Molecular Analysis of the Aureothin Biosynthesis Gene Cluster from *Streptomyces thioluteus* HKI-227; New Insights into Polyketide Assembly

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

vorgelegt dem Rat der Biologisch-Pharmazeutischen Fakultät der

Friedrich-Schiller-Universität Jena

von

Jing He

Geboren am 10.02.1977 in Wuhan, People's Republic of China

Jena, im November 2004

Gutachter

1. Prof. Dr. Susanne Grabley
2. Prof. Dr. Erika Kothe
3. Prof. Dr. Jörn Piel

Tag der Doktorprüfung: 12 January 2005

Tag der öffentlichen Verteidigung: 31 January 2005

Index

Abbreviation

6-dEB	6-deoxyerythronolide
aac(3)IV	apramycin resistance gene
act	actinorhodin
ACP	acyl carrier protein
Am	apramycin
Amp	ampicillin
ARO	aromatase
AT	acyltransferase
ave	avermectin
bla	β-lactamase gene
bp	base pair
cml	chloramphenicol acetyltransferase, synonym for *cat*
CCC	covalently closed circular DNA
CHR	chalcone reductase
CHS	chalcone synthase
CIF	chain initiation factor
CLF	chain-length factor
cos	cohesive end
CYC	cyclase
dam	gene encoding DNA adenine methylase
dcm	gene encoding DNA cytosine methylase
DEBS	6-deoxyerythronolide B synthase
DH	dehydrase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
E	*Escherichia*
EDTA	ethylene diaminetetraacetic acid
*erm*E	erythromycin resistance gene
ery	erythromycin
ER	enoyl reductase
ESI-MS	electro spray ionisation mass spectrometry
FAS	fatty acid synthase
G+C%	percentage of G+C content
HPLC	high performance liquid chromatography
kb	kilobases (pairs)

KR	ketoreductase
KS	ketosynthase
LC-ESMS	liquid chromatography electro spray mass spectrometry
LC-MS	liquid chromatography mass spectrometry
MAT	malonyl-acetyl transferase
MM	minimal medium
NBT	nitro-blue tetrazolium
NRPS	nonribosomal peptide synthase
OD	optical density
ORF	open reading frame
ori	origin of replication
*ori*T	origin of transfer
*pab*AB	p-aminobenzoate biosynthesis gene
PCR	polymerase chain reaction
PEG	polyethylene glycol
PFGE	pulsed field gel electrophoresis
PKS	polyketide synthase
R	resistance
RAPS	rapamycin synthase
rec	gene encoding for recombinase
rep	replicon
rpm	rotations per minute
S	sensitivity
S	*Streptomyces*
SDS	sodium dodecyl sulfate
Spc	spectinomycin
SSC	sodium chloride and sodium citrate solution
sti	stigmatellin
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA buffer
TE	thioesterase
TES	N-Tris(hydroxymethyl) methyl-2-amino ethane sulxonic acid
Thio	thiostrepton
Tris	Tris-(hydroxymethyl)-aminomethane
tsr	thiostrepton resistance gene

A. Introduction

1. Streptomycetes as Producers of Bioactive Secondary Metabolites

Streptomycetes are the most widely studied and well known genus of the actinomycete family. They are ubiquitous in nature and are largely responsible, through the secretion of chemicals called geosmins, for the earthy smell of soil. In general, the genus *Streptomyces* is characterized as a kind of non-motile, filamentous, aerobic and Gram-positive bacteria (Waksman *et al*., 1943).

During the complex life cycle (Fig.1.1) of the streptomycetes from a spore to a colony, elaborate cell division emerges for adaptation to changes of environment (Hodgson *et al*., 1992 and Chater *et al*., 1996). After a suitable germination trigger, streptomycetes form branching filaments of cells, which become a network of strands called a mycelium, similar in appearance to the mycelium of some fungi. However, they are true bacteria - prokaryotic cells - unlike eukaryotic fungal cells. As colonies grow, the substrate hyphae become densely piled up and the growth of aerial hyphae is initiated by responses to nutrient limitation or other physiological stresses (Takano *et al*., 1994 and Chakraburtty *et al*., 1997), and/or to cell density via extracellular signals (Horinouchi *et al*., 1994 and Willey *et al*., 1993). Streptomycetes reproduce and disperse through the formation of spores, called conidia, from sporogenous hyphae, which follows the period of vegetative growth. The sporogenous hyphae form spores by simple cross-wall divisions in the top region of aerial filament. Noticeably, the creation of spores is distinct from the formation of bacterial endospores. The spores are in a semi-dormant stage, which can impart resistance to low nutrient and water availability and survive in soil for long periods (Mayfield *et al*., 1972 and Ensign, 1978).

The genome size of these high G+C content (61-80 mole %) organisms (Wright *et al*., 1992) was accurately estimated by analysis of macrorestriction fragments of the chromosome with pulsed-field gel electrophoresis (PFGE). It is composed of approximately 8,000 kb, twice the size of an *Escherichia coli* chromosome (Kieser *et al*., 1992). The chromosome of most streptomycetes is a linear structure (Lin *et al*., 1993 and Lezhava *et al*., 1995) and carries proteins covalently bound to both free 5' ends. Chromosome linearity probably contributes to the high toleration for large deletions and amplifications in the chromosome. The frequency of spontaneous mutagenesis in streptomycetes is as high as 0.1-1%. Mutagenesis occurs more frequently after treatment with mutagenic agents, or interfering with its replication by cold shock or protoplasting (Cullum *et al*., 1988 and Leblond *et al*., 1990). It is shown by a recent report that a spontaneous mutant of *Streptomyces ambofaciens*, which lost nearly quarter of the chromosome, has a genome size of only 6,500 kb (Leblond *et al*., 1991). Most

S*treptomyces* strains contain linear and/or circular plasmids, nearly all of which are self-transmissible fertility factors. These plasmids can transfer between different S*treptomyces* strain cells by conjugation with 100% efficiency and spread over a diameter of up to 2 mm to result in the formation of "pocks" (Rafii *et al*., 1988 and Wellington *et al*., 1988).

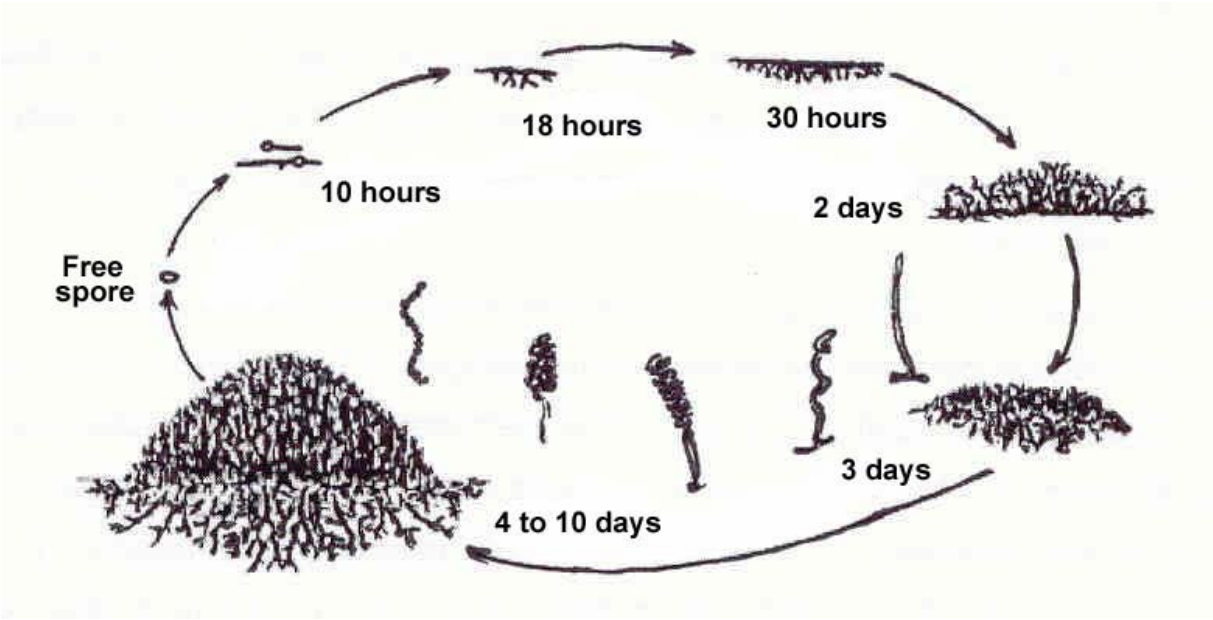


Figure 1.1. The life cycle of *Streptomyces* strains (Kieser *et al*., 2000)**.** Under favourable conditions, one or two germ tubes emerge from a spore and grow by tip extension and branch formation to give rise to substrate mycelium. After about 2-3 days, aerial hyphae grow up to form a spiral syncytium. When the grey spore pigment is deposited, cell wall thickening occurs to generate spores in the top region of aerial mycelium after 4-10 days.

Because streptomycetes inhabit soil, some of them are phytopathogens, known for attacking root vegetables, such as potatoes, beets, radishes, rutabaga, turnips, carrots, and parsnip (Lechevalier, 1988 and Kennedy *et al*., 1980). But comparison with other genus of pathogenic actinomycetes, such as *Mycobacterium*, *Actinomadura*, *Nocardia* and *Actinomyces*, streptomycetes are safe to animals and human beings except *S. somaliensis*, an established human pathogen. Streptomycetes can consume almost anything, including sugars, alcohols, amino acids, organic acids, and aromatic compounds by producing extracellular hydrolytic enzymes. Thus, there is considerable interest in these organisms as agents for bioremediation.

Streptomycetes are most widely known for their ability to synthesize a great number of antibiotics and other classes of biologically active secondary metabolites. It is believed that antibiotics help streptomycetes compete with other organisms in the relatively nutrient-depleted environment of the soil by reducing competition. Actinomycetes produce about two-third of the known microorganism-produced antibiotics and amongst them nearly 80% is

made by members of the genus *Streptomyces*. Many of the metabolites are applied in human and veterinary medicine and agriculture, such as anti-parasitic agents, herbicides, pharmacologically relevant compound (e.g. immunosuppressants), and several enzymes are important in food and other industries.

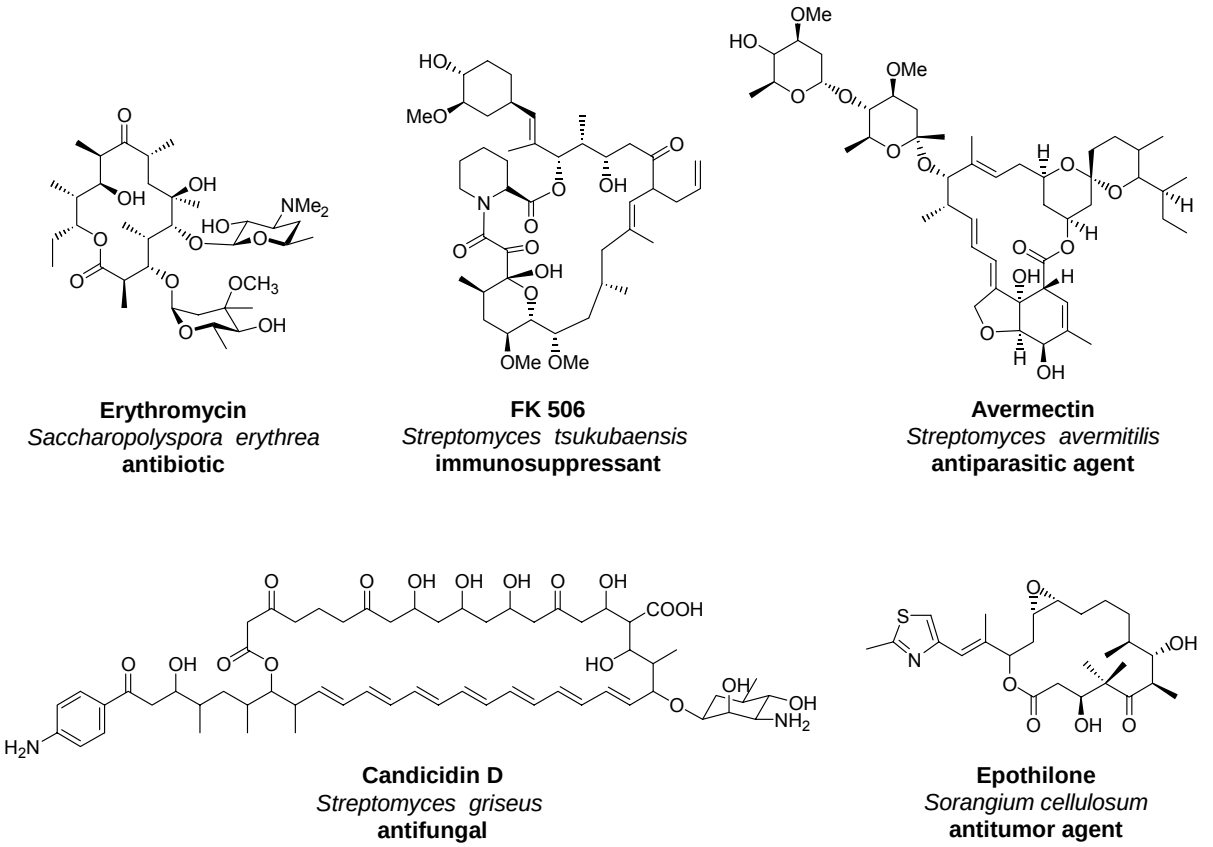
As multicellular differentiated prokaryotic organisms, streptomycetes become a good pattern for the study of complex regulation of gene expression in space and time, morphogenesis, metabolism and the flux of metabolites in developmental biology. And the extensive knowledge of *Streptomyces* genetics, molecular biology and physiology, will provide a valuable background to their use for industrial exploitation and clinic utilization as well as yielding a wealth of fundamental knowledge of prokaryotic genetics with wide implications.

2. Microbial Polyketide Biosynthesis

In order to defend their habitat, microorganisms, especially ones that inhabit soil and marine environments, produce a wide range of chemicals (Berdy, 1987) that are usually called “secondary metabolites” because of the dispensability for the organism’s ontogeny. Some of secondary metabolites were usually named as antibiotics because of their complex structures resulting in the unusually extensive use in pharmacology for treating infectious diseases. Polyketides are one of the most important chemical classes of antibiotics. In addition to possessing a wide range of functional and structural diversity, polyketides are famous for an extremely rich source of biologically active compounds, including antibacterial (erythromycin), antifungal (candicidin), antiparasitic (avermectin), antitumor (epothilone and doxorubicin), immunosuppressive (FK 506) and other biological activities (Walsh, 2003) (Fig.1.2).

2.1 Genetic Contributions to Understanding Polyketide Biosynthesis

In the last 25 years, there has been a revolution in the understanding of the processes responsible for the maintenance, transmission and expression of genetic information at the molecular level. After the discovery of genetic materials, DNA and RNA, the development of genetics goes very fast and its explosion of knowledge also advances many other studies, such as molecular biology, biochemistry and ecology. Before the 1950s, only a small part of antibiotics was isolated. Since the improvement of the methods for isolation and identification of chemical compounds, the discovery of news substances potential for drugs has entered a new stage. Approximately 70-80% of the antibiotics have been discovered in the recent 50 years.



Erythromycin
Saccharopolyspora erythrea
antibiotic

FK 506
Streptomyces tsukubaensis
immunosuppressant

Avermectin
Streptomyces avermitilis
antiparasitic agent

Candicidin D
Streptomyces griseus
antifungal

Epothilone
Sorangium cellulosum
antitumor agent

Figure 1.2. Examples of polyketide metabolites. The structures of polyketides and their biological activities are shown together with the producers.

Soon after methods for gene cloning in different organisms were published, the isolation of genes for antibiotic biosynthesis attracted more interests in order to enrich the understanding of the relationship between chemical structures of compounds and functions of the related biosynthetic proteins. Through rugged efforts in the early genetic research, some gene clusters responsible for the biosynthesis of polyketide antibiotics were identified by comparison the differences in the genome of wild-type strain and mutant blocked in a step in the biosynthesis. In succession, the close linkage between genes for self-resistance to an antibiotic and some of the biosynthetic genes was disclosed and gave more evidence for a most important generalization that most of the biosynthetic genes needed to make a particular antibiotic, and antibiotic self-resistance genes locate together with regulator genes and transport genes in a single cluster (Chater *et al*., 1985 and Keller *et al*., 1997). Analysis of the sequences of different cloned biosynthetic gene clusters revealed the conserved motifs in the same type of synthase enzymes. And these conserved regions can be used as probes to isolate genes for other antibiotics. Up to now, nearly one hundred of the entire gene clusters for polyketide antibiotic biosynthesis were cloned and sequenced and the largest one covers over 100 kb DNA sequence in chromosome (Hu *et al*., 1994 and Campelo *et al*., 2002).

The potential for producing new “hybrid” polyketide structures by engineering novel combinations of polyketide biosynthetic genes from different organisms was demonstrated by many publications (Hopwood *et al*., 1985 and McDaniel *et al*., 1995 and Staunton *et al*., 2001 and Kealey 2003). The study of more polyketide biosynthetic gene clusters will help us understand the mechanism of enzymatic synthesis and open a new field of drug discovery.

2.2 Molecular Diversity of Polyketides

Polyketides are synthesized by multifunctional enzymes called polyketide synthases (PKS) in a manner that closely parallels fatty acid biosynthesis. The core of the PKS function is the synthesis of long chains of carbon atoms through repetitive Claisen condensation reactions by joining together small organic acids, such as acetic and malonic acid, by a ketosynthase. The building units, acetate, propionate, malonate or methylmalonate, are activated into the form of coenzyme A (CoA) esters, such as acetyl-CoA and malonyl-CoA, before involvement in the assembly of the polyketide chain. The most common starter unit, acetyl-CoA, with two carbon atoms, is condensed with a malonyl-CoA, with three carbons, to give a chain of four carbon atoms with loss of one carbon dioxide. Only two carbons are added into the chain in each round of condensation with malonyl-CoA (Fig.1.3A). If the extender unit is methylmalonyl-CoA, the “extra” carbon forms a methyl side branch to keep the original extension speed in the main chain. Selection of different numbers and types of starter and extender units diversifies the structures of the polyketide chains (Hopwood, 2004).

After every Claisen condensation reaction of the fatty acid biosynthesis, the resulting β-keto ester is successively reduced by ketoreduction, dehydration and enoylreduction to give a saturated chain. Conversely, the β-keto groups in the growing polyketide chain may be untreated, partially or fully reduced, yielding different reduction levels (Fig.1.3B). A keto group can be reduced to a hydroxyl group, or successively dehydrated to form a double bond between two adjacent carbon atoms, or finally fully saturated with hydrogen atoms by “enoyl” reduction of the double bond (Fig.1.3C). PKS can omit this cycle or curtail it after some or even all condensation steps. The mode of β-keto processing and the stereochemistry of the hydroxyl groups also contribute to the diversity of polyketide metabolites.

After the carbon chain reaches a specific length, it is released from the PKS by thiolysis or acyltransfer (Wakil, 1989) and cyclized to give different folding patterns. The “tailoring” enzymes, including cyclases, group transferases (e.g. *C*-, *O*-, and *N*-methyltransferases, glycosyltransferases, and acyltransferases), NADP(H)- or FAD(H)-dependant oxidoreductases, and cytochrome P450-type oxygenases (Pfeifer *et al*., 2001), can act on the PKS-derived intermediate to yield the final biologically active product.

Characteristics of the polyketide biosynthetic processing lead to the huge variation among naturally occurring polyketides.

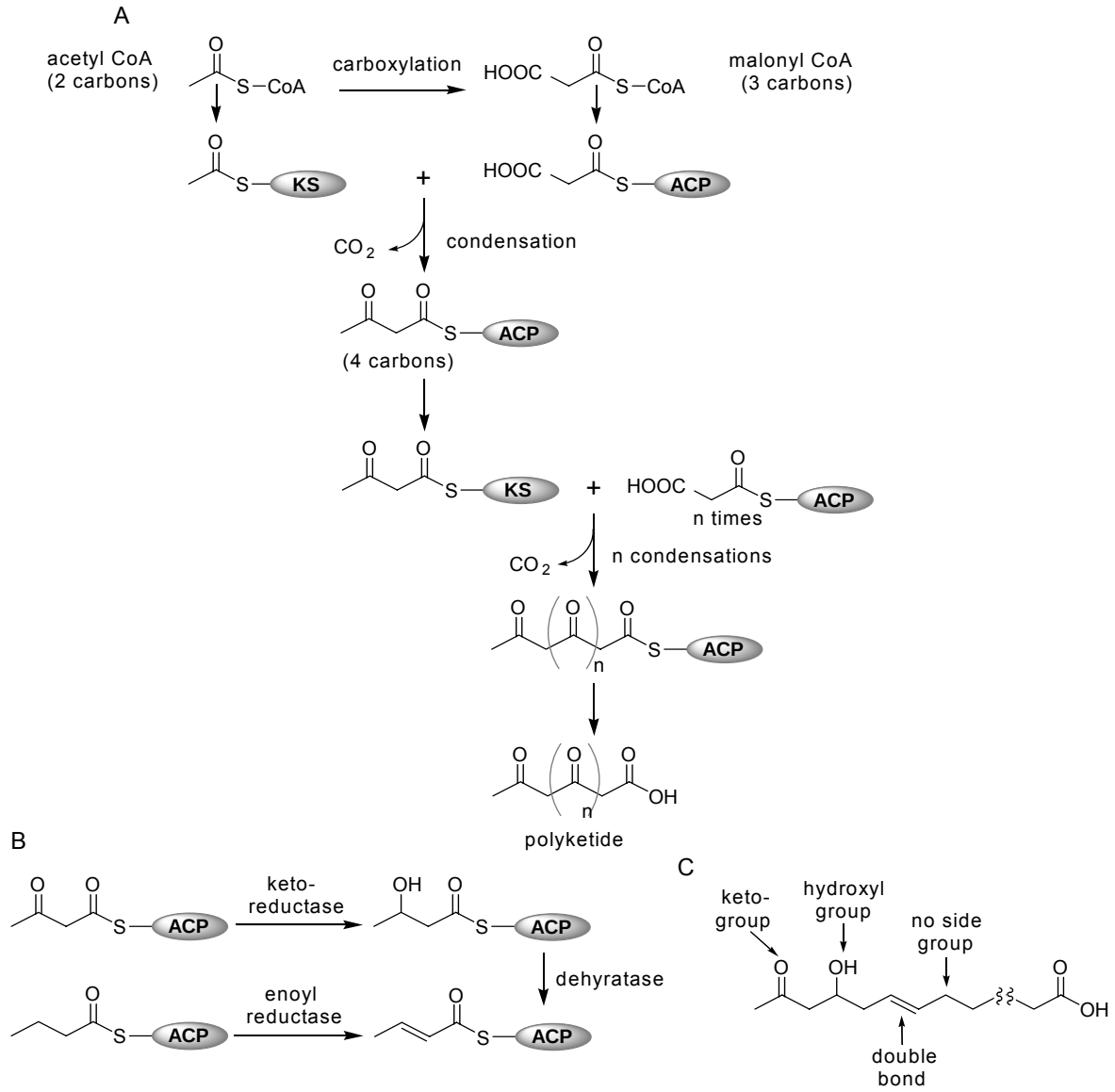


Figure 1.3. The chemistry of polyketide chain assembly (Hopwood, 2004)**.**

A. After activation as the corresponding coenzyme A (CoA) esters, acetic acid and malonic acid are attached, by specific acyl transferases, to components of the polyketide synthase (PKS). The starter unit acetyl-CoA is attached to the active site (cysteine thiol) of the ketosynthase (KS), which catalyzes condensation, and the extender unit malonyl-CoA to a thiol residue of the acyl carrier protein (ACP). One carbon from malonyl-CoA is lost as carbon dioxide during the condensation to yield a four-carbon chain attached to the ACP. After transfer back of the saturated chain from the ACP to the KS, the cycle is then repeated to produce a polyketide chain.

B. The three-step reductive cycle that converts a keto group to a hydroxyl, then to a double bond, and finally to a fully saturated carbon.

C. A complex polyketide containing keto groups, hydroxyl groups, double bonds and fully saturated carbons at different positions.

2.3 Classification of Polyketide Synthases

In analogy to the classification of fatty acid synthases (FAS), PKS have traditionally been defined into two classes. The multifunctional modular systems that are responsible for the biosynthesis of macrolactones, polyenes and polyethers (Bibb *et al*., 1989 and Fernandez-Moreno *et al*., 1992 and Yu *et al*., 1994) were designated type I. The fully dissociable complex of small, discrete monofunctional proteins that catalyze the biosynthesis of bacterial aromatic polyketides is exemplified by the class of type II. In the past decade, as cloning and sequencing of more and more PKS genes, especially the discovery of fungal PKS and plant PKS, the classes of PKS were redefined and enriched by accession of type III PKS and expansion of type I PKS into two subclasses, modular type I PKS and iterative type I PKS (Fig.1.4).

In iterative type II PKS, the active site for each biosynthetic step is encoded in a single gene. There is only one set of a heterodimeric ketosynthase (KSα-KSβ) and an acyl carrier protein (ACP), that have to operate a specific number of times to build a polyketide chain in correct length and subsequent cyclisation, reduction and aromatization are performed by cyclase (CYC), KR and aromatase (ARO), respectively. In some type II PKS, the malonyl-CoA:ACP acyl transferase (MAT), which catalyzes acyl transfer between malonyl-CoA and the ACP, is missing and possibly shared between the PKS and the housekeeping FASs (Revill *et al*., 1995). Type II PKS usually catalyze the biosynthesis of a broad range of polyfunctional aromatic natural products.

Type III PKS are homodimeric synthases of the chalcone synthase (CHS) superfamily known as CHS-like PKS without traditional ACP and KS domains and use acyl CoA substrates directly. Chalcone synthases, the most well-known representatives of this family, are ubiquitous in higher plants and provide the starting materials for a diverse set of biologically important phenylpropanoid metabolites (Schroder 1999). Type III PKS were traditionally associated with plants but recently discovered in a number of bacteria.

Modular type I PKS presenting at bacterial systems are large multifunctional polypeptides arranged in a modular fashion with each module being responsible for one round of chain extension and subsequent β-keto processing. Particularly, each active site in modular type I PKS is used only once during the polyketide biosynthesis.

Iterative type I PKS are responsible for the biosynthesis of fungal metabolites such as 6-methylsalicyclic acid (Shoolingin-Jordan *et al*., 1999) and lovastatin (Staunton 1998). Iterative type I PKS possess only one multidomain protein, in which all the enzyme activities are covalently linked together. The single multifunctional protein is used iteratively to catalyze multiple rounds of chain elongation and appropriate β-keto processing. Notably, in some iterative type I PKS, the one set of catalytic domains is able to vary the reduction level of β-

keto groups during different extension cycles (Kennedy *et al*., 1999 and Hendrickson L *et al*., 1999).

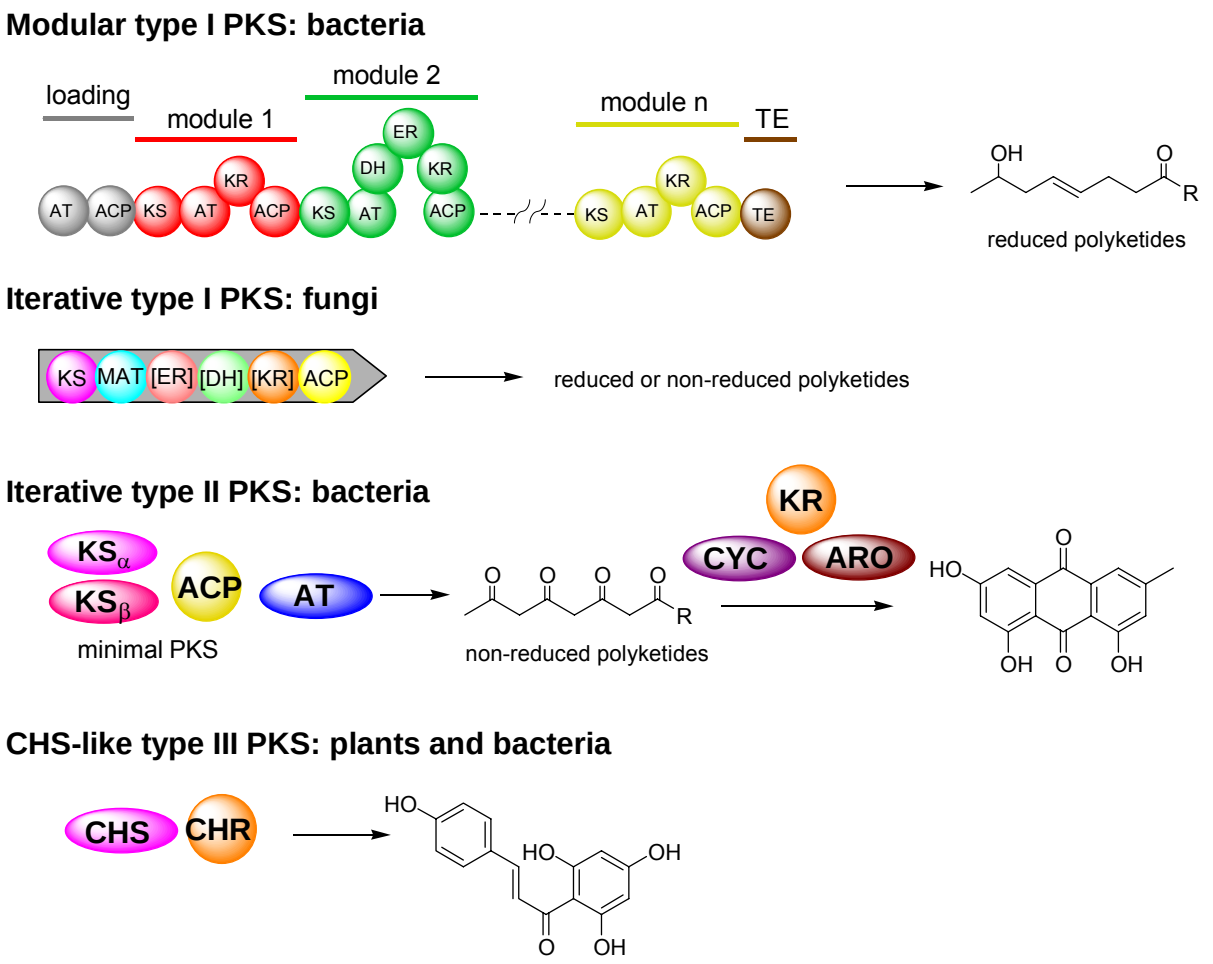


Figure 1.4. Different classes of polyketide synthases. Modular type I PKS consist of multi-domain proteins that form a modular unit for each condensation cycle. In iterative type I PKS, one copy of each active domain, ketosynthase (KS), malonyl-acetyl transferase (MAT), acyl carrier protein (ACP) and optional activites for reduction, ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) are assembled in one protein and iteratively used during the biosynthesis. In contrast, active sites of type II PKS are encoded in different genes and act in an iterative fashion. Type III CHS-like PKS have a simple architecture like CHS (a homodimer of identical KS monomeric domains) with an optional chalcone reductase (CHR). TE, thioesterase; CYC, cyclase; ARO, aromatase.

2.4 Modular Type I Polyketide Synthases

According to their operative mode, PKS can be generally classified into iterative and non-iterative PKS. Some polyketide synthases, such as bacterial aromatic (type II) PKS, fungal and plant PKS, operate in an iterative fashion reminiscent of fatty acid synthases. In these systems, catalytic domains are re-used until the polyketide chain has reached its full length. In stark contrast, modular type I PKS mostly from bacteria, represent giant multienzyme

systems that accommodate distinct active sites for each step of catalysis, and are thus considered as non-iterative. This kind of PKS is usually responsible for the synthesis of macrolide polyketides, such as erythromycin (Cortés *et al*., 1990)(Fig.1.2), oleandomycin (Swan *et al*., 1994), avermectin (MacNeil *et al*., 1994), rapamycin (Schwecke *et al*., 1995), spiramycin (Kuhstoss *et al*., 1996), FK506 (Motamedi *et al*., 1997), tylosin (Gandecha *et al*., 1997) and nystatin (Zotchev *et al*., 2000) (Fig.1.5).

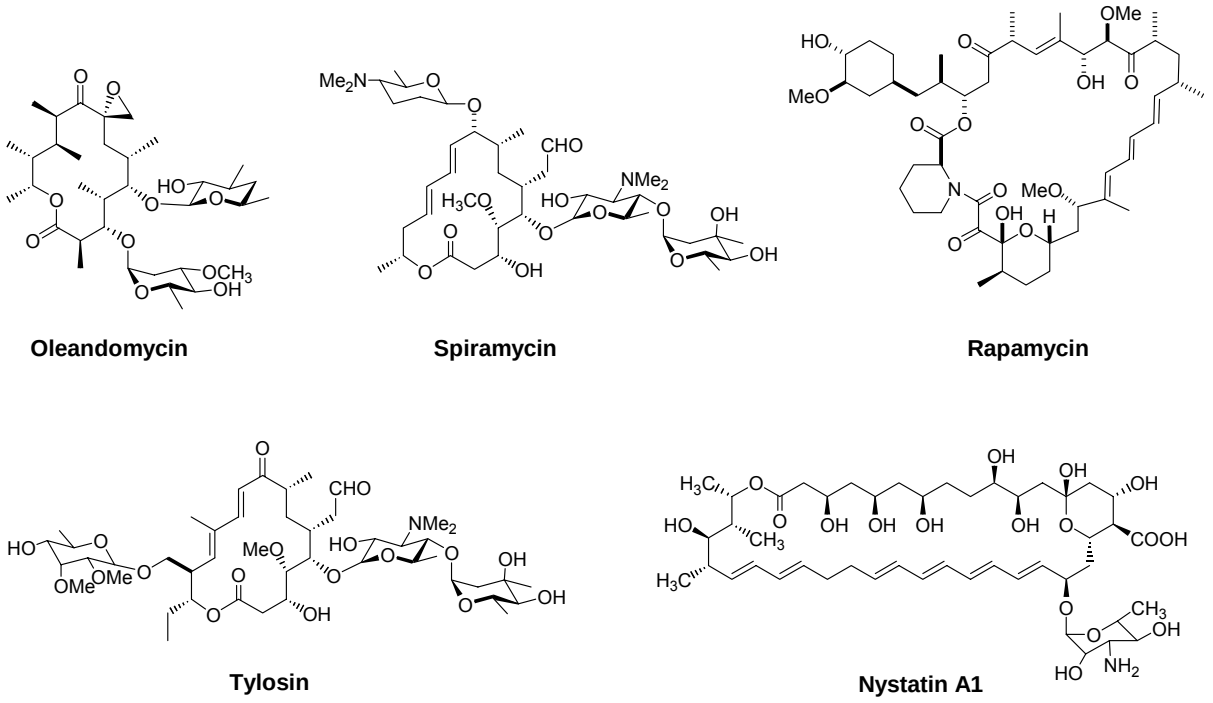


Figure 1.5. Chemical structures of several macrolide polyketides synthesized by modular type I PKS.

In modular type I PKS, catalytic domains are physically grouped into modules that consist of a minimal set of domains for chain propagation, usually a ketosynthase (KS), an acyl transferase, (AT), and ACP. Additional domains may be present for keto group processing, such as ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER). While the first module is usually fused to a loading domain at its *N*-terminus, the last PKS module terminates polyketide assembly with an additional down-loading thioesterase (TE) domain (Cortés *et al*., 1990 and Donadio *et al*., 1991). According to the molecular assembly line character of modular PKS, order and architecture of the modules usually reflect the chain length as well as reduction degree of the resultant polyketide (Cortés *et al*., 1990 and Donadio *et al*., 1991). This one-to-one correspondence has been demonstrated experimentally (Yu *et al*., 1999) and provides the basis for rationally engineering type I PKS biosynthetic pathways.

2.4.1 The Erythromycin Polyketide Synthase

The macrolide antibiotic erythromycin polyketide synthase (Cortés *et al*., 1990 and Donadio *et al*., 1991), produced by a relative of *Streptomyces* called *Saccharopolyspora erythraea*, is the prototype of type I PKS. It catalyzes the decarboxylative condensations of six (2S)-methylmalonyl-CoA extender units with a propionate starter unit to give 6-deoxyerythronolide B (6-dEB), the aglycon (polyketide) moiety of erythromycin. The entire PKS (DEBS, 6-deoxyerythronolide B synthase) comprises three large polypeptides (DEBS1, DEBS2 and DEBS3) that together have 28 distinct active sites (Fig.1.6). The genes (*eryAI*, *eryAII* and *eryAIII*) encoding these three proteins span 32 kb in the erythromycin (*ery*) gene cluster. DEBS1 contains a loading didomain (AT and ACP), which accepts the starter unit, while DEBS3 terminates with a TE domain, which catalyzes the release and cyclisation of the completed polyketide chain. Each of the DEBS proteins contains two functional modules and the sum of six modules corresponds to six condensation reactions needed to build up the polyketide backbone. Module 3 only contains KS, AT and ACP without all reductive domains, so the keto group could survive. Module 4 possesses a full set of reductive activities (DH+ER+KR) and thus a methylene emerges after the fourth round of condensation as expected. Other four modules have only one reductive domain (KR) and hence four β-keto groups of the corresponding elongation cycles are converted into hydroxyl groups, respectively. The correspondence between the catalytic domain composition of consecutive modules in the PKS and the structure of the resulting biosynthetic product provided the basis for the “assembly line” model, in which the program for the production of polyketides is hardwired into the DNA and expressed in a linear array of active domains.

Several successful experiments provided strong evidences for this model. When the KR of module 5 was removed by deletion of the encoding gene, a keto group survived in the resulting erythromycin analogues at the position of the fifth condensation in place of the normal hydroxyl group (Donadio *et al*., 1991). Inactivation of the ER in module 4 by mutagenesis resulted in a double bond at the expected position of the resulting product (Donadio *et al*., 1993). Further evidences were provided by experiments of repositioning the TE in order to cause premature chain release at the diketide, tetraketide and hexaketide stages (Kao *et al*., 1995 and Kao *et al*., 1996 and Böhm *et al*., 1998).

Besides three big open reading frames (ORFs) for the *eryA* polyketide synthase genes, several genes, including a hydroxylase gene (*eryF*) (Haydock *et al*., 1991), *O*-glycosyl-transferase genes (*eryB* and *eryC*) (Dhillon *et al*., 1989 and Haydock *et al*., 1991), a P450 hydroxylase gene (*eryK*) (Stassi *et al*., 1993) and an *O*-methyltransferase gene (*eryG*) (Haydock *et al*., 1991), were also found to be involved in the secondary modifications of polyketide chain to yield biologically active erythromycin at the end.

