Heterologous expression of a lichen polyketide synthase in filamentous fungi

A thesis submitted for the degree of Master of Science

Andrey N. Gagunashvili

Department of Biology
Faculty of Science
University of Iceland

Instructor:
Ólafur S. Andrésson, Ph.D.

External examinator:
Ólafur H. Friðjónsson, Ph.D.

Masters committee:
Ólafur S. Andrésson, Ph.D.
Zophonías O. Jónsson, Ph.D.

Reykjavik, Iceland
October 2007

Declaration

I hereby state that this thesis is written by me and it has not been submitted – neither totally or in part – for a higher degree.

Yfirlýsing

Hér með lýsi ég því yfir að ritgerð þessi er samin af mér og að hún hefur hvorki að hluta né í heild verið lögð fram áður til hærri prófgráðu.

"_____" _____ 2007 _____ Andrey N. Gagunashvili

Abstract

Polyketides are a group of secondary metabolites produced by a wide range of living organisms – bacteria, fungi, plants and animals (O'Hagan, 1991, 1992, 1993, 1995; Rawlings, 1997). These compounds exhibit remarkable structural diversity and possess an extremely broad spectrum of biological activity. Polyketides are of great commercial interest for drug discovery since many of these compounds have desirable pharmaceutical properties and they are a source of novel antibiotics, anti-tumor and anti-cancer agents, as well as cholesterol-lowering drugs.

The biosynthesis of polyketides is catalyzed by large multifunctional enzymes called polyketide synthases (PKSs) that assemble core polyketide molecules from simple starter carboxylic acid precursors and several malonyl-CoA units in a manner similar to fatty acid synthesis.

Although polyketide synthesis is widespread in filamentous fungi and lichen mycobionts (Huneck, Yoshimura, 1996), relatively few PKS genes have been isolated from filamentous fungi and no PKS gene from lichens can be found in the GenBank database except the one from the lichen *Solorina crocea* recently submitted (Gagunashvili, Davíðsson, Andrésson, *unpublished*).

Although lichens produce various pharmaceutically relevant polyketides and have been appreciated in traditional medicines (Huneck, 1999; Müller, 2001), their value has largely been ignored by the modern pharmaceutical industry because slow growth and difficulties in establishing pure cultures do not allow their routine use in most conventional screening processes (Miao *et al.*, 2001). Recently, molecular genetic techniques using PCR, genomic library construction and heterologous expression in surrogate hosts have provided an alternative approach to begin accessing, exploring and harnessing the diversity of polyketide biosynthetic pathways in lichens.

The aim of this work was the surrogate expression of a PKS gene from the lichen *Solorina crocea* in the filamentous fungus *Aspergillus oryzae*, analysis of the product(s) of its expression and determination of the gene structure: exon–intron boundaries, 5′- and 3′-prime ends, the native promoter region and the domain organization of the deduced protein sequence.

Útdráttur

Margvíslegar lífverur – bakteríur, sveppir, plöntur og dýr – framleiða fjölketíð, gjarnan flokkuð sem hliðarafurðir efnaskipta (O'Hagan, 1991, 1992, 1993, 1995; Rawlings, 1997). Þar eð fjölketíð eru fjölbreytt að byggingu og lífvirkni eru þau jafnframt mikilvæg fyrir lyfjaþróun, t.d. sem sýklalyf, krabbameinslyf og til að draga úr kólesteroli í blóði.

Myndun fjölketíða er hvötuð af stórum próteinum með fjölda virknihneppa, svonefndum fjölketíð synþasa ensímum (PKS) sem setja saman fjölketíðgrindur þar sem byrjað er á einfaldri karboxýlsýru (ediksýru, própýlsýru) og malonýl-CoA einingar notaðar til keðjulengingar á svipaðan hátt og við nýmyndun fitusýra.

Enda þótt nýmyndun fjölketíða sé algeng meðal sveppa og fléttusveppa (Huneck, Yoshimura, 1996) þá hafa fá PKS gen verið einangruð úr sveppum og einungis eitt PKS gen úr fléttu er í GenBank gagnagrunninum, PKS gen úr glóðarskóf (*Solorina crocea*) (Gagunashvili, Davíðsson, Andrésson, *óbirt*).

Enda þótt fléttur framleiði margvísleg lífvirk fjölketíðefni sem nýtt hafa verið í lækningaskyni (Huneck, 1999; Müller, 2001) þá hefur lyfjaiðnaðurinn lítt sinnt þeim sökum hversu erfitt er að rækta fléttur og fléttusveppi. Framfarir í erfðatækni svo sem genamögnun, genaflutningur og tjáning í hentugum hýsilfrumum hafa opnað nýja möguleika til að nálgast, skilgreina og nýta margbreytileika nýmyndunar fjölketíða í fléttum.

Markmið þessa verkefnis var að flytja PKS gen úr glóðarskóf í þráðsveppinn *Aspergillus oryzae*, sýna fram á tjáningu gensins og skilgreina byggingu gensins: Innraðir, $5'$ og $3'$ enda RNA umrits, stýrisvæði. Jafnframt voru skoðuð og skilgreind starfræn próteinsvæði í afleiddri amínósýruröð, einkum með samanburði við skyld prótein.

Acknowledgments

I would like to gratefully thank

My supervisor, professor Ólafur S. Andrésson, for accepting me and believing in me, for being flexible and for giving me plenty of independence; Dr. Olga Marfenina, for the strong basis in mycology; Snorri P. Davíðsson, for helping me in my first steps on this project; Dr. Russel J. Cox and Dr. Colin M. Lazarus, for the gift of the plasmids and the fungal strain; Laura M. Halo, for the help with the protocol; Prof. Einar Árnason, for maintenance the analyzer in a perfect state and for Miracloth; Dr. Zophonías O. Jónsson, for general help, advice and for critical reading some parts of my manuscript; Sigríður H. Þorbjarnardóttir, for knowing almost everything in the lab; Prof. Guðmundur G. Haraldsson, for the help with glassware; Katrín Halldórsdóttir, for advices on sequencing and for crystal violet; Hörður Guðmundsson, for introduction to some techniques; my family and friends, for the help and support.

Table of Contents

List of Figures

List of Tables

I

Introduction

1. Fungal polyketide biosynthesis

Polyketides are probably the most abundant group of fungal secondary metabolites exhibiting a high degree of structural diversity and various biological activities (O'Hagan, 1991). Many of these compounds and their derivatives are used as pharmaceuticals, for example the cholesterol-lowering agent lovastatin. On the other hand, many fungal polyketides are notorious for causing severe damage in agriculture, food industry, and human health, being toxins and/or pathogenic factors. Important examples are represented by T-toxin, the fumonisins and the aflatoxins (Bennett, Klich, 2003; Chu, 1991; Sweeney, Dobson, 1999). Furthermore, many polyphenolic polyketide compounds, such as tetrahydroxynaphthalene (THN), YWA1 and norsolorinic acid, are components of fungal pigments.

Common to all fungal polyketides is their biosynthetic origin. They are produced by repetitive Claisen condensations of an acyl-coenzyme A (CoA) starter unit with malonyl-CoA extender units in a manner reminiscent of fatty acid biosynthesis (Figure I.1). One of the major differences between these metabolic pathways is that polyketides can differ in the degree of β-keto processing. Some are not or only partially reduced, resulting in the formation of mono- or polycyclic aromatic compounds. Alternatively, they may be partially or highly reduced, which usually gives rise to linear or macrocyclic, non-aromatic carbon backbones (Rawlings, 1997; Staunton, Weissman, 2001).

The biosynthesis of fungal polyketides is catalyzed by multidomain enzymes called type I polyketide synthases (PKSs) in analogy to multidomain fatty acid synthases (FASs). Fungal polyketide synthases consist of a minimal set of ketosynthase (KS), acyl transferase (AT), and acyl carrier protein (ACP) domains. An ACP domain carries the malonyl-CoA extender units and transiently hold the growing acyl chain. The *apo*-ACP requires post-translational modification to its *holo*-form through the addition of phosphopantetheine

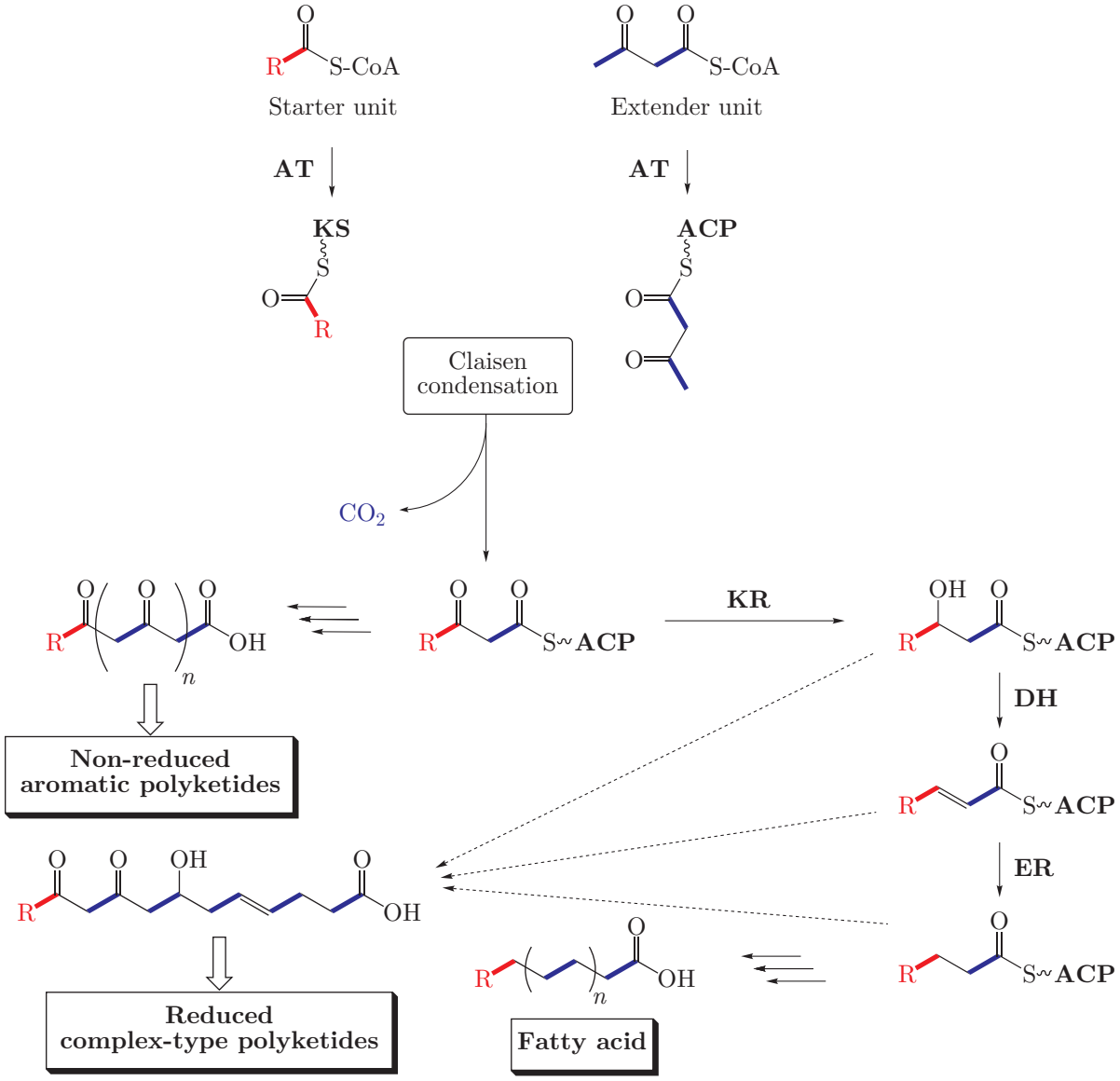


Figure I.1. The basic pathway of fatty acid and polyketide biosynthesis. KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein.

derived from CoA to a conserved serine (Figure I.8) (Kealey *et al.*, 1998). This reaction is carried out by a phosphopantetheine transferase (PPT, also known as *holo*-ACP synthase). An AT domain is responsible for transfer of acyl groups from CoA onto the KS and ACP components. In addition, optional β-keto processing reactions may be catalyzed by ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) domains. Further optional accessory domains are represented by thioesterase (TE)/Claisen cyclase (CYC) (Fujii *et al.*, 2001) and *C*-methyltransferase (*C*MeT) activities (Cox, 2007) (Figure I.2). In contrast to post-PKS tailoring *O*- and *N*-methylation reactions, which are catalyzed

by discrete enzymes after polyketide assembly, *C*-methylation of the polyketide carbon backbone occurs during chain formation by means of the intrinsic fungal *C*MeT domains (Nicholson *et al.*, 2001). It should be noted that in fungal polyketide biosynthesis no methylmalonyl extender units are utilized, as opposed to numerous examples in bacterial polyketide biosynthesis involving modular type I PKS. Fungal iterative type I PKSs can use each active site in an *iterative* way during chain extension, and determine the degree of reduction and *C*-methylation within each elongation round (Fujii *et al.*, 1998; Rawlings, 1999). However, to date it is not known how a single set of catalytic domains of a certain PKS determines chain length, degree of reduction, and timing of *C*-methylation at a particular step in the pathway. It should be mentioned that although most fungal PKSs act in an iterative manner, some only catalyze a single round of elongation, for example the *Aspegillus terreus* LDKS catalyzing the biosynthesis of lovastatin diketide. Thus it is not strictly correct to call them *iterative* type I PKSs (Müller, 2004).

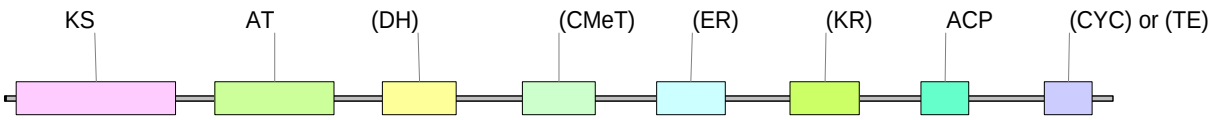


Figure I.2. Generalized domain structure of fungal type I PKS. KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein; CYC, Claisen cyclase; TE, thioesterase. The minimal set of domains is KS→AT→ACP. Optional domains are shown in brackets.

In addition to the chain length, the degree of β-keto processing and cyclization, the large structural diversity of polyketides derives from modification of the polyketide carbon skeleton by alkylation, acylation, and oxygenation, by so-called post-PKS or *tailoring* reactions catalyzed by corresponding enzymes. From analyses of the few PKS gene loci that have been investigated to date it is evident that genes necessary for fungal polyketide biosynthesis are clustered, i.e. PKS genes, genes encoding enzymes involved in tailoring reactions, as well as genes required for regulation and resistance, as exemplified for the lovastatin (Kennedy *et al.*, 1999), compactin (Abe *et al.*, 2002) and aflatoxin (Yu *et al.*, 1995; Bhatnagar et al., 2003) biosynthesis gene clusters (Figures I.28, I.42). Knowledge on the biosynthesis genes and their regulation could provide important information on the ecological role of fungal polyketide metabolites. Furthermore, investigations of fungal PKS genes and their expression could provide a foundation for improving metabolite production and for engineering the biosynthetic machinery with the aim of synthesizing new compounds with desirable properties (Burkart, 2003; Reeves, 2003).

2. Fungal PKSs

Relatively few fungal PKS genes have been isolated and characterized to date, compared to the large number of bacterial PKS genes, mainly of actinomycete origin. However, fungi equal and sometimes exceed bacteria in their polyketide biosynthetic potential. The study of polyketide biosynthesis in fungi has been hampered for a long time by the lack of molecular tools available to fungal researchers. Of late, however, various methods have been developed and proven successful for fungi.

At present, fungal PKSs are almost exclusively represented by iterative type I enzymes. Modular type I and type II systems have never been observed in fungi to date. Genes encoding type III PKSs have been discovered recently in *Aspergillus oryzae*, proving that their distribution is not limited to bacteria and plants (Seshime *et al.,* 2005). Subsequent comparative analysis of fungal genomes allowed identification single putative type III PKSs in *Neurospora crassa* and *Fusarium graminearum*, two each in *Magnaporthe grisea* and *Podospora anserina*, and three in *Phenarocheate chrysosporium*.

The nomenclature of fungal PKSs is rather confused and does not have any conventional rules yet, with many fungal PKS genes and their corresponding proteins being named something like *pks*1 and PKS1, respectively.

According to their domain organization fungal PKSs are grouped into non-reducing, partially reducing, and highly-reducing PKSs. The evolutionary relationships between different types of fungal PKSs have been revealed by phylogenetic studies based on amino acid sequences of their KS domains (Bingle *et al.*, 1999; Kroken *et al.*, 2003; Nicholson *et al.*, 2001). Kroken and coworkers showed that fungal PKSs cluster according to their architecture and the degree of reduction of their products into reducing PKSs with additional reductive domains (KR, ER, DH) and non-reducing PKSs lacking such domains. Reducing PKSs are further divided into four subclades. Non-reducing PKSs are divided into three subclades with some PKSs being basal to clades I and II. This type is also characterized by the gain of Claisen-type cyclase and additional ACP domains. The third main fungal PKS type comprising partially reducing 6-methylsalicylic acid synthases nested as a separate subclade within the clade of bacterial type I PKSs. Based on obtained KS-domain genealogy it was proposed that the ancestral domain structure of type I PKSs was KS→AT→DH→*C*MeT→ER→KR→ACP and all diversity of fungal type I PKSs can be explained by the loss of some ancestral domains and/or gain of novel domains.

The PKSs described in this section are divided into three groups: tetraketide synthases, producing monocyclic aromatic compounds, non-reducing PKSs, producing polyketides with multi-ring aromatic frameworks, and highly-reducing PKSs, giving rise to linear or macrocyclic, non-aromatic carbon backbones.

2.1. Tetraketide synthases

6-Methylsalicylic acid (6-MSA) and orsellinic acid represent a group of single aromatic ring tetraketides and are direct precursors of various fungal metabolites (Figure I.3) (Fujii, 1999; Turner, 1971; Turner, Aldridge, 1983). Their biosynthesis is catalyzed by 6-methylsalicylic acid synthases (MSAS) and orsellinic acid synthases (OAS), respectively, the smallest iterative type I PKSs known with less than 2 000 amino acid residues (Figures I.4, I.9).

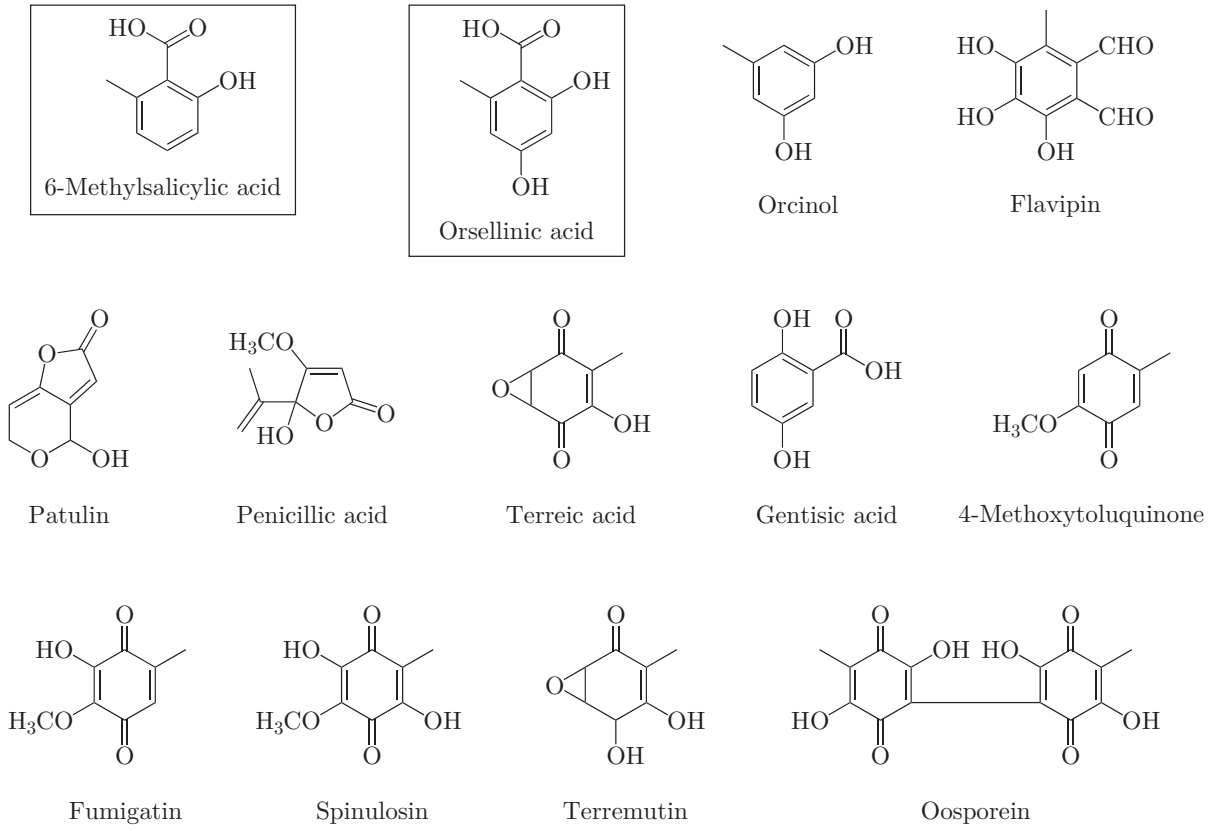


Figure I.3. Structures of 6-methylsalicylic acid, orsellinic acid and selected fungal metabolites derived from them.

2.1.1. 6-Methylsalicylic acid synthase

The detection of MSAS in the cell-free extract of the filamentous fungus *Penicillium griseofulvum*[1] was reported by Lynen and Tada in 1961 and was the first example of polyketide synthase activity discovered in any natural source (Lynen, Tada, 1961). These observations were later confirmed by Light for the same fungal system (Light, 1967).

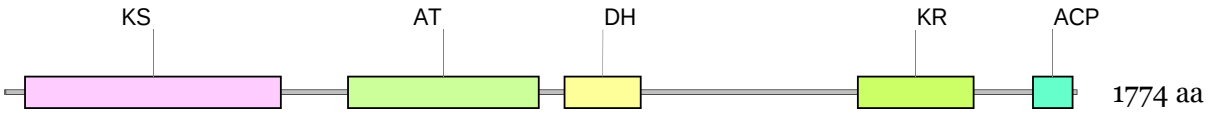
[1]Previously *Penicillium patulum*.

The first purification of MSAS from *P. griseofulvum* was achieved by Lynen and coworkers in 1970 (Dimroth *et al.*, 1970). However, repetition of the enzyme isolation was hampered by its inherent instability. The next successful purification of MSAS was reported only in 1988 from *Penicillium urticae* by improving enzyme stabilization *in vitro* (Lam *et al.,* 1988). Purification and further characterization of MSAS from *P. griseofulvum* was continued by other groups in 1990 (Beck *et al.,* 1990) and 1992 (Spencer, Jordan, 1992). MSAS was found to be a homotetramer with molecular weight of 750 kDa as determined by gel filtration, consisting of 180 kDa subunits as judged by SDS-PAGE analysis.

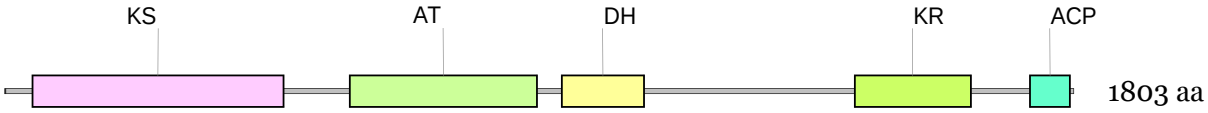
The gene encoding MSAS from *P. griseofulvum* was the first fungal PKS gene cloned by Beck and coworkers in 1990. The MSAS gene was accessed by screening an *E. coli* expression library of *P. griseofulvum* cDNA with antibodies raised against the purified MSAS protein (Beck *et al.*, 1990). The genomic DNA was obtained using the cloned cDNA as a hybridization probe for screening a genomic DNA library. An isolated 7.2-kb genomic clone carried a 5 322-bp coding sequence intervened by a single intron of 69 bp near the 5′ end. The open reading frame encoded a 190 731 Da protein of 1 774 amino acids.

