclinone pathway and tandem incubation with NovM,P,N from the novobiocin pathway." Biochemistry **44**(12): 4949-4956.

Paget, M. S., L. Chamberlin, A. Atrih, S. J. Foster and M. J. Buttner (1999). "Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in Streptomyces coelicolor A3(2)." J Bacteriol **181**(1): 204-211.

Painter, J. and E. A. Merritt (2006). "Optimal description of a protein structure in terms of multiple groups undergoing TLS motion." Acta Crystallogr D Biol Crystallogr **62**(Pt 4): 439-450.

Palumbi, S. R. (2001). "Humans as the world's greatest evolutionary force." Science **293**(5536): 1786-1790.

Pan, H. Y., R. Zhou, G. V. Louie, J. K. Muhlemann, E. K. Bomati, M. E. Bowman, N. Dudareva, R. A. Dixon, J. P. Noel and X. Q. Wang (2014). "Structural Studies of Cinnamoyl-CoA Reductase and Cinnamyl-Alcohol Dehydrogenase, Key Enzymes of Monolignol Biosynthesis." Plant Cell **26**(9): 3709-3727.

Parascandolo, J. S., J. Havemann, H. K. Potter, F. L. Huang, E. Riva, J. Connolly, I. Wilkening, L. J. Song, P. F. Leadlay and M. Tosin (2016). "Insights into 6-Methylsalicylic Acid Bio-assembly by Using Chemical Probes." Angewandte Chemie-International Edition **55**(10): 3463-3467.

Park, S. R., J. A. Yoon, J. H. Paik, J. W. Park, W. S. Jung, Y. H. Ban, E. J. Kim, Y. J. Yoo, A. R. Han and Y. J. Yoon (2009). "Engineering of plant-specific phenylpropanoids biosynthesis in Streptomyces venezuelae." Journal of Biotechnology **141**(3-4): 181-188.

Parks, W. M., A. R. Bottrill, O. A. Pierrat, M. C. Durrant and A. Maxwell (2007). "The action of the bacterial toxin, microcin B17, on DNA gyrase." Biochimie **89**(4): 500-507.

Patrikainen, P., L. Niiranen, K. Thapa, P. Paananen, P. Tahtinen, P. Mantsala, J. Niemi and M. Metsa-Ketela (2014). "Structure-based engineering of angucyclinone 6-ketoreductases." Chem Biol **21**(10): 1381-1391.

Payne, D. J., M. N. Gwynn, D. J. Holmes and D. L. Pompliano (2007). "Drugs for bad bugs: confronting the challenges of antibacterial discovery." Nat Rev Drug Discov **6**(1): 29-40.

Persson, B. and Y. Kallberg (2013). "Classification and nomenclature of the superfamily of short-chain dehydrogenases/reductases (SDRs)." Chemico-Biological Interactions **202**(1-3): 111-115.

Phillips, T., M. Chase, S. Wagner, C. Renzi, M. Powell, J. DeAngelo and P. Michels (2013). "Use of in situ solid-phase adsorption in microbial natural product fermentation development." Journal of Industrial Microbiology & Biotechnology **40**(5): 411-425.

Pojer, F., S. M. Li and L. Heide (2002). "Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics." Microbiology **148**(Pt 12): 3901-3911.

Ponstingl, H., K. Henrick and J. M. Thornton (2000). "Discriminating between homodimeric and monomeric proteins in the crystalline state." Proteins **41**(1): 47-57.

Potterton, E., P. Briggs, M. Turkenburg and E. Dodson (2003). "A graphical user interface to the CCP4 program suite." Acta Crystallogr D Biol Crystallogr **59**(Pt 7): 1131-1137.

Pozzi, R., M. Simone, C. Mazzetti, S. Maffioli, P. Monciardini, L. Cavaletti, R. Bamonte, M. Sosio and S. Donadio (2011). "The genus Actinoallomurus and some of its metabolites." Journal of Antibiotics **64**(1): 133-139.

Rampakakis, E., C. Gkogkas, D. Di Paola and M. Zannis-Hadjopoulos (2010). "Replication initiation and DNA topology: The twisted life of the origin." J Cell Biochem **110**(1): 35-43.

Read, R. J. (1986). "Improved fourier coefficients for maps using phases from partial structures with errors." Acta Crystallogr. Sect. A **42**(MAY): 140-149.

Richardson, M. A., S. Kuhstoss, P. Solenberg, N. A. Schaus and R. N. Rao (1987). "A new shuttle cosmid vector, pKC505, for streptomycetes: its use in the cloning of three different spiramycin-

resistance genes from a Streptomyces ambofaciens library." Gene **61**(3): 231-241.

Robert, X. and P. Gouet (2014). "Deciphering key features in protein structures with the new ENDscript server." Nucleic Acids Research **42**(W1): W320-W324.

Rohr, J., J. M. Beale and H. G. Floss (1989). "Urdamycins, New Angucycline Antibiotics from Streptomyces-Fradiae .4. Biosynthetic-Studies of Urdamycin-a to Urdamycin-D." Journal of Antibiotics **42**(7): 1151-1157.

Rohr, J. and R. Thiericke (1992). "Angucycline group antibiotics." Nat Prod Rep **9**(2): 103-137.

Rohr, J., A. Zeeck and H. G. Floss (1988). "Urdamycins, new angucycline antibiotics from Streptomyces fradiae. III. The structures of urdamycins C and D." J Antibiot (Tokyo) **41**(1): 126-129.

Rossmann, M. G., A. Liljas, C.-I. Branden and L. J. Bansazak (1975). Evolutionary relationships among dehydrogenases. The Enzymes. I. P. D. Boyer. New York, NY, Academic Press. **11:** 61-102.

Ruckert, C., A. Albersmeier, T. Busche, S. Jaenicke, A. Winkler, O. H. Friojonsson, G. O. Hreggviosson, C. Lambert, D. Badcock, K. Bernaerts, J. Anne, A. Economou and J. Kalinowski (2015). "Complete genome sequence of Streptomyces lividans TK24." Journal of Biotechnology **199**: 21-22.

Ruiz, N., B. Falcone, D. Kahne and T. J. Silhavy (2005). "Chemical conditionality: a genetic strategy to probe organelle assembly." Cell **121**(2): 307-317.

Saha, R. P., R. P. Bahadur, A. Pal, S. Mandal and P. Chakrabarti (2006). "ProFace: a server for the analysis of the physicochemical features of protein-protein interfaces." BMC Struct Biol **6**: 11.

Sanchez, C., I. A. Butovich, A. F. Brana, J. Rohr, C. Mendez and J. A. Salas (2002). "The biosynthetic gene cluster for the antitumor rebeccamycin: Characterization and generation of indolocarbazole derivatives." Chemistry & Biology **9**(4): 519-531.

Sawitzke, J. A., L. C. Thomason, M. Bubunenko, X. Li, N. Costantino and D. L. Court (2013). "Recombineering: highly efficient in vivo genetic engineering using single-strand oligos." Methods Enzymol **533**: 157-177.

Schimana, J., H. P. Fiedler, I. Groth, R. Sussmuth, W. Beil, M. Walker and A. Zeeck (2000). "Simocyclinones, novel cytostatic angucyclinone antibiotics produced by Streptomyces antibioticus Tu 6040. I. Taxonomy, fermentation, isolation and biological activities." J Antibiot (Tokyo) **53**(8): 779-787.

Schimana, J., M. Walker, A. Zeeck and P. Fiedler (2001). "Simocyclinones: diversity of metabolites is dependent on fermentation conditions." J Ind Microbiol Biotechnol **27**(3): 144-148.

Schmieder, R. and R. Edwards (2012). "Insights into antibiotic resistance through metagenomic approaches." Future Microbiology **7**(1): 73-89.

Schmutz, E., S. Hennig, S. M. Li and L. Heide (2004). "Identification of a topoisomerase IV in actinobacteria: purification and characterization of ParYR and GyrBR from the coumermycin A1 producer Streptomyces rishiriensis DSM 40489." Microbiology **150**(Pt 3): 641-647.

Schmutz, E., A. Muhlenweg, S. M. Li and L. Heide (2003). "Resistance genes of aminocoumarin producers: two type II topoisomerase genes confer resistance against coumermycin A1 and clorobiocin." Antimicrob Agents Chemother **47**(3): 869-877.

Schoeffler, A. J. and J. M. Berger (2008). "DNA topoisomerases: harnessing and constraining energy to govern chromosome topology." Q Rev Biophys **41**(1): 41-101.

Schultz, J., F. Milpetz, P. Bork and C. P. Ponting (1998). "SMART, a simple modular architecture research tool: identification of signaling domains." Proc Natl Acad Sci U S A **95**(11): 5857-5864.

Sheldrick, G. M. (2008). "A short history of SHELX." Acta Crystallographica Section A **64**: 112-122.

Shin, S. M., J. M. Choi, E. di Luccio, Y. J. Lee, S. J. Lee, S. J. Lee, S. H. Lee and D. W. Lee (2015). "The structural basis of substrate promiscuity in UDP-hexose 4-epimerase from the hyperthermophilic Eubacterium Thermotoga maritima." Archives of Biochemistry and Biophysics **585**: 39-51.

Sissi, C., E. Vazquez, A. Chemello, L. A. Mitchenall, A. Maxwell and M. Palumbo (2010). "Mapping simocyclinone D8 interaction with DNA gyrase: evidence for a new binding site on GyrB." Antimicrob Agents Chemother **54**(1): 213-220.

Sosio, M., F. Giusino, C. Cappellano, E. Bossi, A. M. Puglia and S. Donadio (2000). "Artificial chromosomes for antibiotic-producing actinomycetes." Nat Biotechnol **18**(3): 343-345.

Spellberg, B., R. Guidos, D. Gilbert, J. Bradley, H. W. Boucher, W. M. Scheld, J. G. Bartlett, J. Edwards, Jr. and A. Infectious Diseases Society of (2008). "The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America." Clin Infect Dis **46**(2): 155-164.

Steffensky, M., A. Muhlenweg, Z. X. Wang, S. M. Li and L. Heide (2000). "Identification of the novobiocin biosynthetic gene cluster of Streptomyces spheroides NCIB 11891." Antimicrob Agents Chemother **44**(5): 1214-1222.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar (2011). "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods." Mol Biol Evol **28**(10): 2731-2739.

Tang, G. L., Y. Q. Cheng and B. Shen (2004). "Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase." Chemistry & Biology **11**(1): 33-45.

Theobald, U., J. Schimana and H. P. Fiedler (2000). "Microbial growth and production kinetics of Streptomyces antibioticus Tu 6040." Antonie Van Leeuwenhoek **78**(3-4): 307-313.

Thiara, A. S. and E. Cundliffe (1989). "Interplay of novobiocin-resistant and -sensitive DNA gyrase activities in self-protection of the novobiocin producer, Streptomyces sphaeroides." Gene **81**(1): 65-72.

Thompson, T. B., K. Katayama, K. Watanabe, C. R. Hutchinson and I. Rayment (2004). "Structural and functional analysis of tetracenomycin F2 cyclase from Streptomyces glaucescens - A type II polyketide cyclase." Journal of Biological Chemistry **279**(36): 37956-37963.

Toprak, E., A. Veres, J. B. Michel, R. Chait, D. L. Hartl and R. Kishony (2012). "Evolutionary paths to antibiotic resistance under dynamically sustained drug selection." Nature Genetics **44**(1): 101-U140.

Tosin, M., L. Betancor, E. Stephens, W. M. A. Li, J. B. Spencer and P. F. Leadlay (2010). "Synthetic Chain Terminators Off-Load

Intermediates from a Type I Polyketide Synthase." Chembiochem **11**(4): 539-546.

Tosin, M., Y. Demydchuk, J. S. Parascandolo, C. B. Per, F. J. Leeper and P. F. Leadlay (2011). "In vivo trapping of polyketide intermediates from an assembly line synthase using malonyl carba(dethia)-N-acetyl cysteamines." Chemical Communications **47**(12): 3460-3462.

Trefzer, A., S. Pelzer, J. Schimana, S. Stockert, C. Bihlmaier, H. P. Fiedler, K. Welzel, A. Vente and A. Bechthold (2002). "Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic." Antimicrob Agents Chemother **46**(5): 1174-1182.

Uchida, T., M. Imoto, Y. Watanabe, K. Miura, T. Dobashi, N. Matsuda, T. Sawa, H. Naganawa, M. Hamada, T. Takeuchi and H. Umezawa (1985). "Saquayamycins, New Aquayamycin-Group Antibiotics." Journal of Antibiotics **38**(9): 1171-1181.

Walsh, C. (2003). Antibiotics, American Society of Microbiology.

Wang, G. J., M. K. Kharel, P. Pahari and J. Rohr (2011). "Investigating Mithramycin Deoxysugar Biosynthesis: Enzymatic Total Synthesis of TDP-D-Olivose." Chembiochem **12**(17): 2568-2571.

Wang, Z. X., S. M. Li and L. Heide (2000). "Identification of the coumermycin A(1) biosynthetic gene cluster of Streptomyces rishiriensis DSM 40489." Antimicrob Agents Chemother **44**(11): 3040-3048.

Watve, M. G., R. Tickoo, M. M. Jog and B. D. Bhole (2001). "How many antibiotics are produced by the genus Streptomyces?" Archives of Microbiology **176**(5): 386-390.

Weber, S., C. Zolke, J. Rohr and J. M. Beale (1994). "Investigations of the Biosynthesis and Structural Revision of Landomycin-A." Journal of Organic Chemistry **59**(15): 4211-4214.

Weber, T., K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Bruccoleri, S. Y. Lee, M. A. Fischbach, R. Muller, W. Wohlleben, R. Breitling, E. Takano and M. H. Medema (2015). "antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters." Nucleic Acids Res **43**(W1): W237-243.

Wiegand, I., K. Hilpert and R. E. Hancock (2008). "Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances." Nat Protoc **3**(2): 163-175.

Winter, G. (2010). "Xia2: an expert system for macromolecular crystallography data reduction." J. Appl. Crystallogr. **43**: 186-190.

Wohlkonig, A., P. F. Chan, A. P. Fosberry, P. Homes, J. Z. Huang, M. Kranz, V. R. Leydon, T. J. Miles, N. D. Pearson, R. L. Perera, A. J. Shillings, M. N. Gwynn and B. D. Bax (2010). "Structural basis of quinolone inhibition of type IIA topoisomerases and target-mediated resistance." Nature Structural & Molecular Biology **17**(9): 1152-1153.

Wright, G. D. (2007). "The antibiotic resistome: the nexus of chemical and genetic diversity." Nat Rev Microbiol **5**(3): 175-186.

Xie, Z. P., B. Liu, H. P. Wang, S. X. Yang, H. Y. Zhang, Y. P. Wang, N. Y. Ji, S. Qin and H. Laatsch (2012). "Kiamycin, a Unique Cytotoxic Angucyclinone Derivative from a Marine Streptomyces sp." Marine Drugs **10**(3): 551-558.