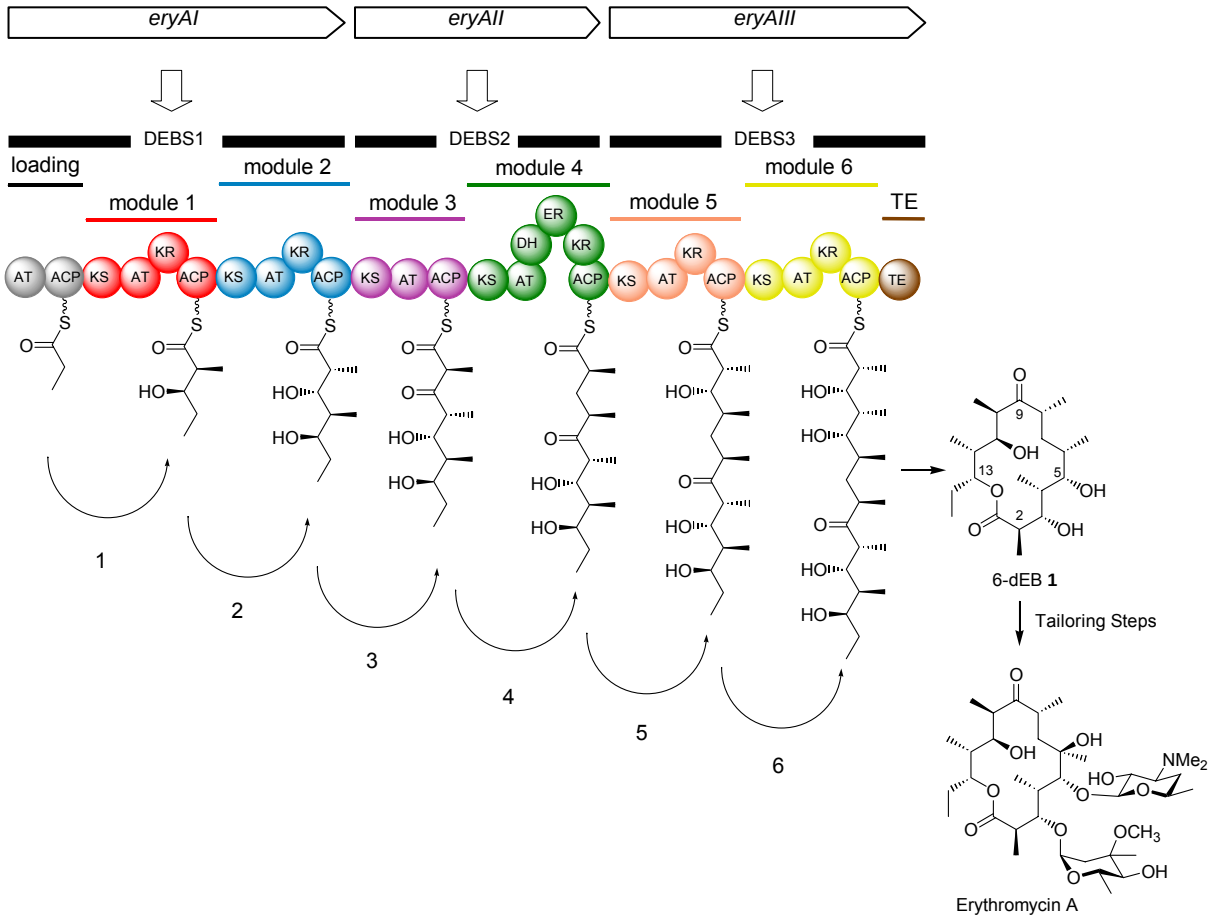


Figure 1.6. Domain organization of the erythromycin polyketide synthase. Each PKS module contains the essential ketosynthase (KS), acyltransferase (AT) and acyl carrier protein (ACP) domains. Some modules also have optional β-keto processing activities, like ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER). TE, Thioesterase.

2.4.2 Genetic Engineering of Modular Type I Polyketide Synthases

The direct correspondence between the catalytic domains of modules in a PKS and the structure of the resulting biosynthetic product allows rational modifications of the polyketide structure by genetic engineering. When targeted changes are made in the genes, expected compounds could be produced by the mutated PKS. The strategies available include deletion, substitution or addition of catalytic activities by manipulation of corresponding genes. Through expression of PKS, exchange of starter units, rearrangement of modules, swapping or inactivating of domains or modification of post-PKS processing, unnatural natural products could be generated for biological and pharmaceutical applications.

Because the development of genetic methods in uncharacterized microorganisms remains a major technological barrier, a *Streptomyces* host-vector system — CH999/pRM5 has been developed for efficient construction and expression of recombinant PKS (McDaniel *et al*., 1993). The host strain CH999 is a *S. coelicolor* A3 (2) derivative, from which the entire set of

actinorhodin (*act*) gene cluster (a type II PKS) except for one, *act*VI-ORFA, at the extreme left-hand end of the cluster, has been deleted and replaced by an erythromycin resistance gene (Caballero *et al*., 1991) and the biosynthesis of another pigmented antibiotic, undecylprodigiosin, was blocked by a stable mutation. A series of low copy number *S. coelicolor* plasmid SCP2* derivatives carrying the desired sets of PKS subunit genes were introduced into CH999 to investigate novel polyketide metabolites. They usually include a ColE1 replicon and a selection marker of β-lactamase gene (*bla*) for genetic engineering in *E. coli*, an appropriately truncated SCP2* *Streptomyces* replicon and a thiostrepton resistance gene (*tsr*) for selection in *Streptomyces*. The particular components of these plasmids are the *actII*-ORF4 activator gene from the *act* cluster, which activates transcription from *act* promoters, and the divergent *actI*/*actIII* promoter pair together with convenient cloning sites for the insertion of different PKS genes downstream of both promoters. For example, the representative plasmid pRM5 (Fig1.7) contains the *act* KS (KSα), CLF (KSβ), ACP, ARO, CYC and KR genes cloned downstream of P*actI* and P*actIII*. Other recombinant PKS can also be introduced in the same position by exchange of the gene sets.

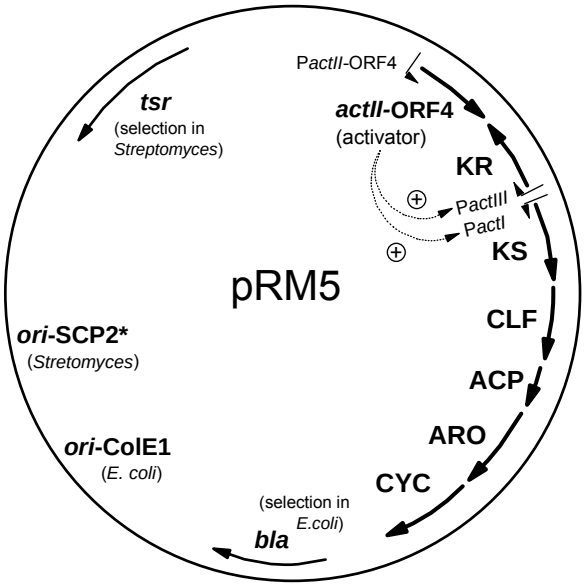


Figure 1.7. Restriction map of the vector pRM5 for expression of recombinant PKS. *ori* SCP2*, origin of replication of SCP2*; *ori* ColE1, ColE1 origin of replication; *bla*, β-lactamase gene from pBR327; *act*, actinorhodin biosynthesis gene; KR, ketoreductase; KS, ketosynthase; CLF, chain-length factor; ACP, acyl carrier protein; ARO, aromatase; CYC, cyclase.

Three big DEBS genes were cloned into pCK7, a derivative of pRM5, in succession and expressed in the host strain *S. coelicolor* CH999. The recombinant strains produced abundant quantities of two polyketides. The major product is 6dEB, while the minor product

was identified as 8,8a-deoxyoleandolide, which apparently originated from an acetate starter unit instead of the propionate starter in 6dEB biosynthesis (Kao *et al*., 1994). A new approach is to use *E. coli* strains for expression of PKS proteins (Gokhale *et al*., 1999). To date, single module portions of DEBS have been expressed successfully in *E. coli* (Khosla, 2000).

By feeding experiments or exchange of loading modules, the structural diversity of polyketides is also enriched by the involvement of different starter units. Under normal conditions *in vivo*, only propionate units are recruited by the DEBS loading module, but in recombinant cells or with the purified PKS *in vitro*, several other acyl-CoA esters are also accepted as starter units (Kao *et al*., 1995 and Wiesmann *et al*., 1995). The additional feeding of alternative carboxylic acids to the fermentation of *S. avermitilis*, has also led to a number of pharmaceutically useful novel avermectins. It seems that the avermectin-producing PKS is unusual in being able efficiently to accept numerous C2-branched carboxylic acids as alternative starter units when they are added to the culture (Chen *et al*., 1989 and Dutton *et al*., 1991). A famous example is the production of doramectin. When fed with cyclohexylcarbonyl CoA, the avermectin PKS uses it as starter unit instead of the isobutyrate starter of avermectin and produces a commercial anthelmintic agent, doramectin (Cropp *et al*., 2000). The characteristic of broader selectivity in avermectin PKS could be transferred to a heterologous PKS by exchanging its loading module. When the avermectin PKS loading module was grafted onto the erythromycin-producing PKS in place of the native loading module, novel erythromycins containing unnatural starter units were obtained (Pacey *et al*., 1998).

Deletion, inactivation and swapping of domains have become the most widely practical engineering techniques in PKS. The choice of extender unit by a modular PKS system is determined by the AT domain of each module. Replacement of AT domains or changing the substrate specificity of the AT domain by site-specific mutation can lead to different extenders involved in polyketide biosynthesis. The methylmalonyl-specific AT domain of DEBS modules 1 and 2 were replaced with three heterologous malonyl-specific AT domains from rapamycin PKS, pikromycin PKS and an uncharacterized PKS. These changes resulted in the production of novel erythromycin derivatives, which lack the methyl group at corresponding sites of the macrolactone rings (Ruan *et al*., 1997).

A promising new approach is to engineer novel combinations of the mutated PKS protein subunits. Each of the three DEBS genes were cloned into three compatible *streptomyces* vectors and the genes were introduced in both their native forms and containing a single mutation including domain swap or point mutation. 46 of such different triple transformants produced detectable levels of polyketide metabolites (Xue *et al*., 1999). By substituting for the AT and beta-carbon processing domains of DEBS with counterparts from different

polyketide PKS, a combinatorial library of polyketides was also constructed and more than 100 macrolide products have been generated by using this simple combinatorial set (McDaniel *et al*., 1999).

An alterative strategy for mutational engineering polyketides is the combinatorial manipulation of larger PKS units, such as the entire modules or the whole protein subunits. For example, the recombination between two homologous KR domain-encoding sequences resulted in the elimination of one complete module from nystatin PKS, consequently yielding a smaller macrolactone (Brautaset *et al*., 2002). Substitution of DEBS module 2 with module 5 from the rifamycin PKS created a hybrid DEBS/Rif PKS, which produces the same 6dEB product as the native PKS. The domain architecture of PKS suggests that modules begin with a KS domain and end in an ACP (Gokhale *et al*., 1999). Some results from two different experiments recently made a new inspect on this view (Ranganathan *et al*., 1999). A DEBS module terminating with an ACP could not communicate effectively with a RAPS (rapamycin synthase) module. Indeed, the DEBS ACP required part of its partner KS in the next module for transfer of the polyketide chain between the modules. When DEBS module 1 was fused to DEBS modules 3 and 6, the investigators found that conservation of the natural linker region between modules 1 and 2 gave a functional multienzyme. All these experiments showed the importance of the linker region between two modules.

After the release from the PKS, the polyketide chain is modified by tailoring reactions. Mixing and matching of the tailoring enzymes also lead to the production of altered structures (Madduri *et al*., 1998 and Zhao *et al*., 1999).

Modular polyketide synthases are potentially valuable for the biosynthesis of novel complex polyketides. From domain and module substitutions to module fusion and subunit complementation can be used for biosynthesizing new polyketides. As the discovery of more and more PKS genes and deeper exploration of the PKS mechanism, genetic engineering with combinatorial PKS catches more attention to product numerous novel polyketides serving as a source for large libraries of unnatural natural products.

2.5 Some Speculations on the Evolution of Iterative and Non-Iterative PKS

The most significant distinction between different types of PKS is the operative mode of the various catalytic sites, *i.e*. PKS act either iteratively or non-iteratively in terms of programming mechanisms. Iterative PKS including iterative type I PKS, type II and III PKS carry a single set of active sites that are used iteratively in successive rounds of chain assembly and reduction, while the active sites of non-iterative PKS such as modular type I PKS act once in the building of an entire carbon chain. Comparison of different PKS, two biochemical features were found to currently correlate with iterative and modular PKS. Iterative PKS exclusively employ malonyl-CoA as extender units and usually make very

simple reductive changes to the β-keto groups of the polyketide chain. Thus, it seems that only a single set of active sites operating in an iterative manner is sufficient for all similar rounds of polyketide extension in the iterative synthases. Whereas, modular PKS use at least a proportion of more complex chain extenders and have the complicated β-keto processing in different elongation cycles of polyketide biosynthesis. Perhaps the complexity of chain extender choices and reduction level choices is the selective pressure for modular PKS to evolve (Hopwood 1997). Up to now, the molecular basis for iterative or non-iterative action of polyketide synthases and their evolutional relationship are still unknown. Studies on the precise mechanisms of more PKS, especially those with different characteristics, will enrich our understanding of the nascence of PKS and their evolutionary progress, and also contribute to the discovery of novel drugs by genetic engineering.

3. Aureothin

Aureothin is a rare nitrophenyl-substituted polyketide metabolite (Fig.1.8) from the soil isolate *S. thioluteus* (Maeda, 1953) with interesting structural and pharmacological features, such as antifungal, insecticidal and weak antitumour properties (Schwartz *et al*., 1976). The homolog neoaureothin, produced by *S. orinoci*, exhibits antifungal and antiviral activities (Kakinuma *et al*., 1976) and the related pyrone, luteoreticulin, isolated from *S. luteoreticuli*, possesses cytotoxic activity (Suzuki *et al*., 1976). Interestingly, a closely related derivative, *N*-acetyl-aureothamine, recently isolated from *S. netropsis* (Taniguchi *et al*., 2000), has been found to be a highly selective agent against *Helicobacter pylori*, the most common cause of chronic gastritis.

Previous feeding experiments with specifically labelled precursors showed that aureothin derives from a mixed polyketide pathway (Cardillo *et al*., 1974 and Yamazaki *et al*., 1975). In fact, the incorporation pattern of ^{13}C-labeled precursors indicated that the aureothin backbone is formally assembled from one acetate and four propionate units, possibly by a type I PKS (Fig.1.8).

Aureothin and its homolog neoaureothin are produced by different *Streptomyces* species, but both of them have many identical features in the chemical structures, such as an unusual nitroaryl moiety, an O-methylated pyrone and a tetrahydrofuran ring. The difference between them is that aureothin has a diene moiety, while neoaureothin has a tetraene system. It is tempting to speculate that the biosynthetic pathways of the two compounds are very similar and compared to aureothin, neoaureothin could have two more propionate units during the polyketide backbone assembly. Molecular analysis of the requisite biosynthetic gene clusters will disclose the evolutionary relationship between aureothin and neoaureothin.

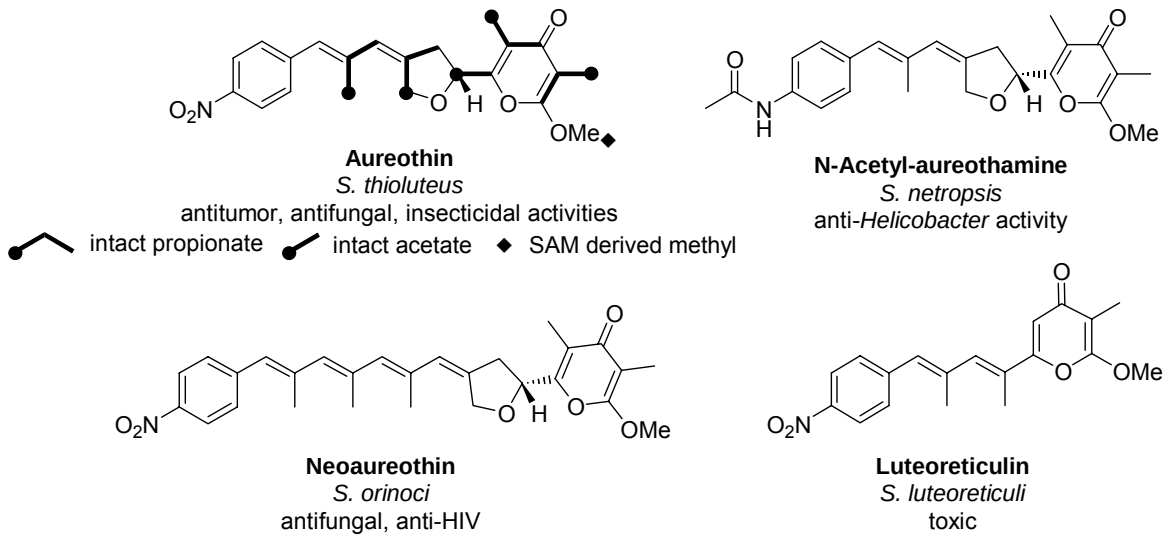


Figure 1.8. Rare nitrophenyl-substituted polyketide metabolites and derivative produced by different *Streptomyces* strains and the results of previous labeling experiments for aureothin.

The structure of the C_7N unit suggests that PABA or a derivative thereof is involved in aureothin biosynthesis as a key intermediate. In polyketide biosynthesis, PABA was found to act as the starter units in two highly similar heptaene macrolides, candicidin (Campelo *et al*., 2002) and FR-008 (Chen *et al*., 2003). Surprisingly, all efforts to incorporate labeled PABA into aureothin failed. Therefore, what is the real starter for the aureothin polyketide chain programming? And which gene(s) is(are) responsible for the synthesis of this rare nitrophenyl moiety? There is great interest to answer these two questions by biological and chemical methods.

Aureothin contains another two unusual moieties, an exomethylene tetrahydrofuran ring and an O-methylated pyrone ring. Up to now, the biosynthetic pathway of this intriguing homochiral exomethylene tetrahydrofuran moiety in polyketide biosynthesis is not clear. Also, the O-methylated pyrone ring is usually found in aromatic polyketides that are normally synthesized by type II PKS (Piel *et al*., 2000). The study of aureothin gene cluster will also help us understanding the biosynthetic mechanism of these unusual structural features.

4. Research Goals

The special characteristics of the aureothin structure (Fig.1.9) motivated the investigation of the aureothin biosynthetic pathway at the biochemical level. The aim of this study was to disclose the whole biosynthetic pathway of aureothin, in particular:

- Primer design and amplification of key genes in aureothin biosynthesis by PCR in order to obtain the probes for cloning the aureothin (*aur*) gene cluster.
- Cloning of the aureothin biosynthesis gene cluster by construction and screening of a genomic library of the aureothin producer, *S. thioluteus*.

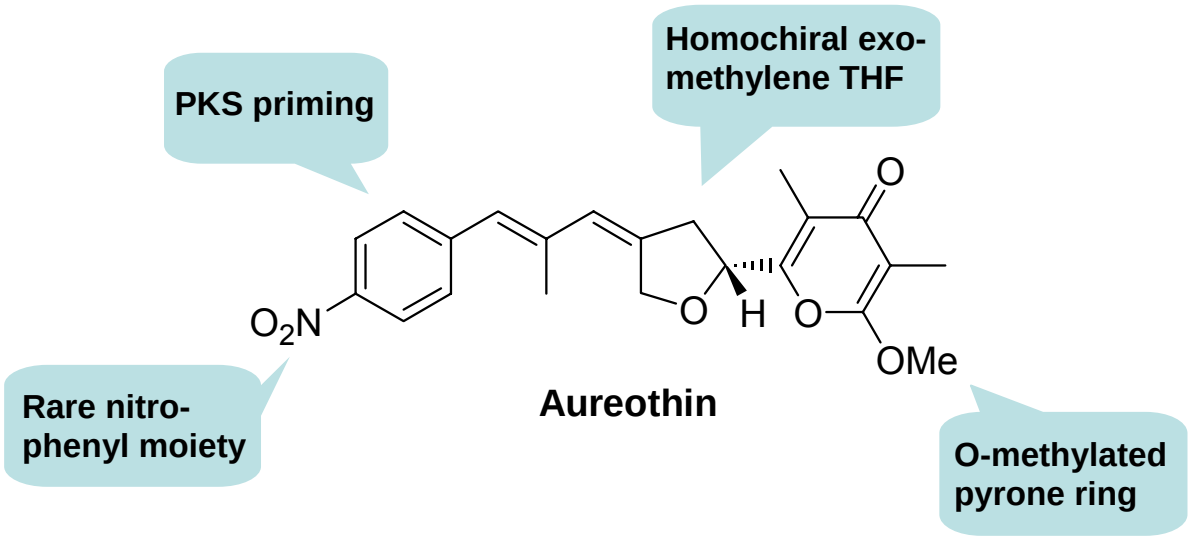


Figure 1.9. Special structural characteristics of aureothin.

- Verification of the identity of the cloned aureothin gene cluster either by inactivation of some essential genes in wild type strains or by expression of the entire gene cluster in a heterologous host.
- Sequence analysis of the cloned region involved in aureothin biosynthesis. Homology analysis with the known genes in databases will indicate the putative functions of some gene products in the aureothin biosynthesis gene cluster.
- Elucidation of the biosynthetic pathway of the rare nitro-phenyl moiety and the starter unit of the polyketide programming in aureothin using both molecular and chemical approaches.
- Disclosure of the enzyme(s) that is(are) responsible for the formation of the aureothin tetrahydrofuran ring by inactivation and heterologous expression of the putative gene(s).
- Detection of the enzyme that introduces the methyl group into the pyrone ring by inactivation and heterologous expression of the putative gene, and investigation of the timing of the tailoring reactions.
- Investigation of the course of the aureothin polyketide chain assembly by genetic engineering of the mutated or hybrid PKS.
- Production of some new aureothin derivatives with enhanced biological activities by engineering approaches.

Molecular analyses and functional studies of the aureothin biosynthesis gene cluster will provide new insights into the machinery and evolution of polyketide synthase and set the basis for engineering novel aureothin derivatives.

B. Materials and Methods

1. Materials

1.1 Media

1.1.1 Media for the Cultivation of *Escherichia coli* Strains

LB Medium (Sambrook *et al*., 1989)
Tryptone 10.0 g, yeast extract 5.0 g, NaCl 10.0 g, agar 15.0 g (optional), purified water 1000 ml

2 × YT Medium
Tryptone 16.0 g, yeast extract 10.0 g, NaCl 5.0 g, purified water 1000 ml

SOB Medium (Hanahan, 1983)
Bacto tryptone 20 g, yeast extract 5 g, NaCl 0.584 g, KCl 0.186 g, pH 7.0

1.1.2 Media for the Cultivation of *Streptomyces* Strains

M10 Medium
Yeast extract 4.0 g, malt extract 10.0 g, glucose 4.0 g, agar 20 g (optional), purified water 1000 ml, pH 7.3

2CM Medium (Yuan, 1983)
NaCl 1.0 g, K_2HPO_4 1.0 g, soluble potato starch 10.0 g, inorganic salt solution* 1 ml, $MgSO_4 \cdot 7H_2O$ 2.0 g, tryptone 2.0 g, $CaCO_3$ 2.0 g, $(NH_4)_2SO_4$ 2.0 g, agar 22.0 g, purified water 1000 ml
Inorganic salt solution:
$FeSO_4 \cdot 7H_2O$ 1.0 g, $MgCl_2 \cdot 6H_2O$ 1.0 g, $ZnSO_4 \cdot 7H_2O$ 1.0 g, purified water 1000 ml

R5 Medium (Thompson *et al*., 1980)
Sucrose 103 g, K_2SO_4 0.25 g, $MgCl_2 \cdot 6H_2O$ 10.12 g, glucose 10.0 g, Difco casaminoacids 0.1 g, trace element solution* 2.0 ml, Difco yeast extract 5.0 g, TES 5.73 g, purified water 1000 ml

Aliquots of 100 ml of the solution were poured into 250 ml Erlenmeyer flasks each containing 2.2 g Difco bacto agar. The flasks were closed and autoclaved. At the time of use, the medium was re-melted and mixed with the following sterile solutions in the order listed:

KH_2PO_4 (0.5%) – 1.0 ml
$CaCl_2 \cdot 2H_2O$ (5 M) – 0.4 ml
L-proline (20%) – 1.5 ml
NaOH (1 M) – 0.7 ml

Trace element solution:
$ZnCl_2$ 40 mg, $FeCl_3 \cdot 6H_2O$ 200 mg, $CuCl_2 \cdot 2H_2O$ 10 mg, $MnCl_2 \cdot 4H_2O$ 10 mg, $Na_2B_4O_7 \cdot 10H_2O$ 10 mg, $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ 10 mg, purified water 1000 ml

Minimal Medium (MM) (Hopwood, 1967)
L-asparagine 0.5 g, K_2HPO_4 0.5 g, $MgSO_4 \cdot 7H_2O$ 0.2 g, $FeSO_4 \cdot 7H_2O$ 0.01 g, glucose (add after autoclaving) 10 g, agar 10 g, purified water 1000 ml, pH 7.0 – 7.2

Soya Flour Mannitol Medium (SFM) (Hobbs *et al*., 1989)
Soya flour 20 g, mannitol 20 g, agar 20 g, tap water 1000 ml, pH 7.2

Yeast extract-malt extract medium (YEME)
Difco yeast extract 3.0 g, Difco bacto-peptone 5.0 g, Oxoid malt extract 3.0 g, glucose 10 g, sucrose 340 g, purified water 1000 ml
After autoclaving, the medium was mixed with the following sterile solutions in the order listed:

$MgCl_2 \cdot 6H_2O$ (2.5 M) – 2 ml (5 mM final)
Glycine (20%) – 25 ml (0.5% final)

Different *Streptomyces* strains may need different concentrations of glycine; 0.5% is best for *S. lividans* and *S. coelicolor*.

Double Strength Germination Medium (Kieser *et al*., 2000)
Difco yeast extract 1 g, Difco casaminoacids 1 g, CaCl2 (5 M) 200 µl, purified water 100 ml

1.2 Buffers and Solutions

1.2.1 Buffers for Plasmid DNA Preparation from *E. coli*

Solution I (*E. coli* cell suspension buffer) (Sambrook *et al*., 2001)
50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0)

The solution was stored at 4°C after autoclaving.

Solution II (*E. coli* lysis buffer) (Sambrook *et al*., 2001)
0.2 M NaOH (fresh diluted from a 10 M stock), 1% (w/v) SDS
The solution was prepared freshly and used at room temperature.

Solution III (neutralizing buffer) (Sambrook *et al*., 2001)

Potassium acetate (5 M) – 60.0 ml
Glacial acetic acid – 11.5 ml
Purified water – 28.5 ml

The resulting solution was 3 M with respect to potassium and 5 M with respect to acetate and stored at 4°C.

TE Buffer

Tris-HCl (2 M, pH 8.0) – 5 ml
Sodium EDTA (0.25 M, pH 8.0) – 4 ml
Purified water – 991 ml

1.2.2 Buffers for Protoplast Transformation of *Streptomyces*

P (Protoplast) Buffer (Okanishi *et al*., 1974; Hopwood *et al*., 1978)
Sucrose 103 g, K_2SO_4 0.25 g, $MgCl_2 \cdot 6H_2O$ 2.02 g, trace element solution 2 ml, purified water 800 ml
Aliquots of 80 ml of the buffer were dispensed into flasks and autoclaved. Before use, the following sterile solutions were added to each flask in order:

KH_2PO_4 (0.5%) – 1 ml
$CaCl_2 \cdot 2H_2O$ (3.68%) – 10 ml
TES buffer (5.73%, to pH 7.2) – 10 ml

T (Transformation) Buffer (Thompson *et al*., 1982)
The following sterile solutions were mixed together:

Sucrose (10.3%) – 25 ml
Purified water – 75 ml
Trace element solution – 0.2 ml
K_2SO_4 (2.5%) – 1 ml

Then the following sterile solutions were added to 9.3 ml of the above mixture:

$CaCl_2$ (5 M) – 0.2 ml

Tris-maleic acid buffer* – 0.5 ml

Tris-maleic acid buffer:

Tris 121.14 g, purified water 1000 ml

The pH number of the buffer was adjusted to 8.0 by adding maleic acid. For use, 3 parts by volume of the above solution was added to 1 part by weight of PEG1000 that was previously sterilised by autoclaving.

TE25S Buffer

Tris-HCl (2 M, pH 8.0) – 12.5 ml
Sodium EDTA (0.25 M, pH 8.0) – 100 ml
Sucrose – 102.69 g
Purified water up to 1000 ml

This stock solution keeps idefinitely after autoclaving. Just before use lysozyme was dissolved in a sample of the solution at a concentration of 2 mg/ml.

1.2.3 Buffers for *N*-Oxidation Assay

Sodium Chloride Solution (0.85%)

Phosphate-Buffered Saline (PBS) (Sambrook *et al*., 2001)

NaCl 8 g, KCl 0.2 g, Na_2HPO_4 1.44 g, KH_2PO_4 0.24 g, purified water up to 1000 ml

The pH number of the buffer was adjusted to 7.0 with HCl.

1.2.4 Buffers for Electrophoresis

50 × TAE Buffer (Hopwood *et al*., 1985)

Tris base 242 g, glacial acetic acid 57.1 ml, EDTA (0.5M, pH 8.0) 100 ml

The concentrated buffer was usually diluted into 1 × TAE as working solution.

6 × Gel-loading Buffer (Sambrook *et al*., 2001)

Sucrose 40 g, purified water 100 ml, bromophenol blue 0.25 g

1.2.5 Solutions for Preparation of Competent *E. coli* Cells by Chemical Method

TFB I

The following sterile solutions were mixed together:

Potassium acetate (5 M) – 0.6 ml

$CaCl_2$ (1 M) – 1 ml
KCl (1 M) – 10 ml
MnCl2 (1 M) – 5 ml
Glycerol (100%) – 15 ml
Purified water – 68.4 ml

TFB II

The following sterile solutions were mixed together:

MOPS (1 M) – 1 ml
$CaCl_2$ (1 M) – 7.5 ml
KCl (1 M) – 1 ml
Glycerol (100%) – 15 ml
Purified water – 75.5 ml

1.2.6 Buffers for Hybridization

Denaturation Solution (for neutral transfer, double-stranded DNA targets only)
NaOH 20 g, NaCl 87.66 g, purified water 1000 ml

Neutralization Buffer

Tris base 60.57 g, NaCl 58.44 g
The pH number of the buffer was adjusted to 7.4 with HCl.

20 × SSC

NaCl 175.32 g, tri-sodium citrate 2-hydreate 88.23 g, purified water 1000 ml
20 × SSC has 3M NaCl and 0.3 M sodium citrate. The concentrated buffer was usually diluted for use as e. g. 1 × SSC, 1.1 × SSC.

Prehybridization and Hybridization Solution

20 × SSC – 12.5 ml
Blocking Reagent – 0.5 g
SDS (10%) – 100 µl
Purified water – 37.4 ml

The solution was prepared at least one hour in advance and blocking reagent was dissolved at 65°C.

Maleic acid Buffer

Maleic acid 11.61 g, NaCl 8.77g, purified water 1000 ml

The pH number of the buffer was adjusted to 7.5 with NaOH.

Washing Buffer

0.3 ml Tween 20 was added to 100 ml of maleic acid buffer.

Detection Buffer

Tris base 12.11 g, NaCl 5.84 g, purified water 1000 ml

The pH number of the buffer was adjusted to 9.5 with HCl.

Blocking Solution

1 g blocking reagent was dissolved in maleic acid buffer under 65°C.

Antibody Solution

Anti-digoxigenin-AP was diluted at 1:5000 (final concentration 150 mU/ml) in blocking solution.

Color-substrate Solution

200 µl of NBT/BCIP was added to 10ml detection buffer.

1.3 Strains and Plasmids

Strains:

Strain	Characteristics	Reference or source
a) *Streptomyces*		
S. thioluteus HKI-227	Aureothin producer	HKI strain collection
S. lividans ZX1	*rec*-46 *pro*-2 *str*-6 HAU3^S	Zhou *et al*., 1988
S. albus	Host strain for expression of PKSs	Fierro *et al*., 1987
HJ11	*S. lividans* ZX1::pHJ11 *aac(3)IV*	This study
HJES	*S. lividans* ZX1::pHJES *aac(3)IV*	This study
HJ41	*S. lividans* ZX1::pHJ41 *tsr*	This study
HJ43	*S. lividans* ZX1::pHJ43 *tsr*	This study
HJSET152	*S. lividans* ZX1::pSET152 *aac(3)IV*	This study
HJWHM4*	*S. lividans* ZX1::pWHM4* *tsr*	This study

HJ46	*S. lividans* ZX1::pHJ46 *aac(3)IV*	This study
HJ47	*S. lividans* ZX1::pHJ47 *aac(3)IV*	This study
HJ48	*S. lividans* ZX1::pHJ48 *aac(3)IV*	This study
HJE01	*S. lividans* ZX1::pHJE01 *aac(3)IV*	This study
HJ56	*S. lividans* ZX1::pHJ56 *aac(3)IV*	This study
HJ60	*S. lividans* ZX1::pHJ60 *tsr*	This study
HJ4772	HJ47/pHJ72 *aac(3)IV tsr*	This study
HJ78	*S. lividans* ZX1::pHJ78 *tsr*	This study
HJ79	*S. lividans* ZX1::pHJ79 *aac(3)IV*	This study
HJ82	*S. lividans* ZX1::pHJ82 *aac(3)IV*	This study
HJ86ab	*S. albus*::pHJ86 *aac(3)IV*	This study
HJ95	*S. lividans* ZX1::pHJ95 *tsr*	This study
HJ4795	HJ47/pHJ95 *aac(3)IV tsr*	This study
HJ97	*S. lividans* ZX1::pHJ97 *aac(3)IV*	This study
HJ98	*S. lividans* ZX1::pHJ98 *aac(3)IV*	This study
HJ110	*S. lividans* ZX1::pHJ110 *tsr*	This study
HJ111	*S. lividans* ZX1::pHJ111 *aac(3)IV*	This study
HJ11143	HJ111/pHJ43 *aac(3)IV tsr*	This study
HJ125	*S. lividans* ZX1::pHJ125 *aac(3)IV*	This study
HJ126	*S. lividans* ZX1::pHJ126 *aac(3)IV*	This study
HJ127	*S. lividans* ZX1::pHJ127 *aac(3)IV*	This study
HJ125ab	*S. albus*::pHJ125 *aac(3)IV*	This study
HJ126ab	*S. albus*::pHJ126 *aac(3)IV*	This study
HJ127ab	*S. albus*::pHJ127 *aac(3)IV*	This study
HJ128	*S. lividans* ZX1::pHJ128 *aac(3)IV*	This study
HJ129	*S. lividans* ZX1::pHJ129 *aac(3)IV*	This study
HJ130	*S. lividans* ZX1::pHJ130 *aac(3)IV*	This study
HJ134	*S. lividans* ZX1::pHJ134 *tsr*	This study
HJ137	*S. lividans* ZX1::pHJ137 *aac(3)IV*	This study
HJ144	*S. lividans* ZX1::pHJ144 *aac(3)IV*	This study
HJ146	*S. lividans* ZX1::pHJ146 *aac(3)IV*	This study
HJ137ab	*S. albus*::pHJ137 *aac(3)IV*	This study
HJ144ab	*S. albus*::pHJ144 *aac(3)IV*	This study
HJ146ab	*S. albus*::pHJ146 *aac(3)IV*	This study
b) *Escherichia coli*		
EPI100™	F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80d/*lac*Z ΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U	EPICENTRE

	*gal*K λ⁻ *rps*L *nup*G	
DH5α	*Sup*E44 Δ*lac*U169 (*Φ*80 *lac*ZΔM15) *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1	Hanahan, 1983
ET12567/pUZ8002	*Rec*F *dam*⁻ *dcm*⁻ *cml str tet km*	MacNeil *et al*., 1992
BW25113/pIJ790	Δ(*araD-araB*)567 Δ*lacZ*4787(::*rrnB*-4) *lacIp*-4000(*lacI*^{Q}) λ⁻ *rpoS*369(Am) rph-1 Δ(*rhaD-rhaB*)568 *hsdR*514 pIJ790: *oriR*101 r*epA*101(ts) *araBp-gam-bet-exo cat*	Gust *et al*., 2003

Usual vectors:

Vector	Size (kb)	Replicon	Characteristic/ Origin	Reference
pSET152	5.5	pUC18	*oriT* RK2 *intΦ*C31 *attPΦ*C31 *lacZα aac(3)IV*	Bierman *et al*., 1992
pBluescript KS-	3.0	ColE1	*lacZα orif*1 *bla*	Alting-Mees *et al*., 1989
pRM5	~27	ColE1 SCP2*	*bla tsr cos*	McDaniel *et al*., 1993
pWHM4*	7.9	pUC19 pIJ101	l*acZ bla tsr* P*ermE**	Matseliukh, 2001

Plasmids/Cosmids:

Construct	Size (kb)	Vector	Insert	Resistance
pPAB1	4.0	pGEM-T Easy	1.0 kb PCR fragment (*pab*AB primer) from *S. thioluteus* (template)	Amp^{R}
pKS1	3.7	pGEM-T Easy	0.7 kb PCR fragment (KS primer) from *S. thioluteus* (template)	Amp^{R}
pPAB181	4.0	pGEM-T Easy	1.0 kb PCR fragment (*pab*AB primer) from cosmid 18E4 (template)	Amp^{R}
pKS181	3.7	pGEM-T Easy	0.7 kb PCR fragment (KS primer) from cosmid 18E4 (template)	Amp^{R}
pHJS18	11.5	pWEB^{TM}	3.5 kb fragment from cosmid 18E4 *Sac*I-digested and self-ligation	Amp^{R}
pHJ1	10.4	SuperCos1	2.5 kb *Bam*HI fragment from pHJS18	Amp^{R}
pHJ2	8.4	SuperCos1	0.5 kb *Bam*HI fragment from pHJS18	Amp^{R}
pHJ3	8.4	SuperCos1	0.5 kb *Bam*HI fragment from pHJS18	Amp^{R}
Cosmid Library	Ca. 48	pWEB^{TM}	Ca. 40 kb fragments from *S. thioluteus*	Amp^{R}
Shotgun Library	Ca. 5.5	pBluescript SK-	2.5 kb fragments from cosmid 18E4	Amp^{R}
pHJ*Bam*HI#	9.7~5.39	pBluescript SK-	*Bam*HI fragments from cosmid 18E4	Amp^{R}
pHJ*Sac*I#	7.08~1.2	pBluescript SK-	*Sac*I fragments from cosmid 18E4	Amp^{R}
pHJ11	Ca. 48	pSET152	Ca. 40 kb *Ssp*I fragment (Insert and Cos site) from cosmid 18E4	Am^{R}
pHJ12	Ca. 48	pSET152	Ca. 40 kb *Ssp*I fragment (Insert and Cos site) from cosmid 18E4	Am^{R}
pHJ13	Ca. 48	pAY1	Ca. 40 kb *Ssp*I fragment (Insert and Cos site) from cosmid 18E4	Am^{R}
pHJ*Eco*72I	~20	pSET152	~20 kb *Eco*72I fragment from pHJ11 and self-ligation	Am^{R}
pHJ31	3.8	pGEM-T Easy	0.8 kb PCR product (the first part of *aurA*) from pHJ*Sac*I3b (template)	Amp^{R}
pHJ32	8.7	pHJ31	4.9 kb *Bsp*1407I-*Nco*I fragment from pHJ11	Amp^{R}

pHJ33	~22	pRM5/*Pac*I-*Eco*RI	5.7 kb *Pac*I-*Nco*I fragment from pHJ32 and 0.7 kb *Nco*I-*Eco*RI fragment from pHJ34	Amp^R, $Thio^R$
pHJ34	3.7	pGEM-T Easy	0.7 kb PCR product (TE domain of *aurC*) from pHJ*Bam*HI1 (template)	Amp^R
pHJ35	10.5	pHGF7505	6.4 kb *Pac*I-*Eco*RI fragment from pHJ33	Amp^R
pHJ36	~21	pRM5/*Hin*dIII-*Eco*RI	7.9 kb *Hin*dIII-*Eco*RI fragment from pHJ35	Amp^R, $Thio^R$
pHJE72I	~35	pKJ1	~20 kb *Eco*72I fragment from pHJ11	Amp^R, $Thio^R$
pHJES	23.35	pSET152	17.85 kb *Eco*72I-*Ssp*I fragment from cosmid 18E4	Am^R
pHJ38	5.76	pBluescript SK-	2.758 kb *Pst*I fragment from cosmid 18E4	Amp^R
pHJ40	12	pHJ35	1.5 kb *Nde*I-*Spe*I fragment from plig2 (from Traitcheva)	Amp^R
pHJ41	~22	pRM5/*Hind*III-*Eco*RI	9.4 kb *Hind*III-*Eco*RI fragment from pHJ40	Amp^R, $Thio^R$
pHJ42	13.9	pHJ40/*Pac*I-*Eco*RI	5.7 kb *Pac*I-*Nco*I fragment from pHJ32 and 2.6 kb *Nco*I-*Eco*RI fragment from pHJ38	Amp^R
pHJ43	~25	pRM5/*Hin*dIII-*Eco*RI	11.3 kb *Hind*III-*Eco*RI fragment from pHJ42	Amp^R, $Thio^R$
pHJ44	~37	cosmid 18E4	1.0 kb and ~10 kb *Xho*I fragments were deletion and self-ligation	Amp^R
PHJ45	~47	cosmid 18E4	1.0 kb *Xho*I fragment was deletion and self-ligation	Amp^R
pHJ46	~36	pSET152	~30 kb *Ssp*I fragment from pHJ44	Am^R
pHJ47	~46	pSET152	~40 kb *Ssp*I fragment from pHJ45	Am^R
pHJ48	35.3	pSET152	29.8 kb *Eco*RV fragment from cosmid 18E4	Am^R
pHJ49	~45	pHJ11	3.4 kb *Sgf*I fragment was deleted from pHJ11	Am^R
pHJ50	~47	pHJ49	2.6 kb *Sgf*I fragment from pMP15 (from Palzer)	Am^R
pHJ56	~19	pHJES/*Kpn*I-*Eco*RI partial	3.2 kb *Kpn*I-*Eco*RI fragment from pHJ33	Am^R
pHJE01	20	pHJES	3.4 kb *Sgf*I fragment was deleted from pHJES	Am^R
pHJ57	4.024	pGEM-T Easy	1.024 kb PCR product (*aurF*) from pHJ*Sac*I3d (template)	Amp^R
pHJ58	4.172	pGEM-T Easy	1.172 kb PCR product (*aurF*) from pHJ*Bam*HI3a (template)	Amp^R
pHJ60	9.1	pWHM4*	1.2 kb *Eco*RI fragment from pHJ58	Amp^R, $Thio^R$
pHJ72	9.75	pWHM4*	1.85 kb *Eco*RI fragment from 9A6 (from Hertweck)	Amp^R, $Thio^R$
pHJ73	18.0	pBluescript SK-	15 kb *Kpn*I fragment from pHJ48	Amp^R
pHJ74	4.04	pGEM-T Easy	1.037 kb PCR product (TE) from pHJ*Bam*HI1 (template)	Amp^R
pHJ75	3.68	pGEM-T Easy	0.677 kb PCR product (Short) from pHJ*Bam*HI1 (template)	Amp^R
pHJ76	3.27	pGEM-T Easy	0.272 kb PCR product (Linker region of module 3 and 4) from pHJ*Bam*HI1	Amp^R
pHJ77	~48	pHJ48	Double crossover mutant (Δ*aurF*) with *aadA* cassette	Am^R, Spc^R
pHJ78	~22	pHJ41/*Pac*I-*Eco*RI	1.0 kb *Pst*I-*Eco*RI fragment from pHJ74 and 5.46 kb *Pac*I-*Pst*I fragment from pHJ41	Amp^R, $Thio^R$
pHJ79	~47	pHJ77	*addA* cassette was deleted from pHJ77	Am^R
pHJ80	5.3	pBluescript SK-	1.276 kb *Bam*HI-*Fsp*I fragment from	Amp^R

			pHJ*Bam*HI1	
			and 1.0 kb *Fsp*I-*Eco*RI fragment from pHJ74	
pHJ81	4.25	pBluescript SK-	0.246 kb *Bam*HI-*Fsp*I fragment from pHJ*Bam*HI5b and 1.0 kb *Fsp*I-*Eco*RI fragment from pHJ74	Amp^R
pHJ82	14.3	pHJ48	3.5 kb *Kpn*I-*Xba*I fragment from pHJ78	Am^R
pHJ83	14.4	pHJ73/*Sgf*I-*Pst*I partial	0.27 kb *Pst*I-*Sgf*I fragment from pHJ76	Amp^R
pHJ84	5.95	pHJ80	0.67 kb *Hind*III-*Xho*I fragment from pHJ75	Amp^R
pHJ85	4.92	pHJ81	0.67 kb *Hind*III-*Xho*I fragment from pHJ75	Amp^R
pHJ86	31.5	pHJ48/*Kpn*I	11.4 kb *Kpn*I fragment from pHJ83	Am^R
pHJ87	8.3	pHJ84	2.36 kb *Sma*I fragment from pIJ4642	Amp^R, Spc^R
pHJ88	7.28	pHJ85	2.36 kb *Sma*I fragment from pIJ4642	Amp^R, Spc^R
pHJ91	3.73	pGEM-T Easy	0.731 kb PCR product (*aurI*) from pHJ*Bam*HI1 (template)	Amp^R
pHJ92	~48	pHJ48	Double crossover mutant (Δ*aurG*) with *aadA* cassette	Am^R, Spc^R
pHJ93	~48	pHJ48	Double crossover mutant (Δ*aurH*) with *aadA* cassette	Am^R, Spc^R
pHJ94	4.53	pBluescript SK-	0.533 kb *Not*I-*Fsp*I fragment from pHJ*Bam*HI4b and 1.0 kb *Fsp*I-*Eco*RI fragment from pHJ74	Amp^R
pHJ95	8.64	pWHM4*	0.74 kb *Eco*RI fragment from pHJ91	Amp^R, $Thio^R$
pHJ96	5.2	pHJ94	0.67 kb *Hind*III-*Xho*I fragment from pHJ75	Amp^R
pHJ97	~47	pHJ48	*addA* cassette was deleted from pHJ92	Am^R
pHJ98	~47	pHJ48	*addA* cassette was deleted from pHJ93	Am^R
pHJ99	7.56	pHJ96	2.36 kb *Sma*I fragment from pIJ4642	Amp^R, Spc^R
pHJ106	4.09	pGEM-T Easy	1.086 kb PCR product (*ery*TE) from *Saccharopolyspora erythraea*	Amp^R
pHJ107	~43	pHJ48	Double crossover mutant (Δ*aurA*) with *aadA* cassette	Am^R, Spc^R
pHJ108	4.43	pGEM-T Easy	1.436 kb PCR product (*aurH*+RBS) from pHJ*Bam*HI2 (template)	Amp^R
pHJ109	4.34	pGEM-T Easy	1.344 kb PCR product (*aurH*) from pHJ*Bam*HI2 (template)	Amp^R
pHJ110	9.4	pWHM4*	1.4 kb *Eco*RI fragment from pHJ108	Amp^R, $Thio^R$
pHJ111	~42	pHJ48	*addA* cassette was deleted from pHJ107	Am^R
pHJ112	4.5	pBluescript SK-	0.533 kb *Not*I-*Fsp*I fragment from pHJ*Bam*HI4b and 1.0 kb *Fsp*I-*Pst*I fragment from pHJ106	Amp^R
pHJ113	4.25	pBluescript SK-	0.246 kb *Bam*HI-*Fsp*I fragment from pHJ*Bam*HI5b and 1.0 kb *Fsp*I-*Pst*I fragment from pHJ106	Amp^R
pHJ114	5.3	pBluescript SK-	1.276 kb *Bam*HI-*Fsp*I fragment from pHJ*Bam*HI1 and 1.0 kb *Fsp*I-*Pst*I fragment from pHJ106	Amp^R
pHJ116	5.2	pHJ112	0.67 kb *Hind*III-*Xho*I fragment from pHJ75	Amp^R
pHJ117	4.9	pHJ113	0.67 kb *Hind*III-*Xho*I fragment from pHJ75	Amp^R
pHJ118	5.95	pHJ114	0.67 kb *Hind*III-*Xho*I fragment from pHJ75	Amp^R
pHJ119	7.56	pHJ116	2.36 kb *Sma*I fragment from pIJ4642	Amp^R, Spc^R

pHJ120	7.28	pHJ117	2.36 kb *Sma*I fragment from pIJ4642	Amp^R, Spc^R
pHJ121	8.3	pHJ118	2.36 kb *Sma*I fragment from pIJ4642	Amp^R, Spc^R
pHJ125	24.4	pHJ48	Double crossover mutant (*aurA*+TE) with *aadA* cassette from pHJ99	Am^R, Spc^R
pHJ126	34.3	pHJ48	Double crossover mutant (*aurB*+TE) with *aadA* cassette from pHJ88	Am^R, Spc^R
pHJ127	37.7	pHJ48	Double crossover mutant (*aurC1*+TE) with *aadA* cassette from pHJ87	Am^R, Spc^R
pHJ128	24.4	pHJ48	Double crossover mutant (*aurA*+*ery*TE) with *aadA* cassette from pHJ119	Am^R, Spc^R
pHJ129	34.3	pHJ48	Double crossover mutant (*aurB*+*ery*TE) with *aadA* cassette from pHJ120	Am^R, Spc^R
pHJ130	37.7	pHJ48	Double crossover mutant (*aurC1*+*ery*TE) with *aadA* cassette from pHJ121	Am^R, Spc^R
pHJ133	6.9	pET-15b	1.2 kb *Nde*I-*Bam*HI fragment from pHJ109	Amp^R
pHJ134	9.4	pWHM4*	1.5 kb *Bgl*II-*Bam*HI fragment from pHJ133	Amp^R
pHJ135	3.328	pGEM-T Easy	0.328 kb PCR product (Linker *aurC1*-*aurC2*) from pHJ*Bam*HI1 (template)	Amp^R
pHJ136	7.8	pHJ73/*Acc*III-*Pst*I partial	0.328 kb *Pst*I-*Acc*III fragment from pHJ135	Amp^R
pHJ137	25	pHJ48/*Kpn*I	4.8 kb *Kpn*I fragment from pHJ136	Am^R
pHJ139	3.688	pGEM-T Easy	0.688 kb PCR product (*Eco*RI-*Afl*III) from pHJ*Bam*HI1 (template)	Amp^R
pHJ140	4.1	pGEM-T Easy	1.100 kb PCR product (*Afl*III-*Xba*I) from pHJ*Bam*HI1 (template)	Amp^R
pHJ141	4.448	pGEM-T Easy	1.448 kb PCR product (*Kpn*I-*Pvu*II) from pHJ*Bam*HI1 (template)	Amp^R
pHJ142	4.936	pGEM-T Easy	1.936 kb PCR product (*Pvu*II-*Xba*I) from pHJ*Bam*HI1 (template)	Amp^R
pHJ143	4.79	pBluescript SK-	0.688 kb *Eco*RI-*Afl*III fragment from pHJ139 and 1.1 kb *Afl*III-*Xba*I fragment from pHJ140	Amp^R
pHJ144	~45	pHJ48/*Bsu*36 I-*Xba*I	1.159 kb *Bsu*36I-*Xba*I fragment from pHJ143	Am^R
pHJ145	6.185	pBluescript SK-	1.249 kb *Kpn*I-*Pvu*II fragment from pHJ141 and 1.936 kb *Pvu*II-*Xba*I fragment from pHJ142	Amp^R
pHJ146	~48	pHJ48/*Bsu*36 I-*Kpn*I partial	3.121 kb *Kpn*I-*Bsu*36I fragment from pHJ145	Am^R

1.4 Antibiotics and Enzymes

Antibiotics:

Name and abbreviation	Stock solution (mg/ml)	Working concentration (µg/ml)				Company
		Streptomyces			*E.coli*	
		Solid	Liquid	Overlay (1 ml ddH_2O)		
Ampicillin (Amp)	100	—	—	—	100	Roth
Apramycin (Am)	50	30	30	1 mg	50	Sigma

Kanamycin (Km)	50	25	25	0.5 mg	50	Sigma
Thiostrepton(Thio)	25	25	5	0.1 mg	—	Calbiochem
Chloramphenicol (Cml)	25	—	—	—	25	Roth
Spectinomycin (Spc)	50	20	—	—	50	Sigma
Nalidixic acid	50	—	—	1 mg	—	Sigma

—: undone or no record.

Enzymes:

Enzyme	Reaction Buffer	Company
*Acc*III	F	Promega
*Afl*III	NEBuffer 3 + BSA	New England Biolads
*Bam*HI	BamH I NEBuffer + BSA	New England Biolads
*Bgl*II	NEBuffer 3	New England Biolads
*Bsu*36I	NEBuffer 3 + BSA	New England Biolads
*Eco*72I	Buffer Tango	Fermentas
*Eco*RI	EcoR I NEBuffer	New England Biolads
*Eco*RV	NEBuffer 3 + BSA	New England Biolads
*Fsp*I	NEBuffer 4	New England Biolads
*Hind*III	NEBuffer 2	New England Biolads
*Kpn*I	NEBuffer 1 + BSA	New England Biolads
*Nco*I	NEBuffer 4	New England Biolads
*Nde*I	NEBuffer 4	New England Biolads
*Not*I	NEBuffer 3 + BSA	New England Biolads
*Pac*I	NEBuffer 1 + BSA	New England Biolads
*Pst*I	NEBuffer 3 + BSA	New England Biolads
*Pvu*II	NEBuffer 2	New England Biolads
*Sac*I	NEBuffer 1 + BSA	New England Biolads
*Sac*II	NEBuffer 4	New England Biolads
*Sma*I	NEBuffer 4	New England Biolads
Ssp I	Ssp I NEBuffer	New England Biolads
*Swa*I	NEBuffer 3 + BSA	New England Biolads
*Xba*I	NEBuffer 2 + BSA	New England Biolads
*Xho*I	NEBuffer 2 + BSA	New England Biolads
T4 DNA Ligase	1 x T4 DNA Ligase Buffer	New England Biolads
Shrimp Alkaline Phosphatase (SAP)	1 x SAP Buffer	Promega
Proteinase K	P buffer or TE25S	Promega
Taq DNA Polymerase	1 x PCR Buffer	Amersham

1.5 PCR Primers

Primer	Generation of	Sequence (5'— 3')
KSF1	*N*-terminal region of KS	MGNGARGCNNWNSMNATGGAYCCNCARCANMG
KSR1	*C*-terminal region of KS	GGRTCNCCNARNSWNGTNCCNGTNCCRTG
PABAF1	*N*-terminal region of *pab*AB	GACAACTACGACWSSTTCACS
PABAR1	*C*-terminal region of *pab*AB	CCTTSRKCTCGTAGCCSAG
PABAF2	*N*-terminal region of *pab*A	GACAACTACGACTCGTTCACC
PABAR2	*C*-terminal region of *pab*A	GACGATCATCAGGTTCTCG
PABAF3	*N*-terminal region of *pab*AB	CTSGGCGTSTGCCTSGGC
PABAR3	*C*-terminal region of *pab*AB	TGTTSGTSAGGCASAYCTCG
TEL1	*N*-terminal region of TE	GCGG*CCATGG*GCGCGAAACTGATCTGCTTCCCG *Nco*I
TER1	*C*-terminal region of TE	GTGGTGGTCATTTTGTTCCTGTCG
PrimerFL1	*N*-terminal region of *aurF*	*GAATTCCATATG*CGAGAAGAGCAGCCGCACT *Eco*RI *Nde*I
PrimerFL2	*N*-terminal region of *aurF*	GAGCGG*GAATTCCATATG*CCACGACACCGCGGG *Eco*RI *Nde*I
PrimerFR	*C*-terminal region of *aurF*	GG*GGATCC*ACGCGGCGTCGGGGTCAACG *Bam*HI
aurFL	Deletion of *aurF*	CCGCGGTGGGCACGGGCGACGAGGTCGGGAGAGA CGATGATTCCGGGGATCCGTCGACC
aurFR	Deletion of *aurF*	CGGGCTCCGGTCGTGGAGGTCGACGCGGCGTCGG GGTCATGTAGGCTGGAGCTGCTTC
aurATEL	*N*-terminal region of TE	CG*CTGCAG*C*TGCGCA*ACCGCCTCGGGGC *Pst*I *Fsp*I
aurATER	*C*-terminal region of TE and inactivation of ACP4	GTG*TCTAGA*CGTGCTTGCGCTGGTGGTGGTC *Xba*I
LinkerL	Linker region of module 3 and 4	GCG*CTGCAG*CTGCGCAACCGGCTCGACG *Pst*I
LinkerR	Linker region of module 3 and 4	*GCGATCGC*GTCGCGTCCCGAGGCGACCA *Sgf*I
LinkerR2	Linker region of module 3 and 4	*TCCGGATCC*GGGTCGTAGAGGCCCTCGAC *Acc*III *Bam*HI
ShortL	Part of module 4	CGC*AAGCTT*CACCCCGTCCTGGCACTGA *Hind*III
ShortR	Part of module 4	GCC*CTCGAG*AGTTCGGGGCGCTCCTCCAC *Xho*I
PrimerA	Sequencing of *aurA*+TE fusion region	GACGACACCCTGGCCGACCGCCTG
PrimerIL	*N*-terminal region of *aurI*	CCCGACACCGACAGGAACAAAATG

PrimerIR	*C*-terminal region of *aurI*	CGCTCCCCCGGTATTTGTCAG
aurGL	Deletion of *aurG*	AGGACGACCCCGCTCGAACCGACAGGGAGGCTCC CGATGATTCCGGGGATCCGTCGACC
aurGR	Deletion of *aurG*	CTTGAAGCCCGTCGAACGACGGGCGGAACGGCCGT ACTATGTAGGCTGGAGCTGCTTC
aurHL	Deletion of *aurH*	CGGCAGCGCGCCCGTACGCCAACGGAAGGATCGA ACATGATTCCGGGGATCCGTCGACC
aurHR	Deletion of *aurH*	GACGGTGGCGCCGCACCCCCGGCGGCCCGGGAAA CGTCATGTAGGCTGGAGCTGCTTC
aurAL	Deletion of *aurA*	GACCGGCCGCCTCACCCTGCCAGGAAGGAAGCGAT GATGATTCCGGGGATCCGTCGACC
aurAR	Deletion of *aurA*	ATCGGGAGCCTCCCTGTCGGTTCGAGCGGGGTCGT CCTATGTAGGCTGGAGCTGCTTC
PrimerHL1	*N*-terminal region of *aurH*	GCCGTTCCGCCCGTCGTTCG
PrimerHL2	*N*-terminal region of *aurH*	*GAATTCCATATG*TCGACCACCGCTCACACCG *Eco*RI *Nde*I
PrimerHR	*C*-terminal region of *aurH*	*GGATCC*ATTCGAGCGTCTTGGCGTCATTGG *Bam*HI
PrimerEryL	*N*-terminal region of *ery*TE	CGGGAGATGACGTCGCAGGAGTTG
PrimerEryR	*N*-terminal region of *ery*TE	GTCGTGGTCATGAATTCCCTCCGCCC
forKpnI	Inactivation of KS4	GCCGCCGCCGAGGATGAAGG
PvuIIrev	Inactivation of KS4	GG*CAGCTG*TGTCGACCGTGATGGCCGGGCCTTC *Pvu*II
PvuIIfor	Inactivation of KS4	A*CAGCTG*CCTCCTCGTCCCTCGTCGCCCTCCAC *Pvu*II
XbaIrev	Inactivation of KS4	*TCTAGA*TCGGGGAGCGGGGTTTCGTC *Xba*I
EcoRIfor	Inactivation of ACP4	*GAATTC*GTGGCGCTGTACCGGCTGTTGG *Eco*RI
AflIIIrev	Inactivation of ACP4	GTCA*ACGCGT*CGAAGCCCAGCTCGGGGAGC *Afl*III
AflIIIfor	Inactivation of ACP4	TCG*ACGCGT*TGACCGCGGTGGACCTGCGCAAC *Afl*III

1.6 Special Devices

	Equipment	Company
Pippete	Research R variabel 3111	Eppendorf
	Mehrkanalpipette Research R multi 3114	Eppendorf
	EasyPet	Eppendorf
	Multipette plus	Eppendorf

Vortex-Genie2		Scientific Industries
Centrifuge	Minizentrifuge "Dual Rotor"	Qualitron Inc.
	Centrifuge 5415 R	Eppendorf
	Centrifuge 5415 D	Eppendorf
	Centrifuge 5810 R	Eppendorf
	Sorvall R RC 5C	Sorvall
pH-Meter		Schott
pH-Eletrode BlueLine 14		Schott
BioPhotometer		Eppendorf
Speedvaccum Concentrator 5301		Eppendorf
Electrophoresis chambers		Amersham and Biorad
Power supply	Electrophoresis Power Supply EPS 301	Amersham
	PowerPac Basic TM	Biorad
Analytical balances	BP3100S	Sartorius
	BP211D	Sartorius
Orbital incubator	Multitron	INFORS
96 well-plate-shaker		Edmund Bühler
Water bath 1083		GFL
PCR machine	Mastercycler Gradient	Eppendorf
	Mastercycler Personal	Eppendorf
CL-1000 Ultraviolett-Crosslinker		UVP
Heat/Agitation plate		Heidolph
Gel documentation analysis	Gene Genius	Synene
Water purification	Purelab TM plus	USF
Incubator	B 6120	Heraeus Instruments
Hybridization oven		Binder
Rotation evaporator		Heidolph
Rocker shaker	Polymax 1040	Heidolph
Blowtorch Burner	Gasprofi 1	WLD-Tec
Digital photographic system	Folio	Severin
Microflow (Biological Safety Cabinet)		NUNC
Thermo incubator	Thermomixer comfort	Eppendorf
	ThermoStat plus	Eppendorf
	Block Heater	Stuart Scientific
UV-lamp		FAUST
Foto-Analysis	Reprostar 3	CAMAG
Camera	HV-C20	Hitachi
E. coli Pluser		Biorad

2. Methods

2.1 Cultivation of *E. coli* Cells

A single colony was always inoculated to 3 ml of LB liquid medium or a LB agar plate and incubated at 37°C overnight for isolating plasmid DNA or detection of CCC plasmid size. For selection of correct clones, different antibiotics were added to media at the appropriate concentration. *E. coli* strain DH5α and strain ET12567 containing the RP4 derivative pUZ8002 served as hosts for routine subcloning and intergeneric conjugation, respectively. *E. coli* cells were concentrated by centrifugation and resuspended in 20% glycerol for preservation at –20°C.

2.2 Growth and Preservation of *Streptomyces* Strains

Streptomycetes were inoculated over the entire surface of the medium on agar in order to obtain confluent cultures. After 10-14 days growth, the grey hydrophobic spores were scraped from the surface of the plate and resuspended in 20% glycerol for preservation at –20°C. To harvest high quality mycelium, spores were inoculated in liquid cultures and incubated for 36-40 hours at 30°C with vigorous shaking.

2.3 Amplification of DNA Fragments by PCR

Typical Polymerase Chain Reaction (PCR) was used for amplification of DNA fragment for cloning, expression or sequencing. Below is a list of reagents for the basic PCR system.

Forward primer (25 μM) – 1 μl
Reverse primer (25 μM) – 1 μl
PCR buffer (10 ×) – 2 μl
dNTPs mix (10 mM for each) – 1 μl
BSA solution (10 mg/ml) – 1 μl
DMSO (100%) – 1 μl
Sterile purified water – 12 μl
Template DNA (100-200 ng/μl) – 0.5 μl
Taq DNA polymerase (5 units/μl) – 0.5 μl
Total reaction volume 20 μl

The nucleic acid sequences were amplified by the cycles of denaturation, annealing and polymerization. The details of PCR program were shown below.

1.	Denaturation:	94°C, 2 min	
2.	Denaturation:	94°C, 30 sec	34 cycles
3.	Primer annealing:	57°C, 30 sec	
4.	Extension: 68°C,	1 min for 1 kb DNA	
5.	Final extension:	68°C, 4 min	

2.4 Purification of DNA Fragments from Solutions or Agarose Gel (Amersham)

GFX PCR DNA and Gel Band Purification Kit (Amersham) can be used to purify DNA (e.g., PCR products, restriction fragments) from solution and from TAE and TBE agarose gel bands.

For purification of DNA from gel bands, the gel slice containing the desired DNA fragment was weighted and treated with the same volume of capture buffer. For example, 300 mg of gel slice was afforded 300 µl of capture buffer. The mixture was incubated at 60°C until the agarose was completely dissolved. After brief centrifugation to collect the sample at the bottom of the tube, it was transferred to one GFX column in a collection tube. After centrifugation in a microcentrifuge at full speed for 30 seconds, the flow-through was discarded and the column was washed with 500 µl of wash buffer by centrifugation at full speed for 30 seconds. The GFX column was transferred to a fresh 1.5 ml microcentrifuge tube and DNA was eluted with 10-50 µl of purified water at room temperature for 1 min. Finally, purified DNA was recovered by centrifugation at full speed for 1 min.

For purification of DNA from solution, the DNA solution (up to 100 µl) was mixed with 500 µl of capture buffer and transferred to the GFX column for centrifugation at full speed for 30 seconds. The following procedures for washing and elution are the same as the above.

2.5 Cloning of PCR Products with the pGEM-T Easy Vetor System (Promega)

The pGEM-T Easy Vector was prepared by cutting with *Eco*RV and adding a 3' terminal thymidine to both ends. These single 3'-T overhangs at the insertion site greatly improve the efficiency of ligation of a PCR product into the plasmid by providing a compatible overhang for PCR products generated by certain thermostable polymerases. The ligation reaction was set up as described below.

Rapid Ligation Buffer (2 ×) – 5 µl	
pGEM-T Easy Vector (50 ng) – 1 µl	
PCR product (~200 ng/µl) – 3 µl	
T4 DNA Ligase (3 units/µl) – 1 µl	
Total volume	10 µl

The reaction was mixed by pipetting and incubated at room temperature for 1 hour or at 4°C overnight to get the maximum number of transformants. Half of the volume of the reaction

was used for transformation of competent *E. coli* cells and transformants were selected by blue/white screening. 40 µl X-gal (20mg/ml) and 40 µl IPTG (200mg/ml) solutions were spread on the agar plate containing appropriate antibiotics prior to the transformation. After incubation at 37°C for overnight, the desired clones which do not contain active β-galactosidase due to the insertion of foreign DNA showed creamy-white, while the transformants only containing vectors were pale blue.

2.6 Preparation High Quality Plasmid DNA from *E. coli* (Birnboim *et al*., 1979)

1.5 ml of the overnight culture was decanted into an Eppendorf tube and spun in a benchtop centrifuge at 13,000 rpm for 1 min. After completely removing the supernatant, the cell pellet was carefully resuspended in 150 µl of solution I. 300 µl of freshly prepared solution II was added and the tube was inverted several times. Then 230 µl of precooled solution III was mixed into the reaction solution and inverted several times without vortexing. After centrifugation at 13,000 rpm for 10 min, the supernatant was removed to a clean tube and 80 µl of neutral phenol/chloroform was added to remove proteins. After centrifugation, the aqueous supernatant was transferred into a new tube containing 800 µl of isopropanol, mixed by inversion and centrifuged for 10 min. The DNA pellet was washed in 70% ethanol, dried under reduced pressure and dissolved in 40 µl TE buffer.

2.7 Introduction of DNA into *E. coli*

2.7.1 Preparation and Transformation of Competent Cells by the Chemical Method

A single colony was inoculated in 3 ml of LB medium for overnight shaking. On the next day, 1 ml of the overnight culture was transferred into 100 ml of fresh LB broth and incubated at 37˚C until the OD_{560} reached 0.4-0.5. *E. coli* cells were harvested by centrifugation at 5, 000 rpm for 6 min at 4°C. The pellet was gently resuspended in 25 ml of ice-cooled TFBI buffer. After centrifugation at 5,000 rpm for 5 min at 4°C, cells were resuspended in 4 ml of TFBII buffer. Aliquots of 100 µl competent cells were dispensed into sterile Eppendorf tubes and stored at –80°C. Up to 20 µl DNA solution was added to one tube containing thawed competent cells. After incubation in ice for ca. 20 min, cells were heat shocked at 37°C for 2 min (or 42°C for 90 sec). The tube was cooled down in ice a little bit and added 0.9 ml of LB broth. After incubation at 37°C for about 1 hour, cells were spread on selective plates for incubation overnight at 37°C.

2.7.2 Transformation of *E. coli* Cells by Electroporation (Chassy *et al*., 1988)

A single colony of *E coli* was inoculated from a fresh agar plate into 50 ml of LB medium and cultured overnight at 37°C with vigorous aeration (200 rpm). 1000 ml of prewarmed LB

medium was inoculated with 25 ml of the overnight bacterial culture and incubated at 37°C with agitation until OD_{600} reached ~0.4. The culture was transferred to several ice-cold centrifuge tubes and cooled down on ice for 15-30 min. The cells were harvested by centrifugation at 2,500 rpm for 15 min at 4°C. The supernatants were decanted and the cell pellets were resuspended in 500 ml of ice-cold pure water. After centrifugation at 2,500 rpm for 20 min at 4°C, the cells were washed three times with 500 ml, 250 ml and 10 ml of ice-cold 10% glycerol in turn. Cells were harvested by centrifugation at 2,500 rpm for 20 min at 4°C and all liquid was carefully removed. Finally, the pellet was resuspended in 2 ml of ice-cold 10% glycerol. Aliquots of 50 µl of the suspension were dispensed into sterile Eppendorf tubes and stored at –80°C. For transformation reaction, a volume of 1-2 µl DNA solution was added to 50 µl of the freshly thawed electrocompetent cells and mixed by pipette for several times. The mixture was transferred to the bottom of an ice-cold electroporation cuvette avoiding bubbles. The outside of the cuvette was dried with tissue before electroporation. A pulse of electricity was delivered to the cells at 25 µF capacitance, 2.5 kV and 200 Ohm resistance. As quickly as possible after the pulse, 1 ml of LB medium was added at room temperature and the mixture was transferred to a sterile Eppendorf tube and incubated with gentle rotation for 1 hour at 37°C. Different volumes of the culture were spread onto a LB agar plate containing appropriate antibiotic for selection of clones.

2.8 Isolation of Genomic DNA from *Streptomyces* (Hopwood *et al*., 1985)

Mycelium from a 25 ml culture was resuspended in the mixture of 4 ml of TE25S buffer and 500 µl of 20 mg/ml lysozyme solution (final concentration of lysozyme is 2 mg/ml). The mixture in a 50 ml tube was incubated at 37˚C for 2-3 hours and turned several times in every 30 min until the suspension became milky. After adding 2 ml of 10% SDS, it was incubated for 30 min at 37˚C again and turned over tubes every 5 min until the suspension became clear. 3 ml phenol/chloroform was added, vortexed 30 seconds and centrifuged for 10 min. The procedure of precipitation with phenol/chloroform was repeated when separation of proteins was not optimal. The upper phase was transferred with “wide mouth” tips into 15 ml tubes containing 700 µl of 3 M unbuffered sodium acetate and 5 ml of isopropanol. The tube was gently swirled and DNA became visible. DNA was spooled onto a sealed Pasteur pipette and washed with 70% ethanol. The DNA pellet was dried on air and dissolved in 500 µl TE at 55˚C for several hours or at 37˚C overnight.

2.9 Plasmid DNA Isolation from *Streptomyces* (Kieser, 1984)

Mycelium from a ~5 ml culture was resuspended in a total volume of ~500 µl TE25S buffer supplemented with 2 mg/ml lysozyme and 5 µg/ml pre-boiled RNase A, and incubated at 37°C for 1 hour. 250 µl of 0.3 M alkaline SDS solution (0.3 M NaOH, 2% SDS) was added

and mixed immediately by vortexing (larger volumes are best mixed using a disposable syringe). Tubes were opened and incubated for 15 min at 70°C (or 30 min at 55°C for plasmids ≥ 20 kb). When the tubes were cooled down to < 37°C, 80 µl of acid phenol/chloroform was added. The mixture was emulsified by vortexing and centrifuged for 10 min. The supernatant (ca. 700 µl) was transferred to a new centrifuge tube containing the same volume of isopropanol and 50 µl unbuffered 3 M sodium acetate, mixed and centrifuged 10 min; all liquid was removed by a pipette and the precipitate was washed with 70% ethanol for 15 min. It was dried at room temperature and dissolved in appropriate volume of TE.

2.10 Introduction of DNA into *Streptomyces*

2.10.1 Protoplast Transformation of *Streptomyces* (Kieser *et al*., 2000)

100 µl of spore suspension (10^{10} spores/ml) was inoculated in 50 ml of YEME medium supplemented with required growth factors and incubated at 30°C for 36-40 hours in an orbital incubator shaker (200 rpm). Mycelium was harvested by centrifugation at 3,000 rpm for 10 min and washed with 30 ml of 10.3% sucrose two times. The mycelium pellet was resuspended in 8 ml P buffer supplemented with 1 mg/ml lysozyme and incubated at 30 °C for 15-60 min. When the supernatant became turbid, the mycelium suspension was gently drawn in and out three times by a 5 ml pipette and incubated for a further 15 min. After adding 5 ml P buffer, the mixture was drawn in and out again to help the release of protoplasts from the mycelium. The protoplast solution was filtered through cotton wool and transferred to a new sterile tube. After gently centrifugating at 2,500 rpm for 7 min, the protoplast pellet was resuspended in 300-1000 ml P buffer. Aliquots of 50 µl protoplast suspension were dispensed into tubes and stored at -80°C. Each tube was used for one transformation reaction. 5 µl of DNA solution was added to 50 µl thawed protoplast solution and mixed immediately with 200 µl T buffer by pipetting up and down several times. The mixture was spread on one dried R5 plate and incubated at 30°C. After 14-20 hours, the plate was overlaid with 1 ml sterile purified water containing appropriate amount of selective antibiotics. Resistant colonies became visible after 3 days of incubation at 30°C.

2.10.2 Intergeneric Transfer of Plasmids from *E. coli* to *Streptomyces* by Conjugation (Flett *et al*., 1997)

The *oriT*-containing construct was introduced into the competent cells of *E. coli* ET12567/pUZ8002 and incubated on LB agar containing 25 µg/ml kanamycin (selection for pUZ8002), 25 µg/ml chloramphenicol (selection for the *dam* mutation), and appropriate antibiotics used to select for incoming plasmid only. A single colony of the transformant was inoculated into 3 ml LB medium containing antibiotics and grown overnight. Next day, 10 µl of

the overnight bacterial culture was transferred to 1 ml fresh LB plus antibiotics and grown at 37°C until OD_{600} reached 0.4-0.6. At the same time, approximately 10^8 *Streptomyces* spores for each conjugation were suspended into 500 µl of TES buffer (0.05 M TES, pH 8.0) and heat shocked at 50°C for 10 min. After cooling down, 500 µl of double strength germination medium was added and the mixture was incubated at 37°C for 2-3 hours. 1 ml of the *E coli* cells was washed three times with fresh LB medium without any antibiotic and mixed with the spores after germination. After brief spin, most of the supernatant was poured off. The pellet was resuspended in the residual liquid and plated on suitable agar medium. After incubation at 30°C for 16-40 hours, the plate was overlaid with I ml water containing 1 mg nalidixic acid and appropriate plasmid selective antibiotics and continuously incubated at 30°C for 3-5 days until potential exconjugants grew up.

2.11 Construction of Cosmid Library (EPICENTRE)

2.11.1 Insert DNA Preparation and End-Repair Reaction

35 µg of the high molecular weight DNA (~0.5 µg/µl) was randomly sheared to 30-45 kb molecules by *Sau*3AI partial digestion and then concentrated by precipitation. The sheared DNA was treated with End-Repair Enzyme Mix to generate blunt ends for cloning into the prepared pWEB vector. The volumes listed below are for one end-repair reaction.

End-Repair Buffer (10 ×) – 8 µl
dNTP Mix (2.5 mM) – 8 µl
ATP (10 mM) – 8 µl
Sterile water – 22 µl
Sheared insert DNA – 30 µl
End-Repair Enzyme Mix – 4 µl
Total reaction volume 80 µl

The reaction was at room temperature for 45 min. After adding gel loading buffer, it was incubated at 70˚C for 10 min to inactivate the End-Repair Enzyme Mix.

2.11.2 Size Selection of Insert DNA

A 1% low melting point (LMP) agarose gel in 1 × TAE buffer was prepared without ethidium bromide. Following completion of electrophoresis, the outer lanes of the gel containing the T7 DNA marker were cut off. The T7 marker lanes were stained with ethidium bromide and the DNA was visible with UV light. The position of the T7 DNA marker was marked by using a pipet tip. The gel was reassembled and a 2-4 mm wide gel slice was excised, which contained insert DNA migrating between the T7 DNA marker. It was transferred into a clean tube and stored at -20˚C.

2.11.3 In-Gel Ligation

The 2-4 mm LMP gel slice was melted at 70˚C for 10-15 min, and then the tube was cooled to 45˚C immediately and incubated for 5 min. In a separate tube at room temperature, the following reagents were combined in the order listed and mixed thoroughly after each addition.

Sterile water – 15.5 µl
Fast-Link™ Ligation Buffer (10 ×) – 5 µl
ATP (10 mM) – 2.5 µl
pWEB Vector – 1 µl
Molten agarose – 25 µl
Fast-Link DNA Ligase – 1 µl
Total reaction volume 50 µl

The mixture was incubated at room temperature for 2 hours. Then the reaction was heated at 70˚C for 10 min to melt the gel mixture and to inactivate the Fast-Link Ligase. 1 µl (1U) of GELase™ Agarose Gel-Digesting Enzyme Preparation was then added and mixed thoroughly (leaving the tube at 45˚C). The tube was kept at 45˚C for 15 min to digest the agarose, and then heated to 70˚C for 10 min to inactivate the GELase enzyme.

2.11.4 *In Vitro* Packaging

The day before performing the packaging reactions, a single colony of EPI100 cells was inoculated into 50 ml of LB broth supplemented with 10 mM $MgSO_4$ and shaken overnight at 37˚C. In the next day, 5 ml of the overnight culture was transferred into 50 ml of supplemented fresh LB broth and shaken at 37˚C until the OD_{600} reached 0.8-1.0. The cells could be stored at 4˚C up to 72 hours.

10 µl of the ligated cosmid DNA was pipetted into a tube containing 25 µl of thawed packaging extract and incubated at 30˚C for 90 min. At the end of the incubation, an additional 25 µl of thawed extract was added to the mixture at 30˚C for an additional 90 min. Subsequently, 500 µl of phage dilution buffer (10 mM Tris-HCl (pH 8.3), 100 mM NaCl, 10 mM $MgCl_2$) and 25 µl of chloroform were added and mixed by gentle vortexing.

To determine the titer of the packaged cosmids, 10 µl of the packaged cosmids was added to 100 µl of prepared EPI100 host cells at room temperature for 20 min and spread on LB-ampicillin selection plates at 37˚C overnight. Count colonies and calculate the titer.

$$\frac{(\#\text{ of colonies})\,(\text{dilution factor})\,(1000\ \mu l/ml)}{(\text{volume of phage plated }(\mu l))} = \frac{(600\ \text{cfu})\,(1)\,(1000\ \mu l/ml)}{(10\ \mu l)} = 6 \times 10^4\ \text{cfu/ml}$$

The total number of clones that could be obtained with the cosmid library was calculated by multiplying the titer (cfu/ml). The number of clones required to ensure that any given DNA sequence will be found in the cosmid library varies with the size of the genome.

$N = \ln (1-P) / \ln (1-f)$

Where P is the desired probability (expressed as a fraction); f is the proportion of the genome contained in a single clone; and N is the required number of cosmid clones. For example, the number of clones required to ensure a 99.9% probability of a given DNA sequence of *Streptomyces* is contained within a cosmid library composed of 40 kb inserts is:

$N = \ln (1-0.999) / \ln (1- (4 \times 10^4 \text{ bases} / 8 \times 10^6)) = -9.21/ -0.005 = 1842$ clones

Finally, a sufficient number of clones were inoculated into 22 ninety-six-well plates with 100 µl of LB medium containing ampicillin. After shaking at 37˚C overnight, 100 µl of 40% glycerol was added to the 100 µl culture of each clone for stored at -80˚C.