A similarity-based approach revealed a relatively low degree of homology with fatty acid synthases (FAS) from yeast and *Penicillium* and significantly higher sequence homology was found between the MSAS and FAS of rat and other vertebrates, especially at their KS, AT, KR and ACP domains located in a similar linear order (Figure I.4). A DH domain was not found since its location had not been determined for any FASs at the time of the study (Beck *et al.*, 1990). The presence of a DH domain in MSAS was established by amino acid sequence alignment with the 6-deoxyerythronolide-B synthase of the erythromycin A biosynthetic gene cluster from *Saccharopolyspora erythraea* (Bevitt *et al.,* 1992). MSAS lacked an ER domain since the biosynthesis of 6-MSA does not require any enoyl reduction. Also the deduced amino acid sequence did not contain a TE domain, which is considered to have a role in control of chain length in fatty acid biosynthesis (Singh *et al.*, 1984).

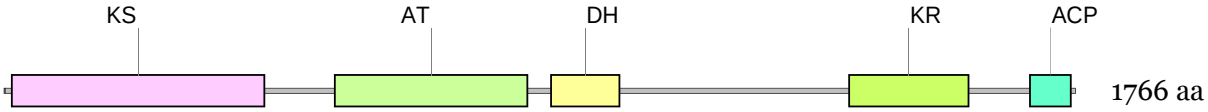
During 6-MSA biosynthesis, the enzyme catalyzes three cycles of chain extension from an acetyl starter and malonyl extender units. The whole reaction is controlled by a yet undetermined cycle-specific mechanism. The polyketide intermediate is attached as a thioester to the cysteine residue of the KS domain of the 4′-phosphopantetheine prosthetic group of the ACP domain of MSAS at all stages of 6-MSAS biosynthesis. In the first cycle the KR and DH domains are not active, but the second condensation is followed by reduction and dehydration of the newly-formed keto group. In the absence of NADPH,



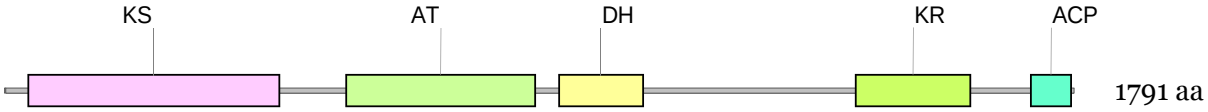
Penicillium griseofulvum MSAS



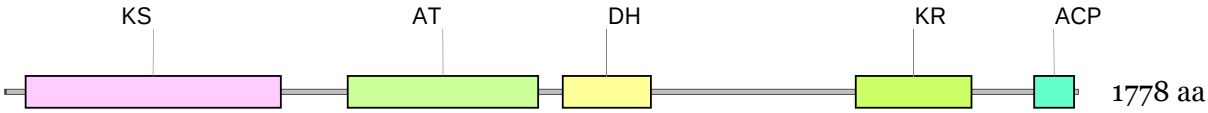
Aspergillus terreus ATX



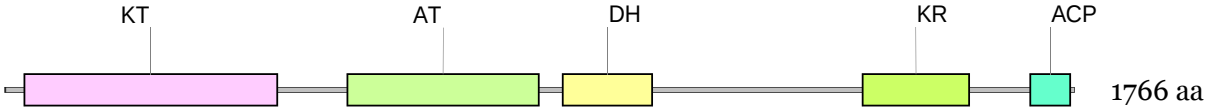
Aspergillus parasiticus PKSL2



Glarea lozoyensis PKS2



Byssochlamys nivea MSAS



Aspergillus ochraceus MSAS

Figure I.4. Domain organization of fungal 6-methylsalicylic acid synthases: MSAS from *Penicillium griseofulvum* (accession numbers CAA39295, P22367; Beck *et al.,* 1990), ATX from *Aspergillus terreus* (accession number BAA20102; Fujii *et al.*, 1996), PKSL2 from *Aspergillus parasiticus* (accession number AAC23536; Feng, Leonard, 1998), PKS2 from *Glarea lozoyensis* (accession number AAX35547; Lu *et al.,* 2005), MSAS from *Byssochlamys nivea* (accession numbers AAK48943; Puel *et al.,* 2007), MSAS from *Aspergillus ochraceus* (accession number AAS98200; Dao *et al., unpublished*). KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; KR, β-ketoreductase; ACP, acyl carrier protein.

the third condensation cycle does not occur resulting in formation of triacetic lactone[2]. After the third cycle the partially reduced polyketide chain undergoes aldol cyclization, dehydration and enolization to give 6-MSA. Aldol-type cyclase and/or aromatase domains presumably necessary for cyclization are absent in MSAS. It is therefore believed that aldol cyclization and aromatization occur spontaneously and the MSAS enzyme might just provide a cavity for the stabilization and appropriate folding of a tetraketomethylene intermediate (Fujii *et al.*, 2004). The absence of a TE domain in MSAS suggests that the final release of the completed tetraketide from the synthase does not occur via hydrolysis. An alternative mechanism involving formation of a ketene intermediate has been proposed, although it has not been verified experimentally (Spencer, Jordan, 1992).

To probe the stereochemistry of events during 6-MSA biosynthesis, (2*R*) and (2*S*) [1-^{13}C,2-^{2}H]-malonyl-CoAs, synthesized *in situ*, were incubated in separate experiments with acetyl-CoA, NADPH and MSAS from *P. griseofulvum*. Mass spectral analysis of 6-MSA formed showed that the hydrogen atoms at C-3 and C-5 were derived from opposite absolute configurations in malonyl-CoA. The absolute stereochemistry of the hydrogen atom at C-3 has been established by incubating the chiral malonyl-CoAs, in separate experiments, with acetoacetyl-CoA as a starter unit to bypass the first cycle of condensation. Mass spectral analysis indicated that the C-3 hydrogen originated from the 2-*pro-R* site of malonyl-CoA. Judging from this result it was stated that the C-5 hydrogen had its origin in the 2-*pro-S* site of malonyl-CoA (Jordan, Spencer, 1990, 1991; Spencer, Jordan, 1990, 1992). The stereochemistry of the NADPH-dependent reduction occurring after the second cycle of condensation was investigated using [4*R*-^{3}H]- and [4*S*-^{3}H]NADPH, and only the *pro-S*-H of NADPH was incorporated in 6-MSA (Schorr *et al.*, 1994). The high degree of stereochemical control clearly indicates that reduction and enolization processes occur under enzymatic control (O'Hagan, 1991).

The function of 6-MSA itself in organisms producing it still remains unclear. Sometimes it is considered to be a phytotoxin and it was found to inhibit the growth of various bacteria and fungi *in vitro*, including those that produce it (Bu'Lock *et al.*, 1966; Venkatasubbaiah, Chilton, 1992; Venkatasubbaiah *et al.*, 1995). 6-MSA is also a precursor of various fungal metabolites, e.g. the mycotoxin patulin which is formed from 6-MSA via oxidative ring opening reactions (Figure I.7) (Bu'Lock, Ryan, 1958; Tanenbaum, Bassett, 1959). Patulin was isolated under different names (clavacin, clavatin, claviformin, expansin, mycoin C, penicidin, terinin and others) from various fungi (*Peni-*

[2]The production of triacetic lactone in the absence of NADPH is not unique to MSAS. It has been detected from several PKSs incubated without NADPH (Bentley, Zwitkowits, 1967), including FASs (Nixon *et al.*, 1968; Yalpani *et al.*, 1969) and appears to a very stable by-product that forms on PKSs when KR activity is absent (O'Hagan, 1991).

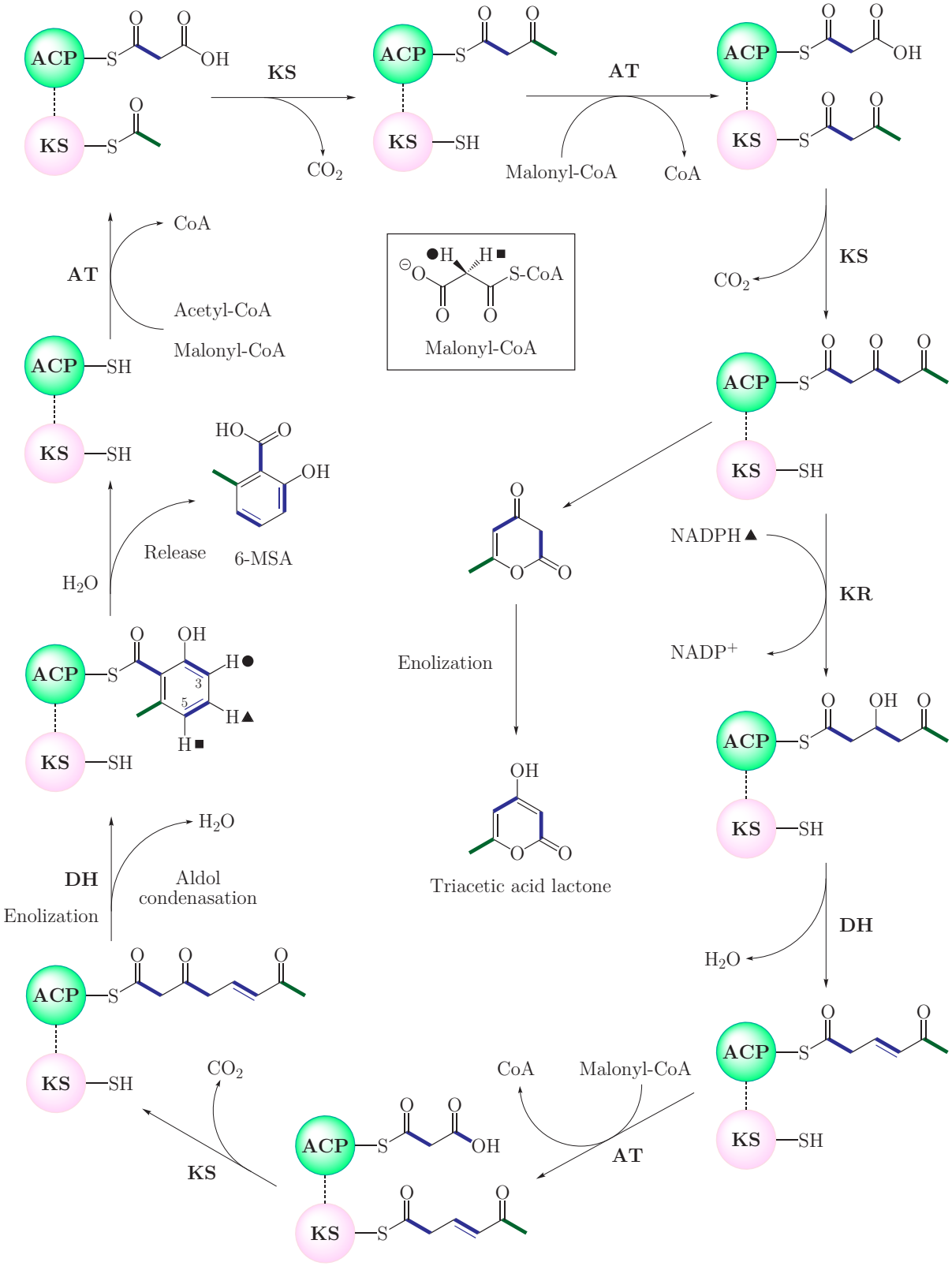


Figure I.5. Biosynthesis of 6-methylsalicylic acid. KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; KR, β-ketoreductase; ACP, acyl carrier protein; –SH and –S– represent a thiol group and a thioester linkage, respectively.

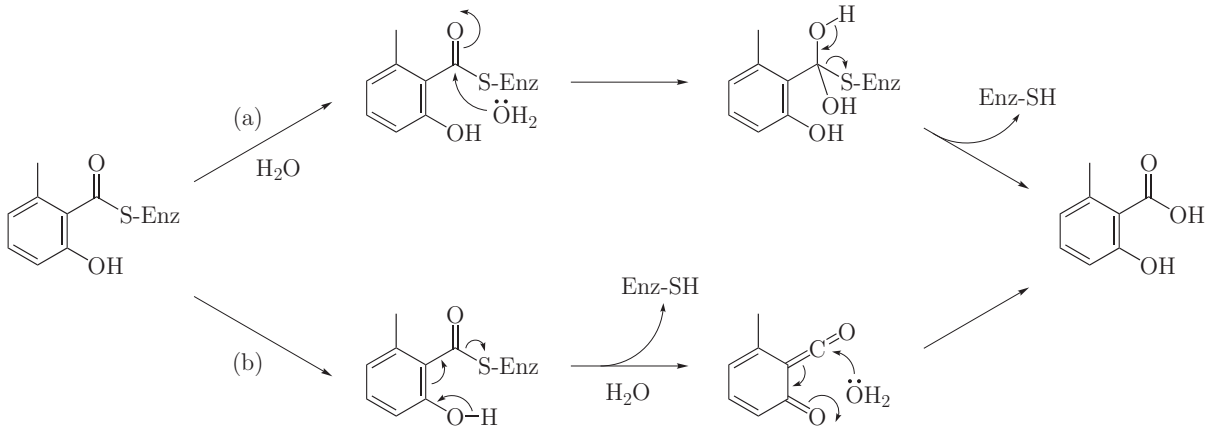


Figure I.6. Release of 6-methylsalicylic acid from the enzyme by thioesterase (a) and ketene formation (b) mechanisms. Enz, enzyme.

cillium, Aspergillus, Byssochlamys and *Paecilomyces spp.*) and has a wide-spectrum antibiotic activity but was abandoned as a pharmaceutical because of its high toxicity to both plants and animals and potent carcinogenicity. This mycotoxin is responsible for illness and death of cattle fed with barley stubble contaminated with *Aspergillus clavatus*, producing patulin. Another patulin producer, *Penicillium expansum*, a causative agent of soft rot mould of apples, can very easily contaminate fresh apple juice (Bennett, Klich, 2003; Davis, Diener, 1987; Hacking, Rosser, 1981; Moss, 2002; Puel *et al.*, 2007; Rice *et al.*, 1977).

Another compound derived from 6-MSA via decarboxylation and oxidation is *p*-toluquione (Figure I.7), detected as a minor metabolite during the study of patulin biosynthesis in *Penicillium urticae* (Bu'Lock *et al.*, 1965, 1969). Related toluquinones were isolated from both ascomycete and basidiomycete fungi. The biosynthetic origin of fungal toluquinones and some other compounds, e.g. flavipin, fumigatin, spinulosin (Figure I.3), was rather controversial because feeding experiments carried out by different researchers (mainly Packter and Pettersson) gave different results, so there was considerable confusion as to whether 6-MSA or orsellinic acid was the real precursor of these compounds. According to Pettersson's suggestion, 6-MSA, rather than orsellinic acid, is involved in the formation of toluquinones which lack the 3-hydroxyl group. It is also worthy to note that *Basidiomycetes* use the shikimate rather than polyketide route for aromatic compounds (Turner, 1971; Turner, Aldridge, 1983).

Genetic homologues of the *P. griseofulvum* MSAS have been found in many fungi such as *Aspergillus clavatus* (accession number XP_001273093; Nierman, 2006), *A. niger* (accession number XP_001402408; Pel *et al.*, 2007), *A. ochraceus* (accession number AAS98200; Dao *et al., unpublished*), *A. oryzae* (accession number BAE65442; Machida

et al., 2005), *A. parasiticus* (accession number AAC23536; Feng, Leonard, 1998) and *A. terreus* (accession number BAA20102; Fujii *et al.*, 1996; accession number AAC49814; Pazoutova *et al.*, 1997; accession number XP_001215453; Birren *et al.*, *unpublished*), *Byssochlamys nivea* (accession number AAK48943; Puel *et al.,* 2007), *Cochliobolus heterostrophus* (accession number AAR90279; Kroken *et al.*, 2003), *Glarea lozoyensis* (accession number AAX35547; Lu *et al.,* 2005), *Penicillium expansum* (accession number AAY88187, partial; White *et al.*, 2006), *P. griseofulvum* (accession number AAB49684; Bingle *et al.*, 1999), *Phaeosphaeria nodorum* (accession number EAT91972; Birren *et al.*, *unpublished*), *Phoma sp.* C2932 (accession number CAB44720, partial; Bingle *et al.*, 1999), *etc.* Cloning of partial cDNA and genomic DNA sequences of the MSAS gene was also reported for *Penicillium urticae*, although these sequences are not available in the GenBank/EMBL database (Wang *et al.*, 1991).

Isolation of numerous secondary metabolites presumably derived from 6-MSA from various fungal species (Turner, 1971; Turner, Aldridge, 1983) suggests that genes encoding MSASs are widely spread amongst fungi.

A *Penicillium freii* DNA fragment encoding a protein (designated PfKS) that shows a high degree of similarity (89.5% identity, 92.1% similarity) to the KS domain of the *P. griseofulvum* MSAS was identified by screening of a cosmid library using a part of the *P. griseofulvum* MSAS gene as a probe (accession number X95884, Nicolaisen *et al.*, 1997). A dot matrix comparison of the nucleotide sequence of the isolated *P. freii* region (approximately 7 kb) and the MSAS gene revealed that the PfKS open reading frame and its downstream region were highly similar to the region encoding the KS domain in the *P. griseofulvum* MSAS gene and, furthermore, to the region of termination of the MSAS gene. Interestingly, the resulting gene harbored only a KS domain and a region similar to the 3′ flanking region of the *P. griseofulvum* MSAS gene but an internal region encoding AT, DH, KR and ACP domains was not present in it. These results suggested that deletion of the corresponding region has occurred or/and that other active domains necessary for the assembly of a polyketide chain are located on other proteins. According to another hypothesis proposed, the isolated DNA fragment was a pseudogene originated from an ancestral PKS or FAS. Expression of the gene on the transcriptional level in its natural host was confirmed by reverse transcriptase PCR and the single intron was correctly spliced. Some correlation between the PfKS transcript level and the polyketide xanthomegnin biosynthesis was observed. Occurrence of no homologous integration events in gene disruption experiments suggested a possible essential role of PfKS in fatty acid biosynthesis or involvement in the biosynthesis of a precursor for both the secondary metabolite xanthomegnin and an essential fatty acid (Nicolaisen *et al.*, 1997).

A MSAS gene was the first fungal PKS gene expressed in a heterologous host. In 1995 Khosla and coworkers reported successful expression of the *P. griseofulvum* MSAS gene in *Streptomyces coelicolor* CH999, an engineered host strain lacking a genomic region encoding actinorhodin biosynthesis, originally developed for functional expression of bacterial PKS genes (Bedford *et al.*, 1995). To allow optimal expression in *S. coelicolor*, the MSAS gene was modified by removing the 69 bp intron, introducing a Shine-Dalgarno sequence complementary to the 3′ region of *S. coelicolor* 16S rRNA and by altering four of the first seven codons to ones more frequently used in *Streptomyces* species (Wright, Bibb, 1992). The expression plasmid used for transformation contained the *P. griseofulvum* MSAS gene under control of the actinorhodin *act*I promoter which is activated at the onset of the stationary phase. Detectable amounts of 6-MSAS were produced (approxi-

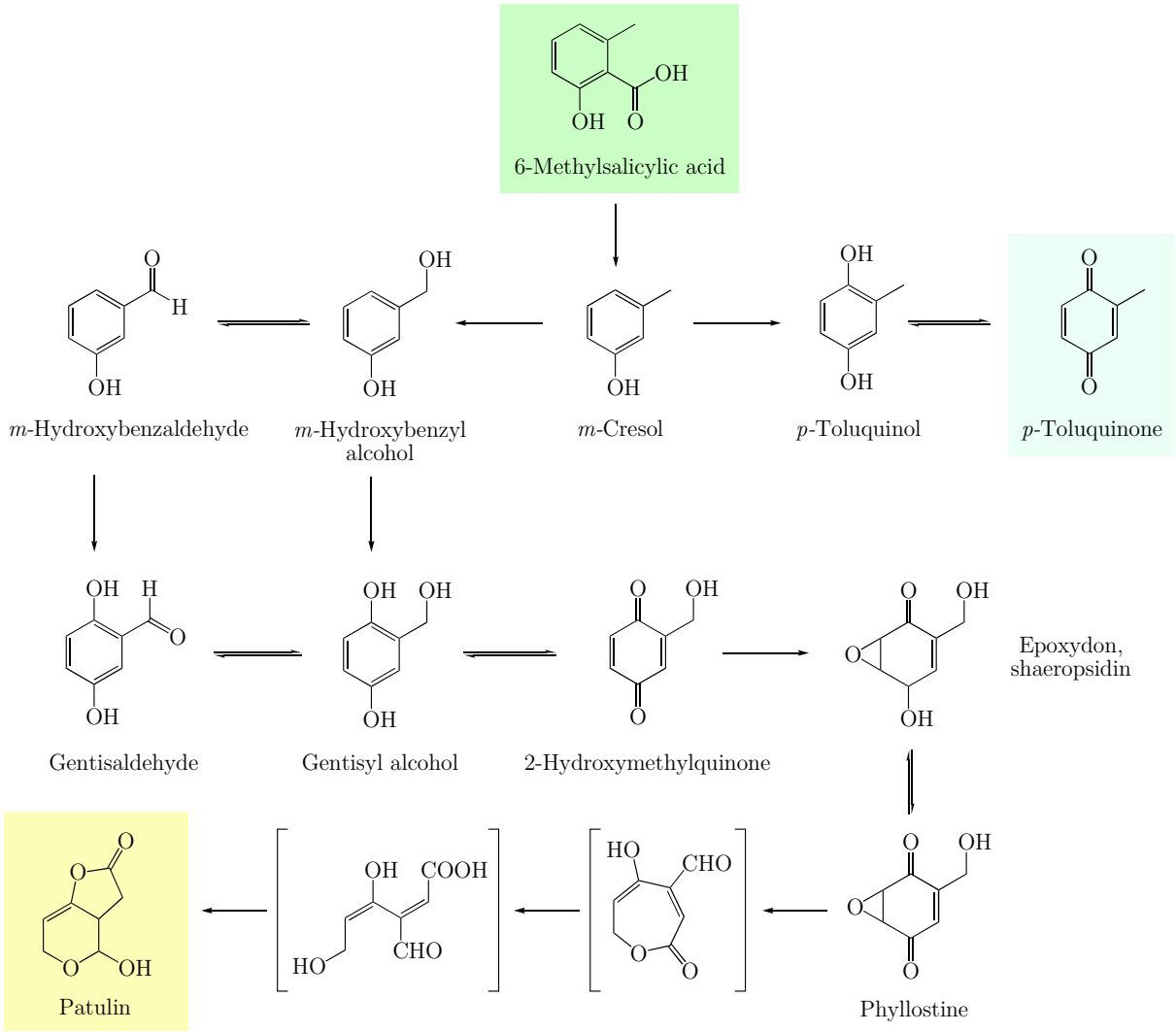


Figure I.7. Biosynthetic route from 6-methylsalicylic acid to the mycotoxin patulin. Hypothetical intermediates are shown in brackets.

mately 67 mg · l^{-1} of solid medium), proving that MSAS was expressed, folded into the correct tertiary structure and its ACP domain was successfully modified by addition of a 4′-phosphopantetheine group in this actinomycete host (Figure I.8) (Bedford *et al.*, 1995).

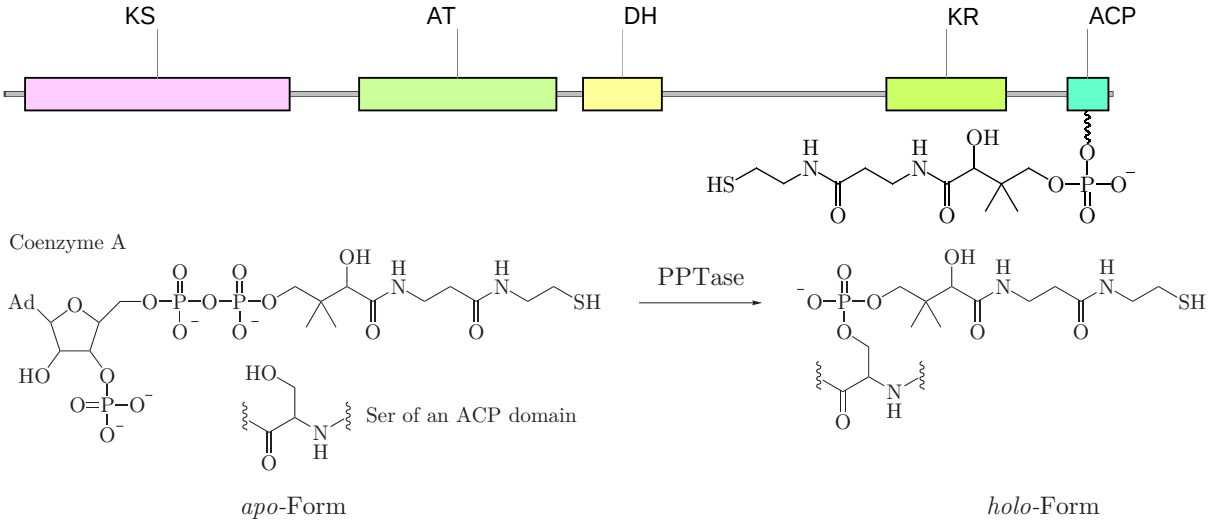


Figure I.8. Phosphopantetheinylated MSAS and proposed mechanism of phosphopantetheinylation. Ad, an adenine moiety; Ser, a serine residue; PPTase, 4′-phosphopantetheinyl transferase.