Xie, Z. P., L. Zhou, L. Guo, X. P. Yang, G. W. Qu, C. J. Wu and S. M. Zhang (2016). "Grisemycin, a Bridged Angucyclinone with a Methylsulfinyl Moiety from a Marine-Derived Streptomyces sp." Organic Letters **18**(6): 1402-1405.

Yamanaka, K., K. A. Reynolds, R. D. Kersten, K. S. Ryan, D. J. Gonzalez, V. Nizet, P. C. Dorrestein and B. S. Moore (2014). "Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A." Proceedings of the National Academy of Sciences of the United States of America **111**(5): 1957-1962.

Yan, X. H., K. Probst, A. Linnenbrink, M. Arnold, T. Paululat, A. Zeeck and A. Bechthold (2012). "Cloning and Heterologous Expression of Three Type II PKS Gene Clusters from Streptomyces bottropensis." Chembiochem **13**(2): 224-230.

Zaburannyi, N., M. Rabyk, B. Ostash, V. Fedorenko and A. Luzhetskyy (2014). "Insights into naturally minimised Streptomyces albus J1074 genome." BMC Genomics **15**(1): 97.

Zheng, X. F., X. Y. Dai, Y. M. Zhao, Q. Chen, F. Lu, D. Q. Yao, Q. Yu, X. P. Liu, C. M. Zhang, X. C. Gu and M. Luo (2007). "Restructuring of the dinucleotide-binding fold in an NADP(H) sensor protein." Proceedings of the National Academy of Sciences of the United States of America **104**(21): 8809-8814.

Zhou, H., Y. Li and Y. Tang (2010). "Cyclization of aromatic polyketides from bacteria and fungi." Nat Prod Rep **27**(6): 839-868.

Zhou, Z., J. Y. Gu, Y. Q. Li and Y. F. Wang (2012). "Genome plasticity and systems evolution in Streptomyces." Bmc Bioinformatics **13**.

Publications

SimC7 is a novel NAD(P)H-dependent ketoreductase essential for the antibiotic activity of the DNA gyrase inhibitor simocyclinone

Martin Schäfer[1], Tung B.K. Le[1], Stephen J. Hearnshaw[2], Anthony Maxwell[2], Gregory L. Challis[3], Barrie Wilkinson[1] and Mark J. Buttner[1]

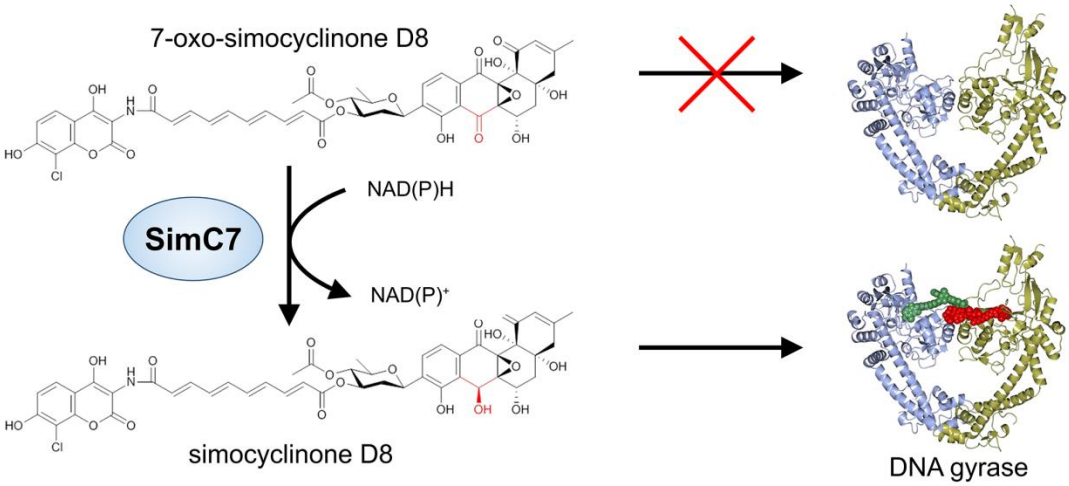
[1] Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom

[2] Department of Biological Chemistry, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom

[3] Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom

http://dx.doi.org/10.1016/j.jmb.2015.03.019

Published in: Journal of Molecular Biology (JMB), Volume 427, Issue 12, 19th June 2015, Pages 2192–2204



Article

SimC7 Is a Novel NAD(P)H-Dependent Ketoreductase Essential for the Antibiotic Activity of the DNA Gyrase Inhibitor Simocyclinone

Martin Schäfer [1], Tung B.K. Le [1], Stephen J. Hearnshaw [2], Anthony Maxwell [2], Gregory L. Challis [3], Barrie Wilkinson [1] and Mark J. Buttner [1]

*1 - **Department of Molecular Microbiology,** John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom*
*2 - **Department of Biological Chemistry,** John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom*
*3 - **Department of Chemistry,** University of Warwick, Coventry CV4 7AL, United Kingdom*

***Correspondence to Mark J. Buttner:** mark.buttner@jic.ac.uk*
http://dx.doi.org/10.1016/j.jmb.2015.03.019
Edited by B. Connolly

Abstract

Simocyclinone D8 (SD8) is a potent DNA gyrase inhibitor produced by *Streptomyces antibioticus* Tü6040. The simocyclinone (*sim*) biosynthetic gene cluster has been sequenced and a hypothetical biosynthetic pathway has been proposed. The tetraene linker in SD8 was suggested to be the product of a modular type I polyketide synthase working *in trans* with two monofunctional enzymes. One of these monofunctional enzymes, SimC7, was proposed to supply a dehydratase activity missing from two modules of the polyketide synthase. In this study, we report the function of SimC7. We isolated the entire ~72-kb *sim* cluster on a single phage artificial chromosome clone and produced simocyclinone heterologously in a *Streptomyces coelicolor* strain engineered for improved antibiotic production. Deletion of *simC7* resulted in the production of a novel simocyclinone, 7-oxo-SD8, which unexpectedly carried a normal tetraene linker but was altered in the angucyclinone moiety. We demonstrate that SimC7 is an NAD(P)H-dependent ketoreductase that catalyzes the conversion of 7-oxo-SD8 into SD8. 7-oxo-SD8 was essentially inactive as a DNA gyrase inhibitor, and the reduction of the keto group by SimC7 was shown to be crucial for high-affinity binding to the enzyme. Thus, SimC7 is an angucyclinone ketoreductase that is essential for the biological activity of simocyclinone.

Introduction

Simocyclinone D8 (SD8) is a natural hybrid antibiotic made by *Streptomyces antibioticus* Tü6040. It consists of a chlorinated aminocoumarin connected to an angucyclic polyketide via a tetraene diester linker and a D-olivose deoxysugar [1] (Fig. 1). SD8 is a potent inhibitor of supercoiling by DNA gyrase, an enzyme that is essential in bacteria but is absent from humans, making it an excellent target for antimicrobial drugs, as illustrated by the clinically important fluoroquinolones [2–4].

All known aminocoumarin antibiotics target bacterial DNA gyrase, but their modes of action differ. The three classical *Streptomyces* aminocoumarin antibiotics—novobiocin, clorobiocin, and coumermycin A1—bind to the GyrB subunit of DNA gyrase and inhibit ATPase activity [5]. In contrast, SD8 binds to the GyrA subunit of DNA gyrase and prevents binding to DNA [6]. SD8 is a bifunctional antibiotic with the aminocoumarin and angucyclic polyketide moieties at either end of the molecule binding to two separate pockets on the DNA-binding interface of GyrA [7,8]. The interaction with GyrA shows positive cooperativity, with the binding of one end of SD8 to its pocket promoting binding of the other end to its pocket. Based on these observations, it has been proposed that binding to GyrA is initiated by the high-affinity angucyclic polyketide, followed by improved binding of the low-affinity aminocoumarin [9]. Because of the bifunctional nature of simocyclinone, removal of either the aminocoumarin or the angucyclic polyketide from the antibiotic reduces its potency as a DNA gyrase inhibitor by about 100-fold compared to SD8, which inhibits DNA supercoiling at sub-micromolar concentrations (IC_{50} ~ 0.1–0.6 μM) [7,9].

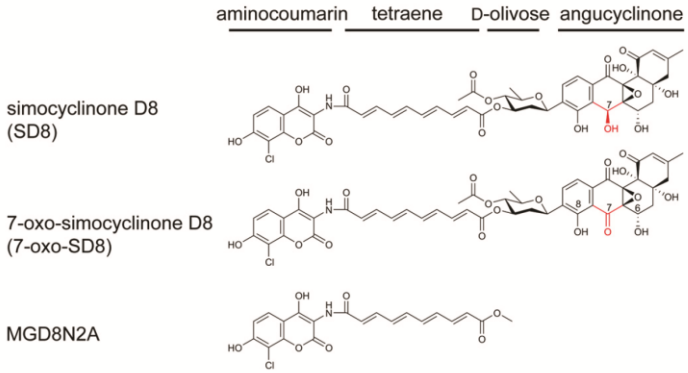


Fig. 1. Chemical structures of simocyclinone SD8, the major product of *S. antibioticus* Tü6040; 7-oxo-SD8, the major product of the *simC7* mutant; and MGD8N2A, a semi-synthetic analog lacking the angucyclinone polyketide moiety (generated by chemical hydrolysis of SD8). The absolute stereochemistry of SD8 was determined from co-crystals with SimR (PDB accession number 2Y30) [14] and *E. coli* DNA gyrase (PDB accession number 4CKL) [8]. The C-6, C-7, and C-8 positions discussed in the text are numbered, and the structural difference between SD8 and 7-oxo-SD8 is highlighted in red.

Based on the intermediates detectable in cultures of *S. antibioticus* Tü6040, it seems that the biosynthesis of SD8 starts with assembly of the angucyclic polyketide, followed by the attachment of the D-olivose deoxysugar, then the tetraene linker, and finally the aminocoumarin (i.e., assembled from right to left as drawn in Fig. 1) [10]. Accordingly, simocyclinone intermediates are classified into four groups (A–D), with A-group intermediates having only the angucyclic polyketide moiety, B-group intermediates having the angucyclic polyketide linked to the D-olivose deoxysugar, and so on [10].

The 49-gene simocyclinone (*sim*) biosynthetic cluster of *S. antibioticus* Tü6040 has been sequenced and a hypothetical biosynthetic pathway has been proposed [11,12]. Although two *sim* cluster transcriptional regulators have been studied [13–16], to date, only one biosynthetic enzyme has been characterized. SimD5 (SimL) catalyzes the presumed last step in the pathway, functioning as an amide-bond-forming ligase that attaches the aminocoumarin to the tetraene linker [17–19].

Trefzer *et al.* proposed that the tetraene linker in SD8 would be the product of a modular type I polyketide synthase, SimC1ABC, working *in trans* with two monofunctional enzymes, SimC6 and SimC7 [11]. One of these monofunctional enzymes, SimC7, a member of the short-chain dehydrogenase/reductase (SDR) superfamily, was proposed to supply the dehydratase activity missing from two modules of the polyketide synthase [11]. Here, we determine the true function of SimC7 experimentally and show that it is an angucyclinone ketoreductase that is essential for the antibiotic activity of simocyclinone.

Results

Heterologous expression of simocyclinones

Although directed mutations have been successfully created in the SD8 producer, *S. antibioticus* Tü6040 [11,12], this strain is particularly challenging to manipulate genetically. As a consequence, we instead chose to clone the *sim* gene cluster and analyze it in a heterologous system. The whole *sim* cluster was isolated on a single phage artificial chromosome (PAC) clone (PAC-12I) after PCR screening of a genomic library generated in *Escherichia coli* according to a recently developed protocol for working with large gene clusters [20]. Sequencing of the 95-kb insert in PAC-12I revealed the presence of the whole minimal *sim* gene cluster, consisting of 49 genes covering an ~72-kb region (Fig. S1), flanked by genomic regions of ~19 kb and 4 kb on either side.

We conjugated PAC-12I from the methylation-deficient *E. coli* strain ET12567 into the heterologous host *Streptomyces coelicolor* M1152. This strain has been engineered for improved heterologous production and analysis of secondary metabolites by deletion of endogenous secondary metabolic clusters, resulting in a strain without antimicrobial activity

and a much simplified extracellular metabolite profile for high-performance liquid chromatography (HPLC) and mass spectrometry analyses [21]. In addition, it carries a point mutation in the *rpoB* gene (encoding the RNA polymerase β-subunit) that causes a pleiotropic increase in the level of expression of secondary metabolic clusters [21]. PAC-12I strongly impaired the growth and sporulation of M1152. In an attempt to overcome this toxicity, we took advantage of a plasmid (pIJ10480 [13]) that expresses the *simEx1* simocyclinone efflux pump gene from the strong constitutive promoter *ermEp**. This plasmid was previously shown to increase the minimum inhibitory concentration (MIC) of SD8 for *Streptomyces lividans* from 2 µg/mL to 65 µg/mL [13]. To improve heterologous production of simocyclinones, we introduced pIJ10480 into M1152, resulting in strain *S. coelicolor* M1152ex1. Overexpression of *simEx1* increased the MIC of SD8 for M1152 from 1 µg/mL to 100 µg/mL.

We introduced PAC-12I and an empty vector control into the optimized heterologous host M1152ex1 by conjugation. Overexpression of the *simEx1* simocyclinone efflux pump gene reduced the toxicity of PAC-12I and M1152ex1 was used for all further work. Strains were grown in a defined production medium consisting of a basic nutrient medium with L-glutamine (5.84 g/L) and glycerol (20 mL/L) as nitrogen and carbon sources, respectively [10]. SD8 production in the heterologous host (24–56 mg/L) was 2- to 5-fold higher than that in the natural producer *S. antibioticus* Tü6040 (9–14 mg/L). No simocyclinones were detectable from M1152ex1 control strains lacking the *sim* gene cluster.