2.12 Construction of a Random Shotgun Library for Sequencing

2.12.1 Random Incision of Cosmid DNA by Sonication

~8 µg cosmid DNA was dissolved in 400 µl of sterile purified water and aliquots of 100 µl cosmid DNA solution were dispensed into 1.5 ml Eppendorf microcentrifuge tubes. First three tubes of sample DNA were used to find the suitable setting conditions of the sonicator (Bandelin). Sizes of sheared DNA fragments were monitored by electrophoresis. When most of the DNA fragments ranged between 0.5-3 kb range, the setting (50% energy, 60% cycle, 10 seconds) was used again for the fourth tube of the DNA sample.

2.12.2 Blunt End-Repair Reaction

Four tubes of sheared DNA solution were mixed together and precipitated with 40 µl of 3 M sodium acetate (pH 5.2) and 800 µl of ethanol. The pellet was dissolved in 30 µl sterile purified water and treated with End-Repair Enzyme Mix (EPICENTRE) to generate blunt ends. The following solutions were combined in the order listed and mixed thoroughly.

End-Repair Buffer (10 ×) – 8 µl
dNTP Mix (2.5 mM) – 8 µl
ATP (10 mM) – 8 µl
Sterile water – 22 µl
Sheared insert DNA – 30 µl
End-Repair Enzyme Mix – 4 µl
Total reaction volume 80 µl

The reaction was incubated at room temperature for 1 hour. After adding gel loading buffer, the mixture was incubated at 70˚C for 10 min to inactivate the End-Repair Enzyme Mix.

2.12.3 Size Selection of Sheared DNA Fragments

Different sizes of sheared DNA were separated by electrophoresis with a 0.8% agarose gel in 1 x TAE buffer without ethidium bromide. After completion of electrophoresis, the agarose

gel was stained with fresh 1% ethidium bromide solution for half an hour until the DNA was visible under UV light. 1.0-1.5 kb fragments were excised with a 2-4 mm wide gel slice and recovered from agarose gel by the GFX PCR DNA and Gel Band Purification Kit (Amersham).

2.12.4 Ligation with Sequencing Vector DNA

The sequencing vector pBluescript SK (-) was prepared in advance by digestion with *Eco*RV and dephosphatization with SAP. The ligation system was established by mixing the following reagents at 4˚C.

T4 DNA Ligation Buffer (10 ×) – 2 µl
pBluescript SK (-)/*Eco*RV/SAP (~100 ng) – 1 µl
Sterile water – 6 µl
Sheared insert DNA (~400 ng) – 10 µl
T4 DNA Ligase – 1 µl
Total reaction volume 20 µl

The reaction was incubated at 14˚C for 12-16 hours.

2.12.5 Transformation of *E. coli* Cells

Aliquots of 2.5 µl ligation product were introduced into *E. coli* DH5α competent cells and the transformants were spread on LB agar plates overlaid with 50 µl of 20 mg/ml IPTG and 50 µl of 20 mg/ml X-gal. After incubation at 37˚C for overnight, only white colonies were selected to check the size of plasmids.

2.13 Southern Hybridization

2.13.1 Capillary Transfer and Fixation of DNA (Sambrook *et al*., 2001)

After completion of electrophoresis, the agarose gel containing DNA samples was stained with ethidium bromide and photographed. The gel was immersed in 0.25 M HCl solution for 2 × 15 min at room temperature with constant gentle agitation until bromophenol blue indicator turned yellow. After washing the gel with purified water for several seconds, it was soaked in denaturation solution for 2 × 15 min with gentle agitation. Then the gel was briefly rinsed in purified water and soaked in neutralization buffer for 2 × 20 min with gentle agitation. Just before transfer, the gel was soaked in transfer buffer (10× –20× SSC) for several minutes with gentle agitation. A nylon membrane (Hybond-N+, Amersham) was cut ~1 mm larger than the gel in each dimension. Two sheets of thick blotting paper were also cut to the same size as the membrane. The membrane was floated on the surface of a dish of purified water until it wetted completely from beneath, and then immersed in 2 × SSC for at least 5 min before use. The following items were put on a smooth, flat surface in the order listed:

A 3 mm Whatman paper saturated with transfer buffer
Nylon membrane prewet and saturated with transfer buffer
3 piece of 3 mm Whatman paper saturated with transfer buffer
Paper towels
A light weight to keep all layers compressed.

The gel was surrounded with plastic wrap or parafilm to prevent the two layers of gel blot paper from coming in contact with each other. The transfer of DNA usually proceeded within 8-24 hours. The DNA was fixed to the membrane by UV-crosslinking (2 min, 120,000 microjoules per cm^2).

2.13.2 Labeling of the Probes (DIG DNA Labeling and Detection Kit, Roche Molecular Biochemicals)

DNA was labeled with Digoxigenin-11-dUTP using a mixture of random hexamers, a dNTP mix containing alkali-labile Digoxigenin-11-dUTP and labeling grade Klenow enzyme according to the manufacture's manual. 10 ng-3 µg DNA was denatured at 99°C for 10 min and quickly chilled into ice. In ice, the followings were mixed and centrifuged briefly.

Freshly denatured probe DNA – 15 µl
Hexanucleotide Mix (10 ×) – 2 µl
dNTP Labeling Mix – 2 µl
Klenow enzyme labeling grade – 2 µl

The reaction was incubated at 37°C for up to 20 hours and stopped by adding 2 µl of 0.2 M EDTA (pH 8.0) and/or by heating to 65°C for 10 min.

2.13.3 Hybridization (DIG DNA Labeling and Detection Kit, Roche Molecular Biochemicals)

The membrane was prehybridized at 65-68°C for several hours in 20 ml hybridization buffer per 100 cm^2 membrane with gentle agitation. Before hybridization, DIG-labeled DNA probe was denatured at 99°C for 10 min and rapidly cooled on ice. The denatured DIG-labeled DNA probe was mixed with pre-heated hybridization buffer (3.5 ml/100 cm^2 membrane) avoiding bubbles, which may lead to background signals. The old prehybridization solution was exchanged for the probe/hybridization mixture. The membrane was incubated with this mixture at least 6 hours to overnight with gentle agitation at 65-68°C.

2.13.4 Immunological Detection (DIG DNA Labeling and Detection Kit, Roche Molecular Biochemicals)

For stringency washes, the membrane was briefly rinsed in ample 2 × SSC, 0.1% SDS for 2 × 5min at 15-25°C under constant agitation and then in 0.5 × SSC, 0.1% SDS (prewarmed to wash temperature) at 65-68°C for 2 ×15min under constant agitation.

After hybridization and stringency washes, the membrane was briefly rinsed for 1-5 min in washing buffer and then incubated for 30 min in 100 ml blocking solution at room temperature. After pouring off the blocking solution, the membrane was incubated again for 30 min in 20 ml antibody solution. After washing with 100 ml washing buffer for 2 × 15min, it was equilibrated in 20 ml detection buffer for 2-5min and then incubated in 10 ml freshly prepared colorsubstrate solution in the dark. When the desired spot or band intensities were achieved, the color reaction was stopped by washing the membrane for 5 min with 50 ml ddH_2O or TE. The results were documented by photocopying the wet filter or by photography.

2.14 Screening the Genomic Cosmid Library by PCR

20 µl of cell suspension from each well in the same vertical row was taken out and mixed together to form a row pool. In one ninety-six-well plate, 12 row pools were built up and the plate pool was composed of 16 µl of cell suspension from each of these row pools. For the *S. thioluteus* genomic library, 22 plate pools from 22 ninety-six-well plates were centrifuged at 13,000 rpm for 2 min. Every cell pellet was washed with purified water twice and then resuspended in 96 µl purified water for use in a PCR reaction with the Triple Master Mix Kit (Eppendorf).

Master-Mix 1:

Forward primer (25 µM) – 4.4 µl
Reverse primer (25 µM) – 4.4 µl
PCR Buffer HiFi (10 ×) – 11 µl
DMSO (100%) – 13.8 µl
Purified water – 48.9 µl
Total volume 82.5 µl

For 20 µl of PCR reaction system, aliquots of 6 µl Master-Mix 1 were put into one 500 µl PCR-tube and mixed with 2 µl of template DNA. The mixture was incubated at 98°C for 1 min and then cooled down on ice.

Master-Mix 2:

PCR Buffer HiFi (10 ×) – 16.5 µl
dNTP Mix (10 mM) – 5.5 µl
Purified water – 141.4 µl
Triple Mastermix Polymerase – 1.6 µl
Total volume 165 µl

After 12 µl of Master-Mix 2 was added to each PCR system, PCR-tubes were immediately transferred to the preheated (94°C) PCR machine to start the programme.

If PCR result was positive in some plate, 12 row pools of this plate were used as templates for PCR screening again. The same strategy was also performed for positive row pools until all positive colonies were detected. The ketosynthase (KS)-specific degenerate primers KSF1 and KSR1 were used to screen the whole library first. One pair of degenerate primers PABAF1 and PABAR1 was used for detection of the presence of *pabAB* homologs in all KS-positive colonies.

2.15 Gene Knock-out by PCR Targeting System (Plant Bioscience Limited)

For each gene disruption, two long PCR primers (58 nt and 59 nt) were designed according to the sequence of the target gene and the antibiotic resistance cassette. Each of the two primers has at the 5' end 39 nt matching the flanking sequence of target gene, and a 3' sequence (19 nt or 20 nt) matching the right or left end of the antibiotic resistance cassette. With these two long primers, the antibiotic resistance cassette (e.g. streptomycin and spectinomycin resistance gene, *aadA*) flanked by FRT sites (FLP recognition targets) was amplified by PCR. After introducing the PCR product into *E. coli* BW25113/pIJ790 (λ Red recombination plasmid) containing the cosmid by electroporation, the target gene was replaced by the extended antibiotic resistance cassette. In the other host strain *E. coli* DH5α/BT340 (temperature sensitive FLP recombination plasmid), the resistance cassette was removed through expression of the FLP-recombinase, yielding an in-frame deletion mutant. The detailed procedures for experiments are described in the manual of the PCR targeting system (Plant Bioscience Limited).

2.16 Feeding Experiments

50 ml of YEME medium was inoculated with 100 µl *Streptomyces* spore suspension (10^{10} spores/ml) and incubated at 30°C for 36-40 hours with shaking. 5 ml of the culture was transferred into 100 ml of fresh M10 medium and continuously incubated at 30°C. After 2 days of growing, when the mycelium became abundant, the culture was fed with different substances dissolved in DMSO or the mixture of DMSO and H_2O and incubated for another 5 days on a rotary shaker.

2.17 *N*-Oxidation Assay

The mycelium was harvested by filtration at an early stage (2 days) of the growth (3-5 mg dry weight of cells per millilitre of the culture medium) and washed with 0.85% sodium chloride solution twice. The washed cells (2 g dry weight) were resuspended in 100 ml of phosphate-buffered saline (PBS) containing 50 µmol of sodium p-aminobenzoate. The mixture was incubated at 30°C with reciprocal shaking for about 6 -12 hours. 10 ml of aliquots of the reaction mixture were filtered to remove cells, acidified with 0.5 ml of 1 M HCl, and extracted with ethylether. The concentrated ether extracts were applied to TLC with chloroform-methanol mixture (9:1).

2.18 Fermentation and Detection of Metabolites

Aureothin and derivatives were extracted from chopped agar plates or from liquid culture broths of wild-type and recombinant strains with 50 ml ethyl acetate. Extracts were concentrated under reduced pressure and redissolved in 500 µl MeOH. Compounds were analyzed and identified by thin-layer chromatography (TLC), high-performance liquid chromatography (HPLC), mass spectrometry (MS) and nuclear magnetic resonance (NMR). HPLC analysis was carried out on a KROMASIL C15 column (Jasco) eluted with methanol/water (75:25) in 20 min at a flow rate of 0.8 ml/min and UV detection at 345nm. ESI-CID-MS^n was recorded on a Finnigan LCQ benchtop mass spectrometer equipped with an electrospray ion source and ion-trap mass analyzer (negative mode).

For large scale fermentation, *Streptomyces* strains were grown with orbital shaking at 30°C in M10 medium supplemented with selective antibiotic. 500 ml of 3 days old seed culture was inoculated in 20 l of fermentation medium in a stainless steel fermenter. After 5 days with constant stirring at 30°C, the broth was centrifuged for separation of the biomass from the liquid. Both mycelium and supernatant were extracted five times with 250 ml ethyl acetate for at least one hour.

C. Results and Discussion

1. Cloning, Sequencing and Heterologous Expression of the Aureothin Biosynthesis Gene Cluster

1.1 Design of the Primers for Cloning

The structure and incorporation pattern of aureothin suggested the involvement of a type I PKS in the biosynthesis of the polyketide moiety. Because of the high similarities of genes encoding isoenzymes, amplification of gene fragments by PCR with conserved degenerate primers has become the preferred method for confirmation the presence of desired genes. Therefore, two ketosynthase (KS)-specific primers KSF1 and KSR1 (Beyer *et al*., 1999) were first employed to amplify KS homologous regions in the genome of the aureothin producer, *S. thioluteus* HKI-227. From the PCR product, the expected 700 bp fragments (Fig.3.2D) were separated from some weak unspecifically amplified fragments of different sizes by electrophoresis and cloned into the pGEM-T Easy Vector for sequencing. Sequence analysis of five single colonies revealed several DNA sequences with high homology to KS genes of modular type I PKS.

In many cases, *Streptomyces* genomes do not only bear one, but several type I PKS gene clusters (Sun *et al*., 2002 and Ikeda *et al*., 2003). In order to target the aureothin (*aur*) biosynthesis gene cluster, several pairs of conserved degenerate primers (Fig.3.1) were also designed for complementary amplification of PABA synthase gene homologs from the *S. thioluteus* genome, one of which may play a key role for aureothin biosynthesis. In primary metabolism, PABA is synthesized from chorismate *via* 4-amino-4-deoxychorismate by the successive action of two individual PABA synthase subunits, PabA and PabB. Conversely, both subunits of PABA synthases from eukaryotes and bacterial secondary metabolism are covalently linked into one translational unit (Campelo *et al*., 2002 and Chen *et al*., 2003). Using total DNA of *S. thioluteus* as template, several PCR products were obtained by PCR with different pairs of primers. By BLAST database searches, two copies of PABA synthase genes were detected in different PCR products. A ca. 300 bp PCR product (Fig.3.2C) showed homology to *pabA* from primary metabolism, while the other ca. 1 kb fragment (Fig.3.2B) had high similarity with the *pabAB* gene of the polyene macrolide candicidin gene cluster (Criado *et al*., 1993 and Campelo *et al*., 2002). Therefore, the bigger fragment of the *pabAB* homolog was used as the second probe for identification of the *aur* gene cluster.

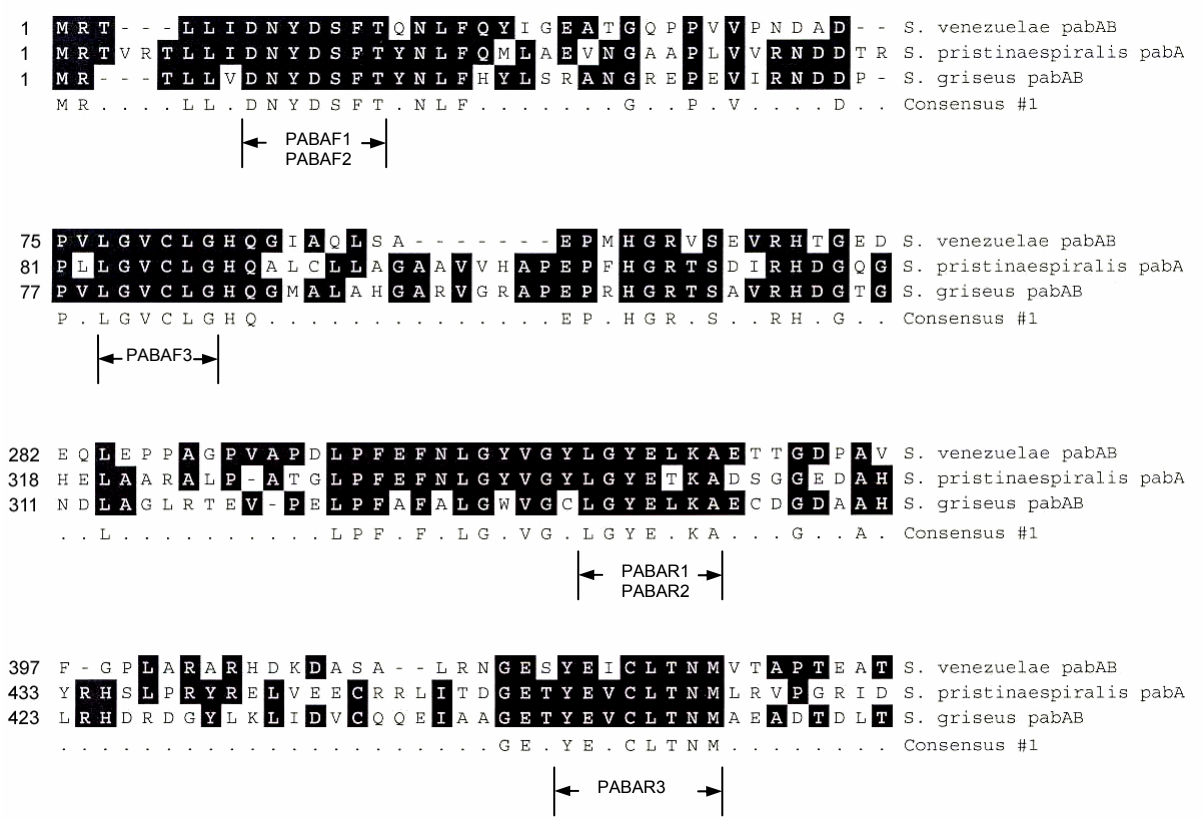


Figure 3.1. Amino acid sequence alignment of some known p-aminobenzoate synthases from *S. venezuelae* (Brown *et al*., 1996), *S. pristinaespiralis* (Blanc *et al*., 1997) and *S. griseus* (Campelo *et al*., 2002). Only several conserved regions used for design of the degenerate primers are shown.

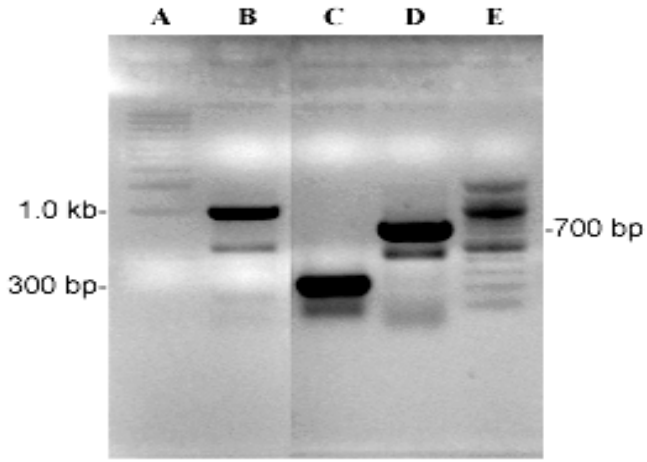


Figure 3.2. Analysis of the PCR products with KS-specific primers and two different pairs of PABA synthase-specific primers, respectively. Lanes A and E contain 1 kb ladder and 100 bp ladder as size standards. Lanes B and C display two PCR products obtained with two different pairs of PABA-specific primers, PABAF1 and PABAR1, PABAF2 and PABAR2, respectively; on lane D the PCR product obtained with KS-specific primers, KSF1 and KSR1, is visible.

1.2 Construction and Screening of a *S. thioluteus* HKI-227 Genomic Cosmid Library

In order to clone the aureothin biosynthesis gene cluster, genomic DNA of *S. thioluteus* HKI-227 was used for construction of a cosmid library. The genomic DNA was partially digested with *Sau*3AI, size-fractionated by gel electrophoresis, yielding fragments with an average size of approximately 40 kb, and blunted using the end-repair mix. After size selection by electrophoresis, the approximate 40 kb DNA fragments were ligated into the supplied blunt-ended pWEB cosmid vector, packaged with lambda packaging extracts and infected the *E. coli* EPI100^{TM} cells (Epicentre) to get enough cosmids for the library. After random detection for the insert size of 10 single colonies, 2112 cosmid clones were inoculated into 22 ninety-six-well plates for preservation of the library at –80°C. The cosmid library was first screened for type I PKS genes by a PCR based approach using one pair of degenerate oligonucleotides, the ketosynthase (KS)-specific primers, KSF1 and KSR1. The PCR product from total DNA was used as DIG-labeled probe for Southern hybridization to confirm the results of screening with PCR. Cosmid mapping, PCR and hybridization experiments revealed the presence of several type I PKS gene clusters in the *S. thioluteus* genome (Fig.3.3 and 3.4).

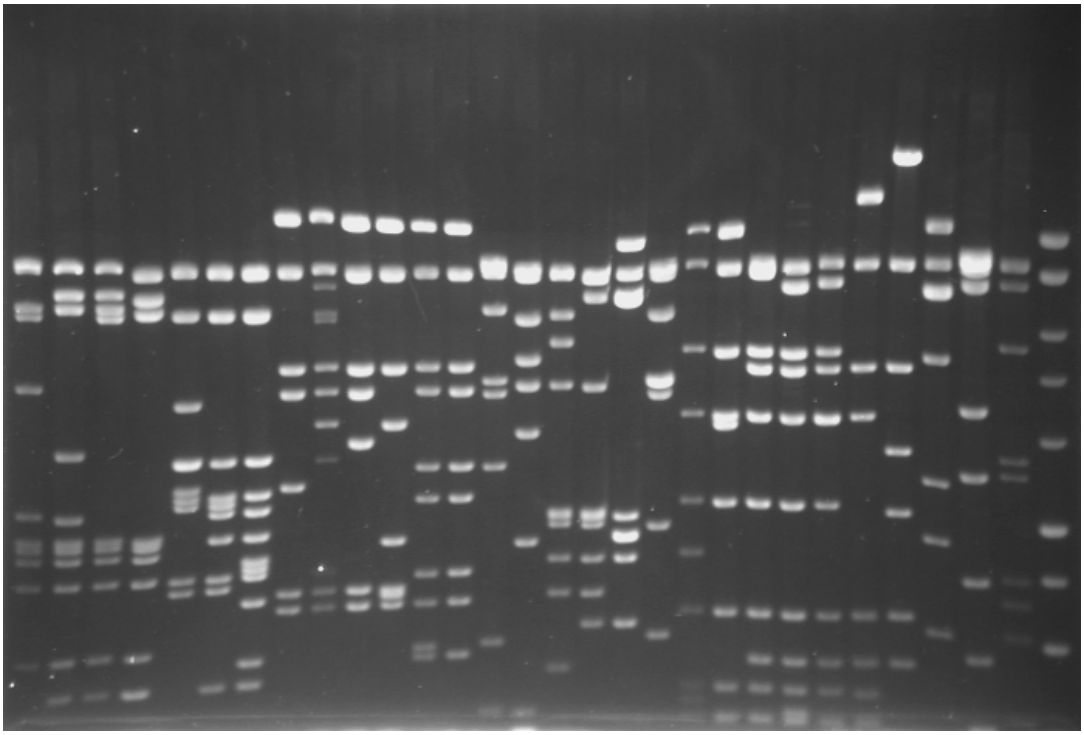


Figure 3.3. Restriction analysis of the positive cosmids obtained by PCR based screening method with KS-specific primers, KSF1 and KSR1. The last lane contains 1 kb ladder as size marker, while the other Lanes display the DNA samples of most KS positive cosmids digested with *Bam*HI.

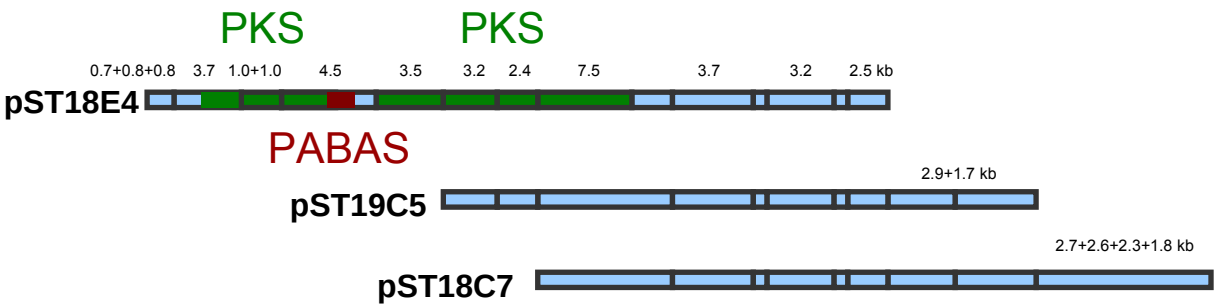


Figure 3.4. Cosmid mapping of pST18E4 and its overlapping cosmids, pST19C5 and pST18C7.

Successively, the PABA synthase (*pabAB*)-specific primers, PABAF1 and PABAR1, based on universally conserved PABA synthase motifs, were used for the complementary screening on a smaller scale of KS positive clones to fish out the putative *aur* genes cluster out of the various *S. thioluteus* type I gene clusters. The PCR product from the genomic DNA was also employed for Southern hybridization. Out of 31 KS positive clones obtained by PCR, only one cosmid, pST18E4, was positive to both of the KS and *pabAB* gene probes by either PCR or hybridization (Fig.3.5).

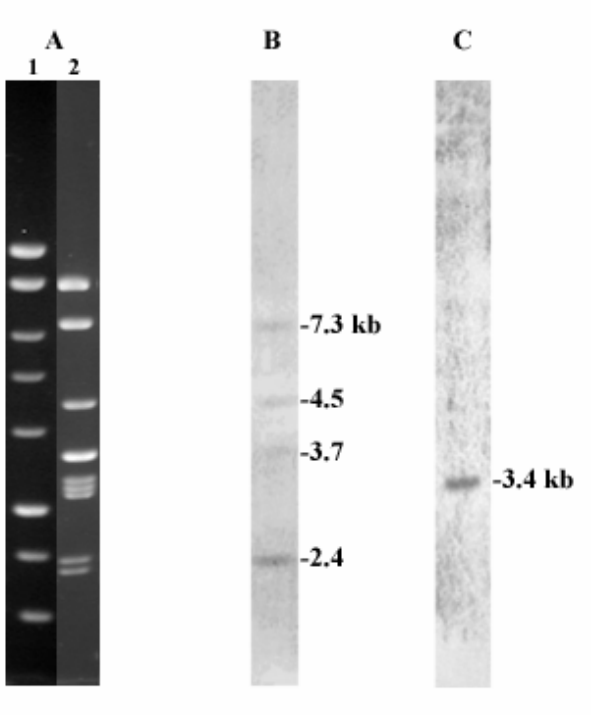


Figure 3.5. Southern hybridization results of cosmid pST18E4 DNA using KS and *pabAB* gene probes. B and C are the color-developed Nylon membranes of Southern transfer of the agarose gel (A) using Dig-dUTP-labeled KS and *pabAB* genes probes, respectively. The DNA sample of pST18E4 (lane 2 in A) was digested with *Bam*HI. Out of all pST18E4 *Bam*HI-digested products, four fragments of 2.4 kb, 3.7 kb, 4.5 kb and 7.3 kb, showed apparently positive to KS probe (B), while only the 3.4 kb fragment gave the positive signal to *pabAB* probe (C). Lane 1 in A contains a sample of 1 kb ladder as size marker.

1.3 Heterologous Expression of the Aureothin Biosynthesis Gene Cluster

To date two general methods are usually employed to verify the functions of cloned gene clusters. One is to inactivate of some important genes in the cluster and this strategy is also usually used to affirm the two boundaries of the gene cluster. The other is to express the entire gene clusters or some functional regions in relational heterologous hosts that are lacking a similar biosynthetic pathway. In order to verify the identity of the aureothin gene cluster by inactivation, various attempts to introduce foreign DNA into *S. thioluteus* HKI-227 have been undertaken, which all failed. Therefore, heterologous expression of the putative gene cluster was attempted. A ca. 40 kb *Ssp*I fragment of cosmid pST18E4, which harbors PKS genes as well as the PABA synthase gene, was recovered including the pWEB *cos* site and ligated into the *Eco*RV site of the integrative *E. coli* - *Streptomyces* shuttle vector pSET152. The resulting cosmid pHJ11 (Fig.3.6) was phage packaged and used for transfection of *E. coli* EPI 100. Restriction mapping proved that the new construct contained all desired fragments as expected. By protoplast transformation, pHJ11 was introduced into *S. lividans* ZX1, and the resulting transformant (*S. lividans* ZX1:pHJ11) was cultivated for five days. After extraction, HPLC and ESI-MS analyses unequivocally showed that aureothin is produced by the transformant at a titer comparable with the wild-type (Fig.3.7).

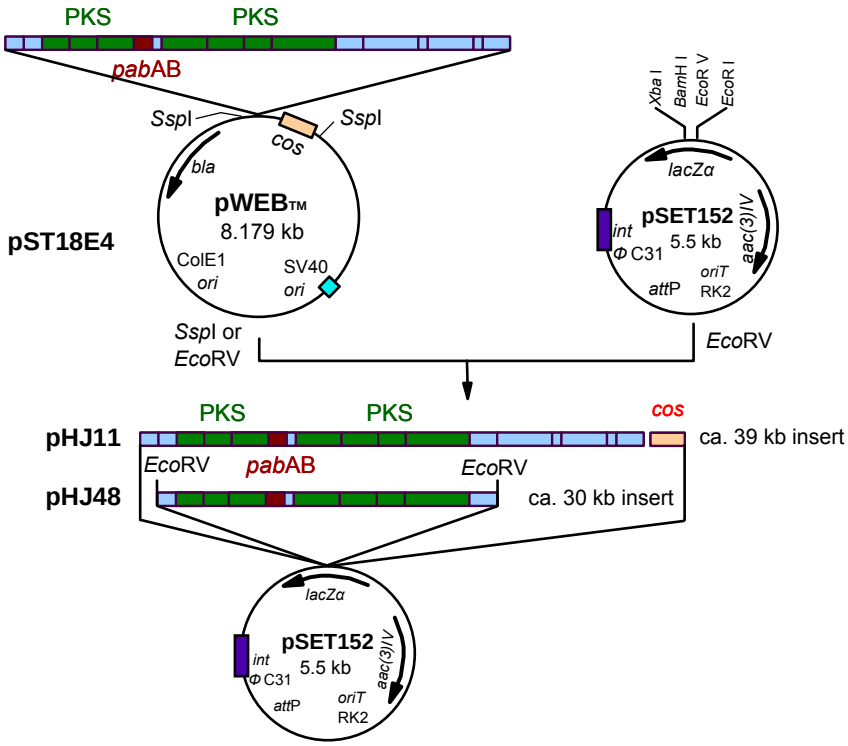


Figure 3.6. Construction of pHJ11 and pHJ48 for heterologous expression of the *aur* gene cluster. Ca. 40 kb *Ssp*I fragment was cut out from pST18E4 and ligated with the blunt ends of shuttle vector pSET152 DNA generated by *Eco*RV digestion to yield the resulting construct pHJ11. For researching the boundaries of the aureothin biosynthesis gene cluster, the 29.8 kb *Eco*RV fragment from pST18E4 was inserted into the *Eco*RV site of vector pSET152 to give the subclone pHJ48 for heterologous expression in *S. lividans* ZX1.

Since the cloned genes were sufficient to confer on *S. lividans* ZX1 the ability to produce aureothin, the cosmid insert of pST18E4 was subjected to both directed and shotgun subcloning and sequencing.

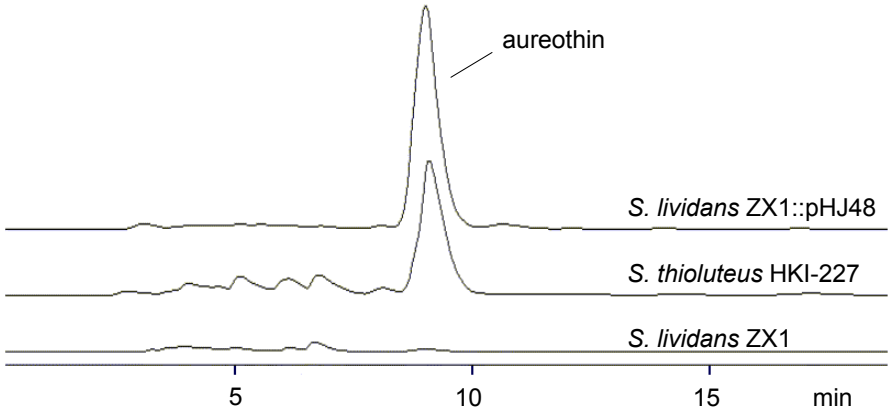


Figure 3.7. HPLC analysis of extracts from the wild-type strain *S. thioluteus* HKI-227, host *S. lividans* ZX1, and transformant *S. lividans* ZX1::pHJ48. The HPLC work was performed by a colleague, Markus Müller.

1.4 Sequence Analysis of the Genomic Region Involved in Aureothin biosynthesis

Cosmid pST18E4 was sonicated, end-repaired by the End-Repair Enzyme Mix (Epicentre), and size-fractionated by agarose gel electrophoresis. Ca. 2 kb DNA fragments were cloned into pBluescript II SK(-) digested by *Eco*RV and end sequenced. Remaining gaps were filled by targeted subcloning and primer walking until completion of the sequence assembly for the whole cosmid. The nucleotide sequence (GenBank accession nr. AJ575648) was analyzed for open reading frames (ORFs) by the FRAME program (Bibb *et al*., 1984). The boundaries of the aureothin gene cluster were determined by subcloning a 30 kb *Eco*RV fragment of pST18E4 into pSET152, yielding the integrative expression plasmid pHJ48 (Fig.3.6). This surprisingly small fragment proved to be sufficient for aureothin biosynthesis by heterologous expression in *S. lividans* ZX1 (Fig.3.7). The deduced gene organization of the gene cluster is graphically presented in Fig.3.8 and the results of the sequence analyses are summarized in Table 3.1.

Putative functions of deduced gene products were assigned based on homologies with known protein sequences in the databases (BLAST and PROSITE searches). Nine genes involved in aureothin biosynthesis were identified and designated *aurA* through *aurI*. In the *aur* operon, all genes are transcribed in one direction except for *aurE*, which is convergently transcribed. Aureothin biosynthesis is probably regulated by the predicted 271 aa gene product of *aurD*, which has 34% identity with AfsR (Horinouchi *et al*., 1990), a global regulatory protein for secondary metabolite formation in *S. coelicolor* A3(2). Downstream of

the *aur* biosynthetic gene cluster, a gene (*aurJ*) encoding a transposase was located. Interestingly, no candidate genes for resistance have been detected.

1.4.1 The Aureothin PKS Genes

In the 27 kb region encoding *aur* biosynthesis, three large ORFs, *aurA-C*, encoding a bacterial type I PKS, were identified. Functional features of the deduced 1906 aa, 2164 aa and 2201 aa gene products were determined by comparing them with known type I PKS systems and detection of signature motifs of the domains. The aureothin PKS shows some unusual characteristics with regard to PKS initiation and termination and, most remarkably, in chain elongation.

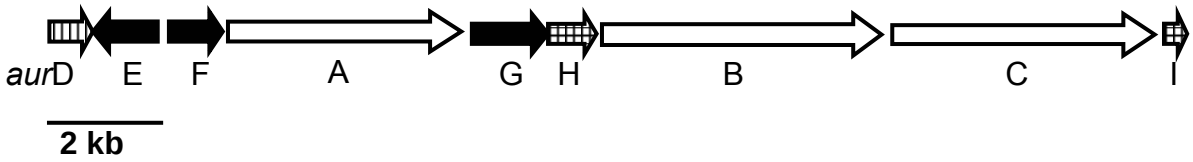


Figure 3.8. Organization of the aureothin biosynthetic gene cluster in *S. thioluteus* HKI-227. Each arrow indicates the direction of transcription and relative sizes of the ORFs deduced from analysis of the nucleotide sequence. White ORFs, type I PKS genes; black ORFs, genes putatively involved in starter unit biosynthesis; checked ORFs, genes encoding tailoring enzymes; hatched ORF, regulator gene.

Table 3.1. Deduced functions of ORFs in the aureothin biosynthetic gene cluster

Protein	Amino Acids	Proposed Function	Sequence Similarity (Protein, Origin)	Identity/ Similarity, %	Protein Accession Number	Reference
AurD	271	Transcriptional activator	AfsR, *Streptomyces coelicolor*	34/45	BAA14186	Horinouchi *et al*., 1990
AurE	505	Acyl-CoA ligase	Feruloyl-CoA synthetase, *Amycolatopsis sp.* HR167	34/50	AX172316	
AurF	481	Amine oxidase	none	none		
AurA Module 1	1906	Polyketide synthase KS AT DH KR ACP	AVES 2, *Streptomyces avermitilis*	45/56	BAA84475	Ikeda *et al*., 1999
AurG	695	PABA synthase	PabAB, *Streptomyces griseus*	44/53	CAC22117	Campelo *et al*., 2002
AurH	406	Cyt P-450 hydroxylase	Cytochrome P-450, *Mycobacterium smegmatis*	36/51	AF102510	Poupin *et al*., 1999
AurB Module 2	2164	Polyketide synthase KS AT ER DH KR ACP	NysC, *Streptomyces noursei*	48/59	AAF71776	Brautaset *et al*., 2000
AurC Module 3 Module 4	2201	Polyketide synthase KS AT ACP KS AT ACP TE	AVES 4, *Streptomyces avermitilis*	45/60	BAA84479	Ikeda *et al*., 1999
AurI	230	O-Methyltransferase	EncK, *Streptomyces maritimus*	36/54	AAF81726	Piel *et al*., 2000

Analysis of AurA reveals a monomodular protein that is assumed to catalyze the first steps in polyketide assembly. The module harbors KS, AT and ACP domains, as well as two keto processing domains (KR and DH). The module arrangement corresponds with unsaturated acyl intermediates, which are in fact the proposed products after the first and second elongation steps. Leadlay and coworkers have reported sequence differences between methylmalonate and malonate specific acyl transferases, which allow the prediction of AT substrate specificity (Haydock *et al*., 1995 and Aparicio *et al*., 1996). In accord with the structure of aureothin, the AT1 signature motif indicates specificity for methylmalonyl-CoA. At its *N*-terminus, AurA shows only very low homology to an ACP domain and the common ACP signature motif (GxDS) is absent, as reconfirmed by repeated sequencing of this region.

AurB is a monomodular enzyme consisting of KS, AT and ACP domains and an additional set of KR, DH and ER, which is capable of performing a full reductive cycle. The module composition and the presence of a malonate-specific AT domain provide strong evidence that this module is involved in the third round of elongation and reduction.

AurC represents a bimodular protein with a thioesterase (TE) domain located at the C-terminus. The TE domain contains characteristic GxSxG and GxH motifs and has high homology to various authentic TE domains from type I PKS. Consequently, the two PKS modules encoded by *aurC* would catalyze the last two Claisen condensations. In agreement with this assumption, each module of AurC is composed of KS, AT and ACP domains only, leaving the β-keto group unreduced. Upon release from the PKS, the resultant β,δ-diketo acid would cyclize – either spontaneously or catalyzed by the TE – to yield the corresponding pyrone moiety. It is noteworthy that in bacteria similar chemistry is only known from polyketides generated from natural and mutant type II PKS rather than type I PKS (Piel *et al*., 2000).