Heterologous expression of the MSAS gene in a fungal host was achieved almost simultaneously, in 1996, by Fujii and coworkers (Fujii *et al.*, 1996). The putative MSAS gene *at*X was isolated from an *Aspergillus terreus* genomic DNA library using the *P. griseofulvum* MSAS gene as a probe. Sequencing of *at*X revealed a 5.5 kb open reading frame intervened by a 70 bp intron near the 5′ end. The atX gene was cloned into the fungal expression vector pTAex3 (Figure II.2) downstream of the α-amylase (Taka-amylase A, *amy*B) promoter and the resulting plasmid was then introduced into the heterologous host *Aspergillus nidulans*. The obtained transformants produced significant amounts of 6-MSA (approximately 330 mg · l^{-1} of culture medium), establishing the function of the ATX polypeptide encoded by *at*X as a MSAS (Figure I.4). The *at*X gene was considered to be silent in its natural host because no transcripts of *at*X were detected by Northern blot analysis of all growth phases of *A. terreus*. However, another MSAS gene identified in *A. terreus* and designated *pks*M was found to be transcribed in the middle of the vegetative growth phase (Pazoutova *et al.*, 1997).

Aspergillus nidulans was used as a host in another successful heterologous expression of a MSAS gene, isolated from a *Glarea lozoyensis* cosmid genomic DNA library using its own previously cloned KS domain as a probe (Lu *et al.*, 2005; Zhang *et al.*, 2003). In contrast to MSAS genes from *P. griseofulvum*, *A. terreus* and *Byssochlamys nivea* that

contain only one intron, the 5373-bp open reading frame of the G. lozoyensis MSAS gene, designated *pks*2, is interrupted by three introns of 51, 50 and 65 bp, clustered near the 5′ end of the gene. To verify the function of the *pks*2 gene, it was transferred into *A. nidulans* under control of the constitutive *trp*C promoter. Transcription of the *pks*2 gene was confirmed by 5′- and 3′-RACE (rapid amplification of cDNA ends) experiments and production of 6-MSA was confirmed by NMR and mass spectrometry analysis.

The first expression of a MSAS from *P. griseofulvum* in *Escherichia coli* and *Saccharomyces cerevisiae* was reported in 1998 by Kealey and coworkers (Kealey *et al.*, 1998). In both heterologous hosts, production of 6-MSA requires coexpression of a heterologous 4′-phosphopantetheinyl transferase (PPTase) to convert the expressed *apo*-PKS to its *holo* form. PPTase transfers the 4′-phosphopantetheine moiety from coenzyme A to the side chain β-hydroxyl group of a conserved serine residue within the PKS, resulting in the conversion of the 4′-phosphopantetheine-dependent ACP from the inactive *apo*-form into the active *holo*-form. To fulfill this requirement the *sfp* gene from *Bacillus subtilis* encoding the surfactin PPTase and the *P. griseofulvum* MSAS gene were coexpressed under the control of the alcohol dehydrogenase 2 (ADH2) and T5 RNA polymerase promoters in *S. cerevisiae* and *E. coli*, respectively. Production of 6-MSA in *E. coli* was both temperature- (30°C) and glycerol-dependent (10% glycerol in ATCC medium 765) and levels of production were low (75 $mg \cdot l^{-1}$ of culture medium) compared to those in the native host. In yeast, however, production levels of 6-MSA were two fold higher (1.7 $g \cdot l^{-1}$ of culture medium) than in the native host.

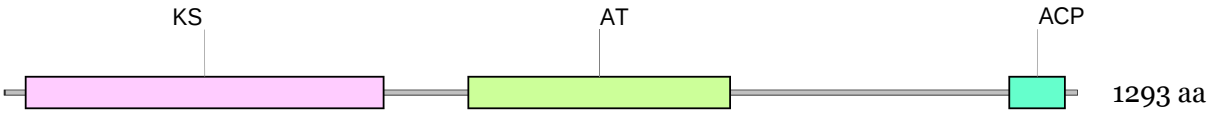
The expression of the MSAS gene from *P. griseofulvum* in tobacco, being to date the only example of fungal PKS gene expression in plants, was reported by Yalpani and coworkers (Yalpani *el al.*, 2001). The aim of their study was to investigate whether 6-MSA can mimic salicylic acid which differs from 6-MSA only by the absence of a methyl group $-CH_3$ and is known to be the native activator of disease resistance in plants. To direct the heterologously produced MSAS to tobacco plastids, the MSAS gene with the intron removed was fused to the chloroplast targeting sequence of the small subunit of ribulose bisphosphate carboxylase from petunia. Expression of the MSAS gene was driven by the strong constitutive promoter SCP1 with the Ω' translation enhancer element from tobacco mosaic virus. The resulting binary expression vector containing a kanamycin resistance marker was successfully introduced into tobacco by *Agrobacterium*-mediated transformation and the kanamycin resistant transformants were selected. Leaves of transgenic tobacco plants expressing the MSAS gene accumulated significant amounts of 6-MSA, primarily in the form of a β-glucoside. Interestingly, the production and accumulation of 6-MSA in transgenic plants correlated with increased levels of defence proteins and

resulted in activated resistance to tobacco mosaic virus. These results demonstrate that certain endogenous plant PPTase appear to be sufficiently flexible to phosphopantetheinylate ACP domains of fungal PKSs and suggest that plants can serve as heterologous hosts for functional expression of other fungal PKS genes.

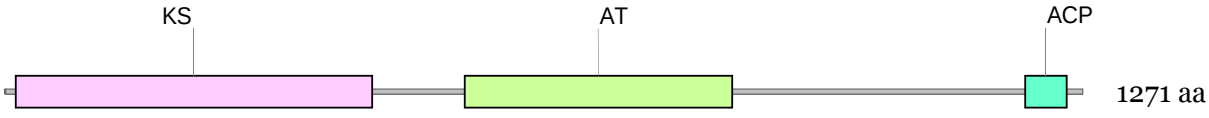
2.1.2. Orsellinic acid synthases

The first isolation of OAS was carried out by Gaucher and Shepherd in 1968 from *Penicillium madriti* (Gaucher, Shepherd, 1968). *In vitro* enzyme activity of OAS was also detected in the cell-free extract of *Penicillium cyclopium* (Woo *et al.*, 1989). Later, OAS was purified from *P. cyclopium* and its size was estimated to be ∼130 kDa (Jordan, Spencer, 1993). The smaller size of the OAS subunit, compared to MSAS (180 kDa), reflects the absence of KR and DH domains in this enzyme, since no ketoreduction and dehydration are necessary for orsellinic acid biosynthesis.

The genes encoding OAS have not yet been cloned from fungi. No hybridization signal with the *P. griseofulvum* MSAS probe was detected in the genomic DNA of orsellinic acid-producing *P. cyclopium* indicating a large sequence divergence between the MSAS and OAS genes (Fujii *et al.*, 1996). So far, OASs have been found only in bacteria, being an exceptional example of bacterial iterative type I PKSs. The *avi*M gene encoding OAS involved in the biosynthesis of the antibiotic avilamycin has been cloned from *Streptomyces viridochromogenes* Tü57 (Figures I.9, I.10) (Gaisser *et al.*, 1997; Weitnauer *et al.*, 2001). The deduced amino acid sequence of AviM, encoded by *avi*M, showed 37% identity to the MSAS from *P. patulum* and heterologous expression of *avi*M in *Streptomyces lividans* and *S. coelicolor* resulted in production of orsellinic acid.



Streptomyces viridochromogenes AviM



Micromonospora echinospora CalO5

Figure I.9. Domain organization of bacterial orsellinic acid synthases: AviM from *Streptomyces viridochromogenes* (accession number AAK83194; Gaiser *et al.*, 1997; Weitnauer *et al.*, 2001), CalO5 from *Micromonospora echinospora* (accession number AAM70355; Ahlert *et al.*, 2002).

Another OAS, CalO5, is involved in the biosynthesis of the enediyne antibiotic and powerful antitumor agent calicheamicin γ_1^I in *Micromonospora echinospora* (Figures I.9, I.10) (AHLERT *et al.*, 2002). CalO5 exhibits striking similarity (65% sequence identity, 93% sequence similarity) to the OAS AviM from *S. viridochromogenes*, thus most likely playing an identical role in the biosynthesis of the calicheamicin aryltetrasaccharide.

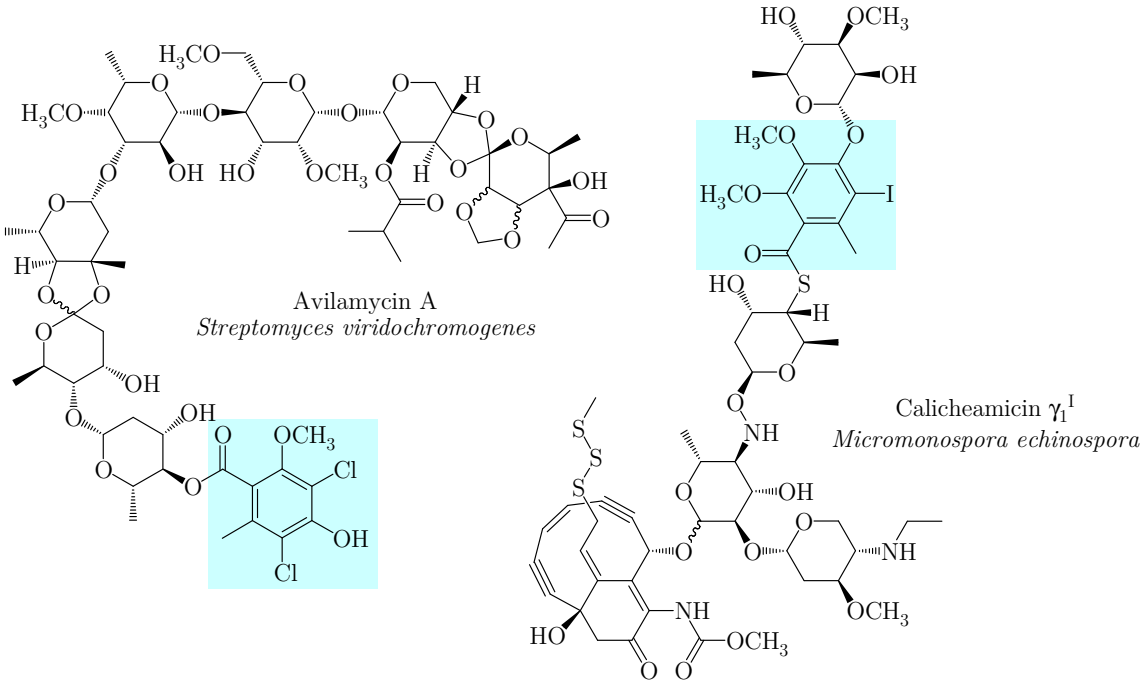


Figure I.10. Structures of avilamycin A and calicheamicin γ_1^I. Orsellinic acid moieties are boxed.

Interestingly, a BLAST search[3] of AviM and CalO5 against all fungal genomes sequenced and available to date in the GenBank/EMBL database and at the Broad Institute gave no OAS homologues in these genomes, supporting a rather enigmatic character of OAS genes in fungi. It is possible that these genes have been lost from fungal genomes during evolution or came via horizontal gene transfer, perhaps from bacteria, and therefore are exceptionally found only in some fungal lineages that are known to produce orsellinic acid and/or its derivatives (This study).

The biosynthetic mode of OAS is more straightforward than MSAS since no reduction and dehydration is required after the second cycle of condensation for the biosynthesis of orsellinic acid. A series of similar studies to those on MSAS were carried out with OAS to probe its stereochemical control. Incubation of (2*S*) and (2*R*) [1-^{13}C,2-^{2}H]-malonyl-CoAs with OAS from *P. cyclopium* in separate experiments revealed identical mass spectra from each experiment, indicating the absence of stereochemical integrity during biosynthesis

[3]http://www.ncbi.nlm.nih.gov/BLAST
http://www.broad.mit.edu/cgi-bin/annotation/fgi/BLAST_page.cgi

of orsellinic acid. It was therefore stated that in contrast to 6-MSAS the enolizations occurring during orsellinic acid biosynthesis are non-stereospecific and non-enzymatically mediated (Woo *et al.*, 1989).

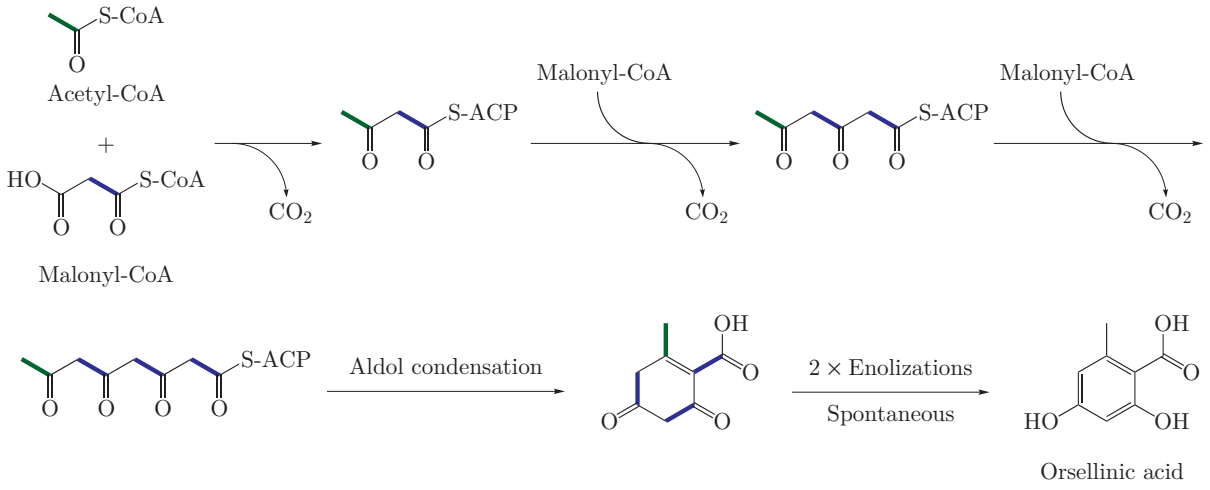


Figure I.11. Biosynthesis of orsellinic acid

Despite its simplicity, orsellinic acid was not isolated until 1958 (Reio, 1958; Mosbach, 1959), although orsellinic acid residues had been known to be constituents of many lichen metabolites (depsides, depsidones, depsones and true dibenzofurans) and its methyl, dimethyl and ethyl derivatives had also been isolated from lichens (Dayan, Romagni, 2001; Huneck, Yoshimura, 1996). Orsellinic acid itself has since been found in some lichens and cultured mycobionts and a number of asco- and basidiomycete fungi. It is rarely produced in high yield probably due to further transformations. Orsellinic acid is a precursor of numerous fungal metabolites, such as orcinol, fumigatin, spinulosin, various quinones, oosporein, phoenicin, penicillic acid and many others (Turner, 1971; Turner, Aldridge, 1983). On analogy with patulin, penicillic acid is formed from orsellinic acid via oxidative ring opening with similar intermediates (Al-Rawi *et al.*, 1974; Axberg, Catenbeck, 1975, 1975a; Bentley, Keil, 1962; Elvidge *et al.*, 1977; Mosbach, 1960). Penicillic acid is considered to a mycotoxin that exhibits wide-spectrum antibiotic (antibacterial, antiviral) and carcinogenic activity (Ciegler *et al.*, 1971; Davis, Diener, 1987).

Interestingly, orsellinic acid, along with isocoumarin, were identified as by-products of heterologous expression of the *Colletrichium lagenarium PKS*1 gene in *Aspergillus oryzae* with the main product being 1,3,6,8-tetrahydroxynaphthalene (T4HN) (Fujii *et al.*, 1999). These results suggest that in some cases orsellinic acid in fungi can originate from derailment of polyketide reactions catalyzed by other, non OAS, PKSs.

2.2. PKSs for non-reduced polyketides

PKSs of this type catalyze biosynthesis of aromatic multi-ring compounds derived from the folding of a non-reduced polyketide chain and have more complex domain organization than MSASs and OASs. These PKSs comprise polypeptides of 2 100–2 200 (up to 2 400) amino acid residues, about 300 amino acids longer than MSASs (Figures I.17, I.22, I.29). Their KS domain is shifted approximately 340 amino acids towards the *C*-terminus, and typically, tandem ACP domains occur at the *C*-terminus. A thioesterase-like domain can also be found at the *C*-terminus of non-reducing PKSs. In the *Aspergillus nidulans* YWA1 PKS this domain was proven to be responsible for catalyzing a Claisen-type cyclization and is therefore sometimes referred to as a Claisen cyclase domain (CYC) (Fujii *et al.*, 2001). Despite of similarity in their fundamental architectures, non-reducing PKSs produce a wide variety of products, reflecting the choice of starter units, chain-lengths, folding and cyclization patters.

A limited number of genes encoding non-reducing PKSs have been cloned and only three of them, namely *Colletotrichum lagenarium* PKS1 (Fujii *et al.*, 1999), *Aspergillus nidulans wA* (Watanabe *et al.*, 1998; Watanabe *et al.*, 1999), and *Aspergillus fumigatus alb*1 (Watanabe *et al.*, 2000), have been expressed in heterologous fungi to investigate their function. Functional analysis of other non-reducing PKS genes was carried out mainly by gene disruption experiments.

2.2.1. Pentaketide 1,3,6,8-tetrahydroxynaphthalene synthases

Depending on folding and cyclization patterns, pentaketide intermediates can give rise to different cyclic products (Figure I.12). Although feeding experiments with ^{13}C-labeled acetate confirmed the pentaketide origin of some aromatic compounds such as sclerin (Figure I.12) (Garson, Staunton, 1976; Garson, Staunton, 1978) and scytalone (Figure I.13) (Aldridge *et al.*, 1974; Sankawa *et al.*, 1977), there is only limited information on pentaketide biosynthesis at the biochemical and molecular genetic level.

Melanins are high molecular weight natural pigments formed by oxidative polymerization of phenolic compounds and usually have a dark brown or black color. In many pathogenic fungi from both *Ascomycota* and *Basidiomycota*, they have been found to be important virulence factors (Bell, Wheeler, 1986; Butler, Day, 1998; Henson *et al.*, 1999; Jacobson, 2000; Langfelder *et al.*, 2003; Wheeler, Bell, 1988).

Most fungal melanins are derived from oxidative polymerization of the precursor molecule 1,8-dihydroxynaphthalene (DHN) and are known as DHN-melanins. Studies on DHN-melanin biosynthesis in pathogenic fungi allowed identification of its pentaketide intermediates, such as scytalone, 1,3,8-trihydroxynaphthalene, vermelone, *etc.* (Figure

I.13). 1,3,6,8-Tetrahydroxynaphthalene (T4HN) is considered to be an initial product of pentaketide PKSs, sometimes referred as T4HN synthases, and serves as a precursor in the biosynthetic route of DHN-melanins in pathogenic fungi (Figure I.13).

The PKS1 gene encoding a T4HN synthase has been cloned from the phytopathogenic ascomycete fungus *Colletotrichum lagenarium*, a causative agent of anthracnose of cucumber and some other members of the family *Cucurbitaeceae* (Kubo *et al.*, 1991). A cosmid DNA library with wild-type *C. lagenarium* genomic DNA was constructed and melanin-deficient albino mutants were transformed with DNA from this library. An 8.4-kb *Bam*HI DNA fragment recovered from one of the melanin-restored transformants was identified as responsible for the melanin producing phenotype. Nucleotide sequencing of the isolated fragment revealed the presence of one open reading frame of 6.7 kb consisting of

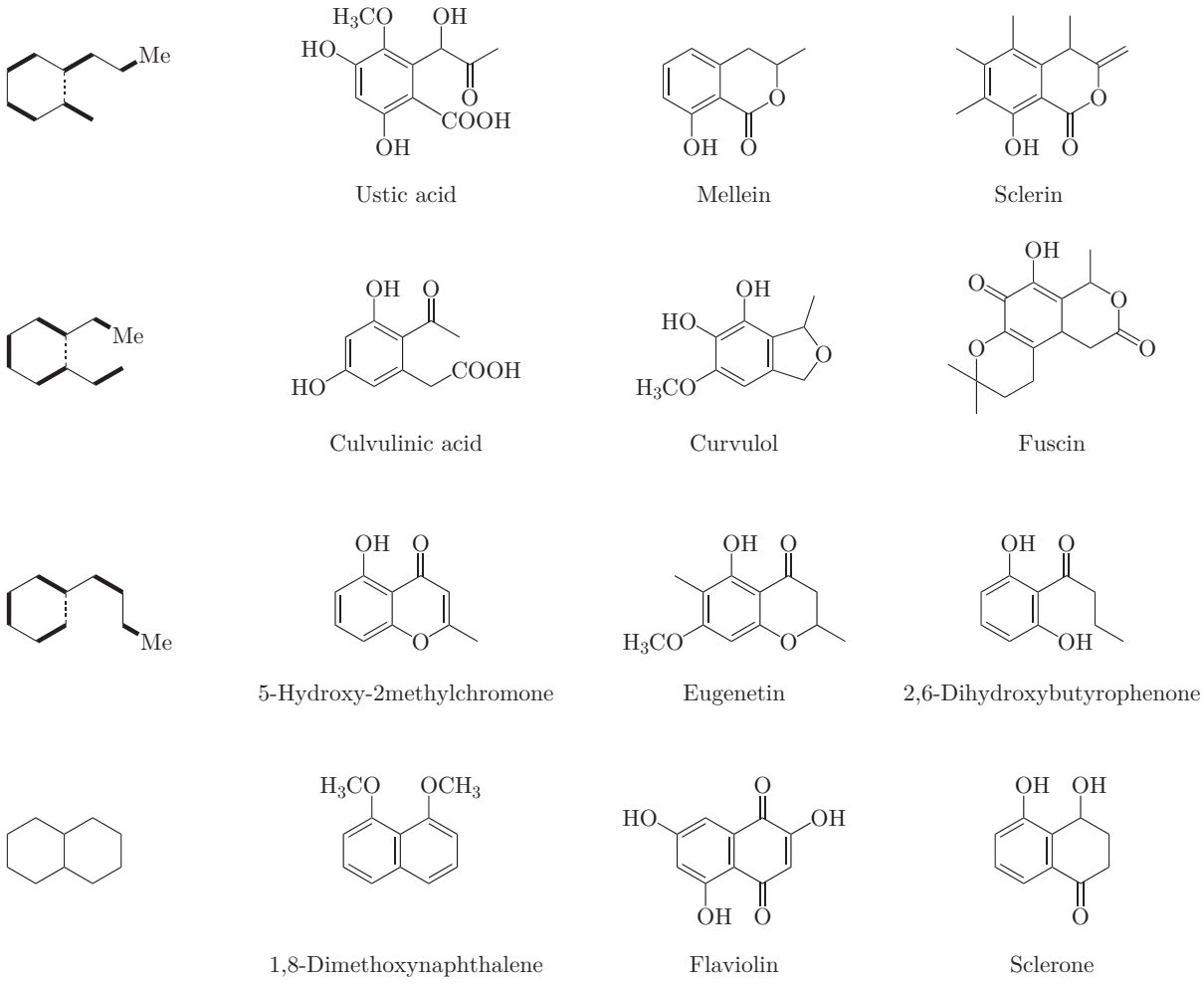


Figure I.12. Folding patterns of pentaketides and their corresponding cyclized products (adapted from Fujii, 1999; Turner, 1971). Me represents a terminal methyl group of the pentaketide chain.

3 exons, intervened by two short introns near the 5′ end. The deduced protein of 2187 amino acids showed significant similarities with other polyketide synthases, particularly with the *Aspergillus nidulans* YWA1, involved in conidial pigmentation (67% sequence identity, 80% sequence similarity), and contained conserved domains in the following order: KS→AT→ACP→ACP→TE (CYC). The *C. lagenarium* PKS1 gene was proposed to encode a polyketide synthase involved in melanin biosynthesis (Takano *et al.,* 1995).