Structural analysis of a novel simocyclinone intermediate

In order to determine the role of SimC7 in simocyclinone biosynthesis, we created an in-frame deletion in *simC7* in PAC-12I using λ-Red PCR targeting, and we introduced the resulting PAC clone (PAC-12IΔC7) into M1152ex1. When fermentation extracts of the resulting *simC7* mutant strain were analyzed by HPLC, the major biosynthetic product had a later retention time than that of SD8 (Fig. 2). The UV–vis spectrum of the new compound was similar to that of SD8 but with reduced absorbance at 275 nm and a slight shift in the absorbance maxima to 245 nm and 361 nm (Fig. S2). Unexpectedly, the major product of the *simC7* mutant had a mass 2 Da lighter than SD8. All of the A-, B-, and C-group simocyclinone intermediates made by the *simC7* mutant were also 2 Da lighter than the equivalent molecules made by the wild-type *sim* cluster, suggesting that the 2-Da difference lies in the angucyclic polyketide and not in the tetraene linker. Based on this observation, we predicted that one of the five hydroxyl groups present in the angucyclic polyketide moiety of SD8 remained as a carbonyl

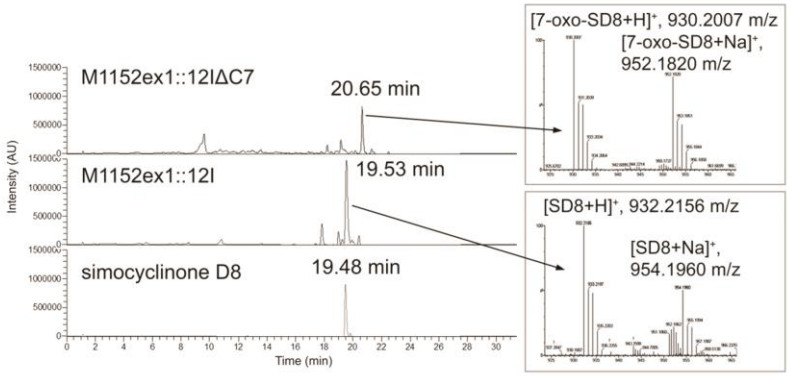


Fig. 2. LC-MS-UV analysis of simocyclinones produced by expression of the *sim* gene cluster in the heterologous host *S. coelicolor* M1152ex1. Chromatograms (200–700 nm) of culture extracts from clones carrying the *simC7* mutant cluster on PAC 12IΔ*simC7* (top) and the complete *sim* cluster on PAC 12I (middle), alongside an isolated standard of SD8 (bottom), are shown. Molecular weights were confirmed by high-resolution mass spectrometry (parent peaks are indicated; full tandem mass spectrometry data can be found in Fig. S3). Cultures were grown for 6 days at 30 °C in production medium.

group in this molecule and that SimC7 acts as a ketoreductase. To test this hypothesis, we purified the new simocyclinone intermediate from 4 L culture and determined its structure (63 mg isolated yield).

The molecular formula of the new compound was determined using high-resolution mass spectrometry, which confirmed that it was missing two hydrogen atoms {*m/z* (electrospray ionization) for $C_{46}H_{40}ClNO_{18}$: calculated, 930.2007; observed, 930.2009; Δ = 0.3 ppm $[M + H]^+$} compared to SD8 {*m/z* (electrospray ionization) for $C_{46}H_{42}ClNO_{18}$: calculated, 932.2163; observed, 932.2156; Δ = −0.8 ppm $[M + H]^+$} (Fig. 2 and Fig. S3). Tandem mass spectrometry analysis revealed similar fragmentation patterns for the new compound and SD8, with all of the ion fragments containing the angucyclic polyketide moiety having *m/z* ratios reduced by 2 Da compared to SD8. Conversely, all ion fragments from the new compound lacking the angucyclic polyketide moiety had identical masses with those derived from SD8 (Fig. 3).

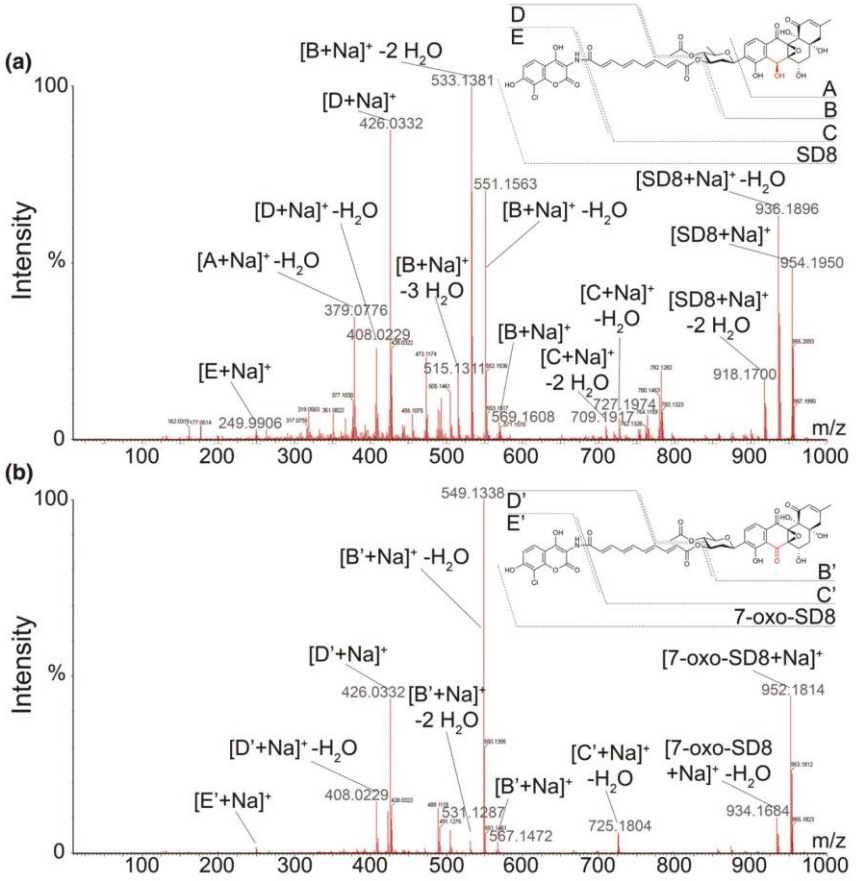


Fig. 3. High-resolution tandem mass spectra for the sodium adducts of (a) SD8 and (b) 7-oxo-SD8. Selected fragments are highlighted, consistent with the mass difference of 2 Da being located in the angucyclic polyketide moiety.

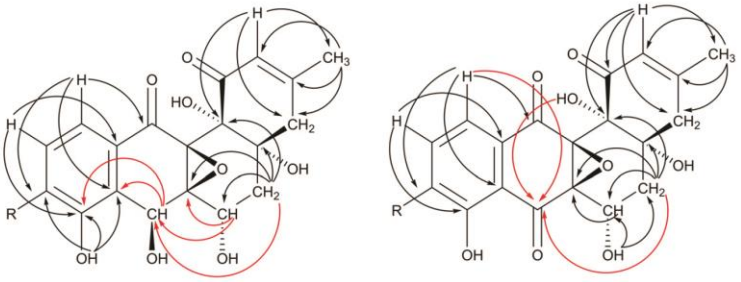


Fig. 4. Selected HMBC correlations for the angucyclic polyketide moiety of SD8 and of the new simocyclinone intermediate 7-oxo-SD8. Key correlations for the C-7 position are depicted in red.

Examination of the nuclear magnetic resonance (NMR) spectra [attached proton test, correlated spectroscopy, total correlated spectroscopy, heteronuclear single quantum coherence, and heteronuclear multiple bond correlation (HMBC)] suggested that the new molecule was structurally similar to SD8. However, the H-6 proton was shifted downfield by 0.4 ppm and no signal could be observed for the H-7 proton, consistent with a structural change in the angucyclic polyketide moiety (Fig. S4 and Table S1). Further analysis of the 2D (*2-d*imensional) NMR data enabled us to determine the position of the structural change (Figs. 4 and 5 and Figs. S5–S8). We confirmed appropriate proton–proton correlations for H-5 and H-6, which were supported by relevant proton–carbon correlations. In addition, we identified a signal shifted downfield for C-7 in the ^{13}C spectrum typical for a carbonyl group (Figs. S9 and S10 and Table S2) and cross-peaks to a C-7 resonance could be identified in the HMBC experiments (Fig. 4 and Fig. S7). From this, we concluded that the product of the *simC7* mutant had a carbonyl group instead of a hydroxyl group at position C-7 in the angucyclic polyketide. The new simocyclinone intermediate was named 7-oxo-SD8 (Fig. 1).

Complementation of the *simC7* mutant

To determine if the *simC7* phenotype was caused solely by loss of SimC7 function, we complemented the mutant with an *in trans* copy of the gene. *simC7* is the last gene of a putative seven-gene operon that starts with *simB7* (Fig. S1). Therefore, we drove expression of *simC7* from its putative native promoter by fusing *simC7* directly to an ~400-bp fragment spanning the intergenic region upstream of *simB7*, and we cloned this fusion into the vector pGM1190. Introduction of this construct restored SD8 production to the *simC7* mutant (data not shown).

SimC7 is an NAD(P)H-dependent ketoreductase that acts on the angucyclic polyketide

The fact that the *simC7* mutant made 7-oxo-SD8 suggested that SimC7 is a ketoreductase involved in biosynthesis of the angucyclic polyketide moiety and is not a dehydratase that acts during formation of the tetraene linker as had been proposed previously [11]. To test this hypothesis, we overexpressed an N-terminally His-tagged version of SimC7 in *E. coli* and assayed the purified recombinant protein for its ability to use 7-oxo-SD8 as a substrate. Both the substrate and the product of the reaction have strong chromophores with only minor differences in their absorbance spectra, and because their spectra also overlap those of $NAD(P)^+$/NAD(P)H, it was not possible to assay SimC7 using spectrophotometric methods. Instead, we followed the reaction by HPLC-UV (with SD8 as a standard) and found that SimC7 readily converted 7-oxo-SD8 into SD8 (Fig. 6b). SimC7 was able to use NADH or NADPH for the reduction (Fig. 6b) but showed a preference for NADPH (data not shown). It was also possible to assay the reaction in the reverse direction, following the oxidation of SD8 to 7-oxo-SD8 by SimC7 (Fig. 6c). This reaction was much slower and required a very high concentration of the NAD^+ cofactor (300 mM). In all cases, the identity of the reaction product was confirmed by tandem mass spectrometry.

The C-7 hydroxyl group is required for the antibiotic activity of simocyclinone

The activity of 7-oxo-SD8 was tested *in vivo* and *in vitro*. Wild-type *E. coli* and other Gram-negative bacteria are resistant to simocyclinones because the compounds cannot penetrate the outer membrane [1]. We therefore employed an *E. coli* strain

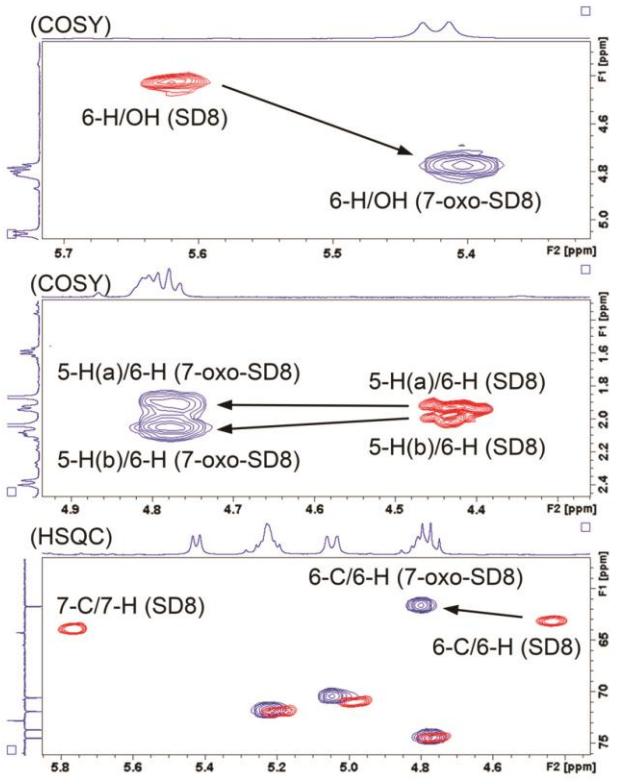


Fig. 5. Shifted peaks and cross-peaks of SD8 (red) and 7-oxo-SD8 (blue) from 2D NMR (correlated spectroscopy and heteronuclear single quantum coherence). 2D NMR maps are shown with projections of proton and carbon signals from 7-oxo-SD8. These are entirely consistent with the introduction of a carbonyl group at the C-7 position.

(NR698) that is sensitive to simocyclinones due to an in-frame deletion in the *imp* (*i*ncreased *m*embrane *p*ermeability) gene [22]. The MIC for SD8 was 0.3 μM. In contrast, NR698 grew in the presence of 7-oxo-SD8 at concentrations up to 17.5 μM, an ~60-fold increase in MIC.

The contrasting antibiotic activities of SD8 and 7-oxo-SD8 against whole cells suggested that the oxidation state of the oxygen at the C-7 position might be important for its ability to block DNA supercoiling by DNA gyrase. We therefore tested the activity of SD8 and 7-oxo-SD8 as DNA gyrase inhibitors *in vitro*. In line with previous studies [7,23,24], we found that SD8 inhibited supercoiling by DNA gyrase with IC_{50} = 0.1–0.5 μM (Fig. 7a). In contrast, 7-oxo-SD8 was almost 3 orders of magnitude less active (IC_{50} = 50–100 μM) (Fig. 7a).

The fluoroquinolone ciprofloxacin links DNA gyrase to its substrate by stabilizing the DNA–protein cleavage complex, leading to a characteristic "cleavage band" on gels (Fig. 7b) [4]. The presence of SD8 blocks DNA cleavage by gyrase by preventing the enzyme from binding to DNA [7]. We tested the ability of 7-oxo-SD8 to abrogate the ciprofloxacin--stimulated cleavage of DNA by gyrase, using SD8 as a control (Fig. 7b) [6]. As we found previously [7], SD8 prevented DNA cleavage at low concentrations (e.g., 0.5 μM), but we found that much higher concentrations of 7-oxo-SD8 (>10 μM) were required for inhibition of cleavage (Fig. 7b).

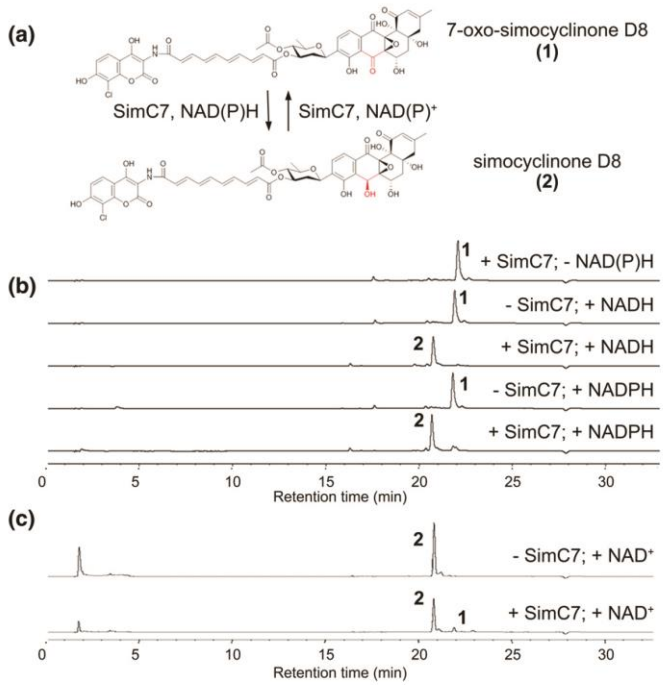


Fig. 6. Ketoreductase activity of SimC7. (a) SimC7 mediates conversion of the novel simocyclinone intermediate 7-oxo-SD8 (1) into SD8 (2). (b) NAD(P)H-dependent conversion of 7-oxo-SD8 into SD8 and (c) NAD(P)$^+$-dependent conversion of SD8 into 7-oxo-SD8, monitored by reversed-phase HPLC. Samples were incubated for 1 h at room temperature. Reactions were stopped by quenching in methanol (1:1) and heat denaturation for 10 min at 100 °C.