In modules 1 to 4, all the PKS domains show typical conserved signature motifs except AT4 (Table 3.2). In contrast to ATs 1-3, which exhibit the GHSxG active site and (methyl)malonyl-CoA specificity motifs, AT4 shows only weak sequence identity to known AT domains, and both the active site and specificity motifs are aberrant. As confirmed by resequencing the AT4 domain with specific primers and the expression cosmid as template, in the typical GHSxG motif, histidine is mutated to arginine and serine is mutated to alanine. The presence of a GRADG motif most likely results in loss of acyltransferase activity. Recently, anomalous and possibly inactive AT domains have also been reported for the pyoluteorin PKS (Nowak-Thompson *et al*., 1997).

Table 3.2. Analysis of the active site and substrate specificity motifs of the *aur* PKS AT1-4 domains.

AT	Active site motif	Specificity motif	Substrate
AT1	**P**AAVV**GHS**Q**G**EIA**A**	RVDVV	Methylmalonyl-CoA
AT2	**P**DLLL**G HS**I**G**ELT**A**	QTGYT	Malonyl-CoA
AT3	**P**GAVV**GHS**Q**G**EIA**A**	RVDVV	Methylmalonyl-CoA
AT4	**P**AFVH**G**RADGEVA**A**	None	Inactive ?

In summary, the *aur* biosynthesis gene cluster encodes two monomodular and one bimodular proteins, which assemble to a tetramodular megasynthase (Fig.3.9). The domain arrangement of modules 2-4 corresponds perfectly to the degree of reduction of the furane and pyrone moieties of aureothin. However, while the architecture of AurA logically accounts for double bond formation, it appears that only one module catalyzes two successive rounds of elongation before transfer of the polyketide chain to module 2. Such an iterative chain elongation would be in clear contrast to the proposed and generally accepted one-to-one correspondence between encoded modularity of the enzyme activities and structure of the resultant products.

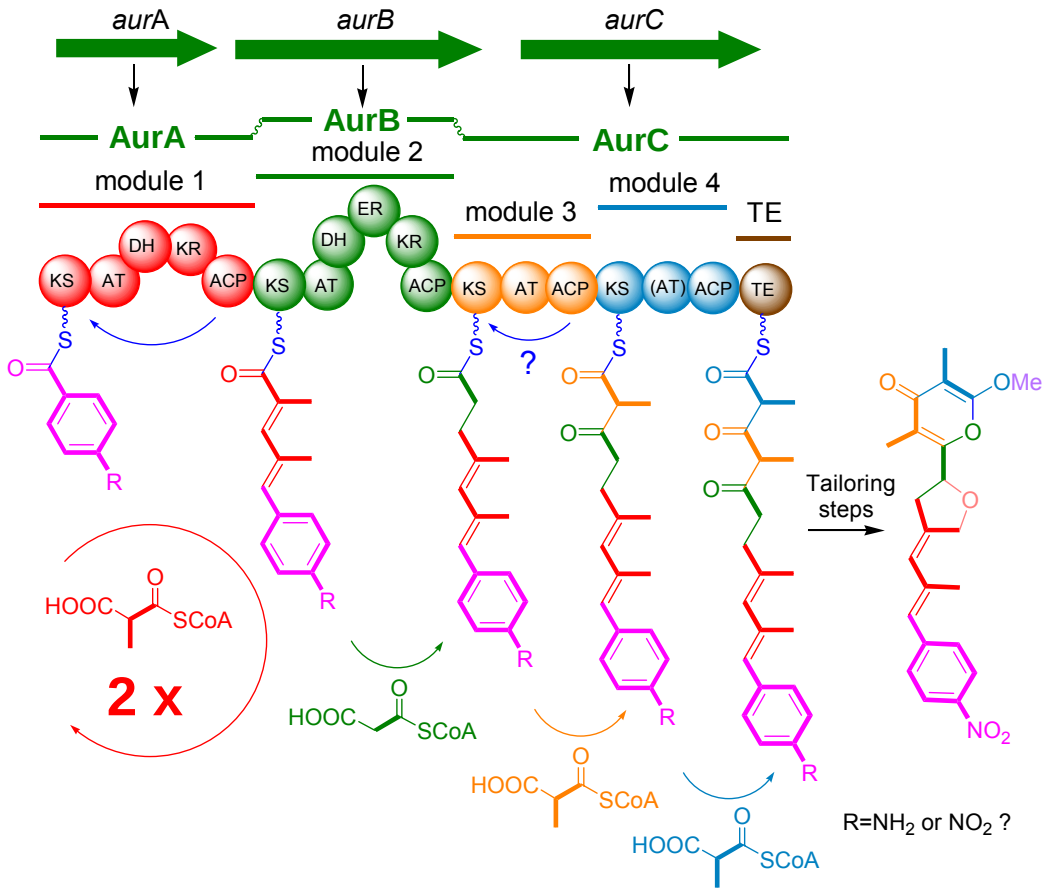


Figure 3.9. Domain architecture of the aureothin polyketide synthase.

1.4.2 Genes Putative Involved in Starter Unit Synthesis and Post-PKS Processing

Besides three large ORFs encoding a type I PKS (*aurA-C*), genes putatively involved in starter unit biosynthesis and post-PKS reactions were detected. By PCR and hybridization experiments, it has been shown that a homolog of *pab*AB, a gene encoding a PABA synthase in *S. griseus* (Criado *et al*., 1993), is located on the aureothin gene cluster. The deduced 695 amino acid gene product of *aurG* has high homology to a group of bifunctional PABA synthases, which are typical for eukaryotes (Edman *et al*., 1993) but have also been implicated in the secondary metabolism of a few *Streptomyces* species (Criado *et al*., 1993 and Brown *et al*., 1996 and Blanc *et al*., 1997). In general, PABA synthases are known to catalyze the conversion of chorismic acid to p-aminobenzoic acid by transamination and subsequent hydrolysis (Green *et al*., 1991). In contrast to the monofunctional enzymes PabA and PabB involved in bacterial folate biosynthesis, the putative *aur* PABA synthase possesses two linked catalytic domains, which are functionally equivalent to PabA and PabB. Homologues of AurG have been found to be involved in the biosynthesis of the polyene macrolides candicidin and FR-008 (Chen *et al*., 2003).

ORF *aurE* presumably encodes an acyl CoA ligase. The deduced 505 aa gene product contains an AMP-binding site and has significant sequence homology to a number of acyl-CoA ligases, such as feruloyl-CoA synthetase from *Amycolatopsis sp.* HR167 (35% identity over 470 amino acids, GenBank accession number AX172316). In analogy to a variety of other unusual starter units (Moore *et al*., 2002), the *aur* PKS primer is probably activated as the corresponding CoA thioester by AurE.

In the deduced gene product of *aurI*, an ORF located downstream of the PKS genes, a probable S-adenosyl methionine-binding motif was identified (PRO-SITE: PS50193) that is characteristic for methyltransferases. Strikingly, AurI shows highest homology (36% identity over 209 amino acids) to EncK from "*Streptomyces maritimus*", which functions as a pyrone O-methyl transferase in enterocin biosynthesis (Piel *et al*., 2000 and Xiang *et al*., 2002). In analogy, AurI is assumed to be responsible for the introduction of the pyrone O-methyl group.

A putative cytochrome P-450 oxygenase encoding gene is located downstream of the PABA synthase gene *aurD* on the aureothin gene cluster, designated *aurH*. The proposed 406 aa gene product has highest homology to a hydroxylase from *Mycobacterium smegmatis*, which is required for piperidine and pyrrolidine utilization (Poupin *et al*., 1999). Two homologous P-450 oxidase representatives involved in polyketide tailoring reactions are the 6-deoxyerythronolide B hydroxylase EryF from *Saccharopolyspora erythraea* (Weber *et al*., 1991), and AveE from *S. avermitilis*, which has been shown to play a role in avermectin biosynthesis (Ikeda *et al*., 1999). This strongly suggests that AurH is implicated in furan ring formation.

Strikingly, the deduced amino acid sequence of the remarkably arginine rich 336 aa gene product of *aurF* has no homology to all known active sites (PROSITE) in databases.

1.5 Discussion

By screening the cosmid library with two special probes, genes involved in the biosynthesis of aureothin, a rare nitroaryl polyketide metabolite from *S. thioluteus* HKI-227, was cloned. Successful production of aureothin in a heterologous host *S. lividans* ZX1 by expression of a 27 kb region bearing the *aur* gene set revealed the identity of the aureothin biosynthesis gene cluster. Detailed sequence analysis of this surprisingly small region revealed several features that are unique in type I PKS catalyzed polyketide assembly, such as a novel priming mechanism. In addition, for the first time it is provided direct evidence for the programmed iterative use of a single PKS module.

1.5.1 PKS Domain Architecture Implicates a Novel Priming Mechanism

For bacterial type I PKS that utilize starter units other than acetate or propionate, two priming mechanisms are known. Free carboxylic acids may be activated and loaded by a nonribosomal peptide synthetase (NRPS)-like adenylation–thiolation (A–T) didomain at the *N*-terminus of the PKS. Alternatively, acyl-CoA ligases may activate the carboxylates as CoA thioesters, which are loaded onto the PKS by the type AT_L /ACP_L loading didomain (Moore *et al*., 2002). Conversely, neither A-T nor AT_L /ACP_L domains could be identified at the *N*-terminus of the first PKS module AurA. A conceivable mechanism could be the activation of starter unit by the putative acyl-CoA ligase AurE and direct transfer onto the PKS. A similar scenario could be considered for CoA ligase – acyl carrier protein (CoL-ACP_L) loading didomains, which have been identified in the candicidin, rapamycin (Schwecke *et al*., 1995) and pimaricin (Aparicio *et al*., 2000) biosynthesis gene clusters. However, since the *N*-terminus of AurA lacks the ACP signature motif for phosphopantheinyl-binding, the priming mechanism of the aureothin PKS remains puzzling, and further biochemical studies are required to understand this unusual set up.

1.5.2 Five Claisen Condensations Are Catalyzed by Only Four PKS Modules

A considerable number of bacterial type I PKS gene clusters have been cloned and sequenced so far. The discovery that the structure of the polyketide metabolite has a one-to-one correspondence to the type I PKS architecture serves as a common model for understanding the programming of modular polyketide synthases and has provided the molecular basis for targeted genetic manipulations. Through detailed sequence analysis and heterologous expression of the full set of aureothin biosynthetic genes it has proven that four PKS modules are sufficient for the catalysis of five elongation and reduction cycles. Amino

acid sequence analysis revealed that module 2, 3 and 4 are responsible for the last three rounds of polyketide chain extension. To synthesize the diene substructure of aureothin, the first two chain extensions both need a module containing KS, AT, ACP, KR and DH activities. In the entire aureothin gene cluster only module 1 could meet this kind of requirement. Module architecture and AT specificity strongly suggests that the first module of the aureothin PKS, AurA, is used twice.

Although this finding is contrary to the principle of co-linearity, it may be rationalized in the context of a few recent observations. In both natural and engineered polyketide-producing organisms, traces of metabolites with C2-elongated polyketide backbones have been detected, which do not accurately correspond to the programmed biosynthetic sequence. Höfle *et al*. have isolated ring-enlarged homologues of epothilone from a *Sorangium cellulosum* So ce90 culture broth (Hardt *et al*., 2001). Leadlay, Staunton and coworkers have investigated a mutant strain of the erythromycin producer *Saccharopolyspora erythraea* that produced two novel octaketides in addition to the heptaketide aglykon DEBS, which is the major metabolite (Wilkinson *et al*., 2000). The occurrence of such side products has been interpreted by Leadlay et al. as an aberrant repeated use of a module, which is referred to as “stuttering”. However, these metabolites are considered as the result of error-prone programming of the requisite polyketide synthases and are thus only formed in trace amounts. In contrast, during aureothin polyketide assembly the iterative action is a programmed event, with exclusive formation of the pentaketide.

At this stage, it can not be ruled out that AurA may be present in two identical copies (or possibly two variants of AurA) docking one behind the other, giving rise to a perfectly collinear pentamodular PKS. Alternatively, the repeated use of AurA may be truly iterative, as, for example, in fungal PKS or type I fatty acid synthases. On the basis of the current model of type I PKS as a helically arranged homodimer (Staunton *et al*., 1996 and Gokhale *et al*., 1998), and considering the domain architecture of AurA, for this model only one mechanism seems plausible: the acyl intermediate would be passed back from the ACP onto the preceding KS domain, as proposed by Leadlay and Staunton for the biosynthesis of minor side products of the erythromycin PKS (Wilkinson *et al*., 2000).

In addition to AurA, it might be possible that iteration also occurs in AurC. As sequence analyses reveal, the AT4 domain is possibly inactive due to the lacking active site serine; furthermore, the specificity region is aberrant. Since module 3 has the same domain arrangement as module 4, one may speculate that both final chain extensions are catalyzed by module 3 alone, according to the mechanism mentioned above.

The heterologous aureothin PKS expression experiments provide the yet strongest body of evidence that iterative use of a module is an inherent property of a natural PKS. At present, it is not possible to deduce the iterative and non-iterative nature of the *aur* PKS modules by

comparison of sequences alone. Furthermore, it is a riddle how this unusual system has evolved. Apparently, modular PKS may have an intrinsic ability to work in an iterative fashion, as recent heterologous expression experiments of individual PKS modules demonstrate (Beck *et al*., 2003), and possibly the line between iterative and non-iterative PKS is more blurred than commonly believed.

2. Functional Analysis of the Aureothin Biosynthesis Gene Cluster

2.1 Biosynthetic Pathway of the Rare p-Nitrobenzoate (PNBA) Starter Unit for Polyketide Synthesis

Besides the results of early feeding experiments, the discovery of *aurG*, a homolog of *pab*AB, also proved the involvement of PABA in the biosynthesis of aureothin. Up to now, there is still no information about how and when the amino group of PABA is converted into the nitro group of the resulting structure aureothin, and which is the real starter unit for polyketide synthesis. In aureothin gene cluster, all gene functions except for *aurF* were implied by homology and fit well into the whole biosynthetic scheme, aside from a gene candidate encoding the *N*-oxidation activity. Surprisingly, the deduced 336 aa gene product (AurF) does not show any sequence homologies to known proteins and not even a known cofactor-binding motif could be determined. It is speculated that *aurF* may be the candidate encoding the *N*-Oxidase. Further functional study will help us to disclose the real role of AurF and the starter unit of polyketide.

2.1.1 Isotope Labelling Experiment

The structure of nitroaryl moiety in aureothin is very suggestive for PABA or a derivative thereof as the starter unit of the polyketide chain assembly. Because so far all efforts to incorporate labelled PABA into aureothin failed, and traces of p-nitrobenzoate (PNBA) was detected in the crude extract of *S. thioluteus* HKI-227 culture by ESI-MS analyses, p-nitrobenzoate (PNBA) was probed as the potential PKS starter. 1 mg d4-PNBA (CDN Isotopes) dissolved in 100 µl sterile aqueous DMSO solution, was administered to 100 ml of 2, 3 or 4 days old M10 liquid cultures of *S. thioluteus* HKI-227. After incubation at 28°C for additional 3 days with vigorous shaking, the cultures were extracted with ethyl acetate. LC-ESMS and HPLC analyses revealed that d4-PNBA was incorporated in the aureothin biosynthesis at a fairly high rate (ca. 20%). The M+4 shift in molecule and daughter ions generated by MS^n experiments (Fig.3.10) indicated that PNBA acts as the aureothin PKS starter unit.

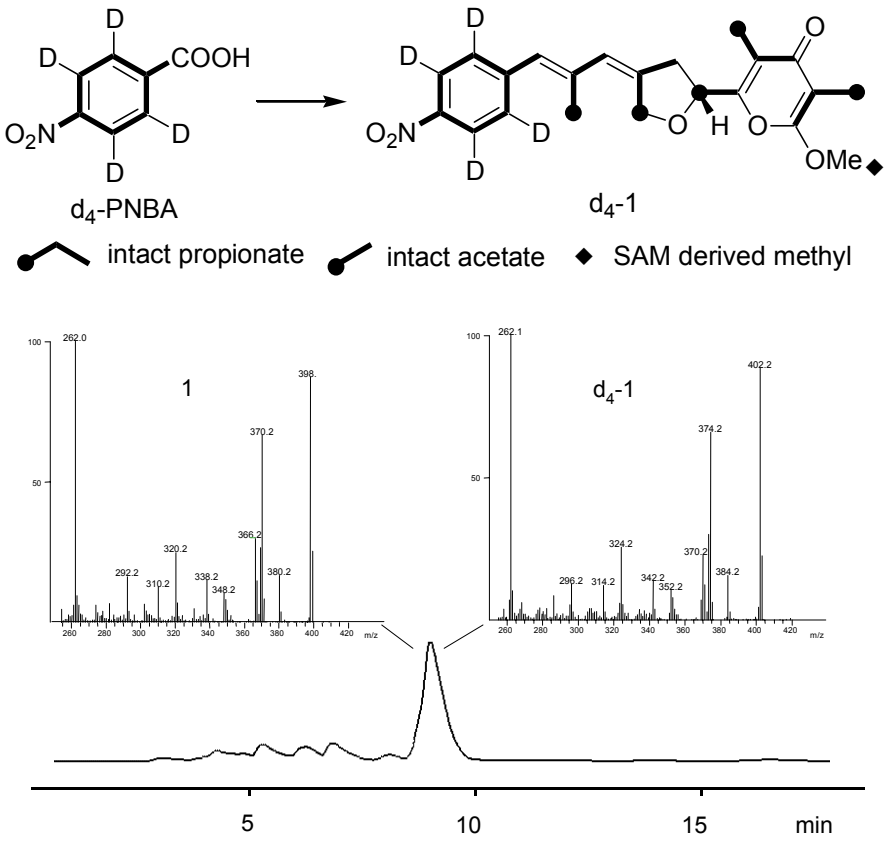


Figure 3.10. HPLC profile of *S. thioluteus* extract and MS-MS analyses of $[1+H]^+$ (m/z=398) and $[d_4\text{-}1+H]^+$ (m/z=402)(insets) obtained by isotope labelling experiments.

2.1.2 Localization of the Putative *N*-Oxygenase Gene Region by an *N*-Oxidation Assay

It was assumed that PNBA would be synthesized in the aureothin biosynthetic pathway by oxidation of PABA, and one PABA synthase gene (*aurG*) was found in the aureothin gene cluster. Strikingly, no obvious gene candidate, which might encode the putative *N*-oxidation activity, could be detected by sequence analysis. From a previous publication (Kawai *et al*., 1965), a biotransformation method was found for an *N*-oxidation assay. In a whole cell approach, mycelium of 2 days old *Streptomyces* culture was harvested, washed and resuspended in PBS buffer. PABA was administered to the cell suspension and the extract of the liquid portion was detected by TLC and MS to monitor the formation of PNBA after reciprocal shaking at 28°C. The major portion of supplied PABA could be readily transformed to PNBA within several hours by the wild type strain *S. thioluteus* HKI-227. In order to screen for the obscure *N*-oxidation activity, several constructs bearing different length of truncated gene sets were generated by successive deletion of the 27 kb *aur* gene cluster from the end (Fig.3.11). All constructs were cloned into the shuttle vector pSET152 as the whole gene cluster and also expressed in the host *S.lividans* ZX1, which is not capable of oxidizing PABA. As monitored by TLC and MS, aureothin production was completely abolished in all truncated mutants, but *N*-oxidation activity was retained (Table 3.3), even in *S.*

lividans::pHJE01, which only harbours a ca. 6 kb genomic fragment. On this 6 kb region, only one ORF (*aurF*) could not be assigned through database searches, which was assumed to be the putative *N*-oxygenase gene.

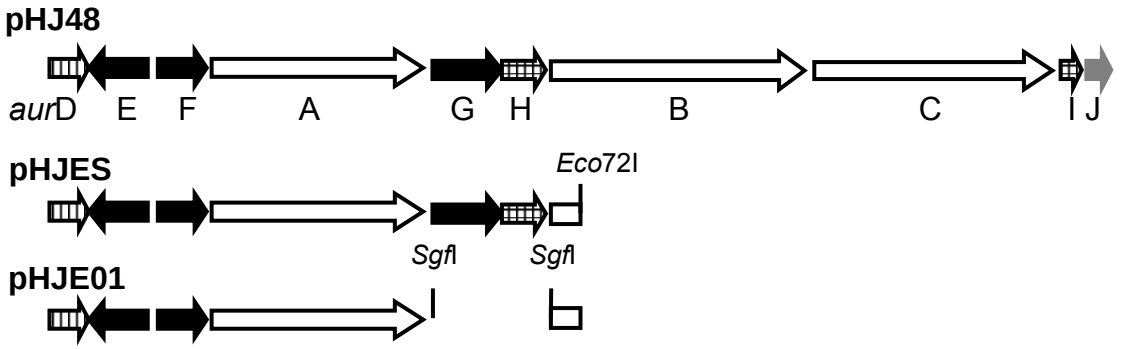


Figure 3.11. Truncated gene sets of the *aur* gene cluster for localization of the novel *N*-oxidase gene. A large *Ssp*I-*Eco*72I fragment from the *aur* gene cluster was ligated with the blunt end of the integrative shuttle vector pSET152 *Eco*RV-digested to give the plasmid pHJES. pHJE01 is a subclone of pHJES, in which the *aurG* and *aurH* genes were eliminated by deletion of the 3.4 kb *Sgf*I fragment with digestion and subsequent self-ligation of the residual large fragment.

Table 3.3. Results of the *N*-oxidation assay and aureothin production of different mutated strains. The host strain *S. lividans* ZX1 and *S. lividans* ZX1::pHJ48 were used as negative and positive control, respectively, for every detection experiment. n.d.: not done.

Strain/Mutant	*N*-Oxidation Activity	Aureothin Production
S. lividans ZX1	–	–
S. lividans ZX1::pHJ48	+	+
S. lividans ZX1::pHJES	+	–
S. lividans ZX1::pHJE01	+	–
S. lividans ZX1::pHJ60	+	–
S. lividans ZX1::pHJ79	–	–
S. lividans ZX1::pHJ79+PABA	n.d.	–
S. lividans ZX1::pHJ79+PNBA	n.d.	+
S. lividans ZX1::pHJ97	+	–
S. lividans ZX1::pHJ97+PABA	n.d.	+
S. lividans ZX1::pHJ97+PNBA	n.d.	+

2.1.3 Heterologous Expression of the Novel *N*-Oxygenase Gene (*aurF*)

The *aurF* gene including the native ribosome binding site (RBS) was amplified by PCR with PrimerFL2 and PrimerFR, and cloned into the pGEM-T Easy Vector for sequencing. The 1.2 kb *Eco*RI fragment containing *aurF* with RBS was recovered from the sequencing vector and ligated into the *Eco*RI site of the *Streptomyces* expression pWHM4* vector, which possesses

a constitutive promoter (P*ermE**). The resulting plasmid, pHJ60 (Fig.3.12), was introduced into *S. lividans* ZX1 by PEG-induced protoplast transformation. The transformant was fermented and the 2 days old mycelium was harvested for the *N*-oxidation assay. Remarkably, the transformant showed *N*-oxidation activity at the same level as the wild type strain (Fig.3.13). The result of the *N*-oxidation assay clearly indicated that the unknown gene *aurF* encodes a novel *N*-oxygenase, which has no homology to any known proteins in the databases. Neither common motifs nor cofactor-binding sites could be identified by database searches. In order to investigate the protein characteristics and cofactors of AurF, two primers, PrimerFL1 and PrimerFR, were designed for amplification of *aurF* from start codon to stop codon. After sequencing, the PCR product was cut out from the sequencing vector by *Nde*I and *Bam*HI digestion and cloned into the corresponding *Nde*I and *Bam*HI sites of two expression vectors pMAL™-c2x or pET-15b, respectively, for protein analyses *in vitro*. Overexpression and purification of the AurF protein are still in progress.

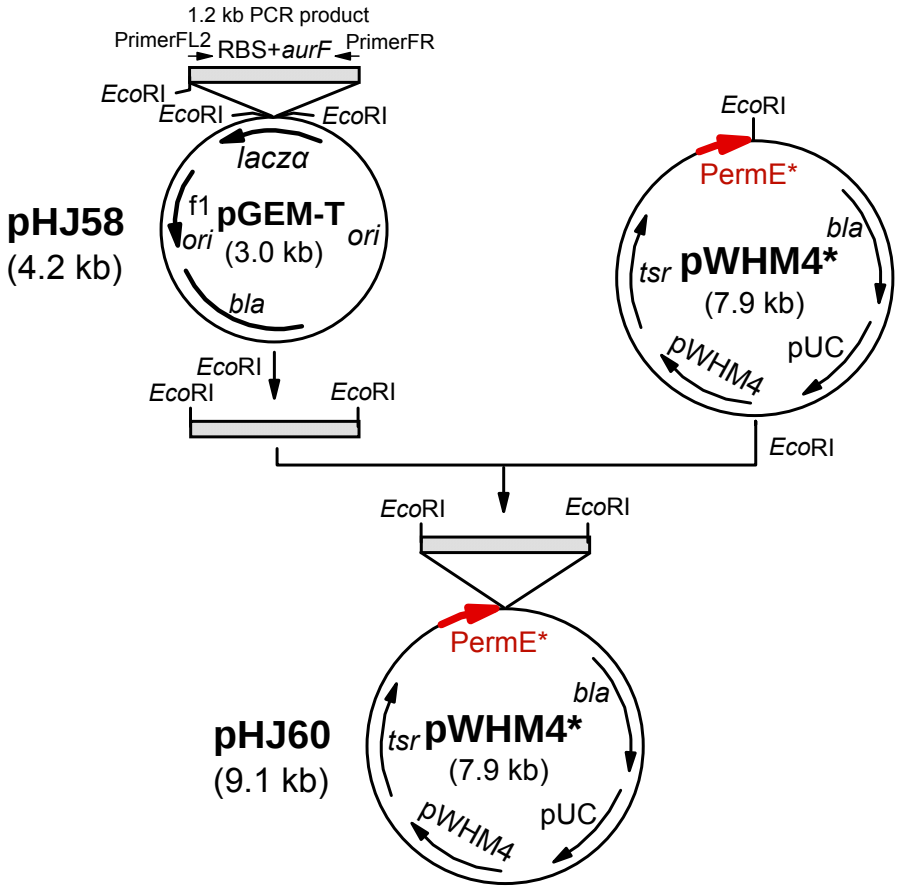


Figure 3.12. Construction of pHJ60 for heterologous expression of the *aurF* gene in *S. lividans* ZX1. The 1.2 kb PCR product including *aurF* gene and its RBS was cloned into pGEM-T Easy Vector to give the plasmid pHJ58. After confirmation by sequencing, the PCR fragment was recovered by *Eco*RI digestion and transferred downstream of the constitutive promoter (P*ermE**) in the *Streptomyces* expression vector pWHM4* to give the resulting plasmid pHJ60.

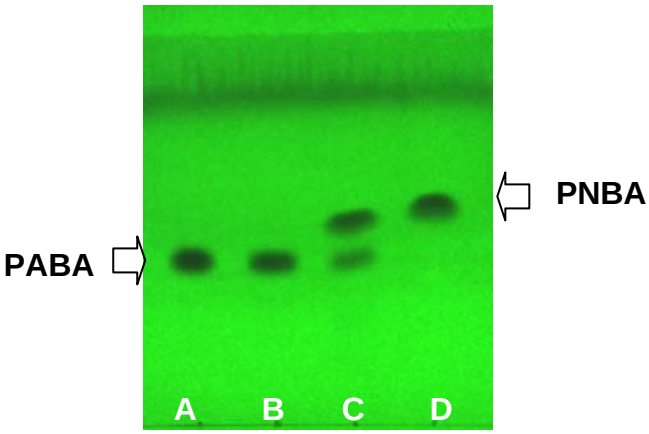


Figure 3.13. Detection of the *N*-oxidation activity of *S. lividans* ZX1/pHJ60 by TLC analysis. The *aurF* gene was heterologously expressed in *S. lividans* ZX1by using expression plasmid pHJ60. A and D are standard samples of PABA and PNBA, respectively. B and C are the results of the biotransformation assay to detect the oxidation of PABA to PNBA by the cell suspension of *S. lividans* ZX1/pWHM4* (negative control) and *S. lividans* ZX1/pHJ60, respectively. Wave length: 254nm.

2.1.4 In Frame Deletion of the *aurF* Gene from the Aureothin Gene Cluster

Further evidence for confirmation of the function of *aurF* gene was obtained by an in-frame deletion of it in cosmid pHJ48, which bears the entire aureothin gene cluster. The *aurF* null mutant was constructed by PCR-targeted mutagenesis using pHJ48 according to the λ red method. Two long PCR primers, aurFL and aurFR were designed for amplification of the extended streptomycin and spectinomycin resistance gene *addA* flanked by FRT sites. First, the extended antibiotic resistance cassette substituted for the *aurF* gene by double cross-over. Subsequently, the cassette was removed through expression of the FLP-recombinase in *E. coli*, yielding an in-frame deletion of *aurF* with the loss of marker gene. The resulting plasmid, pHJ79 (Fig.3.14a) was then introduced into *S. lividans* ZX1 by protoplast transformation. Both *N*-oxidation activity and aureothin production were abolished (Table 3.3) in the transformant. However, when this *aurF* null mutant was fed with synthetic PNBA in the culture, the biosynthesis of aureothin could be fully restored (Fig.3.14b). Meanwhile, PABA was also administered in the mutant culture, but no production of aureothin or a derivative thereof could be detected. The results clearly indicated that *N*-oxidation occurs before the nitroaryl unit is loaded onto the PKS and not during or after polyketide chain propagation.

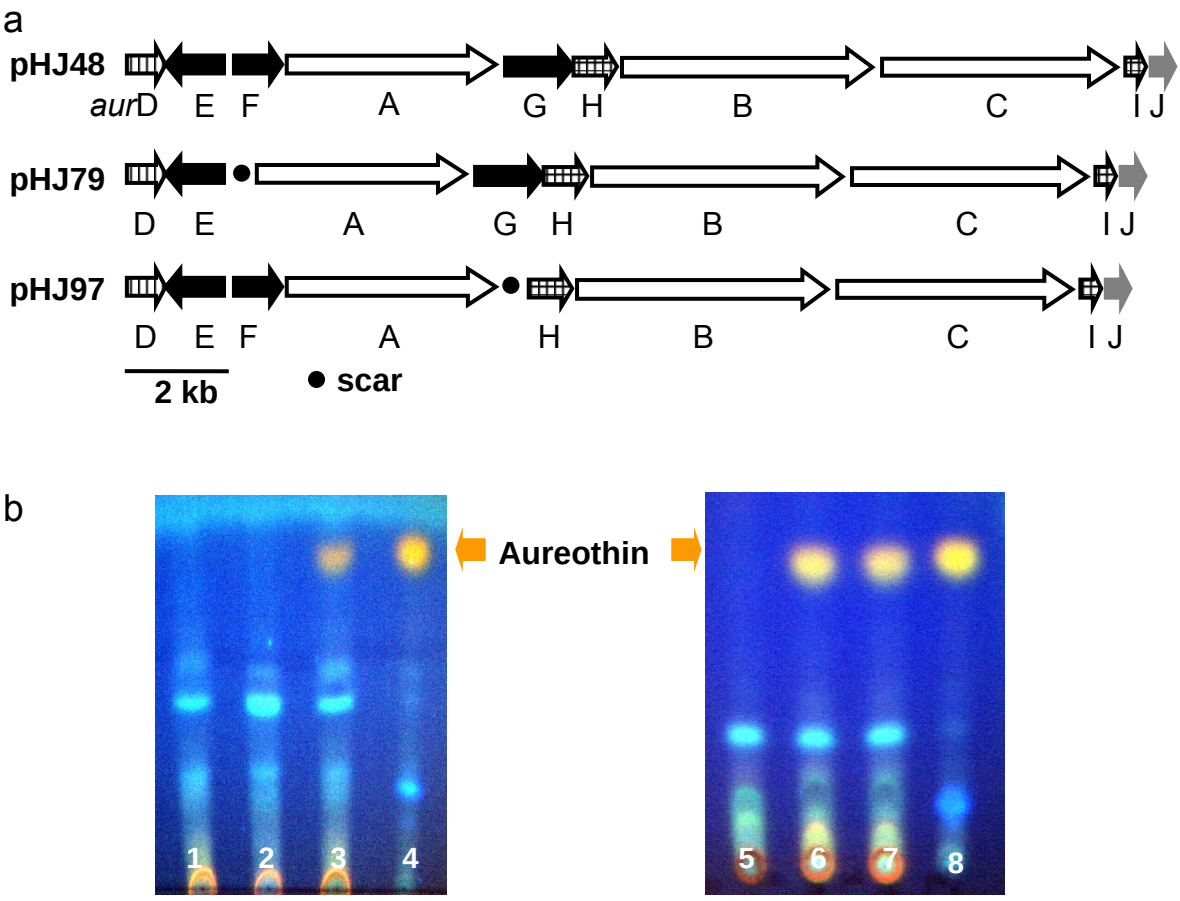


Figure 3.14. a. Organization of the *aur* genes in *aurF* null mutant (pHJ79) and *aurG* null mutant (pHJ97). b. TLC analysis of feeding experiments in *S. lividans*::pHJ79 and *S. lividans*::pHJ97 with PABA and PNBA. 1, *S. lividans*::pHJ79; 2, *S. lividans*::pHJ79+PABA; 3, *S. lividans*::pHJ79+PNBA; 5, *S. lividans*::pHJ97; 6, *S. lividans*::pHJ97+PABA; 7, *S. lividans*::pHJ97 +PNBA; 4 and 8 are the standard samples of aureothin. Wave length: 366nm.

2.1.5 In Frame Deletion of the *aurG* Gene from the Aureothin Gene Cluster

The same strategy was also employed for knocking out the *aurG* gene encoding a putative PABA synthase. The PCR product amplified with two long PCR primers, aurGL and aurGR, was introduced into *E. coli* BW25113/pIJ790 containing pHJ48 with concomitant substitution of the *aurG* gene by the extended cassette. The inserted cassette was removed, yielding an 81 bp “scar” in the preferred reading frame. The resulting plasmid, pHJ97 (Fig.3.14a) was then introduced into *S. lividans* ZX1 by protoplast transformation. The *aurG* null mutant retained the *N*-oxidation activity, but aureothin biosynthesis was abolished (Table 3.3). In the feeding experiments for complementation of the aureothin production, either PABA or PNBA could work in the *aurG* null mutant (Fig.3.14b). Deletion of the *aurG* gene and complementation with different precursors provided more evidence that PNBA and not PABA is the real starter unit for the *aur* polyketide synthase.

2.1.6 Discussion

The biosynthetic origin of the rare nitro aryl moiety from aureothin was elucidated by both molecular and chemical approaches. Incorporation of d_4-PNBA into aureothin in the feeding experiment demonstrated for the first time that PNBA may serve as a polyketide starter unit. Cloning, expression and deletion of *aurF* gene revealed that PABA is transformed into the corresponding nitro compound by an unprecedented type of *N*-oxygenase. Previous results and further heterologous expression and inactivation experiments disclosed the biosynthetic pathway of this unusual primer. The biosynthesis of PNBA involves a PABA synthase (AurG), which converts chorismate acid into PABA, and afterward a novel *N*-oxygenase (AurF), which is responsible for biotransformation from PABA to PNBA (Fig.3.15).

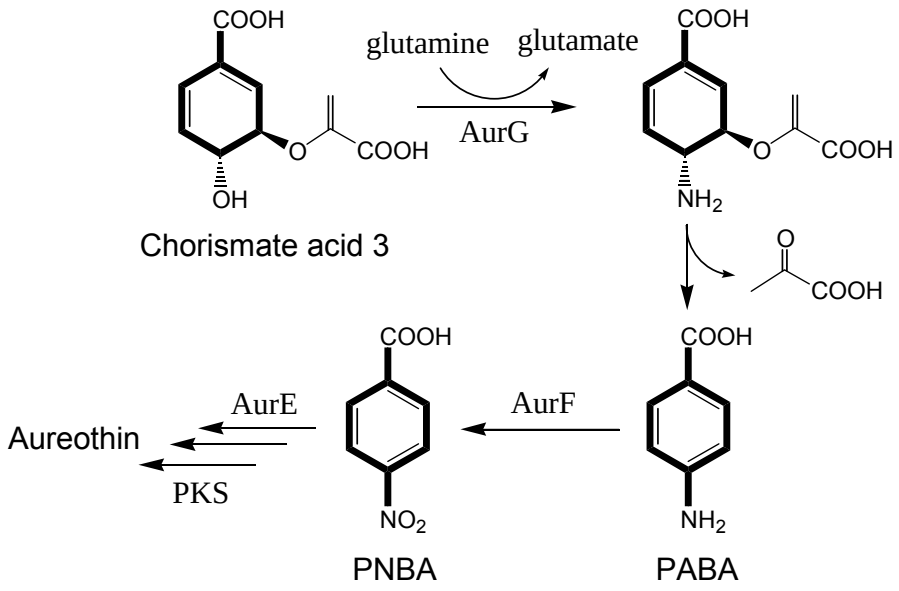


Figure 3.15. Biosynthetic model for the starter unit of aureothin.

Though aromatic nitro compounds widely distribute among various organisms, they are in fact quite rare in nature. Chloramphenicol (He *et al*., 2001) and pyrrolnitrin (Arima *et al*., 1965) are two representatives among the most prominent examples. So far, not much is known about the biosynthesis of nitro aromatic natural products. In a few cases, the possible source of aromatic nitro groups was considered presumably from the substitution reactions (Steglich *et al*., 2002). But it seems that oxidation of primary amines, which probably goes through hydroxylamine and nitroso intermediates stages, is more common. As yet, the only example of an *N*-oxidizing enzyme, which has a Rieske motif, has been implicated in the formation of the nitro group of pyrrolnitrin (Kirner *et al*., 1994). Strikingly, in the aureothin gene cluster, the deduced 336 amino acid gene product of *aurF* has no homology to any known oxido-reductases, and neither common motifs nor cofactor-binding sites could be identified by database searches. Inactivation, expression and complementation experiments

demonstrated the function of AurF, which catalyzes the oxidation reaction from PABA to PNBA. But so far it is not known that how this enzyme works and which are the cofactors for the *N*-oxidation. Functional research of overexpressed AurF *in vitro* may reveal the stereo-structure and nature of the active site and motifs of this protein, and investigate the enzyme properties and kinetics as well as its cofactors.

The vast majority of polyketides is primed with acetate or propionate. However, a number of polyketide synthases utilize alternative starter units, such as short-chain branched fatty acids, alicyclic and aromatic acids, and amino acids. The nature of the primer unit often provides important structural and biological features to the molecule, and variation of the starter unit can significantly alter the activity profile of the natural product (Moore *et al*., 2002). An impressive example for alteration of starter unit is the engineered biosynthesis of doramectin, a highly potent anthelmintic agent obtained by exchanging the avermectin isobutyrate for cyclohexanoate. In order to force PKS to accept the fed non-natural primers, the blocked mutant should be an excellent acceptor strain, in which the formation of natural metabolites is suppressed by either adding a specific inhibitor of the starter biosynthesis, or by deleting particular biosynthetic genes. The tolerance and specificity of different PKS is the determinant for the success of the precursor-directed mutasynthesis (Moore *et al*., 2002). For improving the affinity, the unnatural starter maybe needs to be activated by a CoA ligase or administered as the corresponding *N*-acetyl cysteamine (NAC) thioester. This approach has been successful for generation analogs of erythromycin (Pacey *et al*., 1998), rifamycin (Hunziker *et al*., 1998), enterocin and the wailupemycins (Kalaitzis *et al*., 2003). Here, both of the *aurF* and *aurG* knock-out mutants, in that the biosynthetic pathway of the particular starter unit has been shut down, could be used for mutasynthesis of novel aureothin derivatives by feeding a number of non-natural primers as PNBA surrogates. If exogenously supplied starter units could be accepted by the *aur* PKS, the mutant strains would be capable of catalyzing propagation and modification of polyketide chain to yield aureothin derivatives automatically *in vivo*. In our lab, an aureothin derivative, aureonitrile, with significant cytostatic effect was successfully achieved by employing this strategy (Ziehl *et al*., in press).