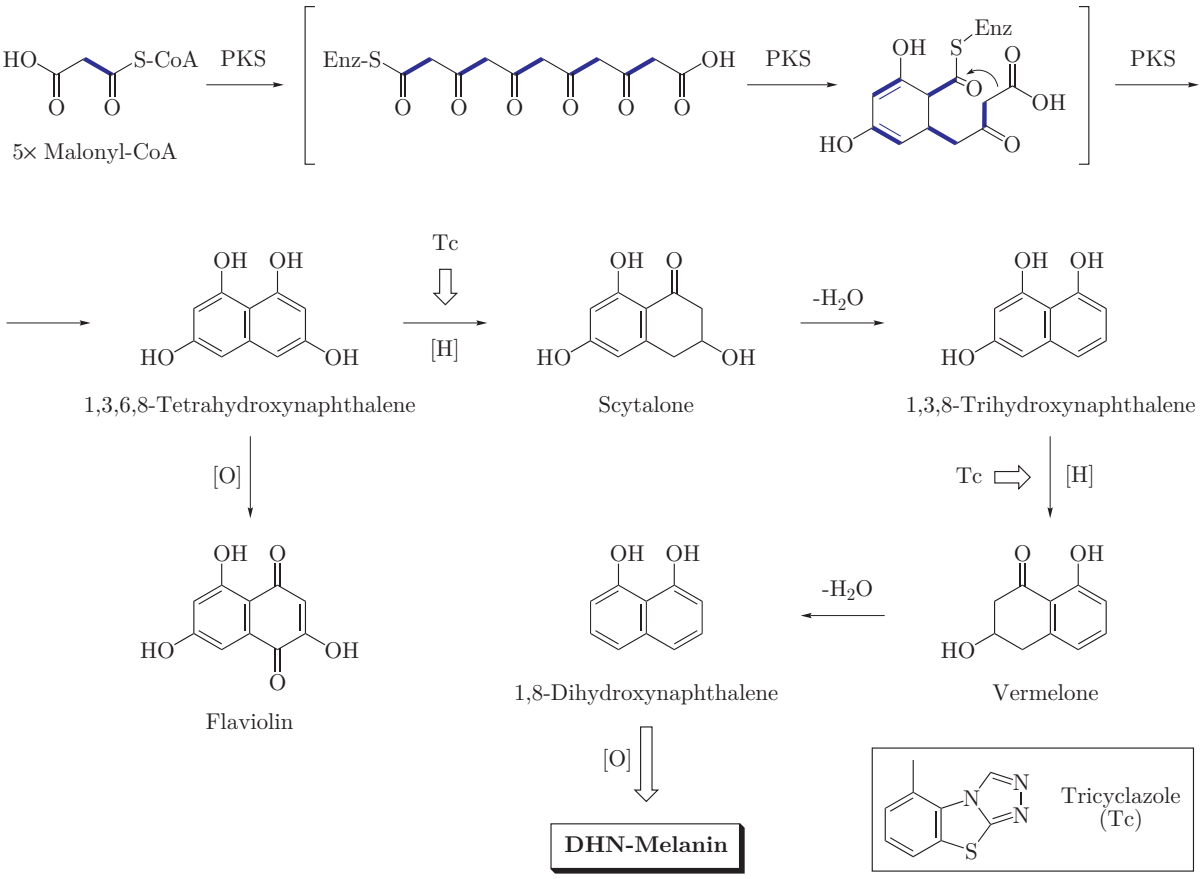


Figure I.13. 1,3,6,8-Tetrahydroxynaphthalene synthase reaction and biosynthesis of DHN-melanin. Enz, enzyme; −S−, thioester linkage; [H], a reduction step; [O], an oxidation step; $-H_2O$, a dehydration step; Tc, a reduction step which can be inhibited by the fungicide tricyclazole (shown boxed).

To identify the product of the *C. lagenarium* PKS1 gene, it was expressed in the heterologous fungal host, *Aspergillus oryzae* (Fujii *et al.*, 1999). Prior to transformation, the PKS1 gene was modified by replacing the 5′ end containing introns with the corresponding cDNA fragment. The re-constructed intronless PKS1 open reading frame was then subcloned into the fungal expression vector pTAex3 (Figure II.2) downstream of the starch inducible α-amylase promoter and the resulting plasmid was then transformed into *A. oryzae*. The *A. oryzae* transformants harboring the *C. lagenarium* PKS1 gene

produced large amounts of brown to black pigments even on minimal medium. The extracted compounds were acetylated to prevent possible polymerization of unstable T4HN and subjected to HPLC and NMR analysis. The physicochemical assay revealed T4HN as a major compound, thus proving that the *C. lagenarium* PKS1 gene encodes a pentaketide synthase for T4HN biosynthesis.

Six possible cyclization patterns of a pentaketide were proposed to form a naphthalene backbone as shown in Figure I.14. Feeding experiments with [1,2-$^{13}C_2$]acetate in *Phialophora lagerbergii*, producing scytalone in high yields, showed that this compound is synthesized as a mixture of two differently labeled species representing different cyclization patterns (Figure I.15) (Sankawa *et al.*, 1977). These results suggest involvement of a symmetrical intermediate compound such as T4HN which can give rise to scytalone by reduction of either aromatic ring with equal probability. The absence of a ^{13}C–^{13}C coupling between the ring junction carbons excludes a folding pattern shown on Figure I.14 in brackets.

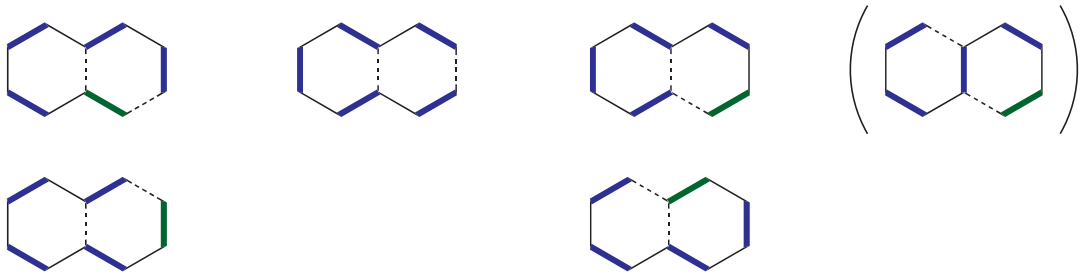


Figure I.14. Possible folding patterns for naphthalene cyclization

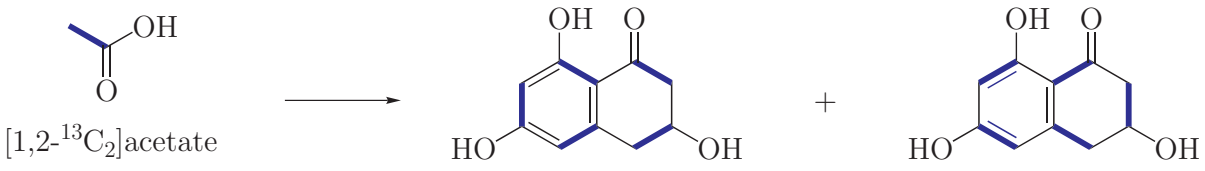


Figure I.15. Labeling patterns of scytalone produced from doubly-labeled acetate

Orsellinic acid and 6,8-dihydroxy-3-methylisocoumarin were isolated as by-products of heterologous expression of the *C. lagenarium* PKS1 gene in *A. oryzae*. Production of 6,8-dihydroxy-3-methylisocoumarin indicated a possible folding and cyclization pattern of a pentaketide chain as shown in Figure I.16. However, isolation of 5-hydroxy-2-methylchromone (Figure I.12) from *Daldinina concentrica* along with DHN (Figure I.13) (Allport, Bu'Lock, 1960) leaves the question of pentaketide cyclization pattern open.

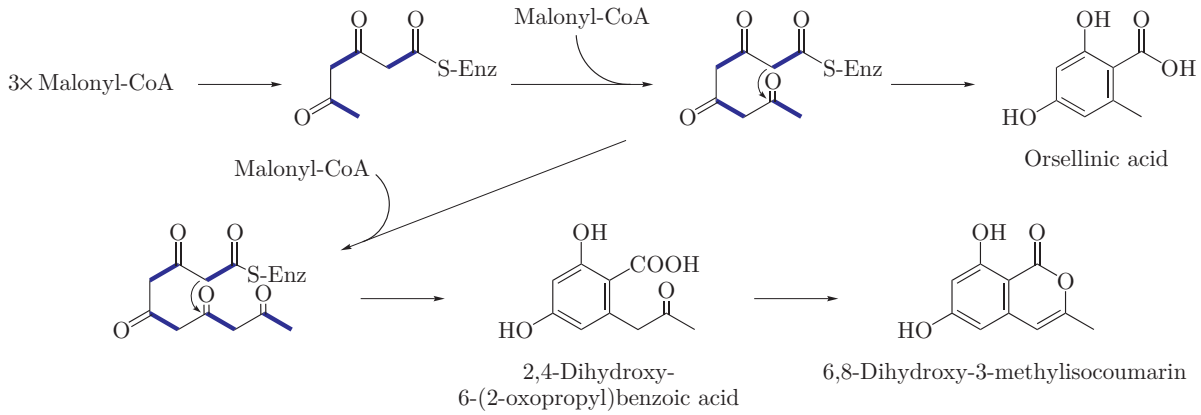


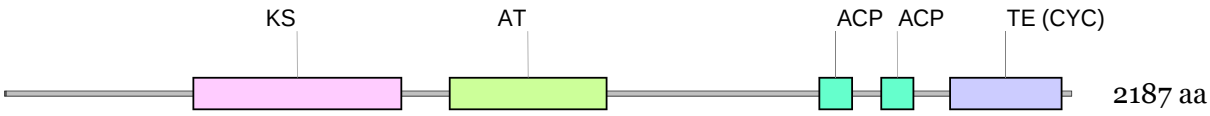
Figure I.16. By-products of heterologous expression of the *C. lagenarium* PKS1 gene in *A. oryzae*. Enz, enzyme; −S−, thioester linkage.

The genes encoding T4HN synthases have been cloned from other melanin-producing fungi, such as *Bipolaris oryzae* (accession number BAD22832, Moriwaki *et al.*, 2004), *Ceratocystis resinifera* (accession number AAO60166; Loppnau *et al.*, 2004; Tanguay *et al.*, 2006b), *Exophiala (Wangiella) dermatitidis* (accession number AAD31436; Feng *et al.*, 2001) and *E. lecanii-corni* (accession number AAN75188; Cheng *et al.*, 2004), *Glarea lozoyensis* (accession number AAN59953; Zhang *et al.*, 2003), *Nodulisporium sp.* (accession number AAD38786; Fulton *et al.*, 1999), *Ophiostoma floccosum* (accession number AAL05883; Wang, Breuil, *unpublished*) and *O. piceae* (accession number ABD47522; Tanguay *et al.*, 2007), *Xylaria sp.* BCC 1067 (accession number AAM93545; Amnuaykanjanasin *et al.*, 2005). In most cases their function was confirmed by gene disruption experiments, resulting in nonmelanizing albino mutants. These genes encode proteins of 2 100–2 200 amino acids that share the same architecture with distinctive two tandem ACP domains and one CYC domain: KS→AT→ACP→ACP→TE (CYC). (Figure I.17).

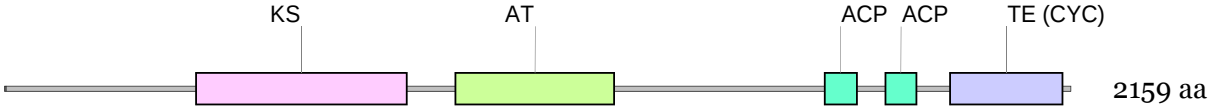
2.2.2. PKS involved in citrinin biosynthesis

Citrinin is a nephrotoxic and bactericidal mycotoxin first isolated from *Penicillium citrinum*. Subsequently, it was identified in many species of the genera *Penicillium* and *Aspergillus*, including some strains of such important food biotechnology species as *Penicillium camemberti* and *Aspergillus oryzae*. More recently, citrinin has also been isolated from *Monascus ruber* and *Monascus purpureus* producing red pigments that are widely used as natural food colorants (Bennett, Klich, 2003).

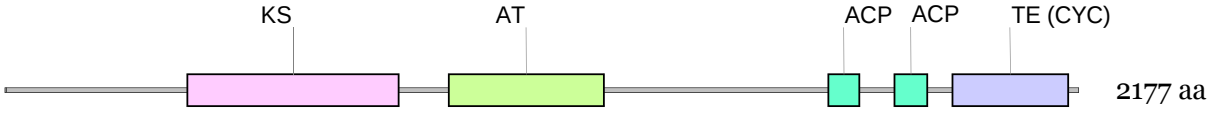
Extensive feeding experiments on citrinin biosynthesis confirmed its polyketide origin (Barber, Staunton, 1980; Colombo *et al.*, 1981; Hajjaj *et al.*, 1999; Sankawa *et al.*, 1983). In *Penicillium* and *Aspergillus*, formation of a pentaketide precursor from one



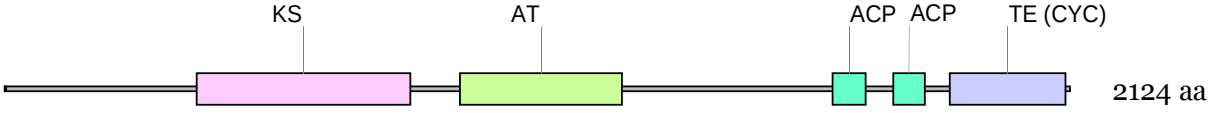
Colletotrichum lagenarium PKS1



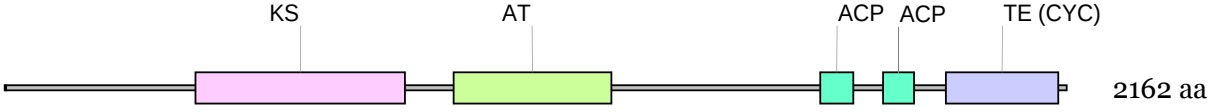
Nodulisporium sp. PKS1



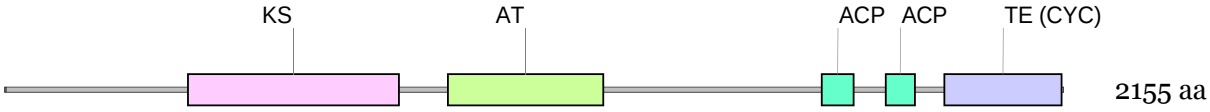
Exophiala dermatitidis WdPKS1



Glarea lozoyensis PKS1



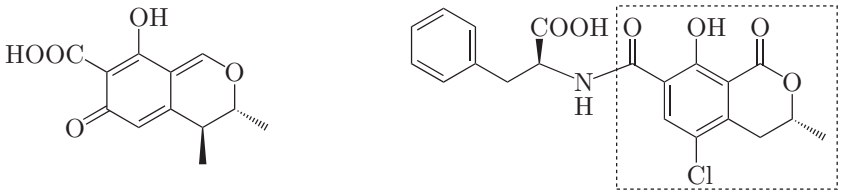
Xylaria sp. PKS12



Bipolaris oryzae PKS1

Figure I.17. Domain organization of selected fungal 1,3,6,8-tetrahydroxynaphthalene synthases: PKS1 from *Colletotrichum lagenarium* (accession number BAA18956; Takano *et al.*, 1995), PKS1 from *Nodulisporium sp.* (accession number AAD38786; Fulton *et al.*, 1999), WdPKS1 from *Exophiala (Wangiella) dermatitidis* (accession number AAD31436; Feng *et al.*, 2001), PKS1 from *Glarea lozoyensis* (accession number AAN59953; Zhang *et al.*, 2003), PKS12 from *Xylaria sp.* (accession number AAM93545; Amnuaykanjanasin *et al.*, 2005), PKS1 from *Bipolaris oryzae* (accession number BAD22832; Moriwaki *et al.*, 2004). KS, β-ketoacyl synthase; AT, acyltransferase; ACP, acyl carrier protein; TE, thioesterase; CYC, Claisen-type cyclase.

acetyl starter unit and four malonyl extender units was proposed. In contrast, biosynthesis of citrinin in *Monascus* species involves a tetraketide precursor arising from the condensation of one acetyl starter unit with three malonyl extender units. This is followed by addition of one more acetate unit to the formed tetraketide. The resulting C_{10} intermediate undergoes *C*-methylation and further modifications to give citrinin.



Citrinin Ochratoxin A

Figure I.18. Structures of the mycotoxins citrinin and ochratoxin A

The PKS gene involved in citrinin biosynthesis was isolated from a *M. purpureus* genomic DNA library (Shimizu *et al.*, 2005) using a fragment, amplified from *M. purpureus* genomic DNA by means of KS degenerate primers (Bingle *et al.*, 1999), as a probe. The full-length PKS gene, named *pksCT*, of 7 838 bp with a single 83-bp intron, encoded a protein of 2 584 amino acids containing conserved KS, AT, ACP domains and a rare *C*MeT domain with an unusual location downsteam of the ACP domain (Figure I.19). The presence of a *C*MeT domain in PKSCT corresponds to the need for a methylation step in the citrinin biosynthetic pathway established for *M. ruber* (Hajjaj *et al.*, 1999). A TE domain, usually found downstream of the ACP domain in aromatic multi-ring PKSs, was not observed in the *M. purpureus* PKSCT.

Monascus purpureus PKSCT

Figure I.19. Domain organization of the polyketide synthase PKSCT involved in biosynthesis of citrinin in *Monascus purpureus* (accession number BAD44749; Shimizu *et al.*, 2005). KS, β-ketoacyl synthase; AT, acyltransferase; ACP, acyl carrier protein; *C*MeT, *C*-methyltransferase.

Transcriptional analysis revealed correlation between citrinin production and the level of *pks*CT mRNA in *M. purpureus*. Homologous recombination between the wild-type allele and a construct containing a truncated *pks*CT resulted in a *pks*CT-disrupted strain of *M. purpureus* that was unable to produce citrinin. A *pks*CT revertant obtained by successive endogenous recombination events in the *pks*CT disruptant recovered its ability

to produce citrinin, establishing that *pks*CT encoded the PKS responsible for citrinin biosynthesis in *M. purpureus*.

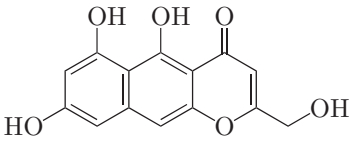
The *M. purpureus* PKSCT has the domain organization that is unique and unusual amongst fungal PKSs characterized so far. PKSs with similar architecture can be identified in genomes of many fungal species but their functions are not characterized yet. They belong to a separate clade III of non-reducing PKSs characterized by a *C*MeT domain located after the ACP domain [KS→AT→ACP→(ACP)→*C*MeT→TE (CYC)] which is an apparent result of domain rearrangement. It was proposed that the genes encoding these PKSs diverged from other non-reducing PKSs after the loss of a *C*MeT domain that was retained in this clade (Kroken *et al.*, 2003).

A structural analog of citrinin, ochratoxin A, consists of a chlorinated isocoumarin derivative linked to L-phenylalanine via an amide linkage (Figure I.18). This compound, produced by several *Aspergillus* and *Penicillium* species, is a nephrotoxin which also displays hepatotoxic, teratogenic, immunosuppressive activities and is considered a potent human carcinogen (Bennett, Klich, 2003). The PKS gene required for ochratoxin A biosynthesis in *Aspergillus ochraceus* was identified by a suppression subtractive hybridization PCR-based approach (SSH-PCR), but only a fragment of it was cloned (O'Callaghan *et al.*, 2003). In another study a part of the ochratoxin A biosynthetic gene cluster was cloned from *Penicillium nordicum* (Karolewiez, Geisen, 2005). Sequencing of the 10 kb fragment revealed the presence of three open reading frames. One open reading frame (partial sequence) encoded a deduced protein that had homology to fungal PKSs. Another open reading frame encoded a putative non-ribosomal peptide synthase and therefore was presumed to be responsible for the formation of an amide linkage between the dihydroisocoumarin moiety and L-phenylalanine. In both studies inactivation of the ochratoxin A PKS genes by disruption resulted in strains with abolished ability to produce ochratoxin A. Cloning of full-length ochratoxin PKS genes has not been achieved yet and no such sequences are available to date in the GenBank/EMBL database.

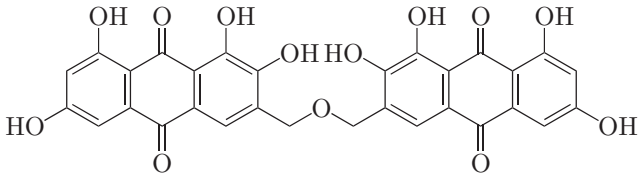
Interestingly, primary analysis of available partial sequences of ochratoxin PKSs obtained from ochratoxigenic *Aspergillus* and *Penicillium* species reveals that these PKSs are not related to each other: ochratoxin PKSs from *Penicillium* show a high degree of homology to fungal MSASs (KS→AT→DH→KR→ACP), while ochratoxin PKSs from *Aspergillus* species tend to be closer to PKSs from the clade I of fungal reducing PKSs with the general domain organization KS→AT→DH→(*C*MeT)→ER→KR→ACP (Kroken *et al.*, 2003). Thus, despite overall similarity in chemical structure, especially in the folding pattern of pentaketide precursors, biosynthesis of citrinin and ochratoxin A does not necessary involve related PKSs.

2.2.3. Heptaketide naphthopyrone synthase

Formation of asexual spores, or conidia, in *Aspergillus nidulans* has been well studied at the genetic level as a model system of multicellular development. Conidia of *A. nidulans* contain in their cell walls a dark green pigment that is not present in other cell types. Pigment is produced as spores mature and it protects them from damage caused by ultraviolet radiation (Wright, Pateman, 1970). The *Aspergillus nidulans* *w*A gene, involved in biosynthesis of this pigment, was identified in the study of differentiation of *A. nidulans* (Pontecorvo *et al.*, 1953). It was observed that the *w*A mutants produced colorless conidia, while the wild type showed green pigmentation in the asexual spores, or conidia. Yellow spores were produced by *y*A mutants and *w*A mutation were epistatic to *y*A mutations (Clutterbuck, 1972). Thus, it was considered that the product of the *w*A gene appeared to be a yellow intermediate, subsequently converted to the mature green spore pigment by a *y*A-encoded laccase (Aramayo, Timberlake, 1990). The structure of the yellow spore pigment intermediate of the related fungus *Aspergillus parasiticus* was determined to be hydroxymethylnaphthopyrone, named parasperone A (Figure I.20), which is considered to be a heptaketide (Brown *et al.*, 1993). Also, the pigment of *A. nidulans* ascospores, that have a reddish-brown color, was characterized as a dimeric hydroxylated anthraquinone, ascoquinone A (Figure I.20) (Brown, Salvo, 1994).



Parasperone A
Aspergillus parasiticus



Ascoquinone A
Aspergillus nidulans

Figure I.20. Structures of parasperone A and ascoquinone A

The *w*A gene was cloned by genetic complementation of the *w*A mutation with a cosmid library of wild-type *A. nidulans* DNA (Mayorga, Timberlake, 1990). Disruption analysis of the *w*A locus confirmed that the gene is required for the synthesis of green spore pigment present in the walls of mature conidia, since *w*A disruptants produced colorless (white) spores. Northern blot analysis detected a 7.5 kb transcript during conidiation, beginning when pigmented spores first appeared. The nucleotide sequence of the *w*A gene indicated that it coded for a 1 986 amino acid polypeptide with a molecular weight of 217 kDa. The predicted WA polypeptide contained KS, AT and two tandem ACP domains. The identified functional domains showed significant sequence similarity with other fungal PKS (MSAS), bacterial PKS (EryA) and mammalian FAS domains. Thus, it

was proposed that the *w*A gene coded for a polyketide synthase involved in the formation of the green spore pigment intermediate (Mayorga, Timberlake, 1992). Later, high similarity with the PKS1 gene of *Colletotrichum lagenarium* was demonstrated (Takano *et al.*, 1995), such as size and domain organization, although the TE (CYC) domain was missing in the first report (Mayorga, Timberlake, 1992).

The first attempt to express the *Aspergillus nidulans w*A gene in heterologous host *A. oryzae* was made by Watanabe and coworkers (Watanabe *et al.*, 1998). In order to identify the product of the WA polyketide synthase, they exploited the pTAex3 expression system (Figure II.2), previously proven to be successful for heterologous expression of the *at*X MSAS gene from *Aspergillus terreus* (Fujii *et al.*, 1996). The 3′-end of the *w*A gene was cut at the *Bam*HI site just after the reported stop codon. Without removing any introns, the gene was cloned into the fungal expression plasmid pTAex3 under control of the starch-inducible *A. oryzae* α-amylase promoter to construct a pTA-wA expression plasmid. The *A. oryzae* transformant harboring pTA-wA produced compounds which have not been detected in the host *A. oryzae* or in *A. nidulans* itself. The newly produced compounds were identified to be citreoisocoumarin and its derivatives, which are made from a heptaketide intermediate with or without reduction in their side chains (Figure I.21).