To investigate the binding of 7-oxo-SD8 to GyrA, we used surface plasmon resonance as described previously [7]. We found that SD8 bound to the GyrA N-terminal domain with a similar affinity to that reported previously, but that only weak, non-specific binding could be seen for the 7-oxo analog (data not shown). Taken together, these data suggest that 7-oxo-SD8 binds gyrase 2–3 orders of magnitude more weakly than SD8.

MGD8N2A is a simocyclinone analog (generated by chemical hydrolysis of SD8) that lacks the angucyclic polyketide (Fig. 1). This analog has been tested previously and found to have greatly reduced activity against DNA gyrase (IC_{50} = 50 µM) in comparison to SD8 (IC_{50} = 0.1–0.6 µM) [7]. We repeated these experiments with MGD8N2A and obtained a similar value (IC_{50} = 25 µM). Thus, comparing the IC_{50} values for 7-oxo-SD8 and MGD8N2A, it is clear that the presence of a carbonyl group at the C-7 position has a similarly negative effect on the activity of simocyclinone as does complete loss of the angucyclic polyketide.

Discussion

To examine the prediction that SimC7 is a dehydratase involved in the biosynthesis of the tetraene linker, we isolated the ~72-kb *sim* gene cluster on a single PAC clone and expressed simocyclinone heterologously in an *S. coelicolor* strain engineered for improved antibiotic production. Deletion of *simC7* from the PAC clone resulted in the production of a novel simocyclinone, 7-oxo-SD8, which unexpectedly carried a normal tetraene linker but was altered in the angucyclinone moiety. We went on to demonstrate that SimC7 is an NAD(P)H-dependent ketoreductase that converts 7-oxo-SD8

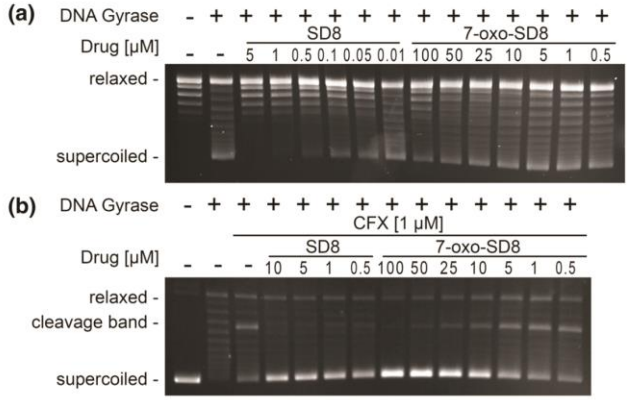


Fig. 7. (a) Supercoiling and (b) cleavage relaxation assays with DNA gyrase in the presence of 7-oxo-SD8 or SD8. The reaction mixtures contained *E. coli* DNA gyrase and varying concentrations of either SD8 or 7-oxo-SD8. For relaxation assays, the mixtures were incubated without ATP and with ciprofloxacin (1 µM).

into SD8 and that reduction of the keto group, catalyzed by SimC7, is essential for the antibiotic activity of simocyclinone.

Use of the *S. coelicolor* M1152 heterologous expression system

Heterologous expression is a convenient tool for structure–function analysis of secondary metabolites, and improved derivatives of *S. coelicolor* and *Streptomyces avermitilis* have recently been developed for this purpose [21,25,26]. Here, we demonstrate the utility of this approach for the manipulation and functional analysis of simocyclinone, an antibiotic naturally made by *S. antibioticus* Tü6040, a strain that is particularly challenging for genetic manipulation. Despite the large size of the *sim* gene cluster (49 genes covering ~72 kb), it was straightforward to isolate a PAC clone carrying the entire cluster and to introduce it by conjugation into an engineered *S. coelicolor* host, where the clone integrated irreversibly and replicated stably as part of the chromosome. The same approach was recently used for the successful heterologous expression in *S. coelicolor* of the 83.5-kb FK506 gene cluster from *Streptomyces tsukubaensis* [20]. The engineered *S. coelicolor* host often gives a greater yield than the natural producer [20,21], and this was also true for simocyclinone. A further advantage of this approach is that genetic modification of the target cluster can be carried out in *E. coli*, giving full access to the advanced recombineering tools available in this host. The explosion of whole genome sequencing has revealed tens of thousands of new secondary metabolic gene clusters [27,28], many in rare and difficult-to-culture actinomycetes for which no genetic tools have been developed. As a consequence, the heterologous expression approach exemplified here is an attractive option for characterizing many of these clusters and is likely to be of increasing importance in the future.

The biochemical function of SimC7

SimC7 was proposed to be a dehydratase involved in the biosynthesis of the tetraene linker [11], but we have demonstrated that SimC7 is actually an NAD(P)H-dependent ketoreductase that reduces a carbonyl group at the C-7 position of the angucyclic polyketide moiety. The fermentation product of the *simC7* mutant, 7-oxo-SD8, was almost inactive as a DNA gyrase inhibitor, showing that reduction of the 7-oxo functional group by the ketoreductase activity of SimC7 is essential for the biological function of simocyclinones. Although SimC7 readily converts 7-oxo-SD8 into SD8, synthesis of the angucyclic polyketide is normally completed before it is linked to the D-olivose sugar and the other moieties of simocyclinone [10]. This means that the natural substrate of SimC7 would be an A-group intermediate (i.e., one having only the angucyclic polyketide moiety) carrying a C-7 carbonyl. This species was detected neither in the native producer *S. antibioticus* [10–12] nor in the

heterologous host carrying the wild-type *sim* cluster, but as expected, a species with the appropriate mass was detected among the intermediates made by the *simC7* mutant. Having determined the true function of SimC7, it remains unclear how the tetraene linker of simocyclinone is assembled and how dehydration takes place.

SimC7 (30 kDa; 284 amino acids) is a member of the SDR family, one of the largest protein superfamilies, having more than 120,000 representatives in the databases. SDR proteins are diverse, with low overall amino acid sequence identity (typically 20–30% in pairwise comparisons), and they are principally characterized by the presence of a predicted pyridine nucleotide-binding Rossmann fold, comprising a parallel β-sheet flanked by three helices on each side. SDR proteins also have diverse biochemical activities, including acting as dehydratases, reductases, dehydrogenases, decarboxylases, and epimerases [29,30]. Despite its function as an angucyclinone ketoreductase, based on sequence identity, SimC7 seems to be most similar to various SDR sugar epimerases, and SimC7 secondary and tertiary structure predictions using the Phyre2 server [31] also show highest similarity to sugar-modifying enzymes, such as epimerases, reductases, and dehydratases. No close homologs in the databases are known to function as ketoreductases.

Although none is closely related to SimC7, several SDR ketoreductases have been identified that act on polyketides. These include SDR proteins that reduce the carbonyl group at position C-9 in the biosynthetic pathways for jadomycin (JadE), actinorhodin (ActKR), and hedamycin (HedKR) [32–34] and the SDR protein that reduces the carbonyl group at position C-6 in the landomycin pathway (LanV) [35]. Close homologs of these enzymes are encoded in the *sim* gene cluster, raising the possibility that SimA6 and SimA9 function to reduce the angucyclinone C-10 and C-6 carbonyl groups, respectively, during the biosynthesis of simocyclinone.

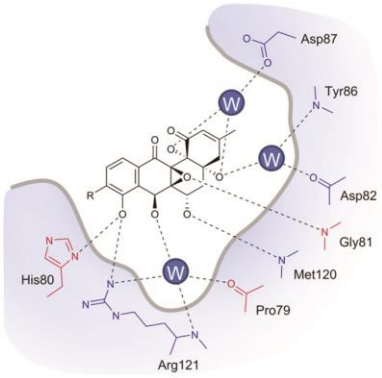


Fig. 8. The angucyclic polyketide binding pocket in GyrA, as revealed in the crystal structure of the GyrA–SD8 complex (PDB accession number 4CKL) [8]. SD8 is contacted by residues from both GyrA subunits; those from one subunit are shown in red and those from the other subunit are shown in blue. For clarity, all hydrogens have been omitted.

SimC7 function is vital for producing simocyclinones with antibiotic activity

Why does such a small structural change render 7-oxo-SD8 effectively inactive as a DNA gyrase inhibitor? The crystal structure of the GyrA–SD8 complex revealed how DNA gyrase binds the angucyclic polyketide [8] (Fig. 8). The C-6 hydroxyl group makes a hydrogen bond to Met120, the C-7 hydroxyl group makes hydrogen bonds to Pro79 and Arg121 (via a water molecule), and the C-8 hydroxyl group makes hydrogen bonds to His80 and Arg121. One speculative possibility is that the presence of a C-7 carbonyl group in 7-oxo-SD8 leads to the formation of intramolecular hydrogen bond with the neighboring C-8 hydroxyl group, thus simultaneously breaking contacts with His80 and a highly coordinated water molecule held in place by Pro79 and Arg121 (Fig. 8). His80 in particular seems to play a crucial role in binding simocyclinone, as the SD8 IC_{50} is increased 230-fold when this residue is mutated to alanine [7]. In addition, the presence of a carbonyl group at C-7 will alter the conformation of the angucyclic polyketide, which may affect other binding interactions with GyrA. Thus, the introduction of a carbonyl group at C-7 is likely to break several hydrogen bonds that secure the angucyclic polyketide in its binding pocket.

Materials and Methods

Strains, plasmids, and oligonucleotides

Strains, plasmids, and oligonucleotides used in this work are shown in Tables S3, S4, and S5, respectively.

Antibiotic selection

Apramycin (50 μg/mL), carbenicillin (100 μg/mL), chloramphenicol (25 μg/mL), hygromycin (40 μg/mL), kanamycin (50 μg/mL), nalidixic acid (25 μg/mL), and thiostrepton (60 μg/mL) were added to solid media as required unless otherwise noted. These concentrations were halved in liquid cultures.

PAC library construction

A genomic library of *S. antibioticus* Tü6040 DNA was constructed by Bio S&T Inc. (Montreal, Canada) as described in Ref. [20], using the phage P1-derived artificial chromosome (PAC) vector pESAC13†, a derivative of pPAC-S1 [36]. The integrative vector pESAC13 conferred kanamycin resistance in *E. coli* and thiostrepton resistance in *Streptomyces*. Genomic DNA was isolated and cloned between the two BamHI sites of pESAC13, replacing the carbenicillin resistance gene. The genomic DNA library consisted of 2688 individual PAC clones with an average insert size of 110 kb (more than 20× genome coverage).

Identification of PAC clones containing the complete *sim* gene cluster

The *S. antibioticus* PAC library was screened with two primer pairs (Table S5) amplifying fragments flanking the *sim* gene cluster. The screening identified six double-positive clones with amplification of both PCR products. One double-positive clone (PAC-12I; Table S4) was sequenced. The insert in PAC-12I contained the minimal 72-kb *sim* gene cluster flanked by genomic regions of ~19 kb and 4 kb on either side (a total insert size of 95 kb), and PAC-12I was chosen for heterologous expression studies.

Deletion of *simC7* from PAC-12I and *in trans* complementation

The *simC7* gene was deleted by PCR targeting according to standard protocols [37,38], leaving an 81-bp in-frame scar. The PAC mutant (PAC-12IΔC7) structure was confirmed by PCR and restriction digestion with EcoRI. For complementation, *simC7* was fused to the *simB7* promoter and cloned into pGM1190. Complementation of *simC7* was confirmed by liquid chromatography mass spectrometry (LC-MS).

Mobilization of PAC clones

PAC clones were conjugated from *E. coli* DH5α into the methylation-deficient *E. coli* strain ET12567 by triparental mating, using *E. coli* TOP10 cells carrying the driver plasmid pR9406 to mobilize the PAC [20]. PAC clones were conjugated from *E. coli* ET12567(pR9406) into *S. coelicolor* M1152 and M1152ex1 as previously described [39]. After 16–20 h of incubation on R2 medium without sucrose, plates were overlaid with thiostrepton and nalidixic acid. Exconjugants were streaked on Mannitol Soy Flour medium with thiostrepton and nalidixic acid. For spore preparation, strains were grown on Instant Mashed Potato Agar (2 g/L Smash® instant mashed potato and 2 g/L LabM agar, made up with tap water).

Production medium and purification of simocyclinones

To optimize production of D-group simocyclinones, we cultured strains for 6 days at 30 °C and 250 rpm in the basal chemically defined medium described in Ref. [10], supplemented with L-glutamine (5.84 g/L) as nitrogen source and glycerol (20 mL/L) as carbon source. The fermentation broth was acidified with HCl (5 M) to pH 4 and simocyclinones were extracted with an equal volume of ethyl acetate. After centrifugation, the organic layer was evaporated *in vacuo*. For the purification of 7-oxo-SD8, the dried extract from 4 L culture medium (~1.5 g with ~15% compound) was dissolved in methanol, filtered through glass wool and purified by normal-phase liquid chromatography on a silica column (Biotage® SNAP cartridge KP-SIL 50 g, 40- to 65-µm particle size, 39 mm × 81 mm) using a gradient elution from CH_2Cl_2-MeOH (95:5) to CH_2Cl_2-MeOH (70:30) over 4 h at a flow rate of 50 mL/h [40]. Fractions of 10 mL were collected and the presence of the compound was verified by thin layer chromatography and LC-MS. Selected fractions were pooled, dried *in vacuo*, and tritiated in acetone. The acetone-soluble fraction was dried and dissolved in methanol for size-exclusion chromatography using Sephadex LH-20 (18- to 111-µm particle size) (Sigma-Aldrich). Following elution using methanol, we collected fractions and confirmed the presence of the compound by LC-MS. Fractions containing 7-oxo-SD8 were combined, the solvent was removed under reduced pressure, and the concentrate was dissolved in a methanol–acetone mix (1:1) and purified further by preparative HPLC. After freeze drying, we obtained ~63 mg of the new simocyclinone derivative in a fine yellow powder (yield, ~27%; purity, >98%).

High-performance liquid chromatography

For analytical HPLC, samples (30 µL) were separated on an HPLC column (Phenomenex Gemini-NX 3u C_{18} 110A, 150 mm × 4.6 mm) using a linear gradient with 0.1% formic acid as mobile phase A and methanol as mobile phase B. The gradient was either 25–95% or 75–95% solvent B over 20 min at a flow rate of 1 mL/min. Absorbance was recorded with a diode array detector at 210 nm, 230 nm, and 360 nm for simocyclinones and 535 nm as reference. For preparative HPLC, samples (350 µL) were injected onto an HPLC column (Phenomenex Gemini-NX 5u C_{18} 110A AXIA Packed, 150 mm × 21.20 mm) with a Phenomenex SecurityGuard Prep Cartridge Holder Kit (21.20 mm) attached to a Dionex UltiMate 3000 HPLC machine. Separation was performed using a linear gradient with 0.1% formic acid as mobile phase A and methanol as mobile phase B. The gradient was 75–95% mobile phase B over 20 min at a flow rate of 21 mL/min, and absorbance was recorded at 360 nm.