In addition to employing the rare PNBA starter unit, the *aur* PKS has some peculiar features, such as the absence of an obvious loading domain, which is usually present in conventional bacterial type I PKS. So far the priming mechanism without any loading domain as in fungal type I PKS or in the *aur* PKS is still unclear. It is speculated that the starter unit is loaded directly onto the KS domain of the first PKS module after activation by the putative acyl-CoA ligase AurE. The investigation for the primer specificity of *aur* PKS and for the substrate specificity of AurE protein will shed light on this unusual priming mechanism in aureothin biosynthesis.

2.2 Post-PKS Modification Reactions

After release from the PKS, the polyketide carbon chains usually undergo further modifications to yield the final biologically active compounds. These additional tailoring reactions are performed by various tailoring enzymes invariably encoded by genes adjacent to PKS genes (Pfeifer *et al*., 2001). In aureothin biosynthesis, it can be assumed that formation of the homochiral five-membered heterocycle and methylation of pyrone ring are post-PKS modification reactions after the polyketide backbone is formed. By sequence analysis, a putative methyltransferase gene (*aurI*) and a putative cytochrome P450 oxygenase gene (*aurH*) were detected in the aureothin gene cluster. Both of genes may be involved in the post-PKS tailoring reactions of aureothin synthesis. Functional studies will disclose the mechanisms of the post-PKS processing.

2.2.1 Knock-Out of the Putative Methyltransferase Gene (*aurI*)

In the insert sequence of cosmid pST18E4, only three *Xho*I sites are present. One is within *aurI*, a gene putatively encoding an O-methyltransferase. The other two are located downstream of the *aur* gene cluster, one of which is within the putative transposase gene (*aurJ*) (Fig.3.16). Elimination of two smaller XhoI fragments, which contain the *aurI* and *aurJ* genes, could yield a knock-out mutant of the putative methyltransferase gene (*aurI*) along with the loss of *aurJ*. The DNA of pST18E4 was digested with *Xho*I and after dilution the digestion product was self-ligated using T4 DNA ligase. Two different new constructs were detected from transformation clones in the *E. coli* DH5α host. The first was named pHJ44, which lost both *Xho*I fragments. And the second was designated pHJ45, in which only the smallest *Xho*I fragment bearing the *aurI* and parts of *aurJ* was discarded. For heterologous expression, the original vector pWEB of these two new constructs was exchanged for the shuttle vector pSET152 by transferring the *Ssp*I inserts of them. The resulting plasmids, pHJ46 (from pHJ44) and pHJ47 (from pHJ45) (Fig.3.16a), were introduced into *S. lividans* ZX1 by protoplast transformation, respectively. Two transformants (*S. lividans* ZX1::pHJ46 and *S. lividans* ZX1::pHJ47) were found to produce the same secondary metabolites as monitored by TLC (Fig.3.16b). Both of them are unable to produce aureothin, but two novel compounds with higher polarity are formed. TLC and LC-MS analyses of the fluorescent fractions of extracts revealed the molecular weight of these two aureothin derivatives, with M=369 g/mol possibly *nor*-deoxyaureothin (main product) and M=385 g/mol possibly *nor*-hydroxydeoxyaureothin (minor product). The structures of them (Fig.3.16b) were fully elucidated by HR-MS and NMR analyses of purified substances after large scale fermentation of *S. lividans* ZX1::pHJ47 (Markus Müller, unpublished). The absence of the O-methyl of the pyrone rings in both new metabolites of the *aurI* knock-out mutant verified the function of AurI. Moreover, the lack of or the incomplete furan ring formation hinted at the

order of the tailoring steps. Apparently, the O-methylation of the pyrone ring takes place before the synthesis of the furan ring.

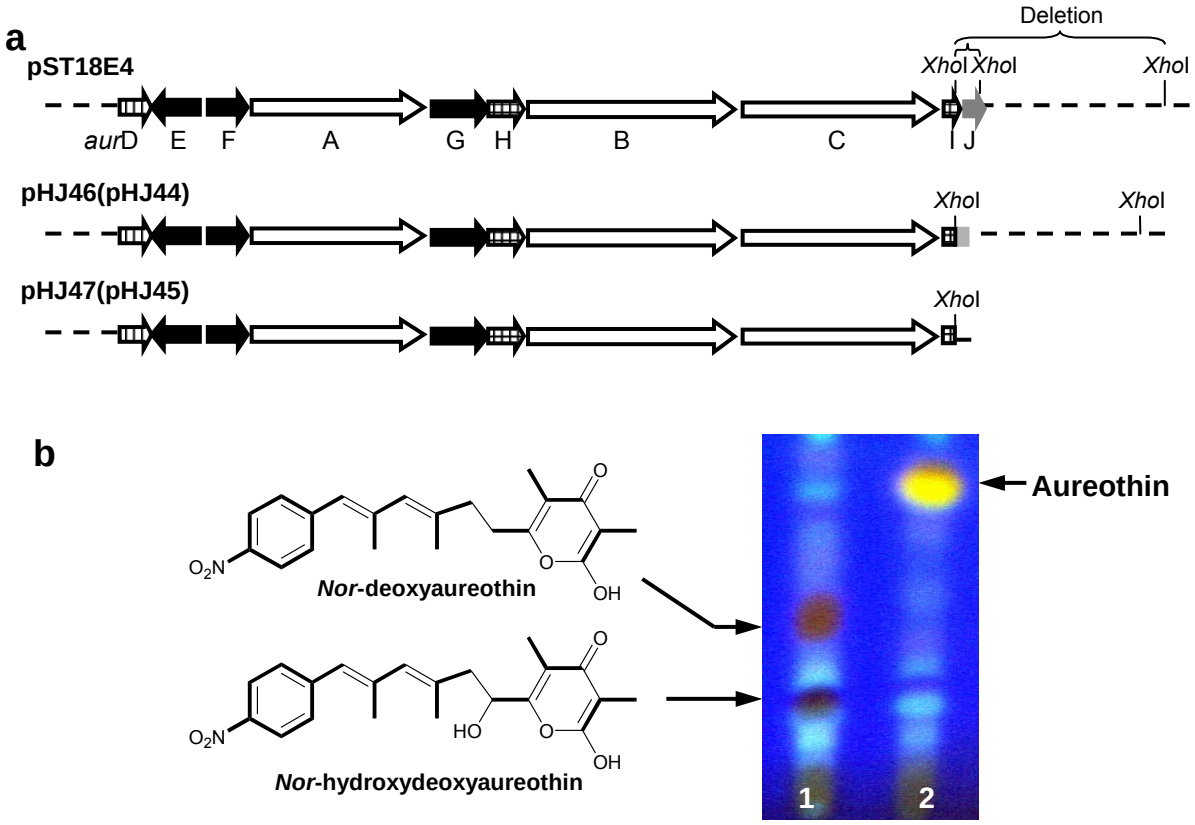


Figure 3.16. a. Organization of the *aur* genes in two *aurI* knock-out mutants (pHJ46 and pHJ47). b. TLC analysis of the extract of *S. lividans* ZX1::pHJ47 (1) with *S. lividans* ZX1::pHJ48 (2) as control. For expressing the mutated constructs in *Streptomyces*, the inserted portions of pHJ44 and pHJ45 were cloned into the *Streptomyces* integrative vector pSET152, yielding pHJ46 and pHJ47, respectively. Wave length: 366nm.

2.2.2 Complementation Experiments of the *aurI* Knock-out Mutant

The *aurI* gene together with its native RBS was amplified by PCR with one pair of primers, primerIL and PrimerIR. After electrophoresis, one PCR fragment of the expected size of ca. 740 bp was recovered from the agarose gel and ligated into the pGEM-T Easy Vector for sequence confirmation. The 0.7 kb *Eco*RI fragment was cut out and cloned downstream of the constitutive promoter P*ermE** in *Streptomyces* expression vector pWHM4*. Finally, the resulting construct pHJ95 (Fig.3.17) was introduced into the *aurI* knock-out mutant (*S. lividans* ZX1::pHJ47) for complementation. In the Δ*aurI* mutant the production of aureothin was fully restored upon co-expression of *aurI.*

Sequence analyses revealed that AurI shows highest homology to EncK, a pyrone O-methyltransferase involved in the enterocin biosynthesis that involves a type II PKS. In order to investigate the similarity and discrepancy of their catalytic activities, the methyltransferase

gene *encK* was also cloned in pWHM4*, yielding plasmid pHJ72 (Fig.3.17). The expression of *encK* in the Δ*aurI* mutant can also complement aureothin biosynthesis, but the complementation in the opposite direction, i.e. expression of *aurI* in the Δ*encK* mutant, failed (Moore *et al*., personal communication).

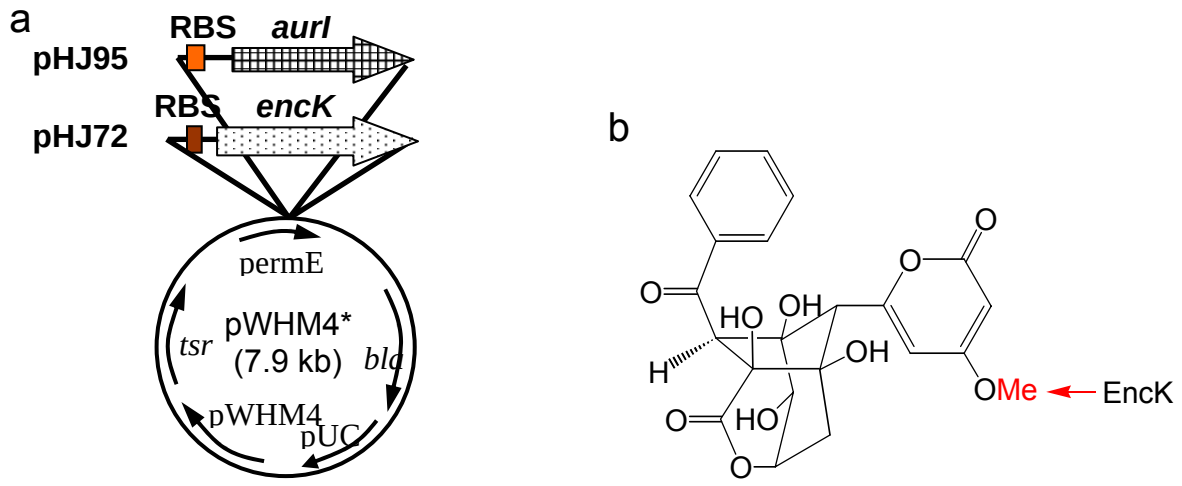


Figure 3.17. a. Expression constructs of *aurI* (pHJ95) and *encK* (pHJ72) genes in *Streptomyces*. b. Chemical structure of enterocin.

2.2.3 Biotransformation Experiments by Heterologous Expression of *aurI*

For investigation of the substrate of AurI, which could be the pyrone ring or the open polyketide chain, the expression plasmid pHJ95 was introduced into *S. lividans* ZX1 for heterologous expression and biotransformation. 5 ml of 2 days old seed culture of *S. lividans* ZX1/pHJ95 was transferred in 100 ml fresh M10 medium. When the mycelium grew rich and strong enough after 2 days with vigorous shaking at 28°C, two substances, *nor*-deoxyaureothin and *nor*-hydroxydeoxyaureothin, were respectively administered to the culture broth in different flasks. At the same time, *S. lividans* ZX1/pWHM4* was also employed in biotransformation experiments as a negative control. After further cultivation of 5 days, the broth was extracted with ethyl acetate for at least one hour. As monitored by LC-MS and HPLC, the control strain did not show any sign of biotransformation, while the strain expressing AurI readily transformed *nor*-deoxyaureothin into deoxyaureothin (Fig.3.18). But for *nor*-hydroxydeoxyaureothin, transformation took place neither in the control strain nor in the AurI expression strain. Biotransformation experiment revealed that *nor*-deoxyaureothin is the substrate of AurI and that the formation of the pyrone ring occurs prior to the methyltion.

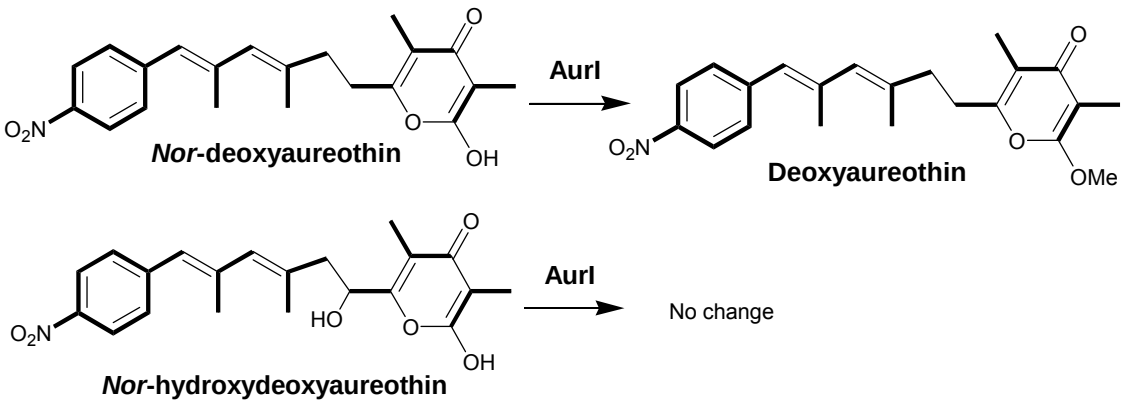


Figure 3.18. Biotransformation experiments by expression of AurI enzyme in *Streptomyces*.

2.2.4 Inactivation of the Putative Cytochrome P450 Oxygenase Gene (*aurH*) by In-Frame Deletion

To prove the involvement of AurH, a putative cytochrome P450 oxygenase, in the unusual oxidative heterocyclization, inactivation of *aurH* gene was performed by PCR-targeting system. As in the deletion of *aurG* and *aurF*, the shuttle cosmid pHJ48 bearing the entire *aur* biosynthesis gene cluster served as suitable starting point for generating the *aurH* knock-out mutant directly in *E. coli* by recombination. Two long primers, aurHL and aurHR, were designed to amplify the extended streptomycin and spectinomycin resistance gene (*aadA*) flanked by FRT sites (FLP recognition targets). After introduction into *E. coli* BW25113/pIJ790 containing pHJ48, the PCR product replaced the *aurH* gene by homologous recombination. The central part of the disruption cassette was subsequently removed by site-specific recombination of the FLP recombinase in *E. coli*. In the resulting cosmid pHJ98 (Fig.3.19a), only the *aurH* gene sequence was excised from the *aur* gene cluster and the deletion was confirmed by restriction mapping. After protoplasts of *S. lividans* ZX1 were transformed with the DNA of pHJ98, the resulting transformant, *S. lividans* ZX1::pHJ98, was cultivated in M10 medium under apramycin selection to maintain the integration of pHJ98 in the chromosome. TLC, HPLC and MS analyses of the crude extract from the mutant broth culture revealed that two other fluorescent metabolites were formed instead of aureothin. On TLC plates, the main product deoxyaureothin appeared in lieu of aureothin with slightly lower polarity, while the minor product hydroxydeoxyaureothin showed much higher polarity. The structures of new metabolites were fully resolved after purification of substances from the culture of a 20 l *S. lividans* ZX1::pHJ98 fermentation by HR-MS and NMR analyses (Markus Müller, personal communication). The fact that both novel compounds already bear methyl groups in the pyrone ring provided further support for the hypothesis of the order of the post-PKS processing. O-methylation is the first post-PKS modification reaction and the second modification is the furan ring formation catalyzed by the

oxygenase. In *aur* gene cluster, only two oxygenase genes, *aurF and aurH*, were detected. Heterologous expression and inactivation experiments have already revealed that AurF is solely involved in the biosynthesis of the PKS primer. Thus, the best candidate gene product for catalyzing the furan ring formation is AurH. Nevertheless, the appearance of the unexpected side product, hydroxydeoxyaureothin, is baffling. In order to eliminate the interference of other CYT P450 genes from other polyketide gene cluster in host strain *S. lividans* ZX1, pHJ98 was also introduced into another host strain, *S. albus*, for expression. TLC, HPLC and LC-MS data revealed that *S. albus*::pHJ98 only produced deoxyaureothin and no hydroxydeoxyaureothin was found (Fig.3.19b), while the control strain *S. albus*::pHJ48 produced aureothin as expected. It indicated that hydroxydeoxyaureothin is an oxygenated derivative of deoxyaureothin by some other oxygenases from the host *S. lividans* ZX1.

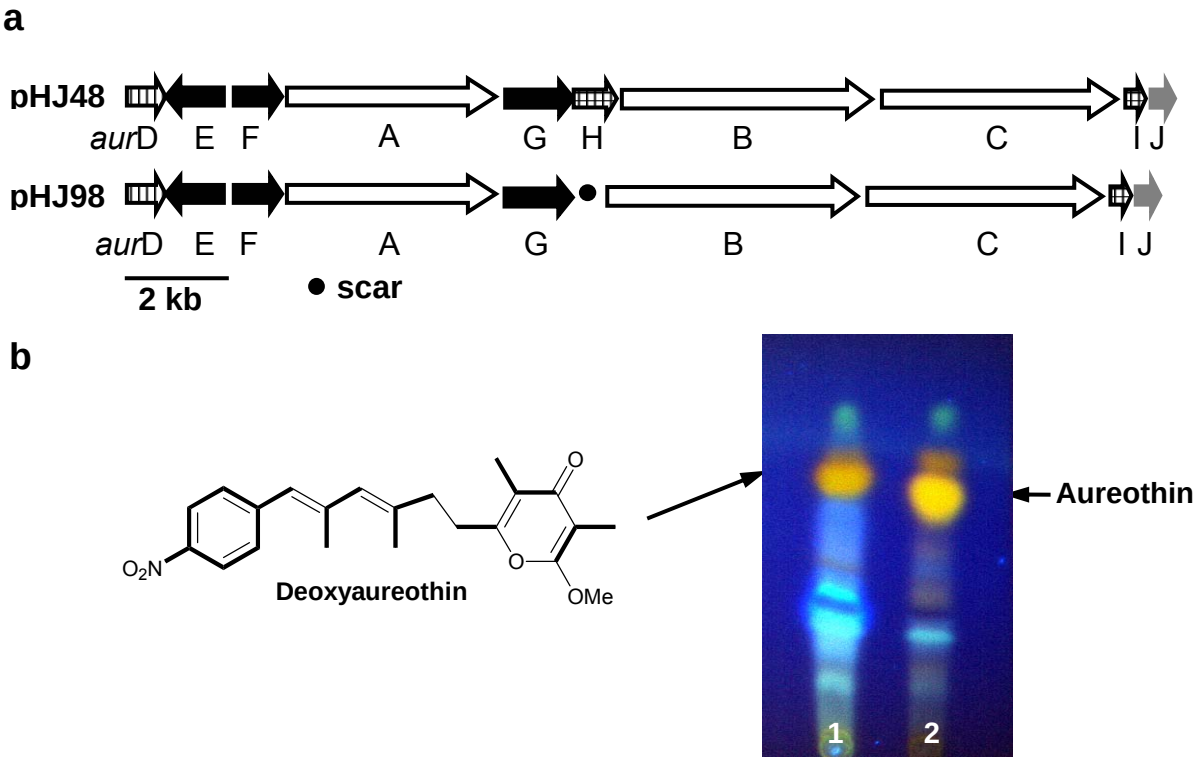


Figure 3.19. a. Organization of the *aur* genes in *aurH* null mutant (pHJ98). b. TLC analysis of the extract of *S. albus*::pHJ98 (1) with *S. albus*::pHJ48 (2) as control. Wave length: 366nm.

2.2.5 Complementation Experiment of the *aurH* Null Mutant

To restore aureothin biosynthesis, AurH was coexpressed in the Δa*urH* null mutant. One pair of primers, PrimerHL1 and PrimerHR, was used to amplify the *aurH* gene together with the native RBS by PCR. After cloning into the pGEM-T Easy Vector for sequencing, the 1.4 kb PCR product was recovered by *Eco*RI digestion and cloned downstream of the constitutive promoter P*ermE** in pWHM4*, yielding the self-replicating expression plasmid pHJ110

(Fig.3.20). When the resulting plasmid pHJ110 was introduced into *S. lividans* ZX1::pHJ98, the mutant was successfully complemented and aureothin biosynthesis was restored.

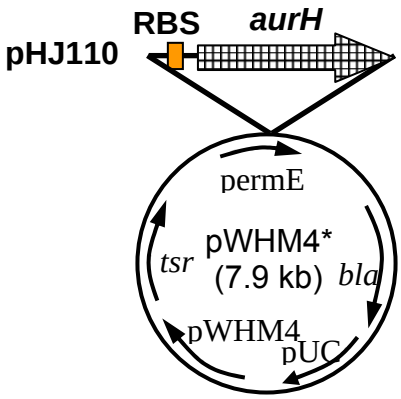


Figure 3.20. Construct of pHJ110 for heterologously expressing *aurH* gene in *Streptomyces*.

2.2.6 Biotransformation Experiments by Heterologous Expression of *aurH*

While inactivation and complementation experiments clearly demonstrated that AurH is required for furan ring formation, it needed to be established if this novel cytochrome P450 oxygenase is indeed sufficient for catalyzing the formation of both C-O bonds of the furan moiety. For this purpose, AurH was heterologous expressed in *S. lividans* ZX1 using expression plasmid pHJ110. When deoxyaureothin was administered to the culture, the AurH expression strain *S. lividans* ZX1/pHJ110 could readily catalyze the formation of furan ring to give the final product, aureothin, as monitored by LC-MS (Fig.3.21). Furthermore, a side product, dehydrodeoxyaureothin, isolated from *S. lividans* ZX1::pHJ48 was also tested in the biotransformation experiment with heterologously expressed AurH, but no transformation occurred.

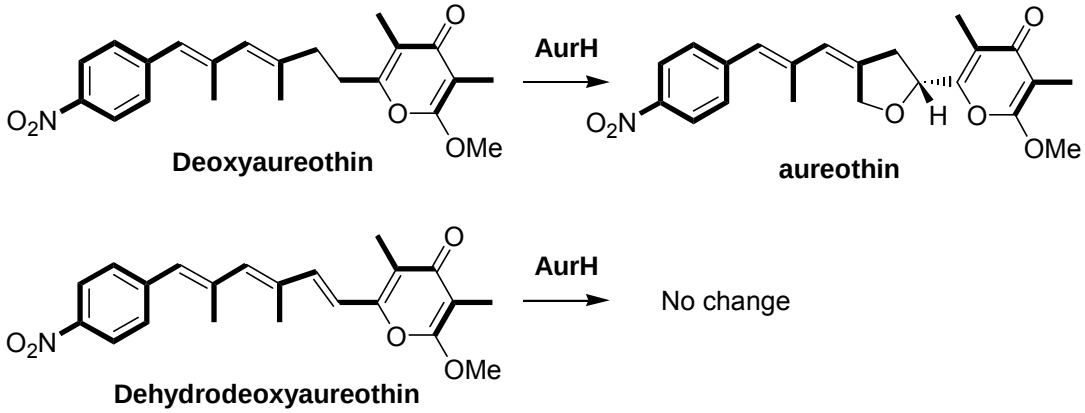


Figure 3.21. Biotransformation experiments by expressing AurH protein in *Streptomyces*.

To study AurH protein *in vitro*, the *aurH* gene was amplified from its start codon to the stop codon by PCR with two primers, PrimerHL2 and PrimerHR. When the PCR product was

cloned into the multiple cloning sites of the *E. coli* overexpression vector pMAL™-c2x, a *malE* signal sequence encoding the *E. coli* maltose binding protein was fused to the *N*-terminal end of the *aurH* gene for purification of the protein. Overexpression and purification of the fusion AurH protein are still in progress.

2.2.7 Discussion

Inactivation, complementation and expression experiments of the different tailoring enzymes, AurH and AurI, provided a lot of information not only on their functions, but also on the timing of the post-PKS processing. The methyltransferase gene *aurI* knock-out mutant produces nor-deoxyaureothin, not nor-aureothin, while the P450 oxygenase gene *aurH* knock-out mutant produces deoxyaureothin. These results clearly revealed a defined order of the tailoring steps. O-methylation is the first modification reaction and synthesis of the furan ring represents the last step in the aureothin pathway, which is catalyzed by a bifunctional cytochrome P450 monooxygenase, AurH (Fig.3.22).

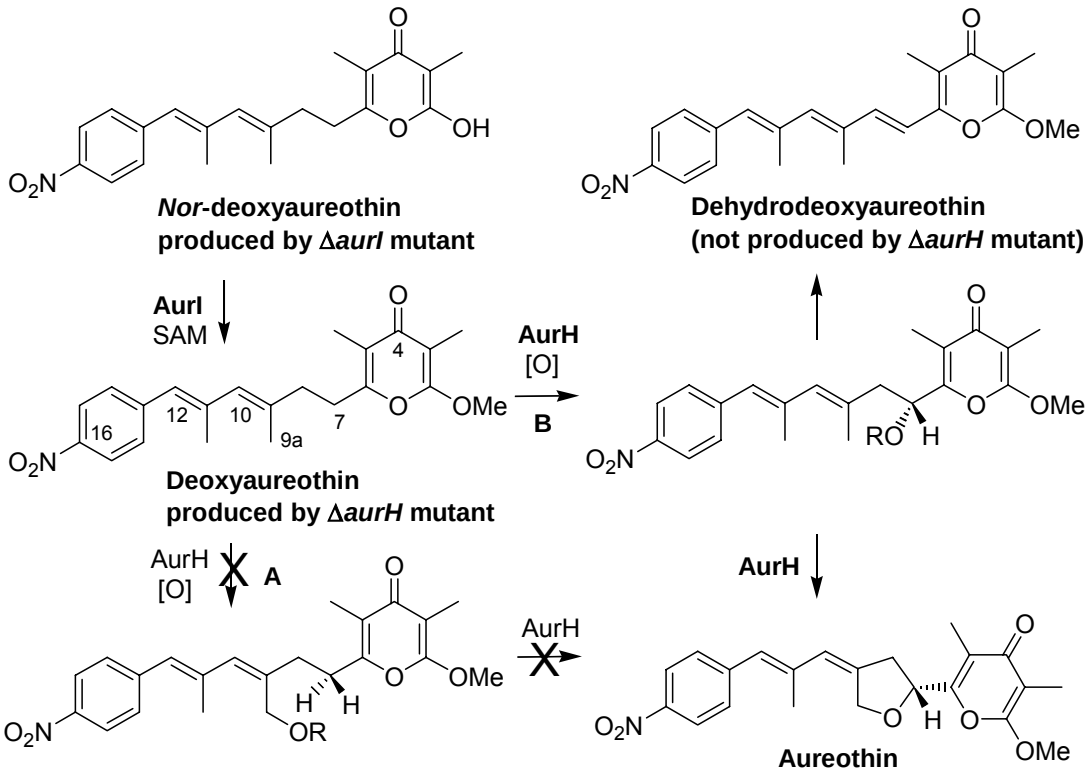


Figure 3.22. Biosynthetic model for tailoring reaction processing. The final ring closure might occur through a radical mechanism or through nucleophilic substitution via a bis-hydroxylated intermediate.

It is obvious that the final modification step occurs with the subsequent oxidation of two non-activated carbons. Thus, two scenarios are generally conceivable for this rare oxidation sequence: either the allylic position (9a-C) is attacked first (route A, Fig.3.22) or the

methylene (7-C) adjacent to the pyrone ring (route B), setting the stereochemistry of the resulting heterocycle at an earlier stage. Route B is significantly supported by the occurrence of a nitrobenzoate-primed polyketide metabolite, dehydrodeoxyaureothin. This side compound is produced by the wild type strain *S. thioluteus* and by *S. lividans* ZX1::pHJ48, which served for heterologous expression of the entire gene cluster, but not by the Δa*urH* null mutant (Markus Müller, personal communication). In general, dehydrodeoxyaureothin may result from an error-prone β-keto processing of the second module (AurB), where the ER domain of module 2 would be accidentally skipped. However, this triene compound could not be detected in the Δa*urH* null mutant, where the identical PKS is employed. It is thus more likely that dehydrodeoxyaureothin is a shunt product from route B, resulting from hydroxylation and subsequent dehydration.

The biotransformation experiment mentioned above revealed that dehydrodeoxyaureothin is in fact not a substrate of AurH. Since in this compound the 7-C methylene is masked, but the 9a-C methyl group is still available, this result could support the model where oxygenation is not initiated at the methyl group. All observations point towards the second path in which AurH first oxidizes the methylene adjacent to the pyrone ring and then promotes the second C-O bond formation, yielding the tetrahydrofuran ring.

To date, a large number of cytochrome P450 genes from various organisms are known. The best studied P450 enzymes play a key role in mammalian liver cells for drug and xenobiotic metabolism. Most cytochrome P450 oxygenases in this superfamily are monofunctional, catalyzing a single hydroxylation or epoxidation of a substrate. Multifunctional cytochrome P450s are rare and yet little explored. For example, the loss of C-14 in steroid biosynthesis (Lamb *et al*., 1999) and the three-step oxidation of *ent*-kaurene to *ent*-kaurenoic acid in *Gibberella fujikuroi* (Ashman *et al*., 1990) have been described as sequential oxidations mediated by single cytochrome P450s at the same carbon atom, whereas the oxidation at multiple sites by a single P450 enzyme has been reported for mammalian steroid biosynthesis (Sun *et al*., 1995 and Nonaka *et al*., 1991) and in some secondary products of fungi, such as gibberellin (Rojas *et al*., 2001). Through inactivation and biotransformation experiments, it demonstrated that AurH is a multifunctional cytochrome P450 monooxygenase catalyzing the formation of the homochiral furan ring of aureothin. To our knowledge this is the first example of proof for a single P450 monooxygenase catalyzing a two-step heterocyclization.

AurH contains only one single heme-binding motif so that both of C-O bond formations must be catalyzed at the same active site. However, the oxidized carbons should be sufficiently close, thus requiring only small changes of the substrate-binding orientation when the enzyme catalyzes the two sequential oxidation reactions. Further studies with the purified AurH protein *in vitro* will help elucidations of this multifunctional enzyme.

During the studies of post-PKS modification, several aureothin derivatives were produced by the mutants. As mentioned earlier, tailoring reactions usually can have a marked influence on the activity profiles of the final metabolites in polyketide synthesis. Therefore, bioassay of these novel compounds would unearth their potentialities for new drug exploitation. So far, only deoxyaureothin was tested for bioassay. The antimicrobial profile of deoxyaureothin is very similar to aureothin, albeit with slightly reduced antifungal activity. Strikingly though, deoxyaureothin exhibits a five-fold increased cytotoxic activity against human K-562 leukemia cells (Markus Müller, personal communication).

2.3 Investigation of the Aureothin PKS

In aureothin polyketide synthase, three PKS genes (*aurA*, *aurB* and *aurC*) encode four PKS modules terminated with a thioesterase (TE) domain. Strikingly, the *N*-terminus of the first module lacks a loading domain and the AT4 domain is supposed to be inactive due to a mutated active site motif. In stark contrast to the accepted principle of co-linearity observed for modular type I PKS, heterologous expression experiment demonstrated in aureothin gene cluster, only four PKS modules are sufficient for catalyzing five rounds of elongation and processing. The domain architecture of the modules suggests that the first module is used twice, giving rise to the diene substructure. Furthermore, because of the aberrant and non-functional AT of module 4, it could be assumed that in analogy to the first module, the isosteric module 3 might operate iteratively, too, in order to accomplish the fifth Claisen condensation. To prove the iterative fashion of AurA, the *aurA* gene could be heterologously expressed and fused with TE domain at the *C*-terminus, which could promote the release of the intermediate after the extension reactions performed by the first module (AurA). Fusion of the *aurA* gene and its downstream gene, *aurB*, could also be an alternative strategy. In addition, the inactivation of some key domains in module 4 would disclose if this module is in fact functional despite its aberrant AT domain and thus reveal if the module 3 is operating in an iterative fashion.

2.3.1 Cloning of the *aurA* Gene in *Streptomyces* Expression Vector pRM5

For functional studies of the first PKS module by obtaining the polyketide intermediate after the Claisen condensations catalyzed by AurA, the *aurA* gene was cloned into the vector pRM5, which has been used previously for recombinant PKS expression in heterologous *Streptomyces* host. pRM5 is a shuttle vector, which contains a replication origin from ColE1 and an antibiotic resistance gene (*bla*) for convenient manipulation of DNA cloning in *E. coli* host, and SCP2* *Streptomyces* replicon and an antibiotic gene (*tsr*) for introduction of exogenous DNA into *Streptomyces*. It has the additional activator gene (*actII*-ORF4) and divergent promoter pair (P*actI*/*actIII*) from the *act* gene cluster for expression of recombinant

PKS. A ca. 800 bp fragment including the *N*-terminal sequence of *aurA* together with the native RBS was amplified by PCR with a pair of primers, ORFL1 and ORFR1, and furthermore a *Pac*I site was introduced upstream of the RBS sequence for cloning. After cloned into pGEM-T Easy Vector for sequencing, the PCR product was recovered by restriction digestion with *Pac*I and *Bsp*1407I together. The second part of *aurA* was covered by a 4925 bp *Bsp*1407I-*Nco*I fragment cut out from pHJ48. This big fragment plus the PCR product covers the majority of the *aurA* gene except that a piece of ACP sequence at the *C*-terminus is missing. In order to enhance the effective release of the polyketide intermediate after the elongation reaction(s) by AurA, the TE domain sequence of module 4 was also amplified by PCR with two primers, TEL1 and TER1. An *Nco*I site was introduced into the *N*-terminus of TE for fusion with AurA. These three components of recombinant *aurA*+TE were assembled and ligated with the larger *Pac*I-*Eco*RI fragment of pRM5 vector to give the plasmid pHJ33. For activation of the starter unit, the acyl CoA ligase gene (*aurE*) is also required. The *act* KR gene region in pHJ33 was replaced by *aurE* by exchanging the *Hind*III-*Pac*I fragments with a subclone of the *aurE* gene, yielding the resulting plasmid pHJ41 (Fig.3.23). After protoplast transformation, the transformant *S. lividans* ZX1::pHJ41 was cultivated in M10 medium and fed with the starter unit PNBA at the second day. By TLC and LC-MS analyses, no intermediate of aureothin biosynthesis could be detected. However, PNBA that was not consumed for initiation of the polyketide chain still stagnated in the extract of the fermentation of the transformant.

To clone the native *aurA* gene, a *Pst*I fragment containing the missing ACP part of *aurA* and its downstream gene *aurG* was cloned into a medi-vector and then recovered by restriction digestion with *Nco*I and *Eco*RI. This renascent fragment was substituted for the TE domain region in pHJ41 to give plasmid pHJ43 (Fig.3.23). The transformant *S. lividans* ZX1::pHJ43 fed with PNBA could not produce any detectable nitrobenzoate-primed polyketide metabolites either.

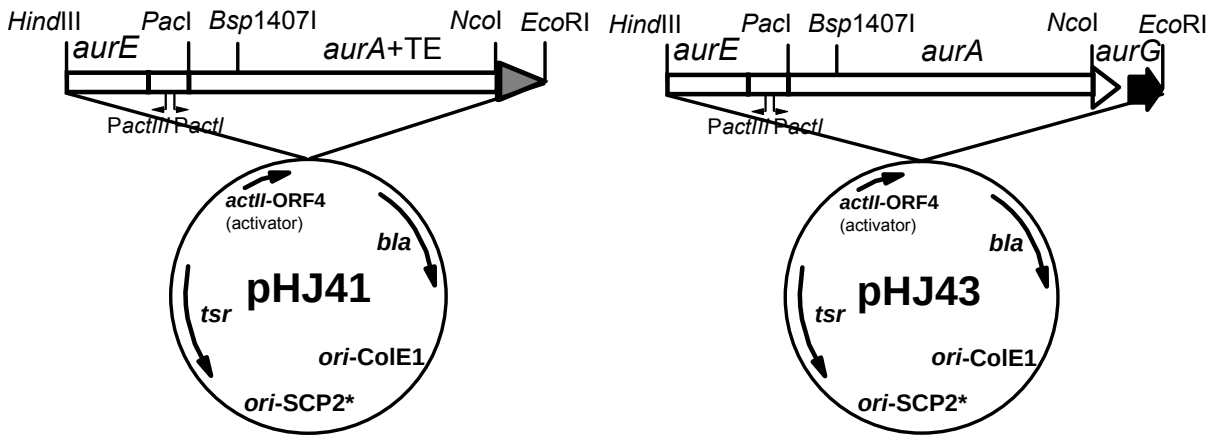


Figure 3.23. Restriction maps of pHJ41 and pHJ43 for expression of the *aurA* gene

2.3.2 Repositioning TE Domain within the Aureothin PKS

One attractive method to gain insight into the course of polyketide assembly would be the identification of pathway intermediates that are anchored to the thio template during chain elongation and processing. In principle, the hydrolytic activity of a *C*-terminally fused thioesterase domain can result in a premature release of the polyketide intermediates, which would enable the isolation and elucidation of all pathway intermediates. In the erythromycin gene cluster, the TE domain has been successfully repositioned to cause a premature chain release (Cortes *et al*., 1995 and Kao *et al*., 1995). The viability of this approach depends critically on the tolerance of the TE to alternative substrates. In the aureothin gene cluster, this method was also adopted to move TE domain in the end of previous PKS module to obtain intermediate after every condensation. In the beginning, the appropriately fused ACP/TE sequence from the *aur* gene cluster was used to replace the corresponding ACP fragment in the end of each PKS module by homologous recombination. A 1038 bp DNA fragment encoding the TE domain and part of the ACP4 domain was amplified by PCR with a pair of oligonucleotides, aurATEL and aurATER. Also, a 672 bp *Hind*III-*Xho*I fragment was amplified as one homologous region of gene replacement constructs by PCR using two primers, shortL and shortR. The DNA sequence of the two PCR products was confirmed by sequencing after cloning into the pGEM-T Easy Vector. Three fragments of a 534 bp *Not*I-*Fsp*I region encoding parts of the KR1 and ACP1 domains (for module 1+TE), a 247 bp *Bam*HI-*Fsp*I region encoding parts of the KR2 and ACP2 domains (for module 1+2+TE), a 1277 bp *Bam*HI-*Fsp*I region encoding parts of the AT3 and ACP3 domains (for module 1+2+3+TE), were used as the other homologous region for homologous recombination, and cloned together with the 1038 bp *Fsp*I-*Eco*RI fragment into *Not*I-*Eco*RI-cut (for module 1+TE) or *Bam*HI-*Eco*RI-cut (for module 1+2+TE and module 1+2+3+TE) pBluescript II SK(-) vector, respectively. The 672 bp *Hind*III-*Xho*I fragment and a 2340 bp *Sma*I fragment containing the streptomycin and spectinomycin resistance gene (*aadA*) for selection of recombinants were inserted into three preceding pBluescript II SK(-) derivatives in succession to give the gene replacement constructs. The whole inserts of three gene replacement constructs, recovered from the pBluescript II SK(-) vector, were purified from agarose gel and introduced into *E. coli* BW25113/pIJ790 containing cosmid pHJ48 by electroporation to perform a double cross-over mediated by the λ Red system. In these three resulting plasmids, pHJ125, pHJ126 and pHJ127 (Fig.3.24), the TE domain has been placed downstream of module 1, 2 and 3, respectively. However, all resulting mutants that express truncated variants of the *aur* PKS were incapable of polyketide production. The reason for the non-working truncated PKS might be the high substrate specificity of the *aur* TE domain.