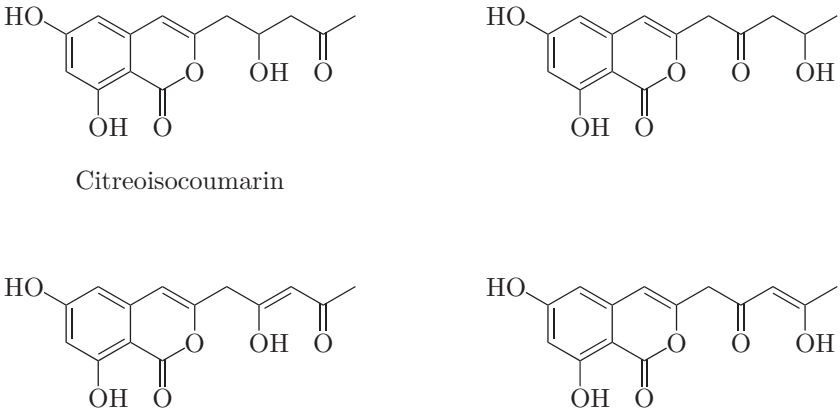
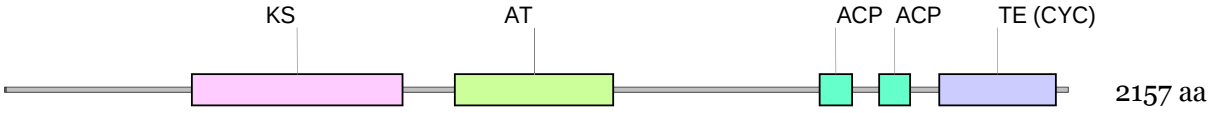


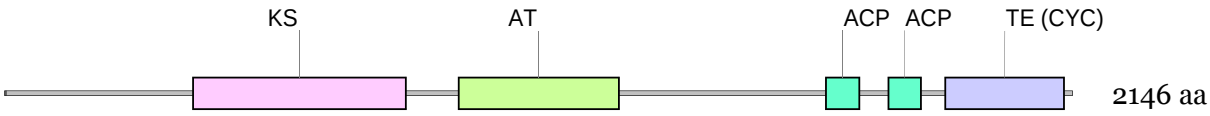
Figure I.21. Structures of citreoisocoumarin and its derivatives

Although it was apparent that production of these compounds was directed by the *w*A gene, there was some ambiguity concerning the length of the polyketomethylene chain and its cyclization pattern, because these compounds did not show a significant yellow color. To solve this discrepancy, re-sequencing of the *w*A gene around the *C*-terminal region was carried out and revealed an error in the original sequence, just before the predicted stop codon. A missing single base caused an apparent frameshift in the deduced amino acid sequence. The WA polypeptide sequence was corrected and appeared to be 171 amino

acids longer (2157 amino acids) (Figure I.22) than first predicted (1986 amino acids). Thus, it was obvious that the previous expression plasmid pTA-wA lacked a part of the WA *C*-terminal region with the TE domain, and the isolated compounds (Figure I.21) were not true spore pigment intermediates.



Aspergillus nidulans WA



Aspergillus fumigatus Alb1

Figure I.22. Domain organization of heptaketide naphthopyrone synthases: WA from *Aspergillus nidulans* (accession number CAA46695; Watanabe *et al.*, 1999) and Alb1 from *Aspergillus fumigatus* (accession number AAC39471; Tsai *et al.*, 1998). KS, β-ketoacyl synthase; AT, acyltransferase; ACP, acyl carrier protein; TE, thioesterase; CYC, Claisen-type cyclase.

To express the correct full-length *w*A PKS gene, a new expression plasmid, designated pTA-nwA, was constructed. The *A. oryzae*/pTA-nwA transformant grown on an agar plate exhibited yellow pigmentation (Figure I.23). This observation was quite striking since the previously obtained pTA-wA transformant with the truncated *w*A gene had white mycelia. From the induction culture medium of the pTA-nwA transformant, a yellow compound was isolated and identified as naphthopyrone YWA1 by physicochemical analysis. Production of citreoisocoumarin and its derivatives was not observed in the *A. oryzae*/pTA-nwA transformant culture (Watanabe *et al.*, 1999). The basic carbon backbone of YWA1 is identical to the asexual spore pigment intermediate

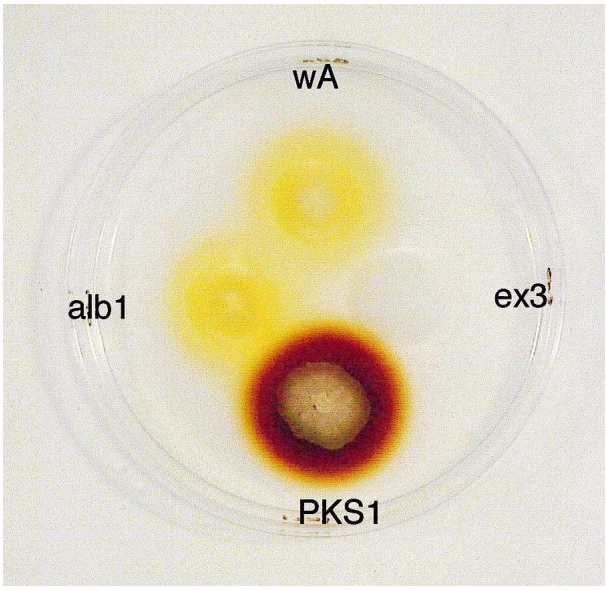


Figure I.23. Phenotype of *A. oryzae* transformants carrying heterologous T4HN and naphthopyrone synthase genes. Potato dextrose agar (PDA). wA, pTA-nwA transformant; alb1, pTA-alb1 transformant; PKS1, pTAPSG transformant with the *C. lagenarium* T4HN synthase gene; ex3, pTAex3 "blank" transformant. Yellow colonies produce YWA1, a brown colony produces DNH melanin. Photo taken from Watanabe *et al.*, 2000.

parasperone A (Figure I.20) from a laccase deficient strain of *A. parasiticus* and to other related pigments such as fonsecin from *Aspergillus fonsecaeus* (Figure I.24) (Galmarini *et al.*, 1962; Galmarini, Stodola, 1965) and rubrofusarin from *Fusarium culmorum* (Figure I.24) (Stout *et al.*, 1961; Tanaka, Tamura, 1961). Thus, naphthopyrone YWA1 was considered to be an actual intermediate of *A. nidulans* conidial spore pigments and the *wA* gene encodes a PKS for heptaketide naphthopyrone YWA1 biosynthesis.

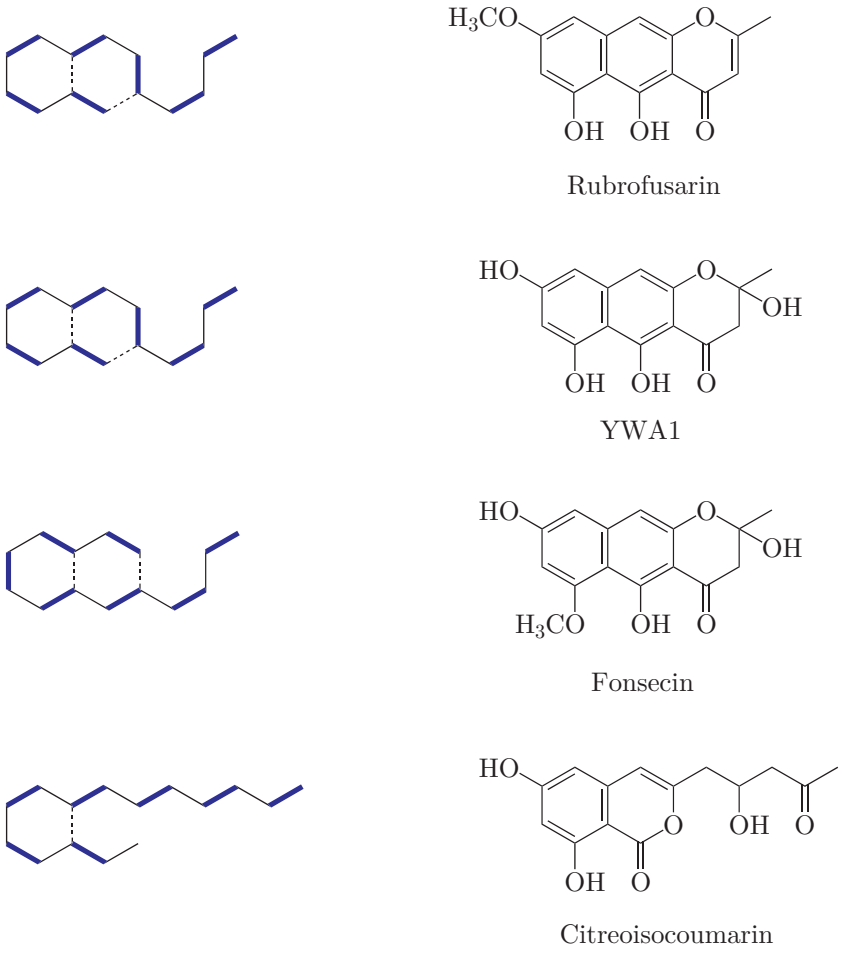


Figure I.24. Folding patterns of different heptaketides

Feeding experiments on fonsecin (Bloomer *et al.*, 1982) and rubrofusarin (Leeper, Staunton, 1984) biosynthesis using [$^{13}C_2$]acetate revealed polyketide chain folding patterns as shown on Figure I.24. Although the observed folding patterns were contradictory to each other, the same folding pattern as that of rubrofusarin was assumed for YWA1 (Fujii, 1999). Comparison between the folding patterns of citreoisocoumarin and YWA1 has given some understanding of the heptaketide chain cyclization. Cyclization and aromatization of the first ring occurs first, in the same manner for both compounds. The second aromatic ring in YWA1 arises from a Claisen-type condensation, but in the case of cit-

reoisocoumarin cyclization *via* Claisen-type mechanism is prohibited and exclusively leads to lactone ring formation (Figure I.24).

Another gene encoding a naphthopyrone synthase has been cloned from a pathogenic fungus *Aspergillus fumigatus*, which causes allergy, noninvasive colonization, or life-threatening invasive pulmonary aspergillosis. *A. fumigatus* produces its bluish-green conidial pigment through a pentaketide pathway similar to the DHN-melanin pathway in brown-to-black fungi (Figure I.25). Genetic and biochemical investigations have revealed that biosynthesis of the conidial pigment in *A. fumigatus* is governed by a six-gene cluster including the gene *alb*1 encoding PKS. Disruption of the *alb*1 gene resulted in the albino strain with altered conidial surface and low virulence to mice comparing to the wild-type. Thus, *alb*1 was considered to encode a T4HN synthase (Tsai *et al.*, 1998; Tsai *et al.*, 1999). However, the Alb1 PKS showed higher sequence similarity to a naphthopyrone synthase WA from *A. nidulans* rather than the *C. lagenarium* T4HN synthase. To clarify the function of Alb1, the *alb*1 gene was expressed in a heterologous host *A. oryzae* under control of α-amylase promoter in the fungal expression system pTAex3 (Figure II.2). The *A. oryzae* transformant, harboring the expression plasmid pTA-alb1, showed yellow pigmentation on agar plates (Figure I.23). The produced compound was identified as the heptaketide naphthopyrone YWA1 instead of pentaketide T4HN as predicted earlier (Watanabe *et al.*, 2000).

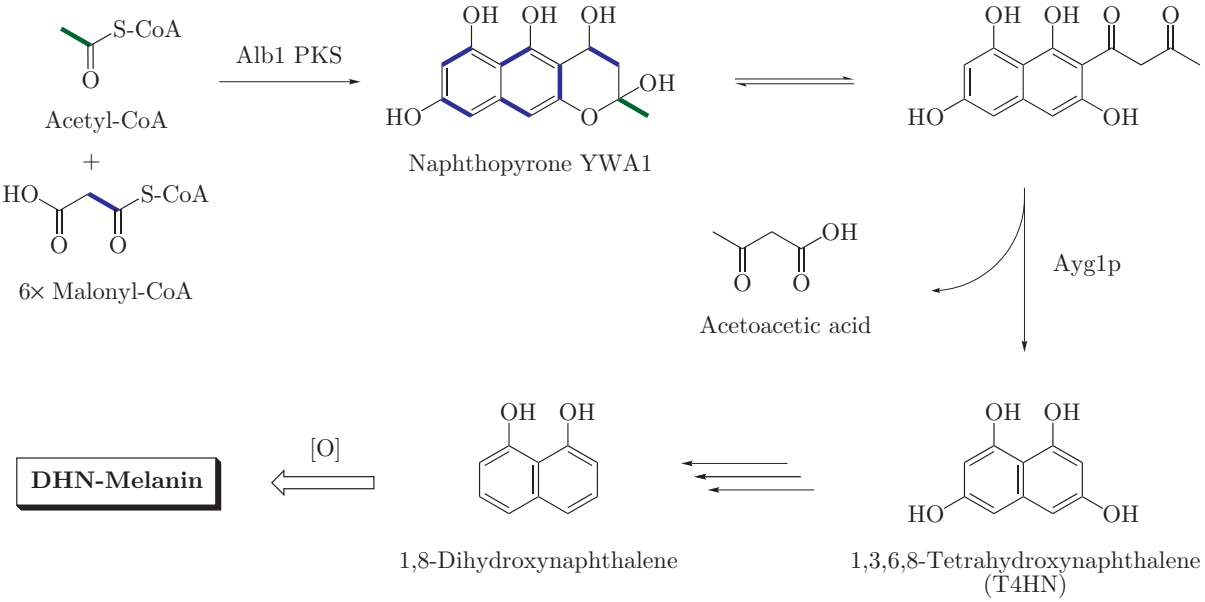


Figure I.25. Proposed biosynthetic pathway for DHN melanin from naphthopyrone YWA1 in *Aspergillus fumigatus*. The heptaketide naphthopyrone YWA1 serves as a precursor of the pentaketide T4HN, which may be converted to DHN-melanin similarly as brown-to-black fungi.

To understand how a heptaketide synthase was used to initiate the biosynthesis of DNH melanin in *A. fumigatus*, the possible involvement of accessory protein(s) in the biosynthetic pathway was explored. A novel protein, named Ayg1p, was shown to be involved in the formation of pentaketide T4HN by chain-length shortening of the heptaketide precursor YWA1 in *A. fumigatus* (Figure I.25). Phenotypic analysis of double gene disruptants showed that Ayg1p catalyzes a novel biosynthetic step downstream of Alb1 PKS. Subsequent genetic and biochemical analyses of the reconstructed strain carrying *alb*1, *ayg*1, or *alb*1+*ayg*1 genes revealed that Ayg1p is essential for biosynthesis of T4HN in addition to Alb1 PKS. Experiments with the cell-free crude Ayg1p protein extract indicated that Ayg1p enzymatically shortened the heptaketide product YWA1 of Alb1 PKS to give T4HN. Thus, it was established that the protein Ayg1p contributes to the participation of a heptaketide synthase Alb1 in a pentaketide pathway via a novel polyketide-shortening mechanism in *A. fumigatus* (Tsai *et al.*, 2001; Fujii *et al.*, 2004).

2.2.4. PKSs involved in aflatoxin biosynthesis

Aflatoxins are a group of mycotoxins produced by several *Aspergillus* species: *A. flavus*, *A. parasiticus*, *A. nomius*, *A. bombycis*, *A. ochraceoroseus*, and *A. pseudotamari*. Production of aflatoxins varies greatly in terms of quality and quantity among different strains within each aflatoxigenic species. The first species, *A. flavus*, is probably the most notable and severe aflatoxin producer.

Aflatoxigenic species of *Aspergillus* are common and widespread in nature. They can easy colonize and contaminate various host crops (cereals, nuts, *etc.*), silage and other feed whenever conditions are favorable for their growth. The toxin can also be found in the milk of animals consuming contaminated feed.

Aflatoxins are infamous for their toxicity, carcinogenic, mutagenic, teratogenic, and immunosuppressive activities. Liver and kidney are the most susceptible organs to the action of aflatoxins. The four major aflatoxins are called B_1, B_2, G_1, and G_2 (Figure I.26) based on their fluorescence under UV light (blue or green) and relative chromatographic mobility during thin-layer chromatography (TLC). Aflatoxin B_1 is the most powerful natural carcinogen known and is usually the major aflatoxin produced by toxigenic strains. It is also the best studied member of this group of compounds. Aflatoxins B_1 and B_2 are typically produced by *A. flavus*, whereas *A. parasiticus* produces aflatoxins G_1 and G_2 as well as B_1 and B_2. Four other aflatoxins M_1, M_2, B_{2a}, G_{2a} which may be produced in minor amounts were subsequently isolated from cultures of *A. flavus* and *A. parasiticus*. A number of closely related compounds, namely aflatoxin GM_1, GM_2, parasiticol and aflatoxicol are also produced by *A. flavus*. Aflatoxins M_1 and M_2 are major hydroxylated

biotransformation derivatives of aflatoxins B_1 and B_2 respectively, originally discovered in milk of livestock that had consumed feed contaminated with aflatoxins (Figure I.26) (Bennett, Klich, 2003; Williams *et al.*, 2004)

Sterigmatocystin (Figure I.27), a biosynthetic intermediate in aflatoxin biosynthesis, is also a toxic, mutagenic and carcinogenic compound but is less potent than aflatoxins. This toxin is synthesized as an end product by numerous ascomycetes, including *Aspergillus versicolor* and *A. nidulans* (Bennett, Klich, 2003).

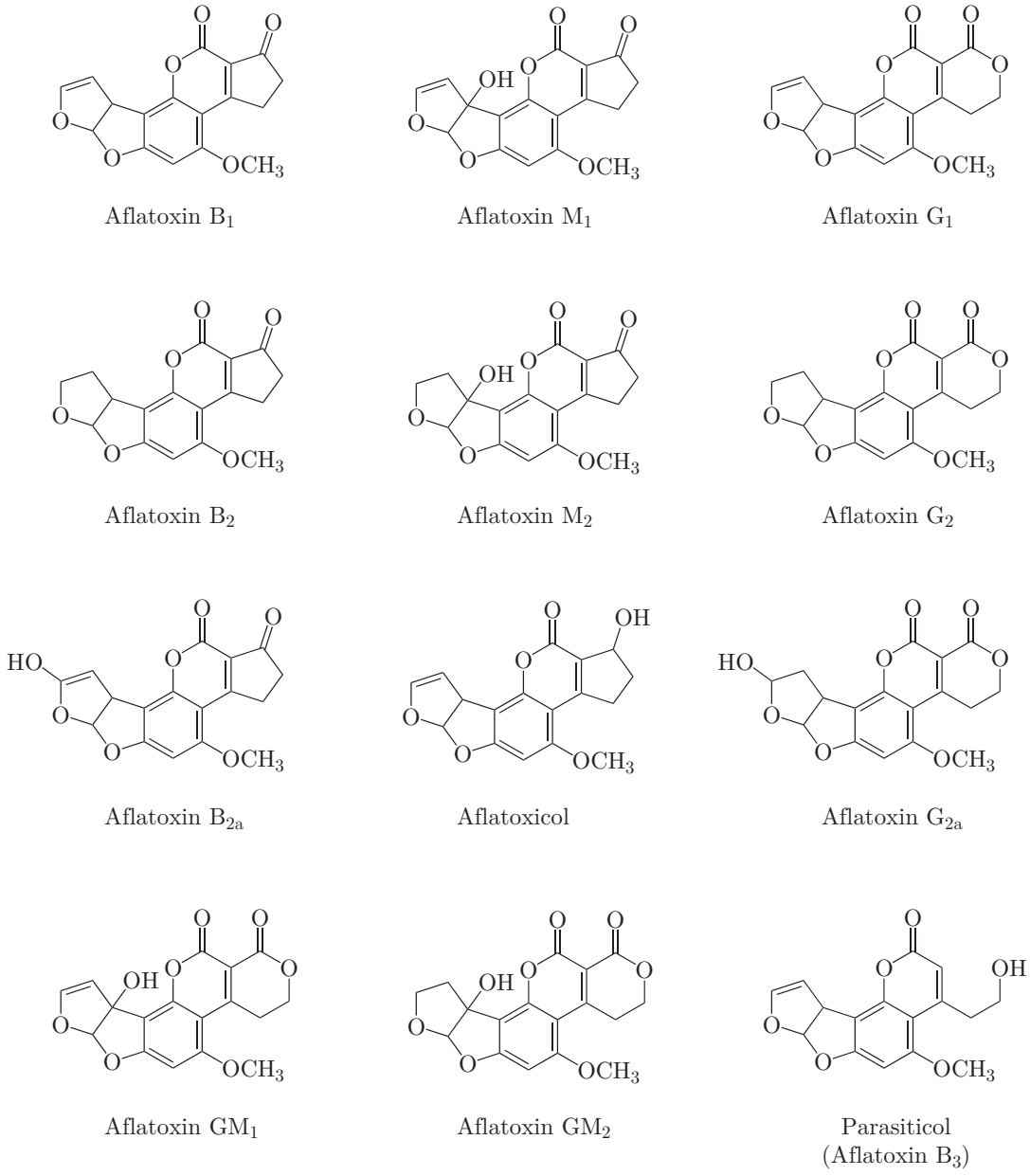


Figure I.26. Structures of selected aflatoxins and related compounds

Because of their notorious properties, biosynthesis of aflatoxins has been studied intensely for almost five decades, since their discovery in the early 1960s, by classical genetics including blocked mutant analysis, feeding experiments with labeled acetates, enzymology, and by molecular genetics.

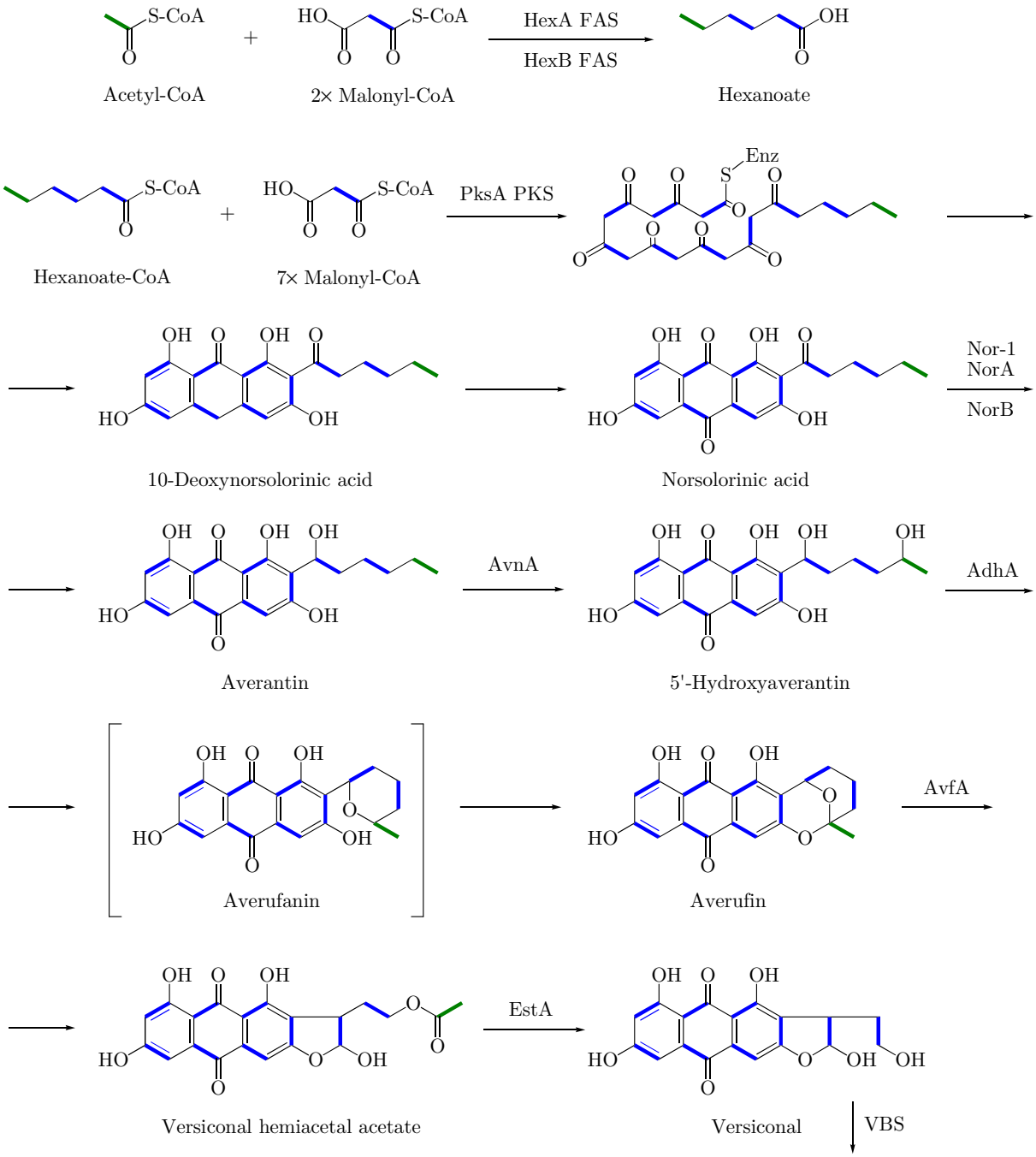


Figure I.27. Biosynthetic pathway of aflatoxins and sterigmatocystins. Abbreviated enzyme names are labeled next to the arrows. For their corresponding genes and functions see Table I.1 and Figure I.28.