Liquid chromatography mass spectrometry

Samples (10 µL) were injected onto an HPLC column (Phenomenex Luna-C_{18} 3u C18(2) 100A, 100 mm × 2 mm) with a Phenomenex SecurityGuard Prep Cartridge Holder Kit (2 mm), attached to a Thermo ion-trap LC-MS (LCQ DECA XPlus). Separation was performed using a linear gradient with 0.1% formic acid as mobile phase A and methanol as mobile phase B. The gradient was 25–95% mobile phase B over 20 min at a flow rate of 1 mL/min.

NMR spectrometry

Spectra were recorded on a Bruker Avance DRX-700 and Avance III 400 spectrometer at 298 K. Data were processed and analyzed using TopSpin v3.2 software (Bruker). Assignments were made from experiments recorded on the DRX-700 spectrometer; ^{1}H coupling constants were determined using experiments from the Avance III 400 spectrometer. Simocyclinones were dissolved in DMSO-d_6.

Overexpression and purification of SimC7

simC7 was PCR amplified from PAC-12I using Q5 DNA polymerase (New England Biolabs) and oligonucleotides that introduced NdeI and BamHI sites (Table S5). The PCR product was cloned into pJET1.2/blunt and then sub-cloned into pET15b vector (Novagen) cut with NdeI and BamHI. The resulting construct, pET15b-NB-C7, was sequenced for errors (MWG Eurofins).

E. coli Rosetta (DE3)pLysS-competent cells (Novagen) were transformed with pET15b-NB-C7 and grown at 25 °C in LB containing chloramphenicol and carbenicillin. One 50-mL overnight culture was used to inoculate two 400-mL main cultures that were grown to an OD_{600} of ~0.4. Protein expression was induced with IPTG added to a final concentration of 0.5 mM and cultures were grown for an additional 16 h at 25 °C. Cells were harvested and lysed by sonication in 40 mL fresh lysis buffer [50 mM Tris–HCl (pH 8.0), 150 mM NaCl, 0.1% (v/v) Triton-X100, 0.5 mg lysozyme, and complete ethylenediaminetetraacetic-acid-free protease inhibitor (Roche)]. Cell debris was removed by centrifugation, the supernatant was incubated with 0.6 mL of Ni-NTA agarose beads (Qiagen) for 30 min, and then the resin was packed into a 5-mL polypropylene column (Qiagen). Unbound proteins were removed with 200 mL wash buffer [50 mM Tris–HCl (pH 8.0), 250 mM NaCl, and 5% (v/v) glycerol] and His-tagged SimC7 was eluted in 2–3 mL elution buffer [50 mM Tris–HCl (pH 8.0), 50 mM NaCl, 5% (v/v) glycerol, and 250 mM imidazole]. The purified protein was dialysed overnight into storage buffer [20 mM Tris–HCl (pH 8.0), 150 mM NaCl, 5% (v/v) glycerol, 1 mM DTT, and 2 mM ethylenediaminetetraacetic acid] and then frozen in liquid nitrogen and stored at −80 °C. Protein concentrations were determined using Bradford reagent (BioRad).

SimC7 ketoreductase assays

Ketoreductase assays were performed in reaction mixtures consisting of substrate (25–200 μM), NAD(P)H (0.3 mM) or NAD^+ (300 mM), Hepes (50 mM, pH 7.2), and enzyme (500 nM). Reactions were initiated with SimC7 and quenched at specific time points with equal volumes of methanol. Stopped reactions were incubated for 10 min at 100 °C and centrifuged to remove precipitated protein (13,000*g* for 10 min at 4 °C). SD8 formation was quantified by analytical reversed-phase HPLC using a linear gradient of 25–95% methanol over 20 min then 95% methanol for 5 min against 0.1% formic acid in water, and absorbance was recorded at 360 nm. SD8 concentrations were calculated by comparison with a standard curve (Fig. S11), which was calculated from a serial dilution of SD8 prepared in methanol (0.25–70 μM) under the conditions described above (7.5–2100 pmol).

Minimum inhibitory concentrations

MICs were determined as described in Ref. [41]. *E. coli* NR698 [22] was grown in Müller-Hinton broth for 16–20 h at 37 °C and bacterial growth was measured photometrically at 600 nm wavelength.

DNA gyrase assays

Assays were performed as described previously [6,8].

Acknowledgements

We would like to thank Agnieszka Ptasińska and Oluwaseyi Shorinola for assistance in the early stages of this work; Hans-Peter Fiedler for the gift of SD8; David Lawson and Stephen Bornemann for helpful discussion; Mervyn Bibb for comments on the manuscript; and Lionel Hill, Clare Stevenson, and Sergey Nepogodiev for help with LC-MS, surface plasmon resonance, and NMR, respectively. This work was funded by a Norwich Research Park Biotechnology and Biological Sciences Research Council (BBSRC) Doctoral Training Program studentship to M.S., by BBSRC grant BB/I002197/1 to M.J.B., and by the Institute Strategic Program on Understanding Plant and Microbial Metabolism Grant to the John Innes Centre from the BBSRC.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2015.03.019.

Received 17 February 2015;
Received in revised form 30 March 2015;
Accepted 31 March 2015
Available online 8 April 2015

Keywords:
simocyclinones;
DNA gyrase;
antibiotics;
short-chain dehydrogenase/reductase (SDR) superfamily;
ketoreductase

Present address: Tung B. K. Le, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

†http://www.ncbi.nlm.nih.gov/nuccore/677282244

Abbreviations used:
BBSRC, Biotechnology and Biological Sciences Research Council; SD8, simocyclinone D8; SDR, short-chain dehydrogenase/reductase; MIC, minimum inhibitory concentration; HMBC, heteronuclear multiple bond

correlation; LC-MS, liquid chromatography mass spectrometry; PAC, phage artificial chromosome.

References

[1] Schimana J, Fiedler H-P, Groth I, Sussmuth R, Beil W, Walker M, et al. Simocyclinones, novel cytostatic angucyclinone antibiotics produced by *Streptomyces antibioticus* Tu 6040. I. Taxonomy, fermentation, isolation and biological activities. J Antibiot (Tokyo) 2000;53:779–87.

[2] Andersson MI, MacGowan AP. Development of the quinolones. J Antimicrob Chemother 2003;51:1–11.

[3] Drlica K, Hiasa H, Kerns R, Malik M, Mustaev A, Zhao X. Quinolones: action and resistance updated. Curr Top Med Chem 2009;9:981–98.

[4] Collin F, Karkare S, Maxwell A. Exploiting bacterial DNA gyrase as a drug target: current state and perspectives. Appl Microbiol Biotechnol 2011;92:479–97.

[5] Gormley NA, Orphanides G, Meyer A, Cullis PM, Maxwell A. The interaction of coumarin antibiotics with fragments of DNA gyrase B protein. Biochemistry 1996;35:5083–92.

[6] Flatman RH, Howells AJ, Heide L, Fiedler H-P, Maxwell A. Simocyclinone D8: an inhibitor of DNA gyrase with a novel mode of action. Antimicrob Agents Chemother 2005;49:1093–100.

[7] Edwards MJ, Flatman RH, Mitchenall LA, Stevenson CE, Le TBK, Fiedler H-P, et al. A crystal structure of the bifunctional antibiotic, simocyclinone D8, bound to DNA gyrase. Science 2009;326:1415–8.

[8] Hearnshaw SJ, Edwards MJ, Stevenson CE, Lawson DM, Maxwell A. A new crystal structure of the bifunctional antibiotic simocyclinone D8 bound to DNA gyrase gives fresh insight into the mechanism of inhibition. J Mol Biol 2014;426:2023–33.

[9] Edwards MJ, Williams MA, Maxwell A, McKay AR. Mass spectrometry reveals that the antibiotic simocyclinone D8 binds to DNA gyrase in a "bent-over" conformation: evidence of positive cooperativity in binding. Biochemistry 2011;50:3432–40.

[10] Schimana J, Walker M, Zeeck A, Fiedler H-P. Simocyclinones: diversity of metabolites is dependent on fermentation conditions. J Ind Microbiol Biotechnol 2001;27:144–8.

[11] Trefzer A, Pelzer S, Schimana J, Stockert S, Bihlmaier C, Fiedler H-P, et al. Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic. Antimicrob Agents Chemother 2002;46:1174–82.

[12] Galm U, Schimana J, Fiedler H-P, Schmidt J, Li SM, Heide L. Cloning and analysis of the simocyclinone biosynthetic gene cluster of *Streptomyces antibioticus* Tu 6040. Arch Microbiol 2002;178:102–14.

[13] Le TBK, Fiedler H-P, den Hengst CD, Ahn SK, Maxwell A, Buttner MJ. Coupling of the biosynthesis and export of the DNA gyrase inhibitor simocyclinone in *Streptomyces antibioticus*. Mol Microbiol 2009;72:1462–74.

[14] Le TBK, Stevenson CEM, Fiedler H-P, Maxwell A, Lawson DM, Buttner MJ. Structures of the TetR-like simocyclinone efflux pump repressor, SimR, and the mechanism of ligand-mediated derepression. J Mol Biol 2011;408:40–56.

[15] Le TBK, Schumacher MA, Lawson DM, Brennan RG, Buttner MJ. The crystal structure of the TetR family transcriptional repressor SimR bound to DNA and the role of a flexible N-terminal extension in minor groove binding. Nucleic Acids Res 2011;39:9433–47.

[16] Horbal L, Rebets Y, Rabyk M, Makitrynskyy R, Luzhetskyy A, Fedorenko V, et al. SimReg1 is a master switch for biosynthesis and export of simocyclinone D8 and its precursors. AMB Express 2012;2:1.

[17] Luft T, Li SM, Scheible H, Kammerer B, Heide L. Overexpression, purification and characterization of SimL, an amide synthetase involved in simocyclinone biosynthesis. Arch Microbiol 2005;183:277–85.

[18] Pacholec M, Freel Meyers CL, Oberthur M, Kahne D, Walsh CT. Characterization of the aminocoumarin ligase SimL from the simocyclinone pathway and tandem incubation with NovM, P, N from the novobiocin pathway. Biochemistry 2005;44:4949–56.

[19] Anderle C, Hennig S, Kammerer B, Li SM, Wessjohann L, Gust B, et al. Improved mutasynthetic approaches for the production of modified aminocoumarin antibiotics. Chem Biol 2007;14:955–67.

[20] Jones AC, Gust B, Kulik A, Heide L, Buttner MJ, Bibb MJ. Phage P1-derived artificial chromosomes facilitate heterologous expression of the FK506 gene cluster. PLoS One 2013;8:e69319.

[21] Gomez-Escribano JP, Bibb MJ. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. Microbial Biotechnol 2011;4:207–15.

[22] Ruiz N, Falcone B, Kahne D, Silhavy TJ. Chemical conditionality: a genetic strategy to probe organelle assembly. Cell 2005;121:307–17.

[23] Oppegard LM, Hamann BL, Streck KR, Ellis KC, Fiedler H-P, Khodursky AB, et al. *In vivo* and *in vitro* patterns of the activity of simocyclinone D8, an angucyclinone antibiotic from *Streptomyces antibioticus*. Antimicrob Agents Chemother 2009;53:2110–9.

[24] Alt S, Mitchenall LA, Maxwell A, Heide L. Inhibition of DNA gyrase and DNA topoisomerase IV of *Staphylococcus aureus* and *Escherichia coli* by aminocoumarin antibiotics. J Antimicrob Chemother 2011;66:2061–9.

[25] Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. Proc Natl Acad Sci USA 2010;107:2646–51.

[26] Komatsu M, Komatsu K, Koiwai H, Yamada Y, Kozone I, Izumikawa M, et al. Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. ACS Synth Biol 2013;2:384–96.

[27] Cimermancic P, Medema MH, Claesen J, Kurita K, Wieland Brown LC, Mavrommatis K, et al. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. Cell 2014;158:412–21.

[28] Medema MH, Cimermancic P, Sali A, Takano E, Fischbach MA. A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. PLoS Comput Biol 2014;10:e1004016.

[29] Kallberg Y, Oppermann U, Persson B. Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. FEBS J 2010;277:2375–86.

[30] Persson B, Kallberg Y. Classification and nomenclature of the superfamily of short-chain dehydrogenases/reductases (SDRs). Chem Biol Interact 2013;202:111–5.

[31] Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 2009;4:363–71.

[32] Kulowski K, Wendt-Pienkowski E, Han L, Yang KQ, Vining LC, Hutchinson CR. Functional characterization of the *jadI*

gene as a cyclase forming angucyclinones. J Am Chem Soc 1999;121:1786–94.
[33] Javidpour P, Das A, Khosla C, Tsai SC. Structural and biochemical studies of the hedamycin type II polyketide ketoreductase (HedKR): molecular basis of stereo- and regiospecificities. Biochemistry 2011;50:7426–39.
[34] Korman TP, Tan YH, Wong J, Luo R, Tsai SC. Inhibition kinetics and emodin cocrystal structure of a type II polyketide ketoreductase. Biochemistry 2008;47:1837–47.
[35] Paananen P, Patrikainen P, Kallio P, Mäntsälä P, Niemi J, Niiranen L, et al. Structural and functional analysis of angucycline C-6 ketoreductase LanV involved in landomycin biosynthesis. Biochemistry 2013;52:5304–14.
[36] Sosio M, Giusino F, Cappellano C, Bossi E, Puglia AM, Donadio S. Artificial chromosomes for antibiotic-producing actinomycetes. Nat Biotechnol 2000;18:343–5.
[37] Gust B, Challis GL, Fowler K, Kieser T, Chater KF. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci USA 2003;100:1541–6.
[38] Gust B, Chandra G, Jakimowicz D, Yuqing T, Bruton CJ, Chater KF. Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. Adv Appl Microbiol 2004; 54:107–28.
[39] Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. Practical *Streptomyces* Genetics. Norwich: The John Innes Foundation; 2000.
[40] Holzenkampfer M, Walker M, Zeeck A, Schimana J, Fiedler H-P. Simocyclinones, novel cytostatic angucyclinone antibiotics produced by *Streptomyces antibioticus* Tu 6040 II. Structure elucidation and biosynthesis. J Antibiot (Tokyo) 2002;55:301–7.
[41] Wiegand I, Hilpert K, Hancock RE. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. Nat Protoc 2008;3: 163–75.