In another attempt, the heterologous TE domain from the erythromycin gene cluster was used to construct analogous variants of the *aur* PKS. The *ery* TE sequence plus a part of the

adjacent ACP region was amplified by PCR with two primers, PrimerEryL and PrimerEryR, from the erythromycin producer, *Saccharopolyspora erythraea*, to substitute the *aur* ACP/TE region and to construct truncated variants of the aur PKS in analogy the above cloning procedure. Unfortunately, all constructs were unable to produce any polyketide metabolites.

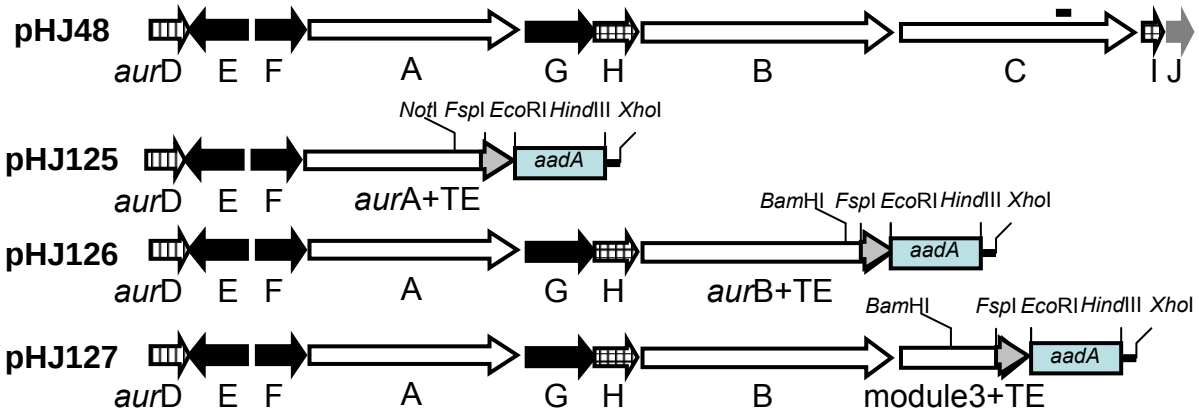


Figure 3.24. Organization of the *aur* genes in recombinant PKS constructed by repositioning the TE domain. Bars indicate a small fragment of module 4 amplified by PCR as the homologous region for recombination.

2.3.3 Fusion of AurA (Module 1) and AurB (Module 2)

Although the *aur* gene cluster encodes only four PKS modules for five rounds of chain extensions, which strongly suggests that module 1 is used twice, the possibility had to be excluded that two copies of AurA might assemble head to tail to yield a perfectly collinear pentamodular PKS. To prove the iterative nature of AurA, the *aurA* and *aurB* genes were fused. In the case of a covalent attachment of the *C*-terminus of AurA (module 1) with the *N*-terminus of AurB (module 2), the occurrence of a twin AurA protein could be ruled out. A similar strategy has recently proven successful for the iterative borrelidin PKS, where genes were fused at the appropriate start/stop codons (Olano *et al*., 2003). Conversely, in this study another strategy was used to accomplish the desired gene fusion. In order to prevent any incompatibility of the linker region, parts of ACP1 and KS2 were replaced with parts of ACP3 and KS4 including the interdomain region. The high similarity of ACP1 and ACP3, KS2 and KS4, appeared ideally suited for the target substitution of partial PKS regions without affecting the functions of the catalytic domains. Several cloning steps were required to yield the desired construct. First, a subclone containing the *aurC* gene was used as template for PCR to amplify the ACP3-KS4 interdomian region, introducing two restriction sites, *Pst*I and *Sgf*I, which were suitable for executing the subsequent substitution. Second, from an *E. coli* – *Streptomyces* shuttle cosmid pHJ48 bearing the entire *aur* biosynthesis gene cluster, a large unique *Kpn*I fragment (ca. 15 kb) was excised and subcloned. After exchanging the biggest *Pst*I-*Sgf*I fragment (3997 bp) with the *Pst*I-*Sgf*I fragment (269 bp) obtained by PCR,

the altered *Kpn*I fragment was reverted to the shuttle cosmid and the correct direction of the insert was ascertained by restriction digestion. In the resulting construct pHJ86, *aurA* and *aurB* were fused in-frame into the same translational unit, while *aurG* and *aurH* were eliminated (Fig.3.25). After introduction of pHJ86 by conjugation into *S. albus*, the transconjugant *S. albus*::pHJ86 was selected for apramycin resistance and cultivated. As the PABA synthase and P450 monooxygenase genes (*aurG* and *aurH*) were eliminated along with the gene fusion, the PABA derived PKS primer PNBA was administered to the culture in order to enable polyketide biosynthesis. TLC, HPLC and MS analyses of the crude extract from a plate culture revealed the formation of a novel metabolite, deoxyaureothin, an aureothin derivative lacking the tetrahydrofurant ring (Fig.3.26), which has already been identified as product of the Δ*aurH* mutant. The production of deoxyaureothin is rationalized by the loss of the *aurH* gene along with the gene fusion. The successful expression of the AurAB fusion protein with concomitant production of deoxyaureothin clearly rules out the possibility that two copies of AurA are forming a twin module. In addition to the domain architecture of AurA, this result provides strong evidence for the iterative use of module 1 during the formation of the diene moiety.

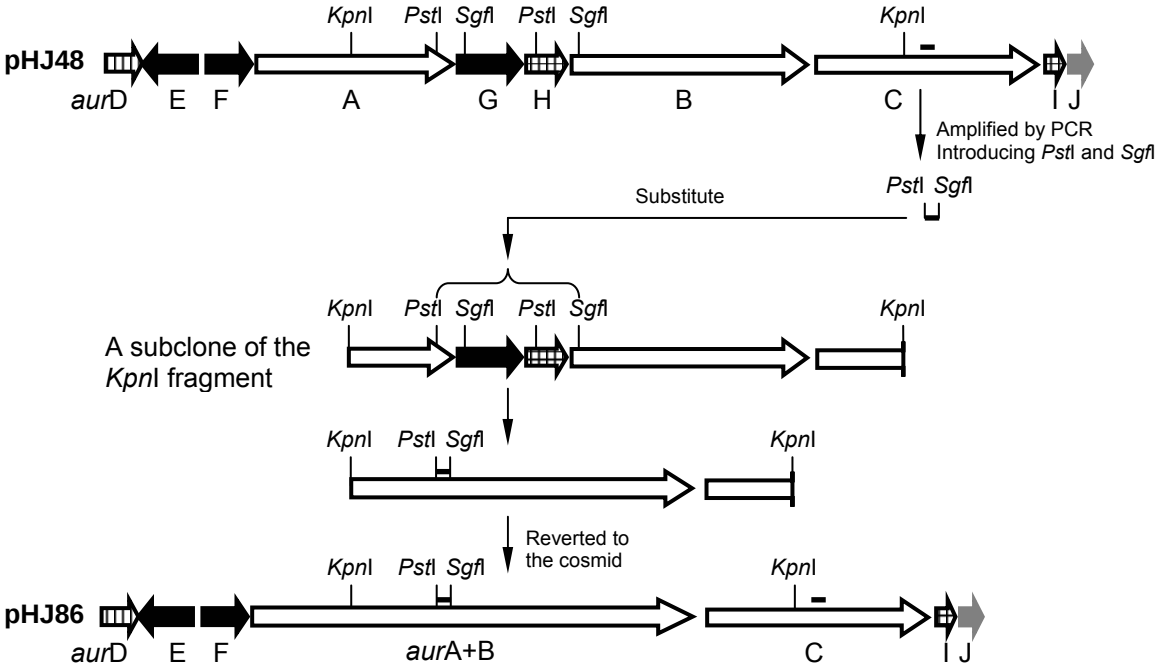


Figure 3.25. Construction of the engineered plasmid (pHJ86) resulting from gene fusion. The ACP3-KS4 interdomain was amplified by PCR with introducing two restriction sites, *Pst*I and *Sgf*I, and substituted for the native *Pst*I-*Sgf*I fragment in a subclone of the *Kpn*I fragment. The altered KpnI fragment was reverted to the shuttle cosmid to give the resulting construct, pHJ86. Bars indicate the ACP3-KS4 interdomain area.

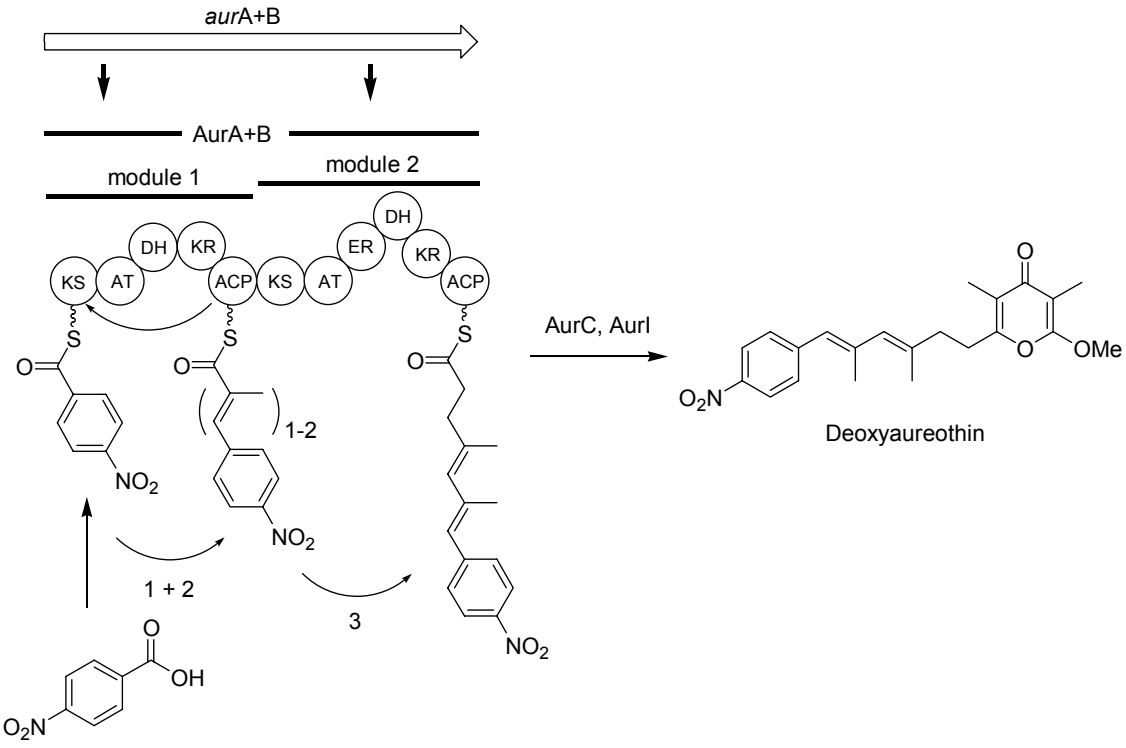


Figure 3.26. Model for deoxyaureothin biosynthesis by a bimodular fusion protein.

2.3.4 Fusion of AurA (Module 1) and AurC (Module 3 and 4)

The successful fusion of *aurA* and *aurB* encouraged another attempt to fuse *aurA* (module 1) and *aurC* (module 3 and 4) together in order to investigate whether AurA could transfer the polyketide intermediate to AurC directly and if so, how many rounds of polyketide elongation AurA could catalyze. By using a related strategy, parts of ACP1 and KS3 were replaced with parts of ACP3 and KS4 including the interdomain region. The ACP3-KS4 interdomain region was amplified by PCR with two primers, LinkerL and LinkerR2 and two restriction sites, *Pst*I and *Acc*III, were introduced into two primers for the following substitution. After exchanging the large *Pst*I-*Acc*III fragment (10607 bp) with the *Pst*I-*Acc*III fragment (326 bp) obtained by PCR, the altered *Kpn*I fragment was ligated into the *Kpn*I sites of the shuttle cosmid pHJ48 and the resulting plasmid, pHJ137, was checked by restriction digestion to verify the correct insert direction. In this new fusion construct, not only *aurG* and *aurH*, but also *aurB* (module 2) were eliminated along with the gene fusion. TLC and MS analyses revealed that the transconjugant *S. albus*::pHJ137 could not produce any nitrobenzoate-primed polyketides even in the presence of administered PNBA. Suppose the assembly line of the polyketide chain was blocked in the resulting shortened PKS, because of the high substrate specificity of AurC or due to an incorrect folding of the mutated PKS.

2.3.5 Site-Directed Mutagenesis of the KS4 and ACP4 Domains

In addition to AurA, it was assumed that iteration also occurs in the third module. As sequence analyses revealed, the AT4 domain shows only weak homology to known AT domains and is possibly inactive due to the lacking active site serine. Since module 3 has the same domain arrangement as module 4, it was tempting to speculate that both final chain extensions are catalyzed by module 3 alone. In order to investigate if module 4 is required for the final chain elongation, site-directed mutagenesis experiments of the KS and ACP active sites were performed. The target regions for mutation were the conserved motifs housing the active site cysteine of KS (VDTACSSS) and the phosphopantheinyl binding site serine of ACP (GFDSL). As previous work on the erythromycin PKS demonstrated, a cysteine to alanine or serine to alanine mutation, respectively, yields inactive PKS domains, but does not affect expression or folding of the enzyme (Thomas *et al*., 2002). Both manipulations in the *aur* PKS were achieved by exchanging gene fragments encoding the KS4 and ACP4 domains with the mutated complements (Fig.3.27).

For inactivation of KS4, a *Pvu*II site was introduced into the active site. The mutated *Kpn*I-*Bsu*36I DNA fragment encoding parts of KS4* and AT4 domains was assembled with two different PCR products, a *Kpn*I-*Pvu*II fragment encoding the *N*-terminal region and a *Pvu*II-*Bsu*36I fragment encoding the *C*-terminal region. In the junction, the KS4 active site cysteine residue was replaced by alanine and a *Pvu*II site was introduced in the region encoding the active site to change the cysteine codon into an alanine codon. After exchanging the naive DNA fragment with the mutated *Kpn*I-*Bsu*36I fragment in cosmid pHJ48, the resulting plasmid, pHJ146, was introduced into protoplasts of *S. lividans* ZX1 by transformation and into *S. albus* by conjugation.

The same strategy was also used for inactivation of the ACP4 domain. Here, an *Afl*III site was introduced at the region encoding the phosphopantheinyl attachment site to change the serine codon into an alanine codon. The mutated *Bsu*36I-*Xba*I fragment encoding parts of AT4*, ACP4 and TE domains was assembled from two PCR products, an *Eco*RI-*Afl*III fragment and an *Afl*III-*Xba*I fragment. In the junction, the ACP active site serine codon was mutated to an alanine codon through introduction of an *Afl*III site. After substituting the native region with the mutated *Bsu*36I-*Xba*I fragment in cosmid pHJ48, the resulting plasmid pHJ144 was introduced into *S. lividans* ZX1 and *S. albus*.

Two transformants, *S. lividans* ZX1::pHJ144 and *S. lividans* ZX1::pHJ146 and two transconjugants, *S. albus*::pHJ144 and *S. albus*::pHJ146, were fermented under the same conditions as the native aureothin producers. However, thorough TLC, HPLC and MS analyses revealed that neither aureothin nor a derivative thereof was formed by any of the mutants. In both cases, inactivation the KS or ACP domains of module 4 clearly resulted in a shutdown of polyketide biosynthesis (Fig.3.27). Consequently, module 4 is in fact required for

the fifth chain elongation of aureothin polyketide production despite its aberrant and probably non-functional AT domain. Thus, an iterative use of module 3 is not taking place.

wt GTC GAC ACC GCC TGC TCC TCG TCC
KS4 V D T A C S S S

⇩

*Pvu*II
*Kpn*I - - - GTC GAC ACa GCt gcC TCC TCG TCC - - - *Bsu*36I
KS4* V D T A A S S S

wt GGC TTC GAC TCG CTG ACC
ACP4 G F D S L T

⇩

*Afl*III
*Eco*RI - - - GGC TTC GAC gCG tTG ACC - - - *Xba*I
ACP4* G F D A L T

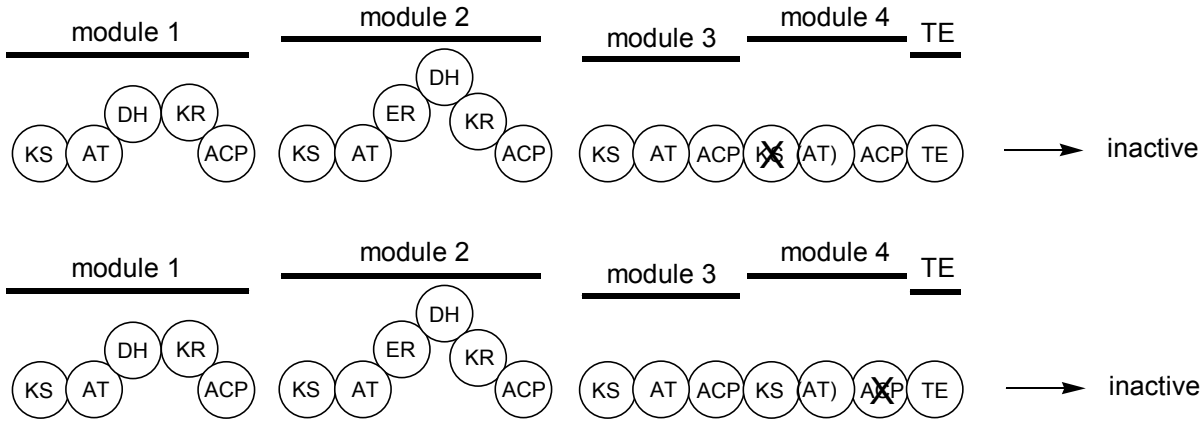


Figure 3.27. Inactivation of KS4 and ACP4 to prove the involvement of module 4.

2.3.6 Construction of an *aurA* Null Mutant and Complementation Experiment

The iterative nature of the first module (AurA) of aureothin biosynthesis has been unequivocally proven by an engineered gene fusion and heterologous expression of the entire gene cluster. However, sequence comparisons of non-iterative and iterative PKS do not reveal what is the reason for the repeated use of the first module. Only AurA is self-sufficient for the iteration or this is the cooperation of AurA and AurB? Perhaps because of

the substrate specificity of AurB, only the polyketide chain resulting from two elongation reactions catalyzed by AurA could enter into the latter extension steps to yield the resultant product, aureothin. In order to investigate the substrate of AurB and prepare a suitable host for engineering hybrid AurA or coexpressing other PKS, a construct which contains the whole aureothin gene cluster except for the *aurA* gene was made by gene deletion.

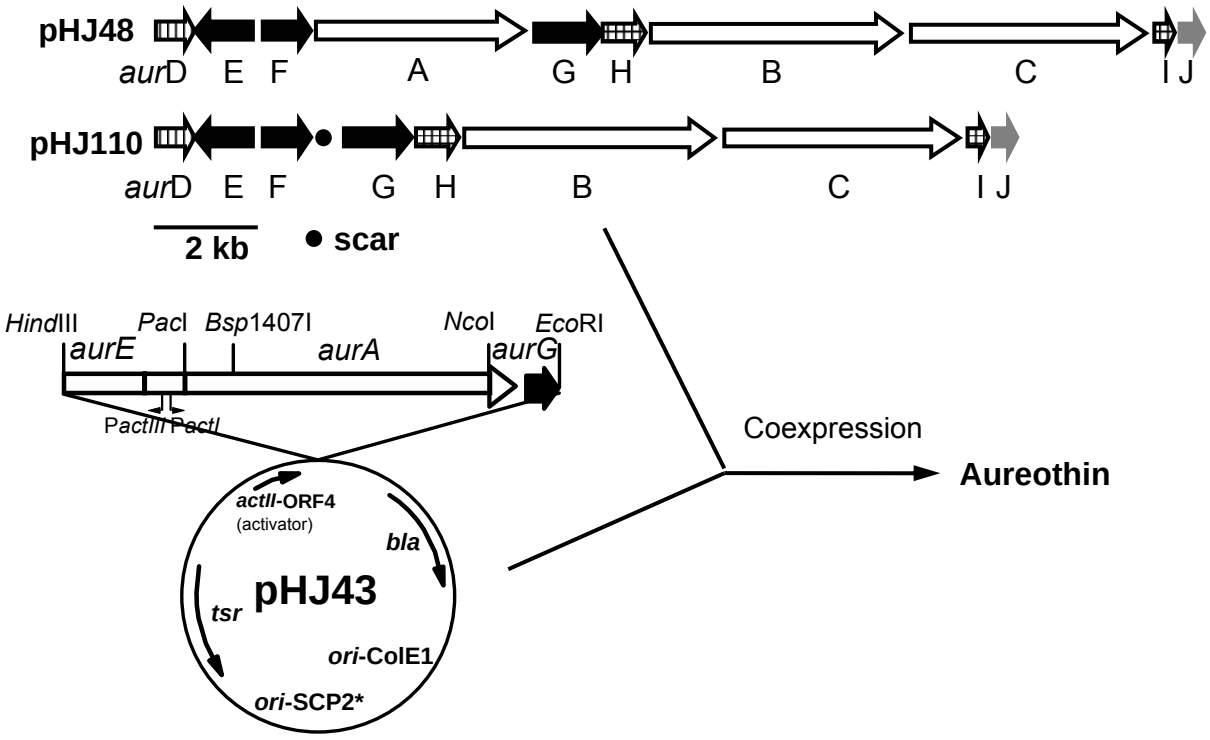


Figure 3.28. Coexpression of the *aurA* null mutant with *aurA* on pHJ43 to restore the aureothin biosynthesis.

The *aurA* null mutant was constructed by PCR-targeting system. Two long primers, aurAL and aurAR, were employed to amplify the extended *aadA* gene flanked by FRT sites. The PCR product was introduced into *E. coli* BW25113/pIJ790 containing cosmid pHJ48 and subsequent gene replacement of the *aurA* gene by the extended antibiotic resistance cassette automatically occurred under the enhanced function of λ Red system. The inserted cassette was then removed through the expression of the FLP recombinase. In the resulting plasmid, pHJ111, the entire *aurA* gene was in-frame deleted except for the start (ATG) and stop (TAG) codons. After introducing pHJ111 into *S. lividans* ZX1 by protoplast transformation, TLC, HPLC and MS analyses confirmed that the transformant *S. lividans* ZX1::pHJ111 could produce neither aureothin nor any derivatives. When pHJ43, in which the *aurA* gene was cloned into a *Streptomyces* self-replication vector (pRM5) for expression, was introduced into the protoplasts of transformant *S. lividans* ZX1::pHJ111, the aureothin biosynthesis was restored (Fig.3.28). The identity of the aureothin metabolite was

undoubtedly proven with standard aureothin as a reference by TLC, HPLC and MS. Construction of the *aurA* null mutant and its complementation experiments are the prerequisite for investigating the substrate specificity of AurB by feeding different synthetic intermediates. In the *aurA* null mutant, only *aurA* was deleted, but other genes in the *aur* gene cluster are still available. If AurB would accept a suitable substrate from feeding experiments, the polyketide biosynthetic pathway would result in the formation of aureothin. By cooperation with a colleague, Nelly Traitcheva, the *norA* gene from the neoaureothin biosynthesis gene cluster, which shows very high identity to *aurA*, was coexpressed in the *aurA* null mutant. The cross complementation with *norA* also restored the aureothin biosynthesis.

2.3.7 Discussion

In contrast to aromatic polyketide synthases, as well as plant and fungal PKS, type I module PKS constitute a processing line along which the polyketide backbone is assembled and processed. Usually the number of modules strictly mirrors the number of elongation and reduction cycles. This one-to-one correspondence of the polyketide metabolite to the modular type I PKS architecture has been accepted as a principle in the programming of modular polyketide synthases and has provided the molecular basis for targeted genetic engineering resulting in natural product derivatives (Staunton *et al*., 2001 and Walsh 2002 and Cane *et al*., 1998). Nevertheless, the PKS programming can be error-prone. For example, the unusual behaviour of “stuttering” and “skipping” phenomena, where a module is used repeatedly or omitted during chain elongation, were observed in the ring-enlarged and ring-contracted side products from large-scale fermentations of the erythromycin (Wilkinson *et al*., 2000 and Thomas *et al*., 2002) and epothilone producer (Hardt *et al*., 2001), respectively. While the erythromycin and epothilone derivatives are obviously the results of an erratic processing of the PKS, recent observations provide growing evidence that iteration or stuttering can in fact be a programmed event encoded by the modular polyketide synthase. The first genomic indication for an iterative use of a single PKS module was reported by the group of R. Müller who found that a gene encoding a PKS module is missing in the stigmatellin (Fig.3.29) gene cluster from *Stigmatella aurantiaca*. Since the lacking PKS gene could not be detected elsewhere in the genome by hybridization, it was concluded that one PKS module would be compensated by the repeated use of a functionally equivalent module, either StiH or StiJ (Gaitatzis *et al*., 2002).

In this work, the aureothin biosynthesis gene cluster was cloned and sequenced. According to the principle of co-linearity, five PKS modules would be required for the assembly of the polyketide backbone. However, sequence analyses and heterologous expression of the entire *aur* gene cluster revealed that four modules are sufficient for catalyzing five rounds of

elongation and processing. Expression of the AurAB fusion protein and the domain architecture unequivocally revealed that the first module is used twice, giving rise to the diene substructure.

Almost simultaneously with these findings, the Salas and Leadlay groups jointly communicated that in borrelidin (*bor*) (Fig.3.29) biosynthesis a single elongation module is used even three times (Olano *et al*., 2003). While classical modular PKS show a one-to-one correspondence between PKS architecture and polyketide product, the novel *sti*, *aur* and *bor* PKS systems clearly breach with the generally accepted principle of co-linearity. However, sequence comparisons of non-iterative and iterative PKS do not point towards the molecular ground for iteration, and to date the reason why PKS act iteratively or non-iteratively remains a mystery.

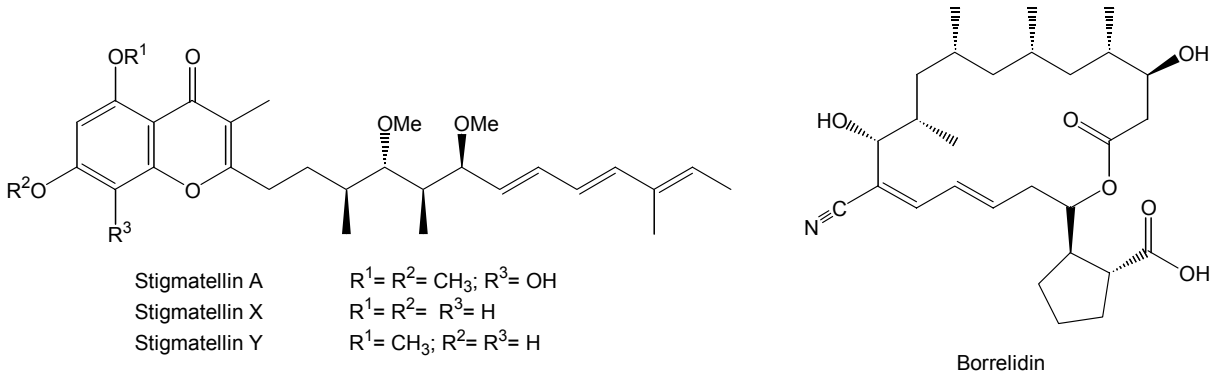


Figure 3.29. Chemical structure of stigmatellin and borrelidin.

For gaining a deeper insight into the intriguing code of the *aur* PKS, several attempts including the expression of the hybrid AurA+TE and the repositioning of the TE domain in the downstream region of each PKS module were made to investigate the course of polyketide assembly. However, all efforts expressing truncated PKS variants failed. This could be explained by a high substrate specificity of the TE or by an incorrect folding of the mutated PKS. Also, the failure of expressing the AurAC fusion protein could be due to the similar reason.

Point mutations of the KS4 and ACP4 domains, which result in a shutdown of polyketide biosynthesis, excluded the possibility of the iteration of module 3. Even if the AT4 domain is aberrant and supposed to be inactive, module 4 is still required for the last elongation step. How can this observation be rationalized? Recently, J. Piel reported on the surprising discovery of a type I PKS from a pederus bacterial symbiont, where all PKS modules are lacking AT domains (Piel, 2002). This novel PKS system is served by iterative acyl transferases *in trans*, which are also encoded in the same gene cluster. Similar findings have been made in other more recent cases, i.e. the leinamycin (Cheng *et al*., 2003) and the mupirucin (El-Sayed *et al*., 2003) PKS. Interestingly, apparently non-functional AT domain

have also been identified in *cis*-AT PKS systems, such as the pyoluteorin PKS (Nowak-Thompson *et al*., 1999 and 2003), as well as the neocarzilin PKS, which represents the most recent example of an iterative modular PKS (Otsuka *et al*., 2004). In these cases, as in the *aur* PKS, it remains a riddle how the anomalous AT domains function or how they might be complemented.

Now, the whole biosynthetic pathway of the aureothin biosynthesis (Fig.4.1) can be clearly portrayed by summarizing all the results from sequence analyses and functional studies of the *aur* gene cluster. First, chorismate acid is converted into PABA by AurG. Subsequently, PABA is oxidized into PNBA by AurF, a novel *N*-oxygenase. The unusual PKS starter unit, PNBA, is activated by AurE and then loaded directly onto the PKS to initiate the polyketide chain elongations. Surprisingly, the first PKS module, AurA, catalyzes two rounds of elongation cycles during the polyketide assembly. After the polyketide chain is released from the TE domain, the first post-PKS tailoring step, *O*-methylation of the pyrone ring, is performed by AurI, and then a bifunctional P450 oxygenase, AurH, catalyzes the formation of the homochiral furan ring, yielding the resultant product, aureothin.

D. Summary

Aureothin is a rare nitroaryl-substituted metabolite from the soil bacterium *Streptomyces thioluteus*, which exhibits a variety of biological activities, including antifungal, antitumoral and insecticidal. Early labelling studies revealed that aureothin is composed of an unknown starter unit, one acetate and four propionate units. By construction and screening of a *S. thioluteus* HKI-227 genomic cosmid library, the aureothin (*aur*) gene cluster was cloned and sequenced. The identity of the *aur* gene cluster was proven by heterologous expression. Detailed sequence analyses revealed that aureothin is synthesized by an unusual modular type I polyketide synthase (PKS), which has several exceptional features, including a novel priming mechanism and the iterative use of one module. Typically, modular type I PKS are arranged in an assembly line fashion, harboring a catalytic domain for each biosynthetic step. In stark contrast to this accepted principle of co-linearity, sequence analyses and heterologous expression of the *aur* gene cluster demonstrated that only four modules are sufficient for catalyzing five rounds of polyketide chain extensions. The engineered fusion of the first two PKS genes and the domain architecture unequivocally revealed that the first module catalyzes two successive cycles of chain extension, one of the first examples of a modular PKS, in which such iteration is required to produce the normal polyketide product. Furthermore, point mutants of the KS4 (ketosynthase) and ACP4 (acyl carrier protein) domains revealed that module 4 is in fact required for the last chain elongation despite its aberrant and probably non-functional acyl transferase (AT) domain. Further functional analyses of the *aur* biosynthesis gene cluster disclosed the whole biosynthetic pathway of aureothin.

The biosynthetic origin of the rare nitro aryl moiety from aureothin was elucidated using both molecular and chemical approaches. Incorporation of d_4-p-nitrobenzoate (d_4-PNBA) into aureothin in a feeding experiment demonstrated for the first time that PNBA may serve as a polyketide starter unit. Cloning, expression and deletion of *aurF* gene revealed that p-nitrobenzoate (PABA) is transformed into the corresponding nitro compound by an unprecedented type of *N*-oxygenase (AurF), which could represent the first member of a new family of oxidizing enzymes.

Furthermore, the order of the tailoring reactions was elucidated by inactivation, complementation and expression of the putative tailoring enzymes. AurI catalyzes the O-methylation of the pyrone ring. AurH is a multifunctional cytochrome P450 monooxygenase catalyzing the formation of the homochiral furan ring of aureothin, which constitutes the last step in the aureothin biosynthetic pathway.

In addition to elucidating the unusual biosynthetic pathway of aureothin, mutants were generated that produce aureothin derivatives, some of which show enhanced biological

activities compared to aureothin, such as deoxyaureothin with increased cytotoxic activity. In addition, the genetic engineering of the *aur* gene cluster has set the basis for the mutasynthesis of aureonitrile, an aureothin derivative with significantly improved cytostatic effects.

Further studies on the detailed mechanism of the unusual iterative modular *aur* PKS system will provide new insights into the molecular basis and the evolutionary relationship of iterative and non-iterative polyketide synthases and set the ground for engineering novel unnatural natural aureothin derivatives.

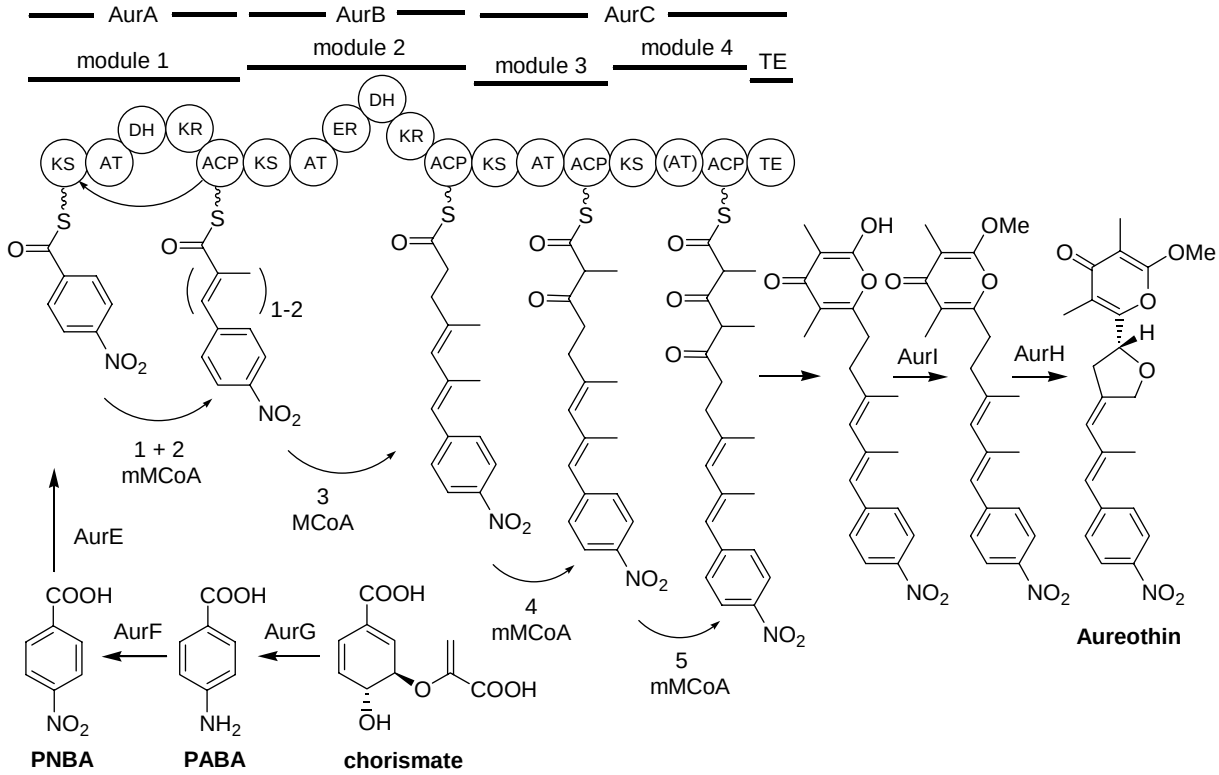


Figure 4.1. Model of non-linear aureothin biosynthesis from chorismate.

E. Zusammenfassung

Aureothin ist ein außergewöhnliches Nitroaryl-substituiertes Polyketid, das von dem Bodenbakterium *Streptomyces thioluteus* produziert wird und eine Vielzahl biologischer Aktivitäten (u.a. antifungisch, zytostatisch, insektizid) aufweist. Klassische Labeling-Studien zeigten, dass Aureothin aus einem unbekannten Starter, einer Acetat- und vier Propionat-Einheiten zusammengesetzt ist.

Zur Klonierung und Sequenzierung des Aureothin (*aur*) Genklusters, wurde eine genomische Cosmidbibliothek von *S. thioluteus* HKI-227 konstruiert und gescreent. Detaillierte Sequenzanalysen des *aur* Genklusters zeigten, dass Aureothin von einer außergewöhnlichen modularen Typ I Polyketidsynthase (PKS) synthetisiert wird. Neben einem neuartigen Priming-Mechanismus und der Nutzung von p-Nitrobenzoat (PNBA) als Starteinheit, ist hier besonders die wiederholte Nutzung eines Moduls der Aureothin-Synthase herauszustellen.

Im Gegensatz zu gewöhnlichen bakteriellen Typ I Polyketidsynthasen, die für jeden Biosyntheseschritt eine katalytische Domäne besitzen, enthält die *aur* PKS für die fünf durchzuführenden Polyketid-Kettenverlängerungen nur vier Module. Durch heterologe Expression des *aur* Genklusters konnte gezeigt werden, dass die gesamte genetische Information für die Aureothin Biosynthese auf diesen vier Modulen vorhanden ist.

Ergänzend zur heterologen Expression konnte durch die Fusion der ersten beider PKS Gene eindeutig demonstriert werden, dass ein Modul zwei aufeinanderfolgende Kettenverlängerungszyklen katalysiert. Die *aur* PKS stellt damit eines der ersten Beispiele für eine iterative oder „stotternde“ bakterielle Typ I PKS dar.

Trotz einer aberranten und wahrscheinlich nicht-funktionalen Acyltransferase (AT)-Domäne, bestätigten Punktmutationen in der KS4 (Ketosynthase)- und der ACP4 (Acylcarrierprotein)- Domäne, dass das Modul 4 für den letzten Kettenverlängerungsschritt verantwortlich ist.

Weitere Funktionsanalysen des Genklusters trugen dazu bei, den Aurethin-Biosyntheseweg näher zu verstehen.

Die biosynthetische Herkunft der seltenen Nitroaryl-Gruppe des Aureothins wurde sowohl durch molekulare als auch durch chemische Methoden untersucht. Der Einbau von d_4-PNBA in Aureothin in Fütterungsexperimente bestätigte zum ersten mal, dass PNBA als Polyketid-Startereinheit genutzt werden kann. Klonierung, Expression und Deletion des *aur*F Gens zeigte, dass p-Aminobenzoesäure (PABA) durch eine bisher unbekannte N-Oxygenase (AurF) in die entsprechende Nitro-Verbindung überführt wird. AurF repräsentiert das erste Mitglied einer neuen Familie von oxidierenden Enzymen.

Weiterhin wurde der Ablauf der Tailoring-Reaktionen durch Inaktivierung, Komplementierung und Expression der Tailoring-Enzyme untersucht. Die O-Methyltransferase AurI katalysiert die

O-Methylierung des Pyronringes. AurH, eine multifunktionale Cytochrom P450 Monooxygenase, katalysiert die Formation des homochiralen Furanringes und damit den letzten Schritt in der Aureothin Biosynthese.

Neben der Funktionsaufklärung des aur Genclusters führten die gezielten Mutanten zur Produktion verschiedener Aureothin-Derivate mit verändertem Wirkprofil, wie z.B. Deoxyaureothin mit erhöhter Zytotoxizität. Zudem bildeten Mutagenesestudien die Grundlage für die Mutasynthese von Aureonitril, einem Aureothin-Derivat mit signifikant erhöhtem zytostatischen Effekt.