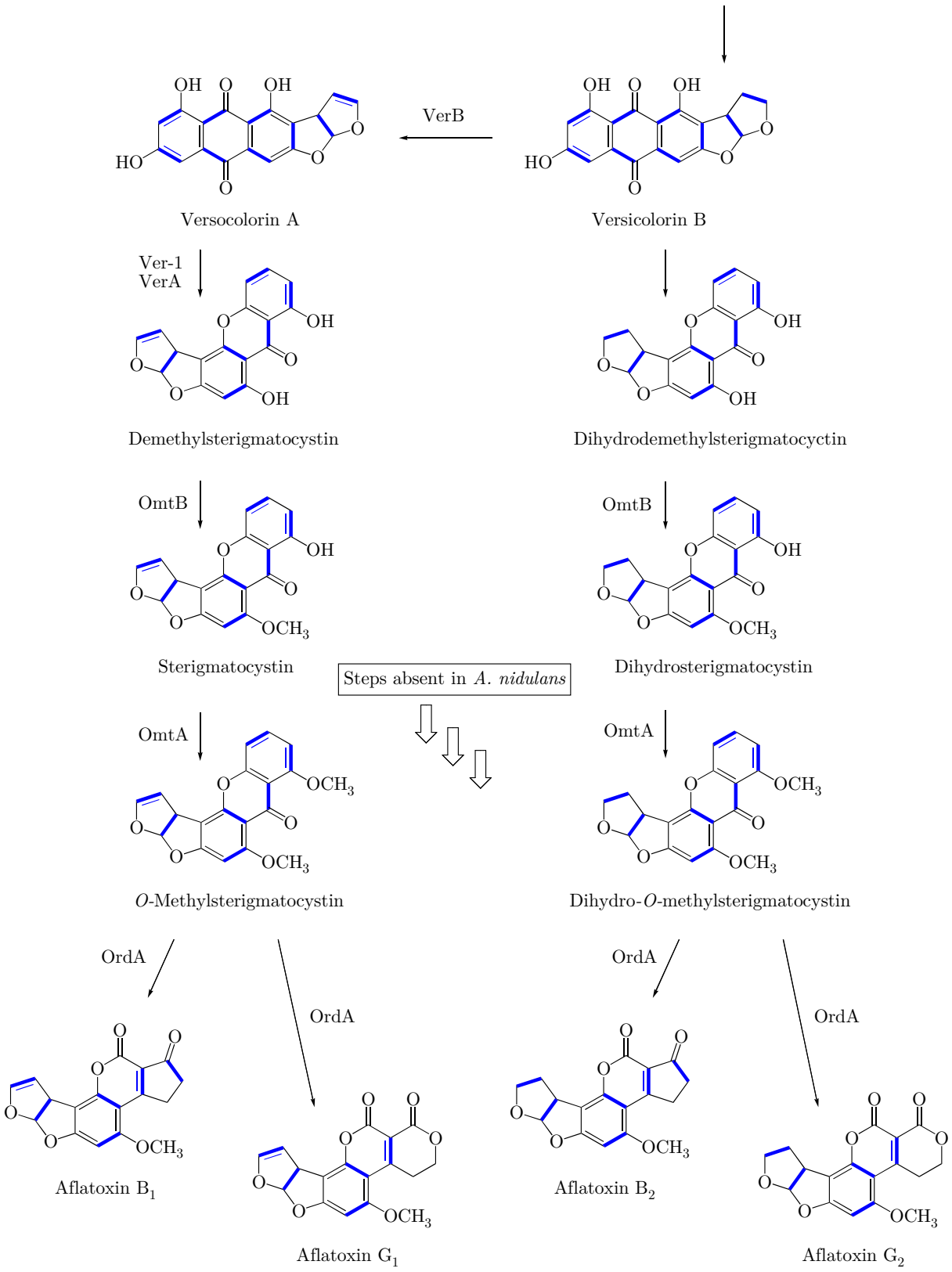


Figure I.27. *Continued*

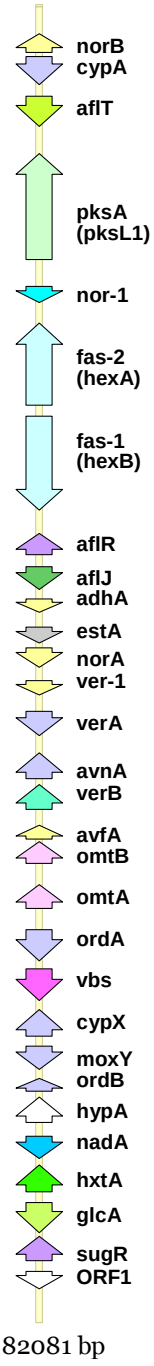


Figure I.28.
The *Aspergillus parasiticus* aflatoxin pathway gene cluster (*nor*B–*hyp*A) with an adjacent sugar utilization gene cluster (*nad*A–ORF1) (accession number AY371490, Yu *et al.*, 2004). For gene functions see Table I.1.

The biosynthetic pathway of aflatoxin (Figure I.27) is unusually long and rather complex. It is governed by a 70 kb aflatoxin gene cluster containing genes for DNA-binding proteins, regulatory proteins, cytochrome P450-type monooxygenases, dehydrogenases, methyltransferases, polyketide and fatty acid synthases and others (Figure I.28, Table I.1). Biosynthesis of aflatoxins and related sterigmatocystins starts with the formation of the heptaketide 10-deoxynorsolorinic acid (Bhatnagar *et al.*, 2003). This anthrone compound with a C_6 side chain is derived from a polyketide pathway and was initially considered to be decaketide. However, hexanoate was later identified to be the starter unit for 10-deoxynorsolorinic acid biosynthesis by feeding experiments using its *N*-acetylcysteamine thioester (Townsend *et al.*, 1984). Norsolorinic acid is the first stable metabolite that can be isolated in aflatoxin biosynthesis and is often referred to as its first polyketide precursor, although this is not strictly correct. No specific enzyme has been linked to the conversion of anthrone to norsolorinic acid and it is thought to involve a monooxygenase (possibly *cyp*A) and a dehydrogenase (possibly *nor*B) (Bhatnagar *et al.*, 2003). Norsolorinic acid then undergoes several enzymatic conversions until aflatoxins are formed (Figure I.27).

Cloning and sequence analysis of the aflatoxin biosynthetic gene clusters from aflatoxigenic fungi, such as *A. flavus*, *A. parasiticus*, and a related cluster for sterigmatocystin biosynthesis from *A. nidulans* revealed the presence of PKS genes and the special FAS genes responsible for the formation of the hexanoyl precursor which serves as a starter unit in anthrone biosynthesis (Brown *et al.*, 1996; Feng, Leonard, 1995; Mahanti *et al.*, 1996; Yu *et al.*, 2004a; Yu *et al.*, 2004b; Yu, Leonard, 1995). No direct confirmation of PKS product, e.g. by heterologous expression, and/or detection of PKS enzyme activity has yet been carried out, although disruption experiments with PKS genes resulted in mutants with abolished ability to produce aflatoxins or any of their biosynthetic intermediates. Homologous PKS genes *pks*A (*pks*L1) and *stc*A(*pks*ST) were cloned from *A. parasiticus* and *A. nidulans*, respectively.

Table I.1. Genes in the *Aspergillus parasiticus* aflatoxin biosynthetic gene cluster (accession number AY371490) and their proposed functions (adapted from Yu *et al.*, 2004). Genes from the adjacent sugar utilization gene cluster are not listed.

Gene	Original name and other names	Encoded enzyme	Proposed function in the pathway
*alf*A	*fas*-2, *hex*A	Fatty acid synthase (α-subunit)	Conversion of malonyl-CoA to hexaketide
*afl*B	*fas*-1, *hex*B, *uvm*8, *fas*-1A	Fatty acid synthase (β-subunit)	Reduction of hexaketide to hexanoyl-CoA
*afl*C	*pks*A, *pks*L1	Polyketide synthase	Conversion of hexanoyl-CoA to polyketide
*afl*D	*nor*-1	Reductase	Conversion of NOR to AVN
*afl*E	*nor*A	NOR reductase/dehydrogenase	Conversion of NOR to AVN
*afl*F	*nor*B, *aad*	Dehydrogenase	Conversion of NOR to AVN
*afl*G	*avn*A, *ord*-1	P450 monooxygenase	Conversion of AVN to HAVN
*afl*H	*adh*A	Alcohol dehydrogenase	Conversion of HAVN to AVF or AVNN
*afl*I	*avf*A, *ord*-2	Oxidase	Conversion of AVF to VHA
*afl*J	*est*A	Esterase	Conversion of VHA to VAL
*afl*K	*vbs*	VERB synthase	Conversion of VAL to VERB
*afl*L	*ver*B	Desaturase	Conversion VERB to VERA
*afl*M	*ver*-1	Dehydrogenase/ketoreductase	Conversion of VERA to DMST
*afl*N	*ver*A	Monooxygenase	Conversion of VERA to DMST

Continued on the next page

Table I.1. (*Continued*)

Gene	Original name and other names	Encoded enzyme	Proposed function in the pathway
*afl*O	*omt*B, *dmt*A	*O*-Methyltransferase B	Conversion of DMST to ST and DHDMST to DHST
*afl*P	*omt*A, *omt*-1	*O*-Methyltransferase A	Conversion of ST to OMST and DHST to DHOMST
*afl*Q	*ord*A, *ord*-1	Oxidoreductase/P450 monooxygenase	Conversion of OMST to AFB_1 and AFG_1 and DHOMST to AFB_2 and AFG_2
*afl*R	*afl*R, *apa*-2, *afl*-2	Transcriptional activator	Pathway regulation
*afl*S	*afl*J	Transcriptional enhancer	Pathway regulation
*afl*T	*afl*T	Potential toxin transporter	Unassigned
*afl*U	*cyp*A	P450 monooxygenase	Unassigned
*afl*V	*cyp*X	P450 monooxygenase	Unassigned
*afl*W	*mox*Y	Monooxygenase	Unassigned
*afl*X	*ord*B	Monooxygenase/oxidase	Unassigned
*afl*Y	*hyp*A	Hypothetical protein	Unassigned

Abbreviations: NOR, norsolorinic acid; AVN, averantin; HAVN, 5′-hydroxyaverantin; AVNN, averufanin; AVF, averufin; VHA, versiconal hemiacetal acetate; VAL, versiconal; VERB, versicolorin B; VERA, versicolorin A; DMST, demethylsterigmatocystin; DHDMST, dihydrodemethylsterigmatocystin; ST, sterigmatocystin; DHST, dihydrosterigmatocystin; OMST, *O*-methylsterigmatocystin; DHOMST, dihydro-*O*-methylsterigmatocystin; AFB_1, aflatoxin B_1; AFB_2, aflatoxin B_2; AFG_1, aflatoxin G_1; AFG_2, aflatoxin G_2.

The *A. parasiticus pks*L1 (*pks*A) gene was identified amongst previously isolated clones, positively correlated with aflatoxin production, by screening with probes amplified from *A. parasiticus* genomic DNA using degenerate primers. Gene disruption experiments resulted in *pksL*1 disruptants which lost both their ability to produce aflatoxins B_1, B_2, and G_1 and the ability to accumulate norsolorinic acid and all other intermediates of the aflatoxin biosynthetic pathway. A *pks*L1 DNA probe detected a 6.6-kb RNA transcript by Northern hybridization analysis. This transcript, associated with aflatoxin production, showed regulated expression that was influenced by growth phase, medium composition, and culture temperature. Both genomic DNA and cDNA of the *pks*L1 gene were sequenced. Sequencing revealed that the 7 021-bp *pks*L1 gene contained a 6 594 open reading frame intervened by 5 introns ranging from 51 to 66 bp. The encoded polypeptide PKSL1 of 2 109 amino acids contained four catalytic domains [KT, AT, ACP, TE (CYC)] (Figure I.29) and showed significant similarity with the *A. nidulans* WA PKS (Figure I.22). On the basis of these results, it was proposed that *pks*L1 encodes the PKS which synthesizes the backbone polyketide and initiates aflatoxin biosynthesis (Chang *et al.*, 1995; Feng, Leonard, 1995).

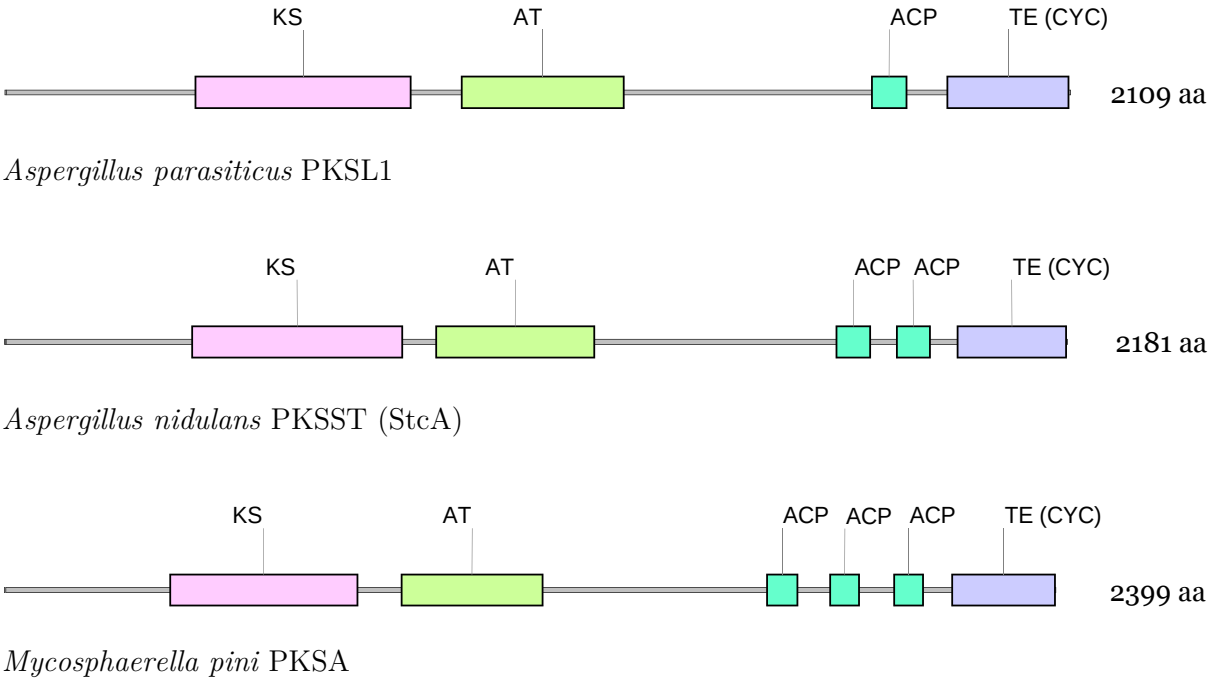


Figure I.29. Domain organization of PKSs involved in biosynthesis of aflatoxin, sterigmatocystin and dothiostromin, respectively: PKSL1 from *Aspergillus parasiticus* (accession number L42766; Feng, Leonard, 1995), PKSST (StcA) from *Aspergillus nidulans* (accession numbers AAA81586 and Q12397; Yu, Leonard, 1995), PKSA from *Mysosphaerella pini* (accession number AAZ95017; Bradshaw *et al.*, 2006). KS, β-ketoacyl synthase; AT, acyltransferase; ACP, acyl carrier protein; TE, thioesterase; CYC, Claisen-type cyclase; SAT and PT domains are not marked.

Townsend's group has recently established the function of the domain lying upstream of the KS domain in PKSL1 as a starter unit:ACP-transacylase (SAT) domain that selects an appropriate starter unit and loads it onto the ACP domain initiating biosynthesis (Crawford *et al.*, 2006). Another domain in PksL1 with unknown function, proposed by the same group and lying between the AT and ACP domains, is thought to be a *product template* (PT) domain, that may be responsible for determining the chain length of the polyketide product and may have a role in control of cyclization (Udwary, 2002; Cox, 2007).

Another PKS involved in biosynthesis of an anthrone precursor was identified in the *A. nidulans* sterigmatocystin pathway gene cluster. Large induced deletion mutants, which did not produce sterigmatocystin or any sterigmatocystin intermediates, were used to identify genes associated with its biosynthesis. Among the transcripts detected within the deletion region, which showed developmental expression with sterigmatocystin production, was a 7.2-kb transcript. Disruption of the gene encoding the 7.2-kb transcript resulted in strains unable to produce sterigmatocystin and all sterigmatocystin intermediates. Sequencing of the region spanning this 6 673-kb gene (*pks*ST) revealed the presence of a 6 546-bp open reading frame intervened by two introns and encoding a deduced PKS polypeptide (PKSST) of 2 181 amino acids with KS, AT, duplicated ACP, and TE (CYC) domains (Yu, Leonard, 1995).

It should be mentioned that *Aspergillus oryzae* and *Aspergillus sojae*, species that are widely used in Asian food fermentations such as soy sauce, miso, and sake, are closely related to the aflatoxigenic species *A. flavus* and *A. parasiticus* and belong to the same *A. flavus* group (Domsch *et al.*, 1980). Although these food fungi have never been shown to produce aflatoxin, they contain homologues of several aflatoxin biosynthesis pathway genes. Deletions and other genetic defects have led to silencing of the aflatoxin pathway in both *A. oryzae* and *A. sojae* and their inability to produce these mycotoxins (Machida *et al.*, 2005; Takahashi *et al.*, 2002; Tominaga *et al.*, 2006; Watson *et al.*, 1999).

Dothistromin, a polyketide toxin with structural similarity to the aflatoxin precursor versicolorin B, is produced by the forest pathogenic fungus *Mysosphaerella pini* (= *Dothistroma septosporum*), a causative agent of needle blight of pines. A PKS gene (*pks*A) involved in dothistromin biosynthesis was isolated from a *M. pini* genomic library using the *A. parasiticus* *pks*L1 gene as a probe. Inactivation of the *pks*A

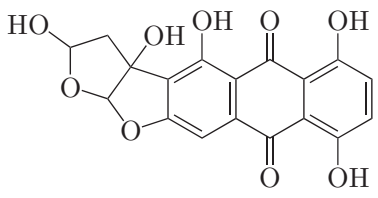


Figure I.30. Structure of dothistromin

gene by disruption resulted in mutants that could not produce dothistromin but that could convert exogenous aflatoxin precursors, including norsolorinic acid, into dothistromin. The mutants also exhibited reduced asexual sporulation compared to the wild type. The *M. pini pks*A gene of 7 304 bp contained a 7 200-bp open reading frame intervened by two introns of 49 and 55 bp and encoding a polypeptide of 2 399 amino acids, rich in alanine (10.55% overall). The predicted PKSA PKS contained one each of KS, AT and TE (CYC) and, interestingly, three ACP domains (Figure I.29) (Bradshaw *et al.*, 2002; Bradshaw *et al.*, 2006; Bradshaw, Zhang, 2006).

2.2.5. *N*-termini of fungal non-reducing PKSs

It was observed that in addition to overall similarity in their domain organization, the *N*-termini of fungal non-reducing PKSs exhibit some degree of homology to each other (Figure I.31). This region, lying upstream of a KS domain and spanning approximately 340 amino acids (∼38 kDa), is distinctive only to non-reducing PKSs and is not present in tetraketide synthases (Section 2.1.) and highly-reducing PKSs (Section 2.3.). *N*-Terminus deletion experiments on the *Aspergillus nidulans* YWA1 have shown that the truncation of as few as 5 amino acids from the *N*-terminus abolished the PKS activity (Fujii *et al.*, 2001). Primary sequence analysis of this region did not reveal any conserved identifiable motifs or homology to enzymes with known function. However, prediction of protein secondary structure and local hydrophobicity exhibited similarity. It was presumed that this region might be important for PKS tertiary or quaternary structure and possibly used as a multimerization domain as it was previously observed in animal FASs (Udwary *et al.*, 2002).

Probing aflatoxin biosynthesis in *Aspergillus parasiticus* it was shown that the norsolorinic acid synthase PKSA (PKSL1) exhibited remarkable preference for hexanoyl-CoA as a starter unit. Judging from these observations it was assumed that starter unit selectivity could be determined by a domain within PKSA. As mentioned, the *N*-terminus region of non-reducing PKSs did not show any sequence similarity to any functionally known enzyme, however, their structure had much in common with malonyl-CoA:ACP-transacylases (MATs) from bacteria. Thus, it was proposed that the *N*-terminus region of non-reducing fungal PKSs acts as a starter unit:ACP-transacylase (SAT) domain that selects an appropriate starter unit and loads it onto the ACP domain initiating biosynthesis (Crawford *et al.*, 2006).

To test whether the *N*-terminus region functions as a SAT domain in non-reducing PKSs a so called *deconstruction* approach was applied. The DNA fragments encoding SAT and ACP domains were amplified from cDNA of the *A. parasiticus* PKSA gene. The

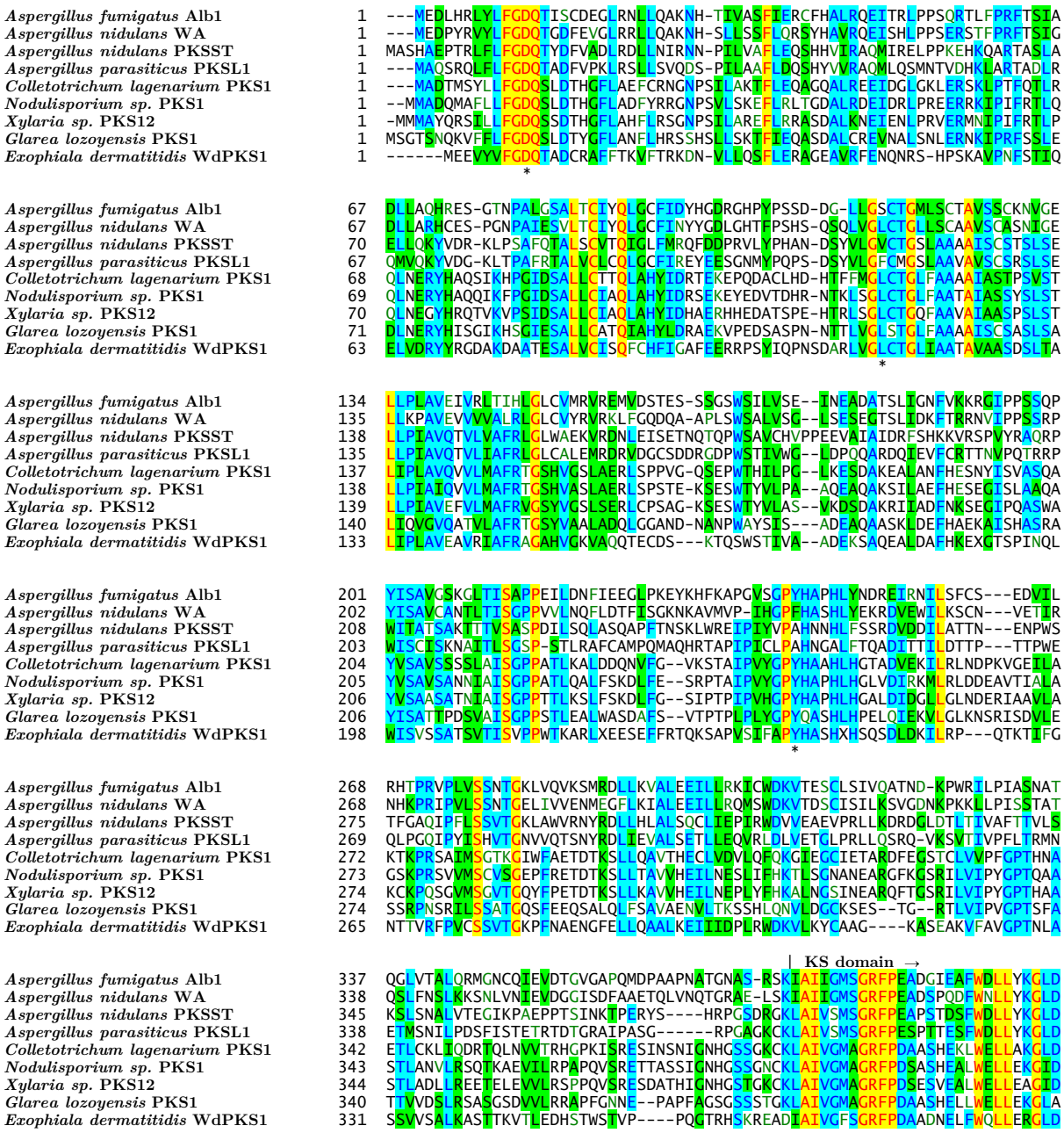


Figure I.31. Amino acid sequence alignment of *N*-terminus regions of selected fungal non-reducing PKSs: Alb1 from *Aspergillus fumigatus* (accession number AAC39471; Tsai *et al.*, 1998), WA from *Aspergillus nidulans* (accession number CAA46695; Watanabe *et al.*, 1999), PKSST from *A. nidulans* (accession number AAA81586; Yu, Leonard, 1995), PKSL1 (PKSA) from *A. parasiticus* (accession number L42766; Feng, Leonard, 1995), PKS1 from *Colletotrichum lagenarium* (accession number BAA18956; Takano *et al.*, 1995), PKS1 from *Nodulisporium sp.* (accession number AAD38786; Fulton *et al.*, 1999), PKS12 from *Xylaria sp.* (accession number AAM93545; Amnuaykanjanasin *et al.*, 2005), PKS1 from *Glarea lozoyensis* (accession number AAN59953; Zhang *et al.*, 2003), WdPKS1 from *Exophiala dermatitidis* (accession number AAD31436; Feng *et al.*, 2001). Putative key amino acids in the active sites are shown with an asterisk.