Substrate-Assisted Catalysis in Polyketide Reduction Proceeds via a Phenolate Intermediate

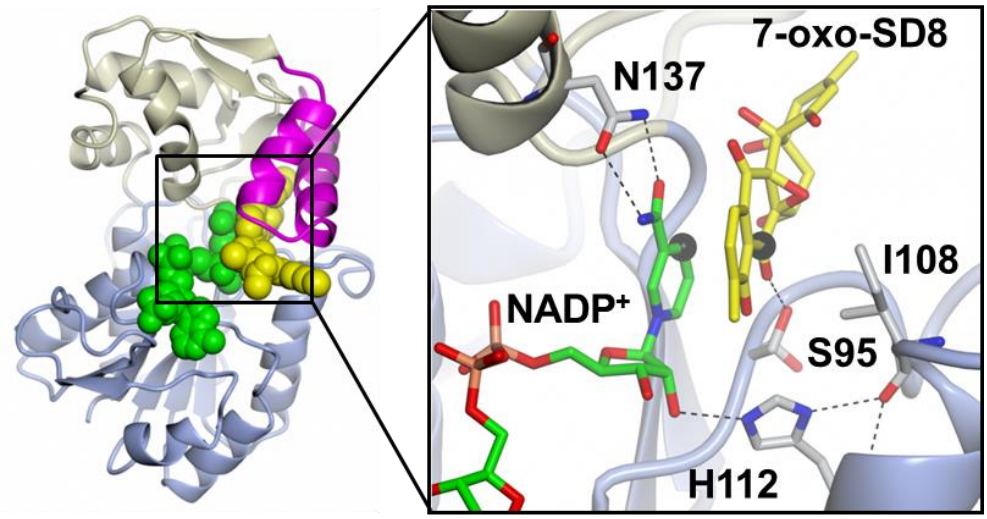
Martin Schäfer[1], Clare E. M. Stevenson[2], David M. Lawson[2], Barrie Wilkinson[1] and Mark J. Buttner[1]

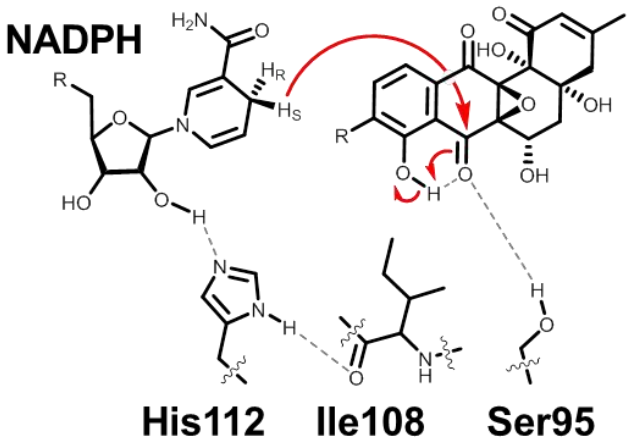
[1] Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, United Kingdom

[2] Department of Biological Chemistry, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, United Kingdom

http://dx.doi.org/ 10.1016/j.chembiol.2016.07.018

Published in Cell Chemical Biology (CCB), Volume 23, Issue 9, 22nd September 2016, Pages 1019-1097



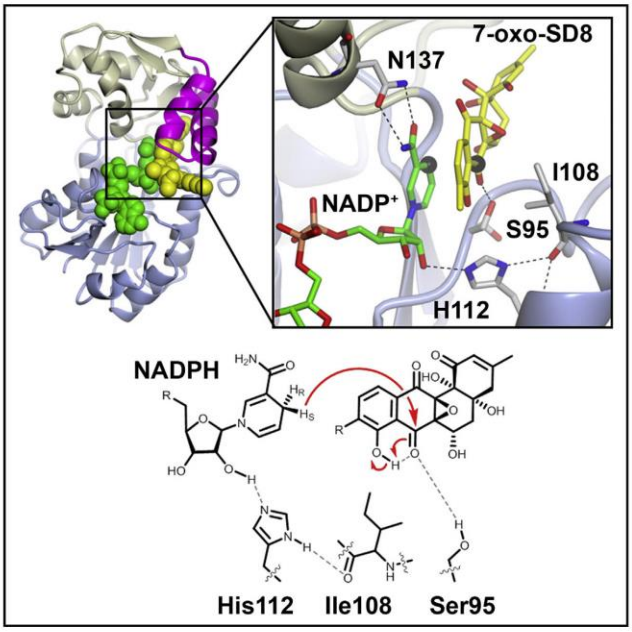


Brief Communication

Cell Chemical Biology

Substrate-Assisted Catalysis in Polyketide Reduction Proceeds via a Phenolate Intermediate

Graphical Abstract



Authors

Martin Schäfer, Clare E.M. Stevenson, Barrie Wilkinson, David M. Lawson, Mark J. Buttner

Correspondence

mark.buttner@jic.ac.uk

In Brief

SimC7 is a polyketide ketoreductase involved in the biosynthesis of the angucyclinone moiety of the DNA gyrase inhibitor simocyclinone. It is structurally distinct from previously characterized ketoreductases, lacking the canonical Ser-Tyr-Lys catalytic triad. Instead, SimC7 catalyzes a substrate-assisted, two-step reaction involving an unusual phenolate intermediate.

Highlights

- SimC7 catalyzes reduction of the C-7 carbonyl of the angucyclinone of simocyclinone
- SimC7 lacks the conserved catalytic triad characteristic of canonical ketoreductases
- SimC7 reduces the C-7 carbonyl via a substrate-assisted, two-step reaction
- This involves intramolecular transfer of a substrate proton to generate a phenolate

Accession Numbers

5L40
5L45
5L3Z
5L4L

Schäfer et al., 2016, Cell Chemical Biology *23*, 1091–1097
September 22, 2016
http://dx.doi.org/10.1016/j.chembiol.2016.07.018

CellPress

Please cite this article as: Schäfer et al., Substrate-Assisted Catalysis in Polyketide Reduction Proceeds via a Phenolate Intermediate, Cell Chemical Biology (2016), http://dx.doi.org/10.1016/j.chembiol.2016.07.018

Cell Chemical Biology
Brief Communication

Substrate-Assisted Catalysis in Polyketide Reduction Proceeds via a Phenolate Intermediate

Martin Schäfer,[1] Clare E.M. Stevenson,[2] Barrie Wilkinson,[1] David M. Lawson,[2] and Mark J. Buttner[1,3,*]
[1]Department of Molecular Microbiology
[2]Department of Biological Chemistry
John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK
[3]Lead Contact
*Correspondence: mark.buttner@jic.ac.uk
http://dx.doi.org/10.1016/j.chembiol.2016.07.018

SUMMARY

SimC7 is a polyketide ketoreductase involved in biosynthesis of the angucyclinone moiety of the gyrase inhibitor simocyclinone D8 (SD8). SimC7, which belongs to the short-chain dehydrogenase/reductase (SDR) superfamily, catalyzes reduction of the C-7 carbonyl of the angucyclinone, and the resulting hydroxyl is essential for antibiotic activity. SimC7 shares little sequence similarity with characterized ketoreductases, suggesting it might have a distinct mechanism. To investigate this possibility, we determined the structures of SimC7 alone, with $NADP^+$, and with $NADP^+$ and the substrate 7-oxo-SD8. These structures show that SimC7 is distinct from previously characterized polyketide ketoreductases, lacking the conserved catalytic triad, including the active-site tyrosine that acts as central acid-base catalyst in canonical SDR proteins. Taken together with functional analyses of active-site mutants, our data suggest that SimC7 catalyzes a substrate-assisted, two-step reaction for reduction of the C-7 carbonyl group involving intramolecular transfer of a substrate-derived proton to generate a phenolate intermediate.

INTRODUCTION

Angucyclin(on)es form the largest group of polycyclic aromatic polyketides, many with anticancer and antibacterial activities (Kharel et al., 2012). They share a polyketide-derived tetracyclic benz[a]anthracene carbon skeleton, but numerous structures are generated by a range of tailoring reactions, including O- or C-linked deoxysugar glycosylation to form angucyclines. In general, these tailoring enzymes are not well understood. An important challenge, therefore, is to define the step catalyzed by each enzyme and to determine their reaction mechanisms. This knowledge is particularly relevant to the rational engineering of angucyclin(on)e biosynthetic pathways for novel therapeutics.

Simocyclinone D8 (SD8) is a potent DNA gyrase inhibitor isolated from *Streptomyces antibioticus* that consists of an angucyclinone connected to a chlorinated aminocoumarin via a D-olivose deoxysugar and a tetraene diester linker (Figure 1) (Schimana et al., 2000; Edwards et al., 2009; Hearnshaw et al., 2014). SD8 is bifunctional, with the angucyclinone and the aminocoumarin at opposite ends of the molecule binding to two distinct pockets on the DNA binding surface of the GyrA subunit of gyrase (Edwards et al., 2009; Hearnshaw et al., 2014), thereby inhibiting DNA supercoiling at submicromolar concentrations (Edwards et al., 2011). Because gyrase is essential in bacteria but absent from humans, it is an attractive target for antimicrobial drugs, as exemplified by the clinically successful fluoroquinolones (Collin et al., 2011).

SimC7 was originally annotated as a dehydratase and predicted to be involved in the biosynthesis of the tetraene linker of SD8 (Trefzer et al., 2002). However, we recently showed that SimC7 is in fact an NAD(P)H-dependent ketoreductase that catalyzes the reduction of a carbonyl to a hydroxyl group at the C-7 position of the angucyclinone (Schäfer et al., 2015). This enzymatic step is essential for antibiotic activity, converting the almost inactive 7-oxo-simocyclinone D8 (7-oxo-SD8; half maximal inhibitory concentration [IC_{50}] ~50–100 μM) into the potent gyrase inhibitor SD8 (IC_{50} ~0.1–0.6 μM) (Schäfer et al., 2015).

Based on the intermediates produced by *S. antibioticus*, it seems that the biosynthesis of SD8 starts with assembly of the angucyclinone, followed by the attachment of the deoxysugar, then the tetraene linker, and finally the aminocoumarin (i.e., SD8 is assembled from right to left in Figure 1) (Schimana et al., 2001). Therefore, the physiological substrate of SimC7 is most likely a 7-oxo angucyclinone intermediate lacking the attached deoxysugar, tetraene linker, and aminocoumarin, an intermediate that is detectable only in *ΔsimC7* mutants (Schäfer et al., 2015). Despite this, the enzyme readily accepts as a substrate the full-length intermediate 7-oxo-SD8, the major product made by *ΔsimC7* mutants (Schäfer et al., 2015).

The angucyclinone moiety of SD8 is synthesized by a type II polyketide synthase (SimA1-3) and multiple tailoring enzymes (SimA4-13, SimC7) that catalyze cyclization, aromatization, oxidation, and reduction reactions. Several ketoreductases of the short-chain dehydrogenase/reductase (SDR) family that act on angucyclinones or related polyketides have been characterized. The reduction of carbonyl groups at the C-6 and C-9 positions of polyketides has been functionally characterized, and the structures of the corresponding SDR enzymes have elucidated their reaction mechanisms and factors determining their stereoselectivity. The ketoreductases LanV and UrdMred act on the C-6 carbonyl group of angucyclic polyketides from the landomycin and urdamycin pathways (Paananen et al., 2013; Patrikainen

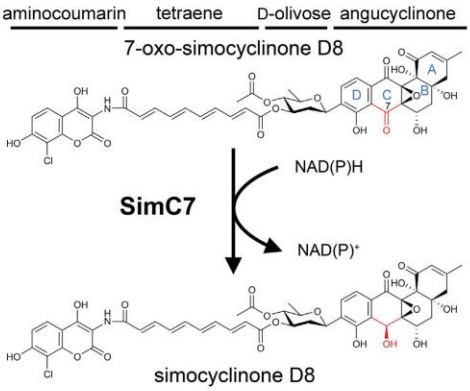


Figure 1. SimC7 Catalyzes the Reduction of 7-oxo-SD8 to Simocyclinone D8

A, B, C, and D denote the four rings of the angucyclinone moiety; the C-7 carbonyl/hydroxyl is highlighted in red.

et al., 2014). In contrast, the ketoreductases ActKR and HedKR act on the C-9 carbonyl group of early intermediates in the actinorhodin and hedamycin polyketide pathways (Javidpour et al., 2011a, 2011b, 2013; Korman et al., 2004, 2008). The LanV, UrdMred, ActKR, and HedKR structures revealed the catalytic Ser-Tyr-Lys triad characteristic of SDR enzymes, where the latter two residues form a YxxxK motif. In these classical SDR proteins, the conserved active-site tyrosine serves as central acid-base catalyst that donates a proton to the substrate. The adjacent lysine residue lowers the pK_a of the tyrosine hydroxyl group and often contributes directly to a proton relay mechanism, and the hydroxyl group of the serine stabilizes and polarizes the carbonyl group of the substrate (Kavanagh et al., 2008).

At the sequence level, SimC7 shares little similarity with any characterized ketoreductase, even with functionally analogous polyketide ketoreductases. The striking differences between the amino acid sequence of SimC7 and those of HedKR, ActKR, LanV, and UrdMred suggested that SimC7 might have a novel catalytic mechanism. To investigate this possibility, we determined the structures of SimC7 alone (apo; 1.6 Å resolution), the binary complex with $NADP^+$ (1.95 Å), and the ternary complex with both $NADP^+$ and 7-oxo-SD8 (1.2 Å) (Tables S1 and S2). Our results reveal that SimC7 is structurally distinct from previously characterized polyketide ketoreductases and, importantly, lacks the canonical SDR Ser-Tyr-Lys catalytic triad (Kavanagh et al., 2008; Kallberg et al., 2010; Persson and Kallberg, 2013). Instead, our data suggest that SimC7 catalyzes a substrate-assisted, two-step reaction for the reduction of the C-7 carbonyl group involving an unusual phenolate intermediate.

RESULTS AND DISCUSSION

Overall Structure of SimC7

SimC7 is made up of two domains, the larger of which is the nucleotide binding domain that adopts a Rossmann fold (Figures 2A, 2B, and S1); the smaller substrate binding domain, characteristic of the so-called extended SDR subfamily (Kavanagh et al., 2008), is mainly α-helical and is largely formed by two insertions in the nucleotide binding domain (between β6 and α6, and between β9 and α10). Notably, the latter insertion contains a "lid" motif consisting of two antiparallel α helices (α8 and α9) that folds over the active site (Figures 2A, 2B, and S1). The substrate binding domain is completed by a short helical segment at the C-terminus of the polypeptide chain.

Overall the apo, binary, and ternary SimC7 structures are very similar (Table S3), with the notable exception of the lid motif, which displays a number of different conformations (Figure 2C). Although the changes are not large (maximum Cα-Cα shift 5.35 Å; Table S3), there is a clear closure of the lid over the bound substrate (Figure 2C), suggesting a role in gating access to the active site and/or substrate capture. Moreover, in the ternary complex the underside of the lid contributes to the tight, highly hydrophobic substrate binding pocket (Figure 3) that provides the necessary environment for catalysis (see below).