Das weitere Studium des detaillierten Mechanismus und der Evolution des ungewöhnlichen modularen *aur* PKS – Systems wird neue Einblicke in das Rätsel des „Stotterns“ gewähren und damit die Grundlage für die biotechnologische Erzeugung neuer Verbindungen bilden.

F. References

Alting-Mees, M.A. and Short, J.M. (1989). pBluescript II: gene mapping vectors. Nucleic. Acids. Res. 17: 9494

Aparicio, J.F., Molnar, I., Schwecke, T., Konig, A., Haydock, S.F., Khaw, L.E., Staunton, J. and Leadlay, P.F. (1996). Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase. Gene 169: 9-16

Aparicio, J.F., Fouces, R., Mendes, M.V., Olivera, N. and Martin, J.F. (2000). A complex multienzyme system encoded by five polyketide synthase genes is involved in the biosynthesis of the 26-membered polyene macrolide pimaricin in *Streptomyces natalensis*. Chem. Biol. 7: 895-905

Arima, K., Imanaka, H., Kousaka, M., Fukuda, A. and Tamura, G. (1965). Studies on pyrrolnitrin, a new antibiotic. I. Isolation and properties of pyrrolnitrin. J. Antibiot (Tokyo). 18: 201-204

Ashman, P.J., Mackenzie, A. and Bramley, P.M. (1990). Characterization of ent-kaurene oxidase activity from *Gibberella fujikuroi*. Biochim. Biophys. Acta. 1036: 151-157

Beck, B.J., Aldrich, C.C., Fecik, R.A., Reynolds, K.A. and Sherman, D.H. (2003). Iterative chain elongation by a pikromycin monomodular polyketide synthase. J. Am. Chem. Soc. 125: 4682-4683

Berdy, J. (1987). CRC Handbook of Antibiotic Compounds: Microbial Metabolites. CRC Press

Beyer, S., Kunze, B., Silakowski, B. and Muller, R. (1999). Metabolic diversity in myxobacteria: identification of the myxalamid and the stigmatellin biosynthetic gene cluster of *Stigmatella aurantiaca* Sg a15 and a combined polyketide-(poly)peptide gene cluster from the epothilone producing strain *Sorangium cellulosum* So ce90. Biochim. Biophys. Acta. 1445: 185-195

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

Bibb, M.J., Findlay, P.R. and Johnson, M.W. (1984). The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene 30: 157-166

Bibb, M.J., Biro, S., Motamedi, H., Collins, J.F. and Hutchinson, C.R. (1989). Analysis of the nucleotide sequence of the *Streptomyces glaucescens tcmI* genes provides key information about the enzymology of polyketide antibiotic biosynthesis. EMBO J. 8: 2727-2736

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic. Acids. Res. 7: 1513-1523

Blanc, V., Gil, P., Bamas-Jacques, N., Lorenzon, S., Zagorec, M., Schleuniger, J., Bisch, D., Blanche, F., Debussche, L., Crouzet, J. and Thibaut, D. (1997). Identification and analysis of genes from *Streptomyces pristinaespiralis* encoding enzymes involved in the biosynthesis of the 4-dimethylamino-L-phenylalanine precursor of pristinamycin I. Mol. Microbiol. 23: 191-202

Bohm, I., Holzbaur, I.E., Hanefeld, U., Cortes, J., Staunton, J. and Leadlay, P.F. (1998). Engineering of a minimal modular polyketide synthase, and targeted alteration of the stereospecificity of polyketide chain extension. Chem. Biol. 5: 407-412

Brautaset, T., Sekurova, O.N., Sletta, H., Ellingsen, T.E., StrLm, A.R., Valla, S. and Zotchev, S.B. (2000). Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei ATCC* 11455: analysis of the gene cluster and deduction of the biosynthetic pathway. Chem. Biol. 7: 395-403

Brautaset, T., Bruheim, P., Sletta, H., Hagen, L., Ellingsen, T.E., Strom, A.R., Valla, S. and Zotchev, S.B. (2002). Hexaene derivatives of nystatin produced as a result of an induced rearrangement within the *nysC* polyketide synthase gene in *S. noursei* ATCC 11455. Chem. Biol. 9: 367-373

Brown, M.P., Aidoo, K.A. and Vining, L.C. (1996). A role for *pabAB*, a p-aminobenzoate synthase gene of *Streptomyces venezuelae* ISP5230, in chloramphenicol biosynthesis. Microbiology 142: 1345-1355

Caballero, J.L., Martinez, E., Malpartida, F. and Hopwood, D.A. (1991). Organisation and functions of the *actVA* region of the actinorhodin biosynthetic gene cluster *of Streptomyces coelicolor*. Mol. Gen. Genet. 230: 401-412

Campelo, A.B. and Gil, J.A. (2002). The candicidin gene cluster from *Streptomyces griseus* IMRU 3570. Microbiol. 148: 51-59

Cane, D.E., Walsh, C.T. and Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations, and mutations. Science 282: 63-68

Cardillo, R., Fuganti, C., Ghiringhelli, D. and Giangrasso, D. (1974). The biosynthesis od aureothin. Tetrahedron 30: 459-461

Chakraburtty, R. and Bibb, M. (1997). The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. J. Bacteriol. 179: 5854-5861

Chassy B.M., Mercenier, A. and Flickinger, J. (1988). Transformation of bacteria by electroporation. Trends. Biotechnol. 6: 303-309

Chater, K.F. and Bruton, C.J. (1985). Resistance, regulatory and production genes for the antibiotic methylenomycin are clustered. EMBO J. 4: 1893-1897

Chater, K.F. and Losick, R. (1996). The mycelial life-style of *Streptomyces coelicolor* A3(2) and its relatives. In Shapiro, J.H. and Dworkin, M. (eds.) Bacteria as Multicellular Organisms. Oxford University Press. Oxford

Chen, S., Huang, X., Zhou, X., Bai, L., He, J., Jeong, K.J., Lee, S.Y. and Deng, Z. (2003). Organizational and mutational analysis of a complete FR-008/candicidin gene cluster encoding a structurally related polyene complex. Chem. Biol. 10: 1065-1076

Chen, T.S. and Inamine, E.S. (1989). Studies on the biosynthesis of avermectins. Arch. Biochem. Biophys. 270: 521-525

Cheng, Y.Q., Tang, G.L. and Shen, B. (2003). Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. Proc. Natl. Acad. Sci. USA. 100: 3149-3154

Cortes, J., Wiesmann, K.E., Roberts, G.A., Brown, M.J., Staunton, J. and Leadlay, P.F. (1995). Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. Science 268: 1487-1489

Cortes, J., Haydock, S.F., Roberts, G.A., Bevitt, D.J. and Leadlay. P.F. (1990). An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. Nature 348: 176-178.

Criado, L.M., Martin, J.F. and Gil, J.A. (1993). The *pab* gene of *Streptomyces griseus*, encoding p-aminobenzoic acid synthase, is located between genes possibly involved in candicidin biosynthesis. Gene 126: 135-139

Cropp, T.A., Wilson, D.J. and Reynolds, K.A. (2000). Identification of a cyclohexylcarbonyl CoA biosynthetic gene cluster and application in the production of doramectin. Nat. Biotechnol. 18: 980-983

Cullum, J., Flett, F. and Piendl, W. (1988). Genetic instability in streptomycetes. Microbiol. Sci. 5: 233-235

Dhillon, N., Hale, R.S., Cortes, J. and Leadlay, P.F. (1989). Molecular characterization of a gene from *Saccharopolyspora erythraea* (*Streptomyces erythraeus*) which is involved in erythromycin biosynthesis. Mol. Microbiol. 3: 1405-1414

Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J. and Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. Science 252: 675-679

Donadio, S., McAlpine, J.B., Sheldon, P.J., Jackson, M. and Katz, L. (1993). An erythromycin analog produced by reprogramming of polyketide synthesis. Proc. Natl. Acad. Sci. USA. 90: 7119-7123

Dutton, C.J., Gibson, S.P., Goudie, A.C., Holdom, K.S., Pacey, M.S., Ruddock, J.C., Bu'Lock, J.D. and Richards, M.K. (1991). Novel avermectins produced by mutational biosynthesis. J. Antibiot. (Tokyo). 44: 357-365

Edman, J.C., Goldstein, A.L. and Erbe, J.G. (1993). Para-aminobenzoate synthase gene of *Saccharomyces cerevisiae* encodes a bifunctional enzyme. Yeast 9: 669-675

El-Sayed, A.K., Hothersall, J., Cooper, S.M., Stephens, E., Simpson, T.J. and Thomas, C.M. (2003). Characterization of the mupirocin biosynthesis gene cluster *from Pseudomonas fluorescens* NCIMB 10586. Chem. Biol. 10: 419-430

Ensign, J.C. (1978). Formation, properties and germination of actinomycetes spores. Ann. Rev. Microbiol. 32: 185-219

Fierro, J.F., Hardisson, C. and Salas, J.A. (1987). Resistance to oleandomycin in *Streptomyces antibioticus*, the producer organism. J. Gen. Microbiol. 133: 1931-1939

Flett, F., Mersinias, V. and Smith, C.P. (1997). High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. FEMS Microbiol. Lett. 155: 223-229

Fujii, I. in comprehensive Natural Products. Sankawa, U. (eds.) Elsevier. Oxford University press. Oxford

Gaisser, S., Trefzer, A., Stockert, S., Kirschning, A. and Bechthold, A. (1997). Cloning of an avilamycin biosynthetic gene cluster from *Streptomyces viridochromogenes* Tu57. J. Bacteriol. 179: 6271-6278

Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blocker, H., Hofle, G. and Muller, R. (2002). The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. J. Biol. Chem. 277: 13082-13090

Gandecha, A.R., Large, S.L. and Cundliffe, E. (1997). Analysis of four tylosin biosynthetic genes from the *tylLM* region of the *Streptomyces fradiae* genome. Gene 184: 197-203

Gokhale, R.S., Hunziker, D., Cane, D.E. and Khosla, C. (1999). Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase. Chem. Biol. 6: 117-125

Gokhale, R.S., Lau, J., Cane, D.E. and Khosla, C. (1998). Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase. Biochemistry 37: 2524-2528

Green, J.M. and Nichols, B.P. (1991). *p*-Aminobenzoate biosynthesis in *Escherichia coli*. Purification of aminodeoxychorismate lyase and cloning of *pabC*. J. Biol. Chem. 266: 12971-12975

Gust, B., Challis, G.L., Fowler, K., Kieser, T. and Chater, K.F. (2003). PCR-targeted

Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc. Natl. Acad. Sci. USA. 100: 1541-1546

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

Hardt, I.H., Steinmetz, H., Gerth, K., Sasse, F., Reichenbach, H. and Hofle, G. (2001). New natural epothilones from *Sorangium cellulosum*, strains So ce90/B2 and So ce90/D13: isolation, structure elucidation, and SAR studies. J. Nat. Prod. 64: 847-856

Haydock, S.F., Dowson, J.A., Dhillon, N., Roberts, G.A., Cortes, J. and Leadlay, P.F. (1991). Cloning and sequence analysis of genes involved in erythromycin biosynthesis in *Saccharopolyspora erythraea*: sequence similarities between EryG and a family of S-adenosylmethionine-dependent methyltransferases. Mol. Gen. Genet. 230: 120-128

Haydock, S.F., Aparicio, J.F., Molnar, I., Schwecke, T., Khaw, L.E., Konig, A., Marsden, A.F., Galloway, I.S., Staunton, J. and Leadlay, P.F. (1995). Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. FEBS Lett. 374: 246-248

He, J., Magarvey, N., Piraee, M. and Vining, L.C. (2001). The gene cluster for chloramphenicol biosynthesis in *Streptomyces venezuelae* ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene. Microbiology 147: 2817-2829

Hobbs, G., Frazer, C.M., Gardner, D.C.J., Cullum, J.A. and Oliver, S.G. (1989). Dispersed growth of *Streptomyces* in liquid culture. Appl. Microbiol. Biotechnol. 31: 272-277

Hodgson, D.A. (1992). Differentiation in Actinomycetes. In Mohan, S., Dow, C. and Cole, J.A. (eds.) Prokaryotic structure and Function: a New Perspective, Society for General Microbiology Symposium. Cambridge University Press

Hopwood, D.A. (1967). Genetic analysis and genome structure in *Streptomyces coelicolor*. Bacteriol. Rev. 31: 373-403

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

Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., Ward, J.M. and Schrempf, H. (1985). Genetic Manipulation of *Streptomyces*-A Laboratory Manual. John Innes Press. Norwich

Hopwood, D.A., Malpartida, F., Kieser, H.M., Ikeda, H., Duncan, J., Fujii, I., Rudd, B.A.M., Floss, H.G. and Ōmura, S. (1985). Production of “hybrid” antibiotics by genetic engineering. Nature 314: 642-644

Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. Chem. Rev. 97: 2465-2497

Hopwood, D.A. (2004). Cracking the polyketide code. PloS. Biol. 2: 0166-0169

Horinouchi, S., Kito, M., Nishiyama, M., Furuya, K., Hong, S.K., Miyake, K. and Beppu, T. (1990). Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2). Gene 95: 49-56

Horinouchi, S. and Beppu, T. (1994). A-factor as a microbial hormone that controls cellular differentiation and secondary metabolism in *Streptomyces griseus*. Mol. Microbiol. 12: 859-864

Hu, Z.H., Bao, K., Zhou, X.F., Zhou, Q., Hopwood, D.A., Kieser, T. and Deng, Z.X. (1994). Repetitive polyketide synthesis modules involved in the biosynthesis of a haptatene macrolide of *Streptomyces hygroscopicus* FR-008. Mol. Microbiol. 14: 163-172

Hunziker, D., Yu, T-W., Hutchinson, C. R., Floss, H.G. and Khosla, C. (1998). Primer unit specificity in rifamycin biosynthesis principally resides in the later stages of the biosynthetic pathway. J. Am. Chem. Soc. 120: 1092-1093

Ikeda, H., Nonomiya, T., Usami, M., Ohta, T. and Omura, S. (1999). Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. Proc. Natl. Acad. Sci. USA. 96: 9509-9514

Kakinuma, K., Hanson, C.A. and Rinehart, K.L. (1976). Spectinabilin, a new nitro-containing metabolite isolated from *Streptomyces spectabilis*. Tetrahedron 32: 217-222

Kalaitzis, J.A., Izumikawa, M., Xiang, L., Hertweck, C. and Moore, B.S. (2003). Mutasynthesis of enterocin and wailupemycin analogues. J. Am. Chem. Soc. 125: 9290-9291

Kao,C.M., Katz, L. and Khosla, C. Engineered biosynthesis of a complete macrolactone in a heterologous host. Science 265: 509-512

Kao, C.M., Luo, G., Katz, L., Cane, D.E., and Khosla, C. (1995). Manipulation of macrolide ring size by directed mutagenesis of a modular polyketide synthase. J. Am. Chem. Soc. 117: 9105-9106

Kao, C.M., Pieper, R., Cane, D.E. and Khosla, C. (1996). Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase. Biochemistry 35: 12363-12368

Kawai, S., Kobayashi, K., Oshima, T. and Egami, F. (1965). Studies on the oxidation of *p*-aminobenzoate to *p*-nitrobenzoate by *Streptomyces thioluteus*. Arch. Biochem. Biophys. 112: 537-543

Kealey, J.T. (2003). Creating polyketide diversity through genetic engineering. Front Biosci. 8: 01-13

Keller, N.P. and Hohn, T.M. (1997). Metabolic pathway gene clusters in filamentous fungi. Fungal Genet. Biol. 21: 17-29

Kennedy, B.W. and Alcorn, S.M. (1980). Estimates of US crop losses to prokaryote plant pathogens. Plant Dis. 64: 674-676

Kennedy, J., Auclair, K., Kendrew, S.G., Park, C., Vederas, J.C. and Hutchinson, C.R. (1999). Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science 284: 1368-1372

Khosla, C. (2000). Natural product biosynthesis: a new interface between enzymology and medicine. J. Org. Chem. 65: 8127-8133

Kieser, H.M., Kieser, T. and Hopwood, D.A. (1992). A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome. J. Bacteriol. 174: 5496-5507

Kieser, T. (1984). Factors affecting the isolation of CCC DNA from *Streptomyces lividans* and *Escherichia coli*. Plasmid 12: 19-36

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A. (2000). Practical *Streptomyces* Genetics. John Innes Press. Norwich

Kirner, S. and van Pee, K.H. (1994). The biosynthesis of nitro compounds: the enzymatic oxidation to pyrrolnitrin of its amino-substituted precursor. Angew. Chem. Int. Ed. Engl. 33: 352-352

Kirner, S., Hammer, P.E., Hill, D.S., Altmann, A., Fischer, I., Weislo, L.J., Lanahan, M., van Pee, K.H. and Ligon, J.M. (1998). Functions encoded by pyrrolnitrin biosynthetic genes from *Pseudomonas fluorescens*. J. Bacteriol. 180: 1939-1943

Kuhstoss, S., Huber, M., Turner, J.R., Paschal, J.W. and Rao, R.N. (1996). Production of a novel polyketide through the construction of a hybrid polyketide synthase. Gene 183: 231-236

Lamb, D.C., Kelly, D.E., Manning, N.J., Kaderbhai, M.A. and Kelly, S.L. (1999). Biodiversity of the P450 catalytic cycle: yeast cytochrome b5/NADH cytochrome b5 reductase complex efficiently drives the entire sterol 14-demethylation (CYP51) reaction. FEBS Lett. 462: 283-288

Leblond, P., Demuyter, P., Simonet, J.M. and Decaris, B. (1990). Genetic instability and hypervariability in *Streptomyces ambofaciens*: towards an understanding of a mechanism of genome plasticity. Mol. Microbiol. 4: 707-714

Leblond, P., Demuyter, P., Simonet, J.M. and Decaris, B. (1991). Genetic instability and associated genome plasticity in *Streptomyces ambofaciens*: pulsed-field gel electrophoresis evidence for large DNA alterations in a limited genomic region. J. Bacteriol. 173: 4229-4233

Lechevalier, M.P. (1988). Actinomycetes in agriculture and forestry. In Goodfellow, M., Williams, S.T. and Mordarski, M. (eds.) Actinomycetes in Biotechnology. Academic Press

Lezhava, A., Mizukami, T., Kajitani, T., Kameoka, D., Redenbach, M., Shinkawa, H., Nimi, O. and Kinashi, H. (1995). Physical map of the linear chromosome of *Streptomyces griseus*. J. Bacteriol. 177: 6492-6498

Lin, Y-S., Kieser, H.M., Hopwood, D.A. and Chen, C.W. (1993). The chromosomal DNA of *Streptomyces lividans* 66 is linear. Mol. Microbiol. 10: 923-933

MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H. and MacNeil, T. (1992).

Analyis of *S. avermitilis* genes required for avermectin biosynthesis utilizing novel integrating vector. Gene 111: 61-68

MacNeil, D.J., Occi, J.L., Gewain, K.M. and MacNeil, T. (1994). Correlation of the avermectin polyketide synthase genes to the avermectin structure. Implications for designing novel avermectins. Ann. N. Y. Acad. Sci. 721: 123-132

Madduri, K., Kennedy, J., Rivola, G., Inventi-Solari, A., Filippini, S., Zanuso, G., Colombo, A.L., Gewain, K.M., Occi, J.L., MacNeil, D.J. and Hutchinson, C.R. (1998). Production of the antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of *Streptomyces peucetius*. Nat. Biotechnol. 16: 69-74

Maeda, K. (1953). A crystalline toxic substance of *Streptomyces thioluteus* producing aureothricin. J. Antibiot. 4: 137-138

Malpartida, F. and Hopwood, D.A. (1984). Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host. Nature 309: 462-464

Matseliukh, A.B. (2001). Genetic transformation of *Streptomyces globisporus* strain 1912: restriction barrier and plasmid compatibility. Mikrobiol. Z. (Ukraine). 63: 15-22

Mayfield, C.I., Williams, S.T., Ruddick, S.M. and Hatfield, H.L. (1972). Studies on the ecology of actinomycetes in soil. Soil Biol. Biochem. 4: 79-91

McDaniel, R., Ebert-Khosla, S., Hopwood, D.A. and Khosla, C. (1995). Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. Nature 375: 549-554

McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M. and Ashley, G. (1999). Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. Proc. Natl. Acad. Sci. USA. 96: 1846-1851

Mochizuki, S., Hiratsu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K. and Kinashi, H. (2003). The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism. Mol. Microbiol. 48: 1501-1510

Moore, B.S. and Hertweck, C. (2002). Biosynthesis and attachment of novel bacterial

polyketide synthase starter units. Nat. Prod. Rep. 19: 70-99

Motamedi, H., Cai, S.J., Shafiee, A. and Elliston, K.O. (1997). Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK506. Eur. J. Biochem. 244: 74-80

Nonaka, Y. and Okamoto, M. (1991). Functional expression of the cDNAs encoding rat 11 beta-hydroxylase [cytochrome P450(11 beta)] and aldosterone synthase [cytochrome P450(11 beta, aldo)]. Eur. J. Biochem. 202: 897-902

Nowak-Thompson, B., Hammer, P.E., Hill, D.S., Stafford, J., Torkewitz, N., Gaffney, T.D., Lam, S.T., Molnar, I. and Ligon, J.M. (2003). 2,5-dialkylresorcinol biosynthesis in *Pseudomonas aurantiaca*: novel head-to-head condensation of two fatty acid-derived precursors. J. Bacteriol. 185: 860-869

Nowak-Thompson, B., Chaney, N., Wing, J.S., Gould, S.J. and Loper, J.E. (1999). Characterization of the pyoluteorin biosynthetic gene cluster of *Pseudomonas fluorescens* Pf-5. J. Bacteriol. 181: 2166-2174

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

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

Olano, C., Wilkinson, B., Moss, S.J., Brana, A.F., Mendez, C., Leadlay, P.F. and Salas, J.A. (2003). Evidence from engineered gene fusions for the repeated use of a module in a modular polyketide synthase. Chem. Commun. 21: 2780-2782

Otsuka, M., Ichinose, K., Fujii, I. and Ebizuka, Y. (2004). Cloning, sequencing, and functional analysis of an iterative type I polyketide synthase gene cluster for biosynthesis of the antitumor chlorinated polyenone neocarzilin in "*Streptomyces carzinostaticus*". Antimicrob. Agents. Chemother. 48: 3468-3476

Pacey, M.S., Dirlam, J.P., Geldart, R.W., Leadlay, P.F., McArthur, H.A., McCormick, E.L., Monday, R.A., O'Connell, T.N., Staunton, J. and Winchester, T.J. (1998). Novel erythromycins

from a recombinant *Saccharopolyspora erythraea* strain NRRL 2338 pIG1. I. Fermentation, isolation and biological activity. J. Antibiot (Tokyo). 51: 1029-1034

Pfeifer, B.A. and Khosla, C. (2001). Biosynthesis of polyketides in heterologous hosts. Microbiol. Mol. Biol. Rev. 65: 106-118

Piel, J., Hertweck, C., Shipley, P.R., Hunt, D.M., Newman, M.S. and Moore, B.S. (2000). Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate *'Streptomyces maritimus*': evidence for the derailment of an aromatic polyketide synthase. Chem. Biol. 7: 943-955

Piel, J. (2002). A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus beetles*. Proc. Natl. Acad. Sci. USA. 99: 14002-14007

Poupin, P., Ducrocq, V., Hallier-Soulier, S. and Truffaut, N. (1999). Cloning and characterization of the genes encoding a cytochrome P450 (PipA) involved in piperidine and pyrrolidine utilization and its regulatory protein (PipR) in *Mycobacterium smegmatis* mc2155. J. Bacteriol. 181: 3419-3426

Rafii, F. and Crawford, D.L. (1988). Transfer of conjugative plasmids and mobilization of a nonconjugative plasmid between *Streptomyces* strains on agar and in soil. Appl. Env. Microbiol. 54: 1334-1340

Ranganathan, A., Timoney, M., Bycroft, M., Cortes, J., Thomas, I.P., Wilkinson, B., Kellenberger, L., Hanefeld, U., Galloway, I.S., Staunton, J. and Leadlay, P.F. (1999). Knowledge-based design of bimodular and trimodular polyketide synthases based on domain and module swaps: a route to simple statin analogues. Chem. Biol. 6: 731-741

Revill, W.P., Bibb, M.J. and Hopwood, D.A. (1995). Purification of a malonyltransferase from *Streptomyces coelicolor* A3(2) and analysis of its genetic determinant. J. Bacteriol. 177: 3946-3952

Rojas, M.C, Hedden, P., Gaskin, P. and Tudzynski, B. (2001). The P450-1 gene of *Gibberella fujikuroi* encodes a multifunctional enzyme in gibberellin biosynthesis. Proc. Natl. Acad. Sci. USA. 98: 5838-5843

Ruan, X., Pereda, A., Stassi, D.L., Zeidner, D., Summers, R.G., Jackson, M., Shivakumar, A.,

Kakavas, S., Staver, M.J., Donadio, S. and Katz, L. (1997). Acyltransferase domain substitutions in erythromycin polyketide synthase yield novel erythromycin derivatives. J. Bacteriol. 179:6416-6425

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. CSH Press

Sambrook, J. and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual. Third Edition. CSH Press

Schroder, J. (1999). Probing plant polyketide biosynthesis. Nat. Struct. Biol. 6: 714-716

Schwartz, J.L., Tishler, M., Arison, B.H., Shafer, H.M. and Omura, S. (1976). Identification of mycolutein and pulvomycin as aureothin and labilomycin respectively. J. Antibiot (Tokyo). 29: 236-241

Schwecke, T., Aparicio, J.F., Molnar, I., Konig, A., Khaw, L.E., Haydock, S.F., Oliynyk, M., Caffrey, P., Cortes, J. and Lester, J.B. (1995). The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. Proc. Natl. Acad. Sci. USA. 92:7839-7843

Shoolingin-Jordan, P.M. and Campuzano, I.D.G. (1999). in Comprehensive Natural Products. Sankawa, U. (eds.) Elsevier. Oxford University press. Oxford

Stassi, D., Donadio, S., Staver, M.J. and Katz, L. (1993). Identification of a *Saccharopolyspora erythraea* gene required for the final hydroxylation step in erythromycin biosynthesis. J. Bacteriol. 175: 182-189

Staunton, J. and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. 18: 380-416

Staunton, J. (1998). Combinatorial biosynthesis of erythromycin and complex polyketide. Curr. Opin. Chem. Biol. 2: 339-345

Staunton, J., Caffrey, P., Aparicio, J.F., Roberts, G.A., Bethell, S.S. and Leadlay, P.F. (1996). Evidence for a double-helical structure for modular polyketide synthases. Nat. Struct. Biol. 3: 188-192

Steglich, W., Fugmann, B., Lang-Fungmann, S., Eds. (2002). Römpp Encyclopedia of Natural Products. Georg Thieme: Stuttgart

Sun, T., Zhao, Y., Nonaka, Y. and Okamoto, M. (1995). Cloning and expression of cytochrome P450(11 beta) of porcine adrenal cortex. J. Steroid. Biochem. Mol. Biol. 52: 227-232

Sun, Y., Zhou, X., Liu, J., Bao, K., Zhang, G., Tu, G., Kieser, T. and Deng, Z. (2002). *'Streptomyces nanchangensis'*, a producer of the insecticidal polyether antibiotic nanchangmycin and the antiparasitic macrolide meilingmycin, contains multiple polyketide gene clusters. Microbiology 148: 361-371

Suzuki, E. and Inoue, S. (1976). Synthesis and geometrical configuration of luteoreticulin, a toxic nitro-containing metabolite of *Streptomyces luteoreticuli Arai*. J. Chem. Soc. [Perkin 1]. (4): 404-407

Swan, D.G., Rodriguez, A.M., Vilches, C., Mendez, C. and Salas, J.A. (1994). Characterisation of a *Streptomyces* antibioticus gene encoding a type I polyketide synthase which has an unusual coding sequence. Mol. Gen. Genet. 242: 358-362

Takano, E. and Bibb, M.J. (1994). The stringent response ppGpp and antibiotic production in *Streptomyces coelicolor* A3(2). Actinomycetol. 8: 1-16

Taniguchi, M., Watanabe, M., Nagai, K., Suzumura, K., Suzuki, K. and Tanaka, A. (2000). Gamma-pyrone compounds with selective and potent anti-*Helicobacter pylori* activity. J. Antibiot (Tokyo). 53: 844-847

Thomas, I., Martin, C.J., Wilkinson, C.J., Staunton, J. and Leadlay, P.F. (2002). Skipping in a hybrid polyketide synthase. Evidence for ACP-to-ACP chain transfer. Chem. Biol. 9: 781-787

Thompson, C.J., Kieser, T., Ward, J.M. and Hopwood, D.A. (1982). Physical analysis of antibiotic-resistance genes from *Streptomyces* and their use in vector construction. Gene 20: 51-62

Thompson, C.J., Ward, J.M. and Hopwood, D.A. (1980). DNA cloning in *Streptomyces*: resistance genes from antibiotic-producing species. Nature 286: 525-527

Walsh, C.T., O'Connor, S.E. and Schneider, T.L. (2003). Polyketide-nonribosomal peptide

epothilone antitumor agents: the EpoA, B, C subunits. J. Ind. Microbiol. Biotechnol. 30: 448-455

Walsh, C.T. (2002). Combinatorial biosynthesis of antibiotics: challenges and opportunities. Chembiochem. 3: 125-134

Waksman, S.A. and Herirei, A.T. (1943). The nomen culture and classification of the *actinomyces*. J. Bacteriol. 46:370-341

Wakil, S.J. (1989). Fatty acid synthase, a proficient multifunctional enzyme. Biochemistry 28: 4523-4530

Weber, J.M., Leung, J.O., Swanson, S.J., Idler, K.B. and McAlpine, J.B. (1991). An erythromycin derivative produced by targeted gene disruption in *Saccharopolyspora erythraea*. Science 252: 114-117

Wellington, E.M.H., Saunders, V.A., Cresswell, N. and Wipat, A. (1988). Plasmid transfer between streptomycetes in soil. In Okami, Y., Beppu, T. and Ogawara, H. (eds.) Biology of Actinomycetes ’88. Japan Scientific Societies Press

Wiesmann, K.E., Cortes, J., Brown, M.J., Cutter, A.L., Staunton, J. and Leadlay, P.F. (1995). Polyketide synthesis in vitro on a modular polyketide synthase. Chem. Biol. 2: 583-589

Wilkinson, B., Foster, G., Rudd, B.A., Taylor, N.L., Blackaby, A.P., Sidebottom, P.J., Cooper, D.J., Dawson, M.J., Buss, A.D., Gaisser, S., Bohm, I.U., Rowe, C.J., Cortes, J., Leadlay, P.F. and Staunton, J. (2000). Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. Chem. Biol. 7: 111-117

Willey, J., Schwedock, J. and Losick, R. (1993). Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by *Streptomyces coelicolor*. Genes & Dev. 7: 895-903

Wright, F. and Bibb, M.J. (1992). Codon usage in the G+C-rich *Streptomyces* genome. Gene 113: 55-65

Xiang, L., Kalaitzis, J.A., Nilsen, G., Chen, L. and Moore, B.S. (2002). Mutational analysis of

the enterocin Favorskii biosynthetic rearrangement. Org. Lett. 4: 957-960

Xue, Q., Ashley, G., Hutchinson, C.R. and Santi, D.V. (1999). A multiplasmid approach to preparing large libraries of polyketides. Proc. Natl. Acad. Sci. USA. 96: 11740-11745

Xue, Y. and Sherman, D.H. (2000). Alternative modular polyketide synthase expression controls macrolactone structure. Nature 403: 571-575

Yamazaki, M., Maebayashi, Y., Kato, F. and Koyama, Y. (1975). Application of ^{13}C-NMR to biosynthetic investigations. II Biosynthesis of aureothin and related nitro-containing metabolites of *Streptomyces luteoreticuli*. Chem. Pharm. Bull. 23: 569-574

Yu, T.W., Bibb, M.J., Revill, W.P. and Hopwood, D.A. (1994). Cloning, sequencing, and analysis of the griseusin polyketide synthase gene cluster from *Streptomyces griseus*. J. Bacteriol. 176:2627-2634

Yu, T.W., Shen, Y., Doi-Katayama, Y., Tang, L., Park, C., Moore, B.S., Richard Hutchinson, C. and Floss, H.G. (1999). Direct evidence that the rifamycin polyketide synthase assembles polyketide chains processively. Proc. Natl. Acad. Sci. USA. 96: 9051-9056

Yuan, L. (1983). Studies on protoplast fusion of the mydimycin producer. Antibiotics (China). 8: 380-387

Zhao, L., Ahlert, J., Xue, Y., Thorson, J.S., Sherman, D.H. and Liu, H-W. (1999). Engineering a methymycin/pikromycin-calicheamicin hybrid: construction of two new macrolides carrying a designed sugar moiety. J. Am. Chem. Soc. 121: 9881-9882

Zhou, X., Deng, Z., Firmin, J.L., Hopwood, D.A., and Kieser, T. (1988). Site-specific degradation of *Streptomyces lividans* DNA during electrophoresis in buffers contaminated with ferrous iron. Nucleic. Acids. Res. 16: 4341-4352

Ziel, M., He, J., Dahse, H.-M. and Hertweck, C. (2004). Mutasynthesis of aureonitrile, an aureothin derivative with significantly improved cytostatic effect. Angew. Chem. In press

Zotchev, S., Haugan, K., Sekurova, O., Sletta, H., Ellingsen, T.E. and Valla, S. (2000). Identification of a gene cluster for antibacterial polyketide-derived antibiotic biosynthesis in the nystatin producer *Streptomyces noursei* ATCC 11455. Microbiology. 146: 611-619

Acknowledgements:

I am very grateful to my supervisor, Dr. Christian Hertweck. When I had any problem in my research work, he always could give me some useful advices and encourage me to overcome difficulties. His suggestions and instruction proved to be invaluable in these three years of my study.

I would like to thank all my colleagues: Markus Müller, Kathrin Jakobi, Sabine Loos, Nelly Traitcheva, Zhongli Xu, Michelle Palzer, Martina Ziehl, Julia Schüman, Robert Winkler, Laila Pamela Partida-Martinez. Without the enthusiastic help of my colleagues, I would not have any academic achievements. I am very happy to work in this talented and kind group.

Especially, I want to say thanks to Markus Müller for his great help in identifying the chemical structures of aureothin derivatives and for supplying purified substances for feeding experiments.

I am thankful to Kathrin Jakobi for her excellent help of translating the summary of my thesis into German.

Thanks Nelly Traitcheva for proofreading and Kerstin Herold for her patient and friendly help explaining the requirements for a Ph. D. thesis.

Thanks to all of my friends, you gave me selfless assistance and support during my stay in Jena.

Finally, I am full of heartfelt gratitude to my husband and parents for their understanding and support in my life.

Curriculum vitae

He Jing

Nationality: People's Republic of China **Gender:** Female	**Date of birth:** Feb.10, 1977 **Place of birth:** Hubei Province, China
Hans-Knoell-Institute for Natural Products Research Department of Bioorganic Synthesis Beutenbergstr. 11a, D-07745 Jena Germany E-mail: hejingjj@yahoo.com	

Education

1. Sept. 1998- Jun. 2001: Huazhong Agricultural University, Wuhan, China
   Master of Molecular Biology and Biochemistry
2. Sept. 1995- Jun. 1998: Huazhong Agricultural University, Wuhan, China
   Bachelor of Bio-technology
3. Sept. 1992- Jun. 1995: No. 3 High School of Wuhan, Wuhan, China
4. Sept. 1989- Jun. 1992: No. 3 Middle School of Wuhan, Wuhan, China
5. Sept. 1983- Jun. 1989: Yaoluti Elementary School, Wuhan, China

Professional data

Nov. 2001- the present: Doctoral student at the Department of Bioorganic Synthesis, HKI, Jena

Oct. 2000- Jun. 2001: Visiting scientist at the Bio-X Life Science Research Center, Shanghai Jiaotong University, Shanghai, China

Sept. 1998- Oct. 2000: Master, Laboratory of Molecular Biology of *Streptomyces*, Huazhong Agricultural University, Wuhan, China

Sept. 1995- Jun. 1998: Bachelor, College of Life Sciences and Technology, Huazhong Agricultural University, Wuhan, China

Jing He

December 3, 2004, Jena

Publications

Original Papers

- Shi Chen, Xi Huang, Xiufen Zhou, Linquan Bai, Jing He, Ki Jun Jeong, Sang Yup Lee, and Zixin Deng: "Organizational and Mutational Analysis of a Complete FR-008/ Candicidin Gene Cluster Encoding a Structurally Related Polyene Complex", *Chem. Biol.* **2003**, *10*, 1065-1076.
- Jing He and Christian Hertweck: "Iteration as Programmed Event during Polyketide Assembly; Molecular Analysis of the Aureothin Biosynthesis Gene Cluster", *Chem. Biol.* **2003**, *10*, 1225-1232. *
- Jing He and Christian Hertweck: "Biosynthetic Origin of the Rare Nitro Aryl Moiety of the Polyketide Antibiotic Aureothin: Discovery of an Unprecedented *N*-Oxygenase", *J. Am. Chem. Soc.* **2004**, *126*, 3694-3695. *
- Jing He, Markus Müller and Christian Hertweck: "Formation of the Aureothin Tetrahydrofuran Ring by a Bifunctional Cytochrome P450 Monooxygenase", *J. Am. Chem. Soc,* in press. *
- Martina Ziehl, Jing He, H.-M. Dahse and Christian Hertweck: "Mutasynthesis of Aureonitril, an Aureothin Derivative with Significant Cytostatic Effect", *Angewandte Chemie*, in press. *
- Jing He and Christian Hertweck: "Functional Analysis of the Aureothin Iterative Type I Polyketide Synthase", *ChemBioChem*, in press. *

* Publications resulting from the Ph. D. work/thesis

Patents

- Christian Hertweck and Jing He: "Amino-Oxidase" aus *Streptomyces thioluteus*.
  Reference mark: 103 35 447.6, 31.07.2003
- Christian Hertweck, Martina Ziehl, Jing He and H.-M. Dahse: "HKI 10429005, neue zytostatische Wirkstoffe, Verfahren zu ihrer Herstellung und ihre Verwendung"
  Reference mark: 10 2004 040 014.8,16.08.2004

Poster

- Jing He and Christian Hertweck: "Cloning, Sequencing and Heterologous Expression of the Aureothin Biosynthesis Gene Cluster from *Streptomyces thioluteus* HKI-227", VAAM Meeting Sekundärmetabolismus der Actinomyceten, Freiburg, 29.09-01.10.2002
- Jing He，Nelly Traitcheva and Christian Hertweck: "Die 'stotternde' Aureothin-Synthase als Modellsystem für die Evolution metabolischer Diversität in Streptomyceten", 1st SPP 1152 Colloquium "Evolution of Metabolic Diversity", IPB Halle, 30-31.01.2003
- Christian Hertweck and Jing He: "Die molekulare Basis der Aureothin-Biosynthase — Entdeckung einer stotternden Polyketidsynthase in *Streptomyces thioluteus*", 15. Irseer Naturstofftage 26-28.02.2003

- Jing He，Nelly Traitcheva, Markus Müller, Robert Winkler and Christian Hertweck: “Non-Linear Polyketide Biosynthesis in the Aureothin and Neoaureothin Pathways”, Frontiers in Bioorganic and Medicinal Chemistry, (H.G. Floss Symposium), University of Washington, Seattle, August 2004

Oral Presentation

- Jing He, Markus Müller, Nelly Traitcheva, Robert Winkler, Martina Ziehl and Christian Hertweck: “Functional Analysis of the Aureothin Biosynthesis Gene Cluster”, VAAM Meeting International Meeting on the Biology of Bacteria Producing Natural Compounds, Jena, 26.09-28.09.2004

Selbständigkeitserklärung

Ich erkläre hiermit, daß ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel und Literatur angefertigt habe.

Jena, den 3. Dezember 2004

Jing He