The alignment was obtained using AlignX, a component of Vector NTI Advance v.10.3 Sequence Analysis Software (Invitrogen). The amino acid residues in the alignment are color-coded according to the following scheme: red on yellow, completely conserved; blue on cyan, consensus derived from block of similar residues; green on white, residue weakly similar to consensus residue; black on white, non-similar to consensus residue; black on green, consensus derived from majority residue.

amplified fragments were cloned into an expression vector with a *N*-terminal His_6 tag and separately expressed in *E. coli*. To determine whether the SAT domain could catalyze the transfer of the hexanoyl starter unit to the ACP, the isolated ACP[4] and SAT domain proteins were probed with [1-^{14}C]hexanoyl-CoA. The experiment revealed that the SAT domain successfully bound and transferred the radioactive C_6 fatty acid unit onto the ACP domain. Subsequent probing of various starter units to determine the rates of their transfer showed that hexanoyl and octanoyl starter units were accepted preferentially, with hexanoyl having the fastest rate.

On the basis of experiments carried out with the *A. parasiticus* norsolorinic acid synthase, it was established the *N*-terminus region of this PKS acts as a hexanoate-CoA:ACP-transacylase domain that selects hexanoate starter units and transfer them onto the ACP domain, leading to the production of norsolorinic acid, a major precursor in aflatoxin biosynthesis. In addition to other metabolites, biosynthetically related to aflatoxin, such as sterigmatocystin and dothiostromin, which clearly exploit hexanoate as a starter unit, other PKS starter units are known. For example, in the biosynthesis of T4HN, where the starter unit is not obvious in the final product, malonyl-CoA has been shown *in vitro* to be both the starter and extender unit.

Based on crystal structure elucidation and mechanistic experiments, bacterial MATs have been shown to possess a conserved GH**S**XG motif with the active site serine (S), which covalently binds to malonyl-CoA from solution with participation of a conserved downstream histidine residue forming a catalytic dyad (Keatinge-Clay *et al.*, 2003; Serre *et al.*, 1995). The malonyl unit is then transferred to the phosphopantetheinylated ACP to prepare the enzyme for chain extension. Sequence alignment of the *N*-termini of fungal non-reducing PKSs clearly shows a cysteine (C) residue in a conserved GX**C**XG motif (Figure I.31). The *N*-termini of fungal non-reducing PKSs also contain a conserved histidine residue, which corresponds to the active site histidine in MATs supporting a transthioesterification mechanism through the formation of a cysteine-histidine dyad. Docking studies[5] with MAT and ACP from *Streptomyces coelicolor* suggested the involvement of a glutamine residue (Q) in creating an oxyanion hole for stabilizing acyl transfer (Keatinge-Clay *et al.*, 2003). This glutamine corresponds to a glutamine residue in a conserved FGD**Q**(S/T) motif found in the *N*-termini of the fungal non-reducing PKS presented on Figure I.31. This observation supports an analogous mechanism of association with the ACP domain.

[4]In a phosphopantetheinylated *holo*-form.

[5]Docking studies are computational techniques for the exploration of the possible binding modes of a substrate to a given receptor, enzyme or other binding site (McNaught, Wilkinson, 1997).

2.3. PKSs for highly-reduced polyketides

Fungal highly reducing PKSs produce reduced complex compounds such as T-toxins (Figure I.33), fumonisins (Figure I.36), lovastatin (Figure I.39) and squalestatin (Figure I.47). Polyketide chains of highly reduced PKS products vary greatly in their state of reduction and dehydration as well as in chain-length. All highly reducing PKSs studied to date are polypeptides of over 2 500 amino acids in length and therefore the largest iterative PKSs known. Fungal iterative type I PKSs of this class have a domain organization quite similar to that of mammalian FASs, possessing DH, ER, and KR domains in addition to KS, AT, and ACP domains (Cox, 2007; Fujii *et al.*, 2004). Those PKSs that do not have an ER domain contain a roughly equivalent length of sequence with no known function. In reactions catalyzed by FASs, the β-ketoacyl moiety formed by each round of condensation is converted to a fully saturated acyl chain by successive KR, DH and ER reactions. On the other hand, fungal highly reducing PKSs control the extent of reduction at each round of chain extension according to the length of the polyketide chain growing on the ACP domain resulting in a highly reduced polyketide backbone. A unique functional domain found in some highly reducing PKSs is the *C*-methyltransferase (*C*MeT) responsible for transfer of a methyl group from *S*-adenosyl-L-methionine to a polyketide chain. A *C*MeT domain is not present in FASs and other types of PKSs, except for the clade III of non-reducing PKSs (Kroken *et al.*, 2003). This domain was also found in some modular bacterial PKSs such as the epothilone PKS from the myxobacterium *Sorangium cellulosum* (Julien *et al.*, 2000; Molnár *et al.*, 2000; Tang *et al.*, 2000). A generalized domain architecture of fungal highly reducing PKSs is shown in Figure I.32.

As with the non-reducing and partially reducing fungal PKSs, many highly reducing PKS genes are known from the numerous sequenced fungal genomes, but only a limited number of gene sequences has been associated with the production of identified compounds. However, in the few cases where both a PKS and its chemical product are known, some progress has been made in understanding function and programming (Cox, 2007).

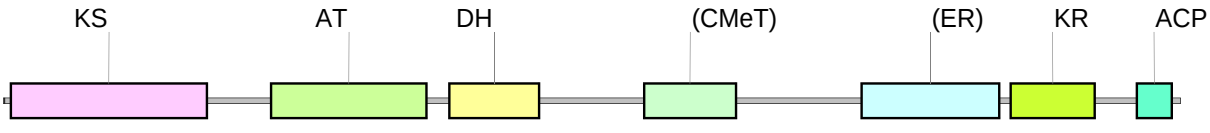


Figure I.32. Generalized domain organization of fungal highly reducing PKS. KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein. Optional domains are shown in brackets.

2.3.1. PKSs involved in the production of T-toxins

T-toxins and related PM-toxins (Figure I.33) are linear long-chain (C_{33} to C_{45}) polyketol compounds isolated from the ascomycete fungi *Cochliobolus heterostrophus*[6] (Kono, Daly, 1979; Kono *et al.,* 1980, 1981) and *Phyllosticta maydis* (Kono *et al.,* 1983; Danko *et al.,* 1984), respectively.

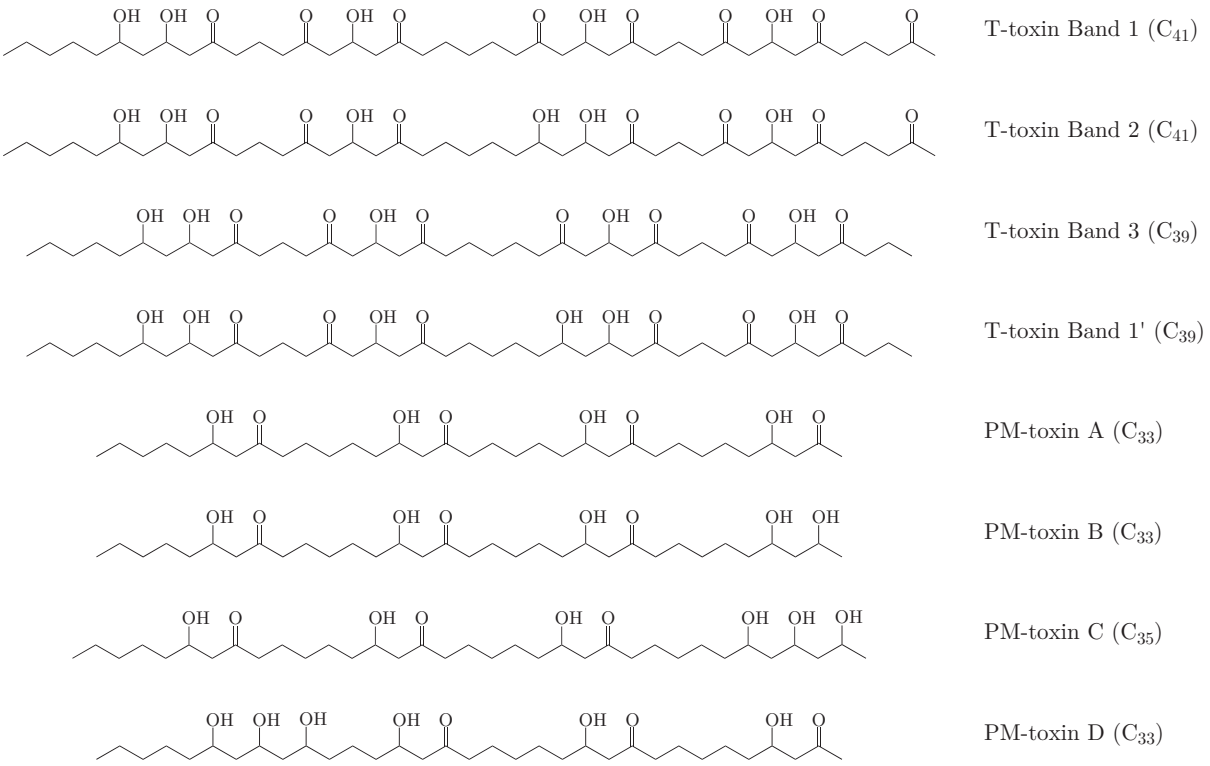


Figure I.33. Structure of T-toxins and PM-toxins. Only four major components for each toxin complex are presented. T-toxins are designated as bands according to their abundance, PM-toxins are designated as letter according to their mobility on TLC plates.

C. heterostrophus exists in two forms: race T, which produces T-toxins, and race O, which does not produce T-toxin. Both races are pathogenic to maize. However, race T is highly virulent to maize containing Texas male sterile (T) cytoplasm because this cytoplasm is uniquely sensitive to T-toxin, causing Southern corn leaf blight (Figure I.34). In contast, race O is weakly virulent on maize in general (Turgeon, Lu, 2000; Yoder, 1980; Yoder *et al.*, 1997).

Previous studies (Kodama *et al.*, 1999; Tzeng *et al.*, 1989, 1992) revealed that the genome of race T contains approximately 1.2 Mb of DNA not found in race O and that the *Tox*1 locus, which determines ability to produce T-toxin, maps to this genomic region. Conventional genetic methods, including Mendelian analysis, in which segregation

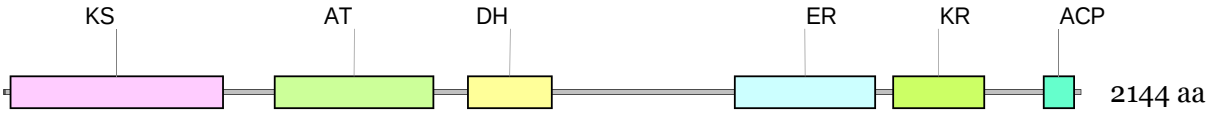
[6]Previously named *Helminthosporium maydis*; anamorph *Bipolaris maydis.*

Figure I.34. Maize plants (*Zea mays*) affected with Southern corn leaf blight caused by *Cochliobolus heterostrophus*. Photos courtesy by Prof. Jerald K. Pataky, Department of Crop Sciences, University of Illinois.

of pathogenicity types and host-specific toxin production were examined in progenies of crosses between naturally occurring race T and O isolates led to the prediction that *Tox*1 is a single locus (YODER, GRACEN 1975). However, molecular and genetic analyses of T-toxin-deficient race T mutants have shown that genetic control of T-toxin production maps at two loci (*Tox*1A and *Tox*1B) on two different race T chromosomes which are reciprocally translocated in race T with respect to their counterparts in race O (KODAMA *et al.*, 1999; TZENG *et al.*, 1989, 1992).

The *C. heterostrophus PKS*1 gene at *Tox*1A and the *DEC*1 gene encoding a decarboxylase at *Tox*1B (ROSE *et al.*, 2002) are necessary for T-toxin production. *PKS*1 was accessed by tagged mutations using the restriction enzyme-mediated integration (REMI) procedure resulting in a T-toxin negative phenotype and its locus flanking the tag was then identified and recovered (LU *et al.*, 1994; YANG *et al.*, 1996). Although there was evidence that additional genes were required for T-toxin production, efforts to clone them had been frustrated because the genes were located in highly repeated, A+T-rich DNA. To overcome this difficulty, ligation specificity-based expression analysis display (LEAD), a comparative amplified fragment length polymorphism/gel fractionation/capillary sequencing procedure, was applied to cDNAs from a near-isogenic pair of race T ($Tox1^+$) and race O ($Tox1^-$) strains. This led to discovery of *PKS*2, a second PKS-encoding gene that maps at *Tox*1A and is required for both T-toxin biosynthesis and high virulence to maize (BAKER *et al.*, 2006). Thus, the carbon chain of each T-toxin family member most likely is assembled by the action of two PKSs, which produce two polyketides, one of which may act as the starter unit for biosynthesis of the mature T-toxin molecule. Both PKSs are typical highly reducing PKSs, of which *C. heterostrophus* PKS1 also appears to possess an inactive *C*MeT domain (Figure I.35).

Cochliobolus heterostrophus PKS1



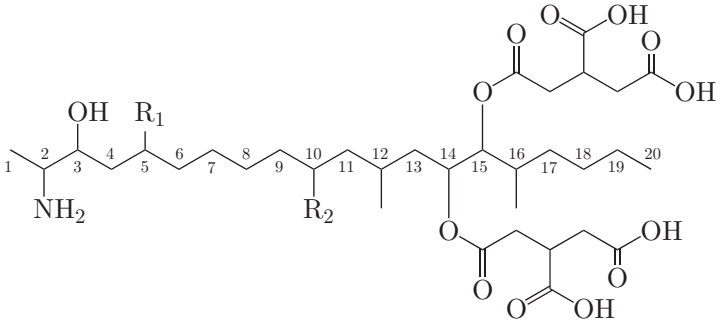
Cochliobolus heterostrophus PKS2

Figure I.35. Domain organization of PKSs involved in T-toxin biosynthesis in *Cochliobolus heterostrophus*: PKS1 (accession number AAB08104; YANG *et al.*, 1996) and PKS2 (accession number ABB76806; BAKER *et al.*, 2006). KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein. The inactive *C*MeT domain of PKS1 is not shown.

2.3.2. PKS involved in the production of fumonisins

Fumonisins are polyketide mycotoxins produced by the maize pathogen *Gibberella moniliformis* (= *G. fujikuroi* mating population A; anamorph *Fusarium verticilloides*, = *F. moniliforme*) and several related species from *G. fujikuroi* complex such as *F. proliferatum*, *F. nygamai* and *F. napiforme* (Nelson *et al.*, 1993). These compounds, similar in structure to the sphingolipid intermediate sphingamine, disrupt sphingolipid metabolism via inhibition of the enzyme sphinganine *N*-acyltransferase (Wang *et al.*, 1991). This disruption may account for some of the fumonisin-induced mycotoxicoses in human and animals such as human esophageal cancer, leucoencephalomalacia in horses, pulmonary edema in swine, and liver cancer in rats (Howard *et al.*, 1999; Marasas, 1996; Merrill *et al.*, 1997; Nelson *et al.*, 1993).

Although there are a number of different groups of fumonisins (Figure I.36), fumonisin B_1 is the most abundant in nature making up about 70% of the total fumonisin content. Fumonisin B_1 consists of a linear C_{20} backbone with an amine group at C-2, methyl groups at C-12 and C-16, tricarballylic ester groups at C-14 and C-15, and hydroxyl groups at C-3, C-5, and C-10. Feeding experiments have revealed that C-3 and C-20 of the fumonisin B_1 backbone are derived from acetate via the polyketide pathway (Blackwell *et al.*, 1994), the amine group and C-1 and C-2 are derived from alanine (Branham, Plattner, 1993), the methyl groups are derived from methionine via *S*-adenosyl-L-methionine (Platter, Shackelford, 1992), and the oxygen atoms attached directly to the backbone are derived from molecular oxygen (Caldas *et al.*, 1998). The tricarballylic moieties may be derived from glutamic acid via the citric acid cycle (Blackwell *et al.*, 1996).

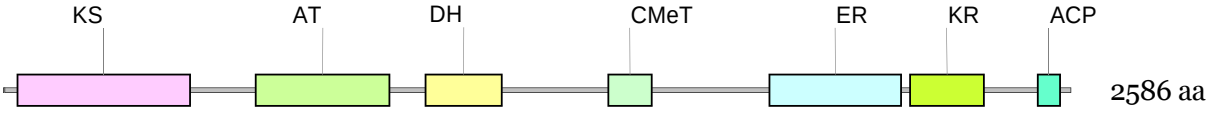


	R_1	R_2
Fumonisin B_1	OH	OH
Fumonisin B_2	OH	H
Fumonisin B_3	H	OH
Fumonisin B_4	H	H

Figure I.36. Structures of B-series fumonisins

The fumonisin biosynthetic genes of *G. moniliformis* are clustered (Proctor *et al.*, 2003; Seo *et al.*, 2001). The *FUM*1 (previously known as *FUM*5) gene encoding the PKS responsible for the production of the highly reduced polyketide backbone of fumonisin has been isolated from a *G. moniliformis* genomic cosmid library (Proctor *et al.*, 1999).

The library was screened with a PCR fragment, amplified from a cDNA template using KS degenerate primers (Table I.3) (Keller *et al.*, 1995), as a probe. Isolated genomic clones carried a 7.82 kb open reading frame interrupted by five introns. The deduced amino acid sequence of FUM1p contains six functional domains characteristic of highly reducing PKSs: KS→AT→DH→*C*MeT→ER→KR→ACP (Figure I.37).



Gibberella moniliformis FUM1p

Figure I.37. Domain organization of *Gibberella moniliformis* FUM1p involved in the biosynthesis of fumonisin (accession number AAD43562; Proctor *et al.*, 1999). KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein.

The function of the *G. moniliformis FUM*1 gene has been established by knockout experiments (Proctor *et al.*, 1999), but heterologous expression of this gene has not been reported yet. However, FUM1p has been the target for the first highly reducing PKS directed domain swap experiments. In initial work, Du and coworkers described the development of an approach to exchange regions of FUM1p with homologous regions of *C. heterostropus* PKS1. In effect, the KS domain from *C. heterostropus* PKS1 was inserted into FUM1p. This domain swap appeared to have no effect on either levels of production of fumonisin B_1, or on programmed events during its biosynthesis (Zhu *et al.*, 2006). Replacement of the whole of *FUM*1 by the *C. heterostropus* PKS1 gene in *G. moniliformis* resulted in no production. However, as it is now known that two PKSs are required for T-toxin production (Baker *et al.*, 2006), these results are not surprising.

Further work, however, in which the FUM1p KS domain was swapped with that of the *Aspergillus terreus* LDKS did result in significant reprogramming of the PKS (Zhu *et al.*, 2007). The new chimeric PKS produced two unusual dihydroisocoumarins (Figure I.38). These compounds appear to be the result of the condensation of two separate polyketide chains – in each case a tetraketide with a pentaketide. It was proposed that the LDKS and FUM1p KS domains are, at least partially, involved in the molecular interactions which help specify chain length and use of reductive domains.

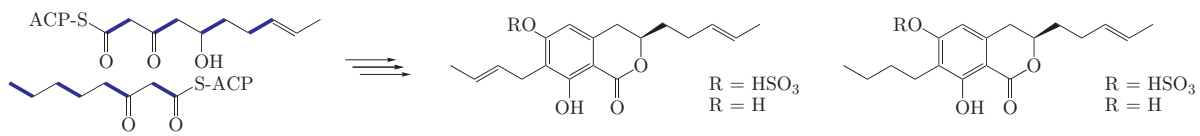


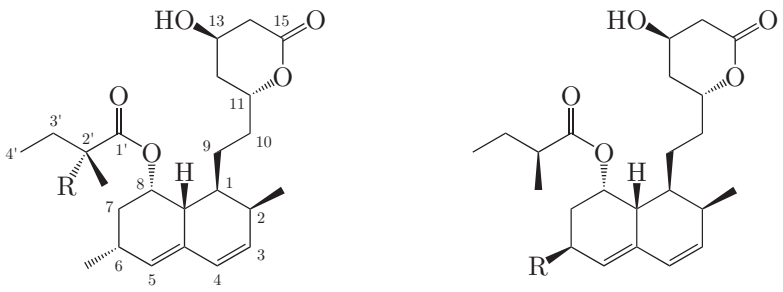
Figure I.38. New compounds produced by *Gibberella moniliformis* after the KS domain of FUM1p was replaced with the LDKS KS. Adapted from Zhu *et al.*, 2007.

2.3.3. Lovastatin and compactin synthases

Lovastatin (Figure I.39) is a secondary metabolite isolated from cultures of *Aspergillus terreus* (Moore *et al.*, 1985; Wagschal *et al.*, 1996). Lovastatin is also produced by *Monascus* species (Endo, 1979; Su *et al.*, 2003) and certain basidiomycete fungi such as *Pleurotus ostreatus* (oyster mushroom) and closely related *Pleurotus* species (Gunde-Cimerman *et al.*, 1993; Gunde-Cimerman, Cimerman, 1995). This compound, also known as mevinolin, monacolin K, and Mevacor™, is a potent anticholesteremic agent (Alberts *et al.*, 1980) and a precursor for synthesis of the widely-prescribed drug, simvastatin (Zocor™) (Figure I.39) (Hoffman *et al.*, 1986; Mauro, MacDonald, 1991).

A structural 6-desmethyl analogue of lovastatin, compactin (Figure I.39), also known as ML-236B or mevastatin and produced notably by *Penicillium citrinum* (Chakravarti, Sahai, 2004), can be converted by microbial oxidation to the 6-β-hydroxy derivative, pravastatin (Pravachol™) (Figure I.39) (Endo *et al.*, 2004; McTavish, Sorkin, 1991; Serizawa *et al.*, 1983).

Lovastatin and related compounds are known as potent inhibitors of (3S)-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase, the rate-limiting enzyme in the biosynthesis of cholesterol, thus exhibit strong cholesterol lowering activity and have been developed for clinical treatment of hypercholesterolemia (Alberts *et al.*, 1980; Endo, Hasumi, 1993; Igel *et al.*, 2001; Sutherland *et al.*, 2001).



R=H, Lovastatin
R=CH_3, Simvastatin

R=H, Compactin
R=OH, Pravastatin

Figure I.39. Structures of lovastatin and related compounds

This class of compounds is represented by reduced complex molecules consisting of a conjugated decene ring system and a (2R)-2-methylbutyryl side chain joined with an ester linkage (Figure I.39). Isotopic feeding experiments revealed that the decene ring core of lovastatin is a nonaketide that undergoes Diels-Alder cyclization to a hexahydronaphthalene ring system. The (2R)-2-methylbutyryl side chain is a simple diketide. The lovastatin molecule possesses two additional methyl groups at positions 6 and 2′, which are

derived from *S*-adenosyl-L-methionine (Chan *et al.*, 1983; Endo, 1979; Moore *et al.*, 1985; Wagschal *et al.*, 1996; Witter, Vederas, 1996; Yoshizawa *et al.*, 1994). The structure and origin of compactin are identical to the ones of lovastatin apart from the C-6 methyl group which is absent in compactin (Abe *et al.*, 2002; Wagschal *et al.*, 1996).