Structural Homologs of SimC7

Structures annotated as SDR proteins (PFAM family PF00106) are prevalent in the PDB, with more than 600 entries. To look for structural homologs of SimC7, we carried out a structure-based similarity search using the DALI server (Table S4). Strikingly, characterized angucyclinone ketoreductases ranked very low in the search, the closest match being LanV (Figure S2A), which was the 166th ranked hit after filtering for sequence redundancy (Table S4). Instead, the two most structurally similar proteins to SimC7 were quinone oxidoreductase (QOR2) from *Escherichia coli* (PDB: 2ZCV) (Kim et al., 2008a) (Figure S2A) and triphenylmethane reductase (TMR) from *Citrobacter* sp. KCTC 18061P (PDB: 2VRB) (Kim et al., 2008b). QOR2 and TMR share with SimC7 the ability to reduce substrates with extensively conjugated pi systems but have roles in the detoxification of xenobiotics rather than the biosynthesis of natural products. The majority of the closest structural homologs of known function are involved in sugar biosynthesis, many of them epimerases.

SimC7 and its closest structural homologs all fall into the extended SDR subfamily of proteins, characterized by having two distinct domains that together form a partially occluded active-site pocket at their junction (Figure S2A). In contrast, the other structurally characterized polyketide ketoreductases such as LanV are more distantly related and belong to the classical SDR subfamily (Kavanagh et al., 2008), in which three insertions within the core Rossmann fold motif delineate a more accessible active-site cavity but do not constitute a well-defined substrate binding domain (Figure S2A).

Substrate Binding

In the ternary complex with substrate, determined at 1.2-Å resolution, the angucyclic ring system of 7-oxo-SD8 binds adjacent and parallel to the nicotinamide ring of the cofactor (Figures 2D, 2E, and S3), where it adopts a relatively planar conformation differing from the conformations seen in the DNA gyrase-SD8 and SimR-SD8 complexes, in which the A ring of the angucyclinone in SD8 is tilted upward toward the epoxide (Hearnshaw et al., 2014; Le et al., 2011) (Figure S2B). This planar conformation is most likely enforced by the shape of the very constricted

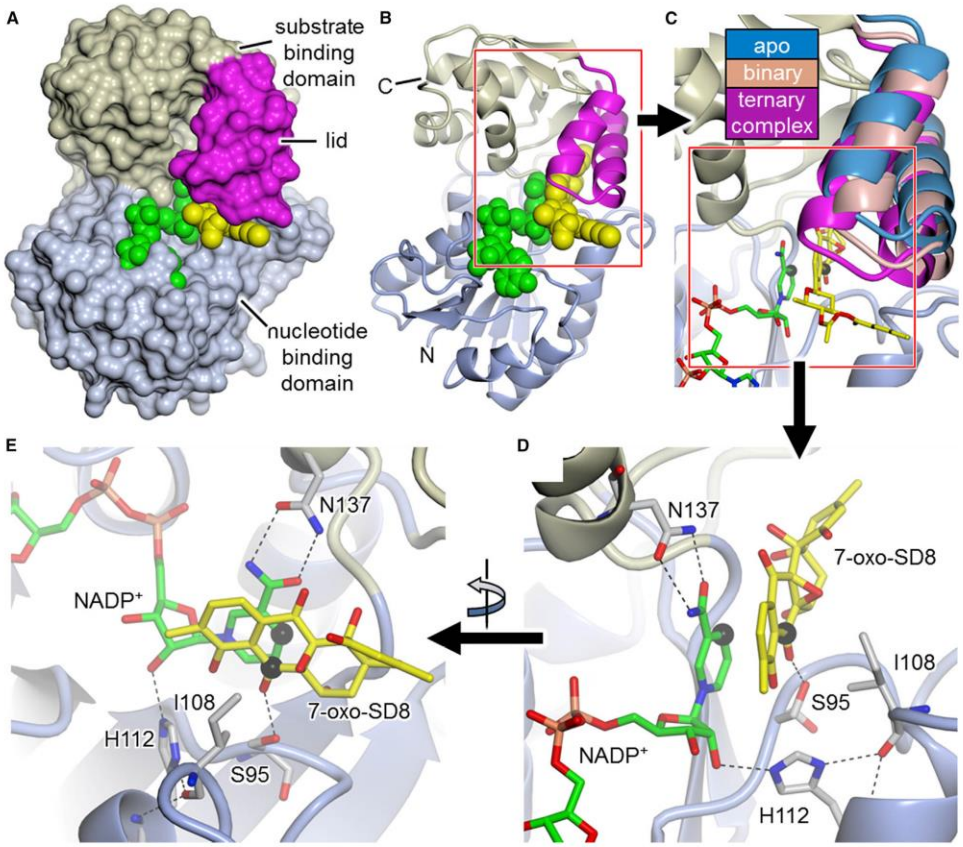


Figure 2. Crystal Structure of SimC7

(A and B) SimC7 displayed as (A) a molecular surface and (B) cartoon representation with the nucleotide binding domain, the substrate binding domain, and the lid motif shown in pale blue, beige, and magenta, respectively. Ligands are represented as van der Waals spheres with the 7-oxo-SD8 substrate shown in yellow and the cofactor shown in green.

(C) Close-up showing conformational changes in the lid between the apo (chain A, form 2; blue), the binary complex (salmon), and the ternary complex (magenta); the core protein structure and ligands shown represent the ternary complex. C-4 of the cofactor and C-7 of the substrate are highlighted by black spheres, showing that C-7 of the substrate is exactly positioned 3 Å from C-4 of the nicotinamide ring, poised for direct hydride transfer.

(D and E) Orthogonal close-ups showing the active site of the ternary complex including the Ser95-Ile108-His112 "catalytic triad" residues, and Asn137, which is important in maintaining the *syn* conformation of the cofactor. For clarity, only the angucyclic polyketide moiety of the substrate is shown. Hydrogen bonds are shown as dashed lines. (E) is also reproduced as a stereo image in Figures S3A and S3B, which also show a difference electron density map calculated from the final model after simulated annealing refinement with the substrate omitted.

See also Figures S1–S4 and Tables S1–S3.

and highly hydrophobic substrate pocket (Figure 3). Within this hydrophobic pocket, 7-oxo-SD8 is bound only by a single direct hydrogen bond between the side chain of Ser95 and the C-7 carbonyl oxygen of the angucyclinone moiety, which may help to position the latter exactly above the C-4 position of the nicotinamide ring, where it is poised for direct hydride transfer (highlighted by black spheres in Figures 2C–2E, 3B, and S3). As mentioned above, the natural substrate for SimC7 is likely to be a 7-oxo angucyclinone intermediate lacking the deoxysugar, the tetraene linker, and the aminocoumarin. Consistent with this, only the angucyclinone moiety is buried in the active site of SimC7. Roughly half of the tetraene linker is visible in the electron

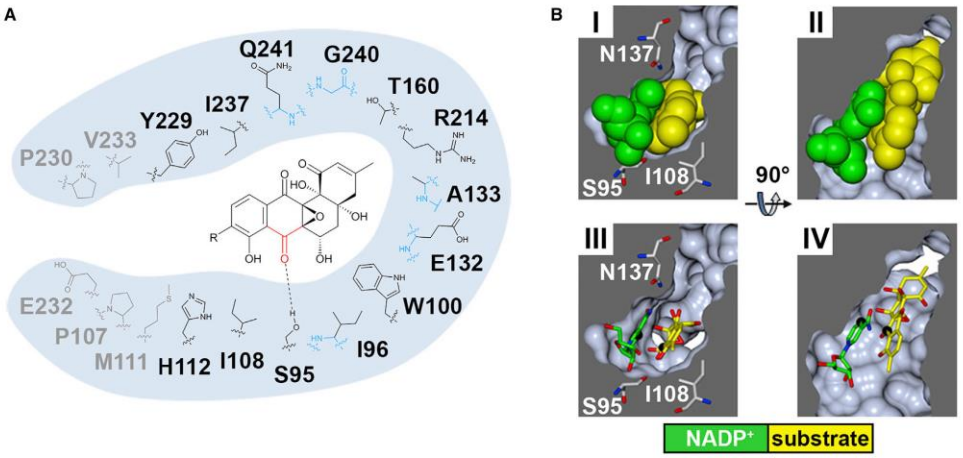


Figure 3. SimC7 Has a Very Hydrophobic and Constricted Substrate Binding Pocket

(A) The substrate is bound by only one direct hydrogen bond between the C-7 carbonyl group (red) in the angucyclinone moiety and the side-chain hydroxyl of Ser95. This interaction may assist in positioning the substrate and facilitating the reaction. Interacting residues are shown in black (side-chain interactions) or blue (backbone interactions). The hydrophobic residues shown in gray line the entrance to the substrate pocket but do not interact directly with the bound angucyclinone. Note that one face of the pocket is formed by the cofactor itself (not shown). In the natural SimC7 substrate R=H, in the substrate used here R includes the deoxysugar, tetraene linker, and the aminocoumarin.

(B) Orthogonal cross-sections through the active-site pocket, revealing how tightly cofactor (green) and substrate (yellow) are bound. For clarity, only the nicotinamide ribosyl moiety of the cofactor and the polyketide moiety of the substrate are shown. In (I) and (III) the view corresponds roughly to that shown in Figure 2D, whereas (II) and (IV) show the view from above relative to Figure 2D. In the lower panels, C-4 of the cofactor and C-7 of the substrate are highlighted by black spheres, showing that C-7 of the substrate is exactly positioned 3 Å from C-4 of the nicotinamide ring, poised for hydride transfer.

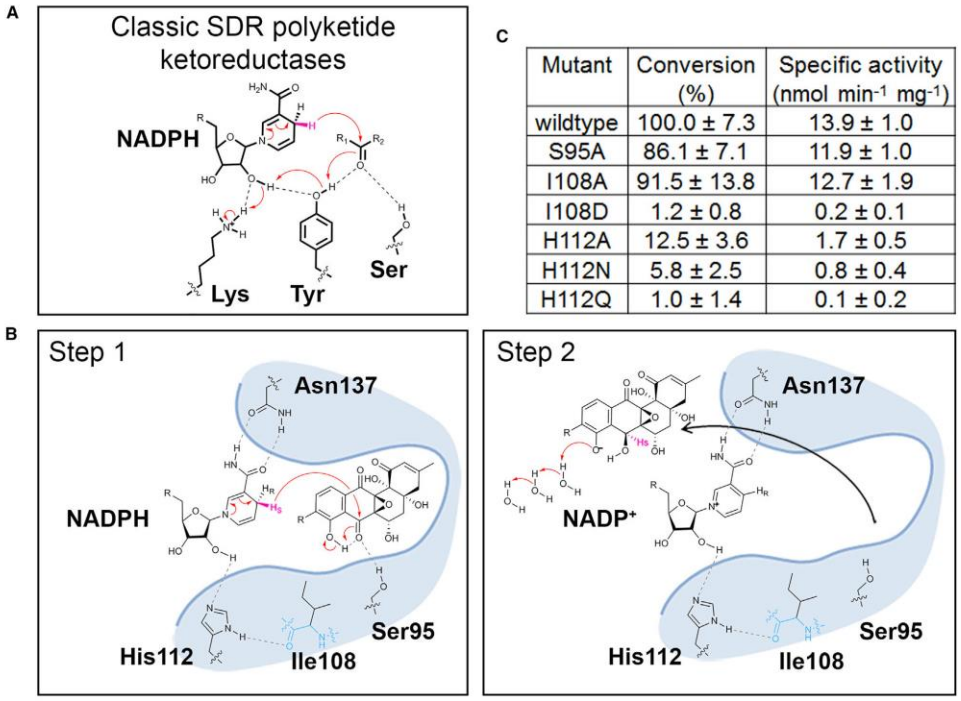
See also Figures S1–S3.

density, projecting away from the protein surface, and the aminocoumarin ring is not resolved at all (Figure S3).

A Novel Catalytic Mechanism for a Polyketide Ketoreductase

SimC7 lacks the Ser-Tyr-Lys catalytic triad characteristic of canonical SDR proteins (Figure 4 and Table S4). While the serine is conserved (Ser95), the other two residues (i.e., the YxxxK motif) are not, being instead replaced by Ile108 and His112, respectively. This Ser-Ile-His triad is unlike any described for the five subfamilies of SDRs defined by Kavanagh et al. (2008). Particularly surprising is the absence of the tyrosine residue that acts as the acid-base catalyst in the classical SDR mechanism (Figure 4A). Inspection of the structure of the ternary complex shows that none of the five tyrosine residues in SimC7 is sufficiently close to C-7 of the angucyclinone ring system of the substrate to play a direct role in catalysis. Furthermore, the structure also shows that there is no alternative residue that could act as an acid-base catalyst. Consequently, SimC7 must perform ketoreduction of 7-oxo-SD8 via a novel mechanism.

Based on the structure of the ternary complex of the enzyme with $NADP^+$ and 7-oxo-SD8, we propose a simple two-step mechanism for SimC7 that does not depend on catalytic residues in the protein, but rather takes advantage of the specific properties of the substrate itself, and is thus a novel example of substrate-assisted catalysis (Dall'Acqua and Carter, 2000). In the first step, the hydrophobic environment of the substrate binding pocket and the juxtaposition of the quinone-like C ring and the phenyl-like D ring of the angucyclinone favor the formation of an intramolecular hydrogen bond between the proton on the C-8 hydroxyl group and the oxygen of the neighboring C-7 carbonyl group (Figure 4B). This enhances the polarization of the latter such that the electrophilicity of C-7 is increased, making it a good acceptor for direct hydride transfer from the 4-*pro-S* position of the nicotinamide ring. Crucially, the hydride donor and acceptor carbon atoms are only 3.0 Å apart in the crystal structure. The C-7 hydroxyl group is then formed by internal proton transfer from the neighboring C-8 hydroxyl group, generating a phenolate intermediate in which the negative charge on the C-8 oxygen atom is stabilized by the aromatic D ring. In the second step of the reaction, the phenolate intermediate leaves the substrate binding pocket and the proton required to reinstate the C-8 hydroxyl group is recovered by abstraction from bulk water, which is not possible within the confines of the active site (Figure 4B). Exchange of the negatively charged reaction intermediate is most likely accelerated by repulsion from the hydrophobic active-site cavity. Finally, the direct hydride attack from below the angucyclic polyketide unambiguously explains the 7*S* stereochemistry of simocyclinones. In support of this proposed mechanism, molecular modeling predicts the existence of the key intramolecular hydrogen bond between the C-8 hydroxyl group and the C-7 carbonyl group of 7-oxo-SD8, and an



C

Mutant	Conversion (%)	Specific activity (nmol min^{-1} mg^{-1})
wildtype	100.0 ± 7.3	13.9 ± 1.0
S95A	86.1 ± 7.1	11.9 ± 1.0
I108A	91.5 ± 13.8	12.7 ± 1.9
I108D	1.2 ± 0.8	0.2 ± 0.1
H112A	12.5 ± 3.6	1.7 ± 0.5
H112N	5.8 ± 2.5	0.8 ± 0.4
H112Q	1.0 ± 1.4	0.1 ± 0.2

Figure 4. Canonical SDR Ketoreduction and the Novel SimC7 Reaction Mechanism

(A) In canonical SDR proteins the conserved active-site tyrosine serves as central acid-base catalyst that donates a proton to the substrate. The adjacent lysine residue lowers the pK_a of the tyrosine hydroxyl group and often contributes directly to the proton relay mechanism; the hydroxyl group of the serine stabilizes and polarizes the carbonyl group of the substrate.

(B) SimC7 has an atypical catalytic triad comprised of Ser95, Ile108, and His112. In the first step of the proposed SimC7 mechanism, the C-7 carbonyl group of the substrate (7-oxo-SD8) is reduced by transfer of the 4-pro-*S* hydride of the cofactor onto the C-7 carbon of the substrate. This transfer from below the C ring results in the characteristic 7-*S* stereochemistry of SD8. Ketoreduction at position C-7 is completed by an intramolecular proton transfer from the neighboring C-8 hydroxyl group of the angucyclinone; the resultant negative charge on the latter is stabilized by the adjacent aromatic ring system (ring D in Figure 1). In the second step, the C-8 phenolate intermediate regains a proton from bulk water after leaving the substrate binding pocket. In the natural SimC7 substrate R=H, in the substrate used here R includes the deoxysugar, tetraene linker, and the aminocoumarin. Note that there are no water molecules in the active-site pocket that could contribute to the reaction mechanism. In the ternary complex, the nearest water to O-7 of the angucyclic polyketide is ~5.5 Å away, and the nearest water to O-8 is ~4.9 Å away. Due to steric constraints within the pocket, neither could approach the substrate oxygen atoms without either a repositioning of the substrate or a conformational change in the protein.

(C) Enzymatic activities of SimC7 active-site mutants. Standard errors are indicated for three independent experiments.

See also Figures S3 and S4; Table S4.

equivalent intramolecular hydrogen bond is observed in the small-molecule crystal structure of panglimycin, a closely related polyketide (Fotso et al., 2008). Furthermore, molecular modeling also predicts that the C-8 hydroxyl group will have the most acidic and exchangeable proton in the angucyclic polyketide, with an estimated pK_a of 6.9–7.7, and therefore could readily transfer to the neighboring C-7 oxygen at the end of step 1. Indeed, there is no other possible proton donor (neither protein nor water derived) sufficiently close to O-7 to perform this role.

Based on their structures (Figure S4A), there are four other angucyclinones in which a SimC7-like mechanism might generate a C-7 hydroxyl group: panglimycin, elmycin, grisemycin, and kiamycin (Fotso et al., 2008; Xie et al., 2012, 2016). However, the biosynthetic gene clusters for these molecules have yet to be reported, and it is not known whether a SimC7-like enzyme is involved in their synthesis. Most other angucyclin(on)es have a carbonyl group at C-7.

Mutagenesis of the SimC7 Active Site

To investigate the potential roles of the SimC7 "catalytic triad" residues in the proposed reaction mechanism, we mutagenized Ser95, Ile108, and His112 (Figure 4B). In the wild-type enzyme,

the hydroxyl group of Ser95 could aid catalysis by helping to bind and correctly orient the substrate, and by providing additional polarization to the C-7 carbonyl group via a hydrogen bond, the latter role being consistent with the function proposed for the structurally equivalent Ser/Thr residues in the classical SDR mechanism (Figure 4A) (Kavanagh et al., 2008; Kallberg et al., 2010; Persson and Kallberg, 2013). However, neither effect would appear to be crucial for activity, as a S95A mutant showed 86% substrate conversion relative to the wild-type (Figure 4C). Therefore, SimC7 must be able to orient 7-oxo-SD8 for catalysis without the necessity for the hydrogen bond from Ser95, and this likely arises because the active-site cavity of the binary complex closely matches the shape of the substrate (Figure 3B).

Ile108 contributes to the hydrophobic surface that forms one side of the active-site cavity and has a nonspecific role in helping to trap the angucyclinone group against the nicotinamide ring of the cofactor (Figure 3B). Unsurprisingly, an I108A substitution had almost no effect, whereas changing it to aspartate (I108D) abolished activity (Figure 4C). In the latter case, the introduction of a negative charge would interfere with the hydrophobic environment and could disrupt the intramolecular hydrogen bond in the substrate, both being necessary for catalysis.

Finally, mutation of His112 (H112A, H112N, and H112Q) strongly reduced or abolished enzyme activity (Figure 4C), consistent with its key role in binding and positioning the cofactor via a hydrogen bond to the 2′-hydroxyl of the nicotinamide ribosyl moiety (Figures 2D, 2E, and 4B).

SIGNIFICANCE

The sequence of SimC7 is distinct from previously characterized polyketide ketoreductases, and the structural data reported here suggest that it catalyzes a novel substrate-assisted, two-step reaction for the reduction of the C-7 carbonyl group. This mechanism involves the intramolecular transfer of a substrate-derived proton to generate a phenolate intermediate, negating the need for proton transfer from a canonical SDR active-site tyrosine. Like SimC7 (Ser-Ile-His), the two closest structural homologs, TMR (Tyr-Leu-His) and QOR2 (Leu-Leu-His), also have unusual active-site triads. Thus SimC7, TMR, and QOR2 share an I/LxxxH motif and have substrates with extensively conjugated pi systems. No enzyme-substrate complex crystal structures have been described for TMR or QOR2 and no firm proposals exist for their catalytic mechanisms, but their substrate structures and the data provided here suggest they are also likely to employ noncanonical mechanisms. Our data therefore point to members of the extended SDR sub-family having I/LxxxH active site motifs as a source of new biochemistry.

EXPERIMENTAL PROCEDURES

For a full explanation of the experimental protocols, see Supplemental Experimental Procedures.

Protein Overexpression and Purification

Point mutants of *simC7* were generated by PCR-based site-directed mutagenesis. All constructs were verified by sequencing. Proteins were expressed in *E. coli* as N-terminally His-tagged fusions, purified by nickel-affinity chromatography and assayed by high-performance liquid chromatography as described previously (Schäfer et al., 2015). The structural integrity of purified proteins was verified using circular dichroism.

Protein Crystallization and Structure Determination

SimC7 was labeled with selenomethionine (SeMet) by metabolic inhibition, and crystals of both native and SeMet proteins were grown by vapor diffusion (crystals of binary and ternary complexes were obtained by cocrystallization). Crystals were harvested and flash-cooled in liquid nitrogen. The native crystals did not require further cryoprotection, while the SeMet-labeled crystals were cryoprotected by supplementing the crystallization solution with 25% (v/v) glycerol. All X-ray data were collected at the Diamond Light Source. For de novo structure determination, a single-wavelength anomalous dispersion dataset was collected at the Se *K* X-ray absorption edge for an SeMet-labeled SimC7 crystal. Native and SeMet data were combined to solve the structure of SimC7 in complex with $NADP^+$. All other structures were determined by molecular replacement using the latter as a template.

ACCESSION NUMBERS

The accession numbers for the SimC7 structures reported in this paper are PDB: 5L40 (apo form 1), 5L45 (apo form 2), 5L3Z (binary complex), and 5L4L (ternary complex).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.chembiol.2016.07.018.

AUTHOR CONTRIBUTIONS

All authors designed the experiments. M.S. and C.E.M.S. performed the experiments. All authors analyzed the data. M.S., D.M.L., and M.J.B. wrote the paper, and all authors made revisions.

ACKNOWLEDGMENTS

We thank Hans-Peter Fiedler for providing simocyclinone D8, and the staff of the Diamond Light Source (Oxford, UK) for access to beamlines I03, I04, and I04-1 under proposal MX9475. This work was funded by a BBSRC studentship to M.S., by BBSRC grant BB/I002197/1 to M.J.B. and D.M.L., and by BBSRC Grant BB/J004561/1 to the John Innes Centre.

Received: June 9, 2016
Revised: July 17, 2016
Accepted: July 26, 2016
Published: September 8, 2016

REFERENCES

Collin, F., Karkare, S., and Maxwell, A. (2011). Exploiting bacterial DNA gyrase as a drug target: current state and perspectives. Appl. Microbiol. Biotechnol. *92*, 479–497.

Dall'Acqua, W., and Carter, P. (2000). Substrate-assisted catalysis: molecular basis and biological significance. Protein Sci. *9*, 1–9.

Edwards, M.J., Flatman, R.H., Mitchenall, L.A., Stevenson, C.E.M., Le, T.B.K., Clarke, T.A., McKay, A.R., Fiedler, H.-P., Buttner, M.J., Lawson, D.M., and Maxwell, A. (2009). A crystal structure of the bifunctional antibiotic, simocyclinone D8, bound to DNA gyrase. Science *326*, 1415–1418.

Edwards, M.J., Williams, M.A., Maxwell, A., and McKay, A.R. (2011). Mass spectrometry reveals that the antibiotic simocyclinone D8 binds to DNA gyrase in a "bent-over" conformation: evidence of positive cooperativity in binding. Biochemistry *50*, 3432–3440.

Fotso, S., Mahmud, T., Zabriskie, T.M., Santosa, D.A., Sulastri, and Proteau, P.J. (2008). Angucyclinones from an Indonesian *Streptomyces* sp. J. Nat. Prod. *71*, 61–65.

Hearnshaw, S.J., Edwards, M.J., Stevenson, C.E., Lawson, D.M., and Maxwell, A. (2014). A new crystal structure of the bifunctional antibiotic simocyclinone D8 bound to DNA gyrase gives fresh insight into the mechanism of inhibition. J. Mol. Biol. *426*, 2023–2033.

Javidpour, P., Das, A., Khosla, C., and Tsai, S.C. (2011a). Structural and biochemical studies of the hedamycin type II polyketide ketoreductase (HedKR): molecular basis of stereo- and regiospecificities. Biochemistry *50*, 7426–7439.

Javidpour, P., Korman, T.P., Shakya, G., and Tsai, S.C. (2011b). Structural and biochemical analyses of regio- and stereospecificities observed in a type II polyketide ketoreductase. Biochemistry *50*, 4638–4649.

Javidpour, P., Bruegger, J., Srithahan, S., Korman, T.P., Crump, M.P., Crosby, J., Burkart, M.D., and Tsai, S.C. (2013). The determinants of activity and specificity in actinorhodin type II polyketide ketoreductase. Chem. Biol. *20*, 1225–1234.

Kallberg, Y., Oppermann, U., and Persson, B. (2010). Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. FEBS J. *277*, 2375–2386.

Kavanagh, K.L., Jörnvall, H., Persson, B., and Oppermann, U. (2008). The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. Cell. Mol. Life Sci. *65*, 3895–3906.

Kharel, M.K., Pahari, P., Shepherd, M.D., Tibrewal, N., Nybo, S.E., Shaaban, K.A., and Rohr, J. (2012). Angucyclines: biosynthesis, mode-of-action, new natural products, and synthesis. Nat. Prod. Rep. *29*, 264–325.

Kim, I.K., Yim, H.S., Kim, M.K., Kim, D.W., Kim, Y.M., Cha, S.S., and Kang, S.O. (2008a). Crystal structure of a new type of NADPH-dependent quinone oxidoreductase (QOR2) from *Escherichia coli*. J. Mol. Biol. *379*, 372–384.

Kim, M.H., Kim, Y., Park, H.J., Lee, J.S., Kwak, S.N., Jung, W.H., Lee, S.G., Kim, D., Lee, Y.C., and Oh, T.K. (2008b). Structural insight into bioremediation of triphenylmethane dyes by *Citrobacter* sp. triphenylmethane reductase. J. Biol. Chem. *283*, 31981–31990.

Korman, T.P., Hill, J.A., Vu, T.N., and Tsai, S.C. (2004). Structural analysis of actinorhodin polyketide ketoreductase: cofactor binding and substrate specificity. Biochemistry *43*, 14529–14538.

Korman, T.P., Tan, Y.H., Wong, J., Luo, R., and Tsai, S.C. (2008). Inhibition kinetics and emodin cocrystal structure of a type II polyketide ketoreductase. Biochemistry *47*, 1837–1847.

Le, T.B.K., Stevenson, C.E.M., Fiedler, H.-P., Maxwell, A., Lawson, D.M., and Buttner, M.J. (2011). Structures of the TetR-like simocyclinone efflux pump repressor, SimR, and the mechanism of ligand-mediated derepression. J. Mol. Biol. *408*, 40–56.

Paananen, P., Patrikainen, P., Kallio, P., Mäntsälä, P., Niemi, J., Niiranen, L., and Metsä-Ketelä, M. (2013). Structural and functional analysis of angucycline C-6 ketoreductase LanV involved in landomycin biosynthesis. Biochemistry *52*, 5304–5314.

Patrikainen, P., Niiranen, L., Thapa, K., Paananen, P., Tähtinen, P., Mäntsälä, P., Niemi, J., and Metsä-Ketelä, M. (2014). Structure-based engineering of angucyclinone 6-ketoreductases. Chem. Biol. *21*, 1381–1391.

Persson, B., and Kallberg, Y. (2013). Classification and nomenclature of the superfamily of short-chain dehydrogenases/reductases (SDRs). Chem. Biol. Interact. *202*, 111–115.

Schäfer, M., Le, T.B.K., Hearnshaw, S.J., Maxwell, A., Challis, G.L., Wilkinson, B., and Buttner, M.J. (2015). SimC7 is a novel NAD(P)H-dependent ketoreductase essential for the antibiotic activity of the DNA gyrase inhibitor simocyclinone. J. Mol. Biol. *427*, 2192–2204.

Schimana, J., Fiedler, H.P., Groth, I., Sussmuth, R., Beil, W., Walker, M., and Zeeck, A. (2000). Simocyclinones, novel cytostatic angucyclinone antibiotics produced by *Streptomyces antibioticus* Tu 6040. I. Taxonomy, fermentation, isolation and biological activities. J. Antibiot. (Tokyo) *53*, 779–787.

Schimana, J., Walker, M., Zeeck, A., and Fiedler, H.-P. (2001). Simocyclinones: diversity of metabolites is dependent on fermentation conditions. J. Ind. Microbiol. Biotechnol. *27*, 144–148.

Trefzer, A., Pelzer, S., Schimana, J., Stockert, S., Bihlmaier, C., Fiedler, H.-P., Welzel, K., Vente, A., and Bechthold, A. (2002). Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic. Antimicrob. Agents Chemother. *46*, 1174–1182.

Xie, Z., Liu, B., Wang, H., Yang, S., Zhang, H., Wang, Y., Ji, N., Qin, S., and Laatsch, H. (2012). Kiamycin, a unique cytotoxic angucyclinone derivative from a marine *Streptomyces* sp. Mar. Drugs *10*, 551–558.

Xie, Z., Zhou, L., Guo, L., Yang, X., Qu, G., Wu, C., and Zhang, S. (2016). Grisemycin, a bridged angucyclinone with a methylsulfinyl moiety from a marine-derived *Streptomyces* sp. Org. Lett. *18*, 1402–1405.