Extensive biosynthetic studies on lovastatin with fungi, such as *Aspergillus terreus*, showed that its assembly *via* a polyketide pathway involves initial formation of dihydromonacolin L (Figure I.40), the established first stable intermediate of lovastatin biosynthesis (Treiber *et al.*, 1989; Nakamura *et al.*, 1990). The cooperative activity of lovastatin nonaketide polyketide synthase (LNKS) and enoyl reductase encoded by the *lov*B and *lov*C genes, respectively (Figure I.42, Table I.2), carry out approximately 35 steps necessary to generate dihydromonacolin L (Hendrickson *et al.*, 1999; Kennedy *et al.*, 1999). The *lov*B gene encoding LNKS was identified from an *A. terreus* cDNA expression library screened with antibody raised against a ∼250-kDa polypeptide normally associated with lovastatin production. The cDNA clone was then used as a probe to isolate a genomic clone from an *A. terreus* genomic DNA library. The 11.6-kb LNKS open-reading frame contains seven short introns and encodes a deduced protein of 3 038 amino acids with a molecular weight 335 kDa (Hendrickson *et al.*, 1999). The LNKS consists of fundamental KS, AT, DH, and KR domains and also carries a *C*MeT domain in the central region and a domain at its *C*-terminus homologous to the peptide synthetase elongation domain of nonribosomal peptide synthase (see the following for detailed discussion) (Figure I.43).

Although a potential ER domain can also be observed in LNKS, close inspection of its sequence suggests that the NADPH binding pocket of the ER domain might be disfunctional and inactive, explaining the lack of observed ER activity. The *lov*C gene encoding enoyl reductase lies downstream from *lov*B gene in the *A. terreus* lovastatin biosynthetic gene cluster (Figure I.42, Table I.2). In the absence of the LovC protein, which imparts enoyl reductase activity to the complex, LNKS generates truncated pyrones, which was confirmed by heterologous expression of the *lov*B gene alone in *Aspergillus nidulans* (Kennedy *et al.*, 1999). Structures of isolated pyrones suggested that they are related to the expected nonaketide formed during normal biosynthesis, for example, the methylation has occurred at the same position, but reductions, specifically enoyl reduction, and later keto reductions, have not occurred resulting in premature termination of chain extension (Figure I.41). By contrast, the coexpression of *lov*B and *lov*C genes in *A. nidulans* resulted in production of the expected dihydromonacolin L.

Dihydromonacolin L then undergoes oxidative conversion with introduction of the second double bond into the decalin (decahydronaphthalene) system until monacolin J is

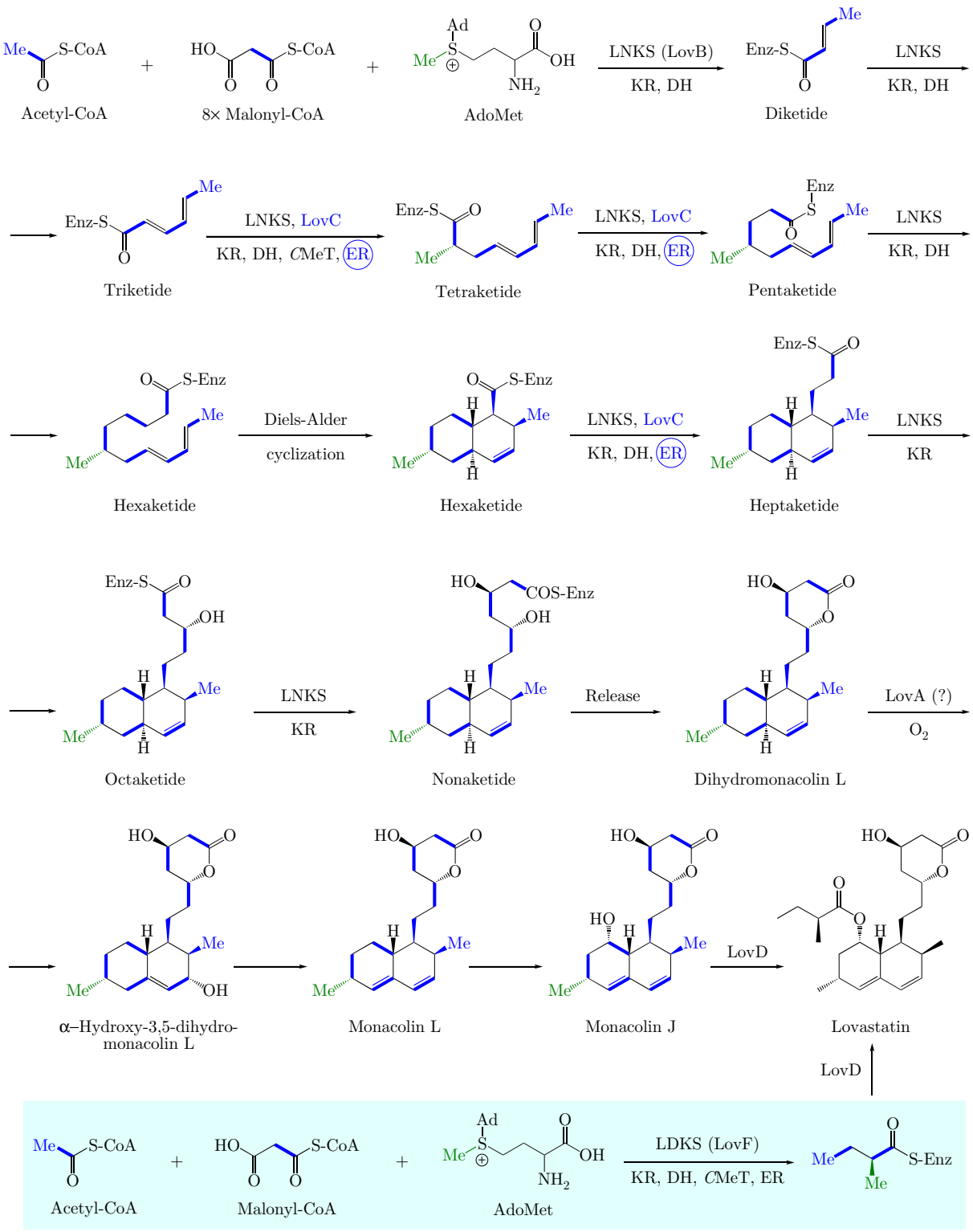


Figure I.40. Biosynthesis of lovastatin. The boxed region shows the biosynthesis of a 2-methylbutyryl side chain. AdoMet, *S*-adenosyl-L-methionine; LNKS (LovB), lovastatin nonaketide synthase; LDKS (LovF), lovastatin diketide synthase; DH, dehydratase; *C*MeT, *C*-methyltransferase; ER, enoyl reductase; KR, *β*-ketoreductase. The enoyl-reductase activity shown in blue and circled belongs to the LovC protein.

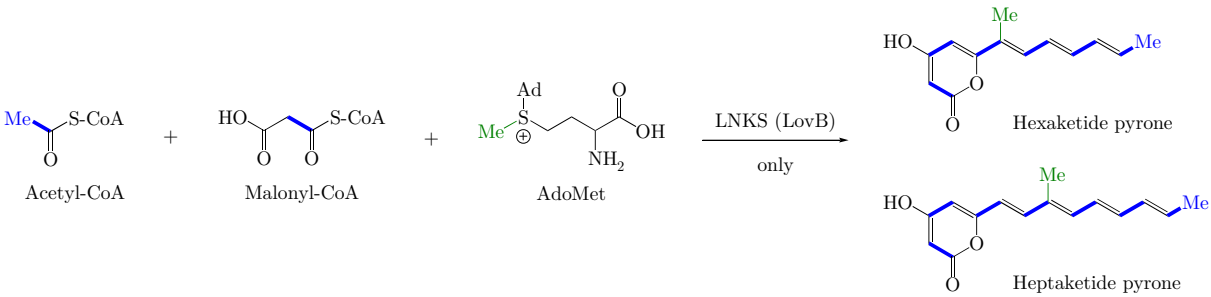
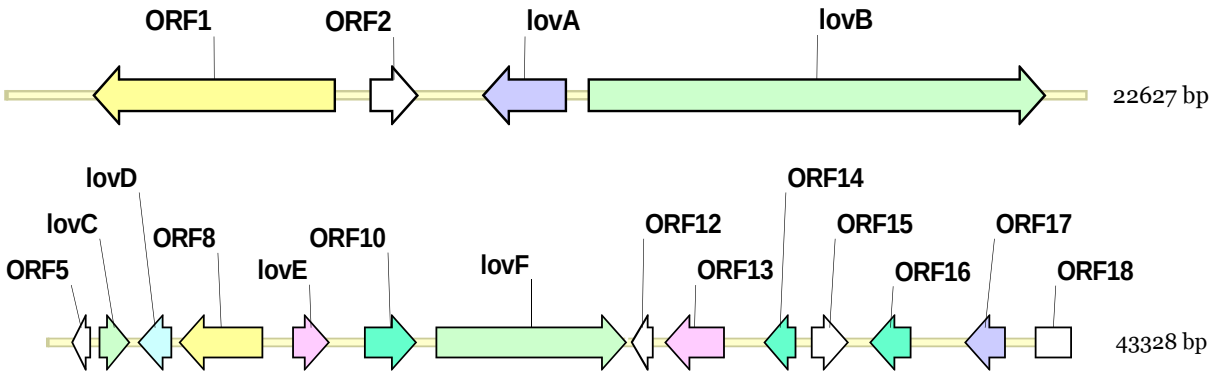
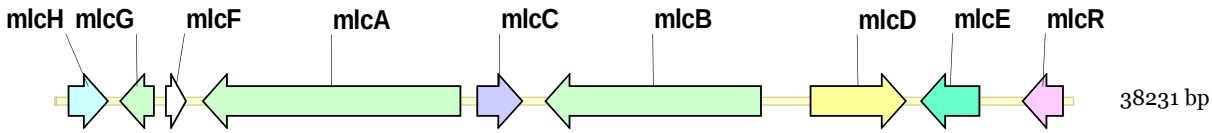


Figure I.41. Formation of pyrones catalyzed by LNKS alone in the absence of LovC.

formed. This process presumably involves the LovA protein (Figure I.42, Table I.2) that exhibits sequence homology to P450 monooxygenase and is clearly required for lovastatin production (Sorensen *et al.*, 2003).



Lovastatin biosynthetic genecluster from *Aspergillus terreus*



Compactin biosynthetic genecluster from *Penicillium citrinum*

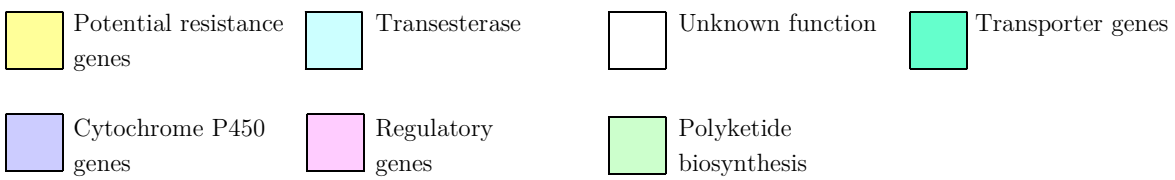


Figure I.42. Organization of lovastatin and compactin biosynthetic gene clusters: lovastatin biosynthetic gene cluster from *Aspergillus terreus* (accession numbers AF141924 and AF141925; Kennedy *et al.*, 1999), compactin biosynthetic gene cluster from *Penicillium citrinum* (accession number AB072893; Abe *et al.*, 2002).

Table I.2. Homologous genes in the lovastatin and compactin biosynthetic gene clusters (Abe *et al.*, 2002; Kennedy *et al.*, 1999).

Lovastatin biosynthetic gene	Compactin biosynthetic gene	Encoded putative enzyme/activity[a]	Protein sequence identity, %[b]	Protein sequence similarity, %
*lov*B	*mlc*A	Polyketide synthase	64	77
*lov*F	*mlc*B	Polyketide synthase	61	76
*lov*A	*mlc*C	P450 monooxygenase	73	86
ORF8	*mlc*D	HMG-CoA reductase	42	58
ORF10	*mlc*E	Efflux pump	70	81
ORF5	*mlc*F	Oxidoreductase	57	70
*lov*C	*mlc*G	Enoyl reductase	72	83
*lov*D	*mlc*H	Transesterase	75	87
*lov*E	*mlc*R	Transcription factor	39	51

[a]The assignment of putative activity to predicted polypeptides is based on their homology to proteins in the databases.
[b]The degrees of sequence identity and sequence similarity were obtained using BLAST 2 Sequences Software (Tatusova, Madden, 1999; *http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi*)

The (2*R*)-2-methylbutyryl side chain is produced by a separate PKS: the *lov*F gene encodes a 2532-amino acid protein with a molecular weight of 277 kDa designated lovastatin diketide synthase (LDKS). LDKS is closely related to LNKS and contains KS, AT, DH, *C*MeT, ER, KR and ACP domains, but unlike LNKS it does not have a PSED domain and its ER domain appears to be active (Figure I.43). The function of *lov*F was confirmed by gene disruption experiments which resulted in a mutant strain with abolished ability to produce lovastatin; instead, the strain accumulated a metabolite that was identified as monacolin J (Kennedy *et al.*, 1999). LDKS is unusual and exceptional among fungal PKSs with known products because it is not iterative - the enzyme catalyzes only a single round of chain extension resulting in the diketide. In this respect, LDKS closely resembles a single module of bacterial modular PKSs.

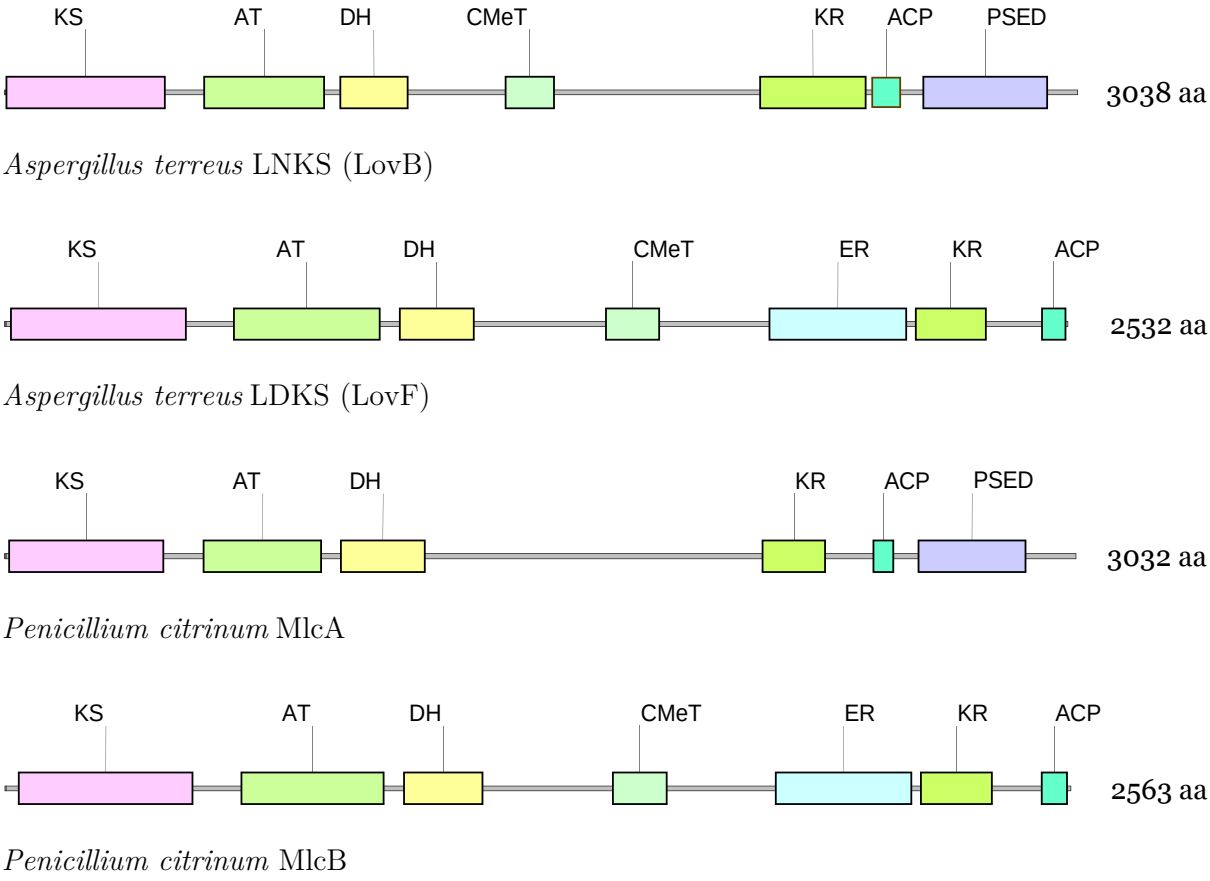


Figure I.43. Domain organization of PKSs involved in lovastatin and compactin biosynthesis: lovastatin nonaketide synthase LNKS (LovB) from *Aspergillus terreus* (accession number Q9Y8A5; Hendrickson *et al.*, 1999), lovastatin diketide synthase LDKS (LovF) from *A. terreus* (accession number AAD34559; Kennedy *et al.*, 1999), compactin nonaketide synthase MlcA from *Penicillium citrinum* (accession number BAC20564; Abe *et al.*, 2002), compactin diketide synthase MlcB from *P. citrimun* (accession number BAC20566; Abe *et al.*, 2002). KT, *β*-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; ER, enoyl reductase; KR, *β*-ketoreductase; ACP, acyl carrier protein; PSED, peptide synthetase elongation domain. Inactive domains are not shown.

The produced (2*R*)-2-methylbutyryl chain is then used by the LovD protein to directly acylate the C-8 hydroxyl of monacolin J and yield lovastatin. Thus, LovC acts as a specialized acyl transferase/transesterase that joins together the two polyketide components of lovastatin (Kennedy *et al.*, 1999). An enzyme with this activity has been partially purified from *A. terreus* and found by SDS-PAGE to have a molecular weight of 46 kDa, the same as that predicted for the LovD protein (413 amino acids). This enzyme was found to catalyze the transfer of the (2*R*)-2-methylbutyryl side chain from lovastatin to monacolin J, and neither methylbutyrate nor its CoA derivative could substitute lovastatin in this reaction. It was therefore assumed that LovD required as substrates (2*R*)-2-methylbutyrate bound to the LDKS protein and monacolin J. This idea was supported by the fact that LDKS lacks an obvious product release domain, such as a TE domain, and also implied that LovD docks with LDKS during a transesterification reaction (Kennedy *et al.*, 1999). Heterologous expression of the *A. terreus lov*D gene in *E. coli* resulted in a recombinant protein that transferred various acyl groups from CoA to the C-8 hydroxyl of monacolin J (Xie *et al.*, 2006).

The compactin (ML-236B) biosynthetic gene cluster was cloned from *Penicillium citrinum* (Figure I.42, Table I.2) (Abe *et al.*, 2002). Two genes from this cluster, *mlc*A and *mlc*B, encoded PKSs with deduced lengths of 3032 amino acids (molecular weight 334 kDa) and 2563 amino acids (molecular weight 280 kDa), respectively. Gene disruption of *mlc*A and *mlc*B resulted in mutant strains unable to produce compactin. A comparison between the genomic DNA and cDNA sequences revealed that the *mlc*A and *mlc*B genes contained ten and six introns, respectively. The MlcA and MlcB proteins showed high homology to LNKS and LDKS, respectively (64% and 61% sequence identity) (Table I.2). Both MlcA and MlcB contained conserved KS, AT, DH, KR, and ACP domains (Figure I.43). A PSED domain was only identified in MlcA and an ER domain was only identified in MlcB. The *C*MeT domains found in LNKS and LDKS were also identified in MlcA and MlcB, but the *C*MeT domain in MlcA contains critical mutations and therefore was considered inactive. This explains the fact that MlcA produces a nonmethylated polyketide – a 6-desmethyl analogue of the lovastatin nonaketide.

As mentioned earlier, an interesting feature of LNKS and MlcA nonaketide synthases is the presence of a condensation domain typical for non-ribosomal peptide synthases (NRPS) (von Dören, Kleinkauf, 1997; Zuber, Marahiel, 1997). These domains are located in the *C*-terminal regions of LNKS and MlcA, downstream of ACP domains, and sometimes referred to as PSED (Peptide-Synthase Elongation Domain)-like domains. In NRPSs such domains catalyze the peptide linkage formation between two amino acid residues attached as thioesters to adjacent thiolation domains. Condensation domains (C)

of NRPSs contain a highly conserved HHXXXDG motif with the active-site histidine (H) residue, which is considered to serve as a general base promoting nucleophilic attack of the amino group of the attaching amino acid to the carbonyl carbon of the growing polypeptide chain (Marahiel *et al.*, 1997). It has been shown that mutations in the HHXXXDG motif abolish activity suggesting this is part of the active site. This conserved motif can be observed in both LNKS and MlcA in their *C*-termini which also show overall similarity with the conserved condensation domain of NRPSs in sequence and in comparable length, ∼450 amino acids (Figure I.44).

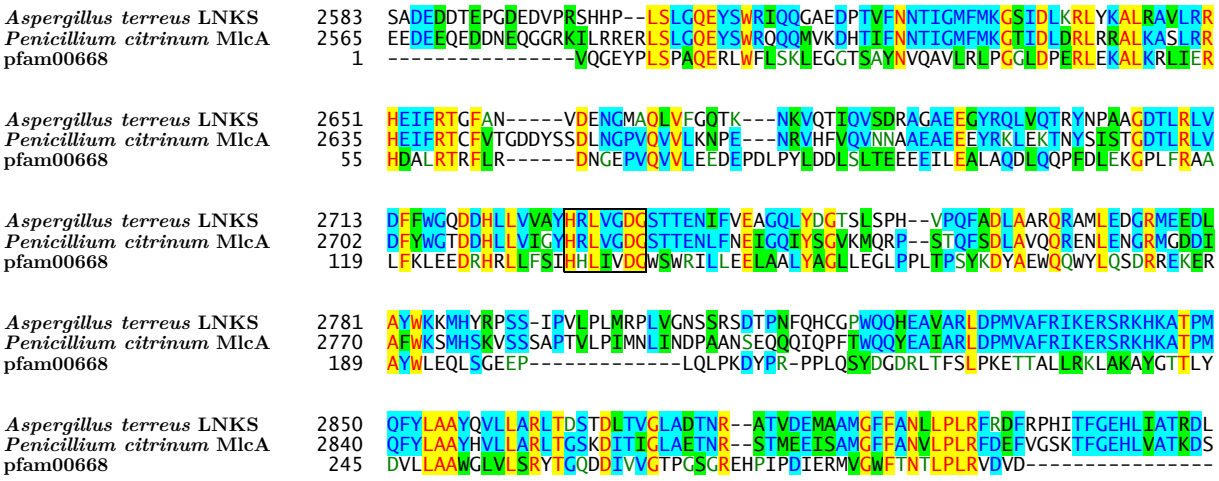


Figure I.44. Amino acid sequence alignment of the *Aspergillus terreus* LNKS and *Penicillium citrinum* MlcA PSED-like domains with a consensus for the conserved condensation domain (pfam00668). A conserved active site motif (HHXXXDG) is shown boxed.

The alignment was obtained using AlignX, a component of Vector NTI Advance v.10.3 Sequence Analysis Software (Invitrogen). The amino acid residues in the alignment are color-coded according to the following scheme: red on yellow, completely conserved; blue on cyan, consensus derived from block of similar residues; green on white, residue weakly similar to consensus residue; black on white, non-similar to consensus residue; black on green, consensus derived from majority residue.

The PSED domain was also found to be closely related with domains in chloramphenicol acetyltransferase and dihydrolipoyl transacetylase (de Crécy-Lagard *et al.*, 1995). These enzymes possess a conserved HXXXDG motif and catalyze the transfer of an acetyl group onto a hydroxy or a thiol moiety. Although the role of a *C*-terminal PSED-like domain in LNKS and MlcA still remains unclear, a similar role in esterification has been proposed (Fujii *et al.*, 2004). The transesterification reaction between diketide and nonaketide moieties is catalyzed by a separate enzyme LovD/MlcH. It was therefore assumed that the probable function of the domain could be as a lactonizing thioesterase releasing dihydromonacolin L and thus acting as an alternative to the TE domain common to type I PKSs and NRPSs, presumably by either activating water as a nucleophile or the C-11 hydroxyl (Figure I.45) (rather than the nitrogen of an aminothiolester as activated

by most C domains of NRPSs (Cox, 2007). This hypothesis was also supported by the assumed function of RapP, a specialized multi-domain protein with similarity to PSED involved in the formation of the ester and amide bonds to pipecoolic acid in rapamycin biosynthesis (Molnár *et al.*, 1996).

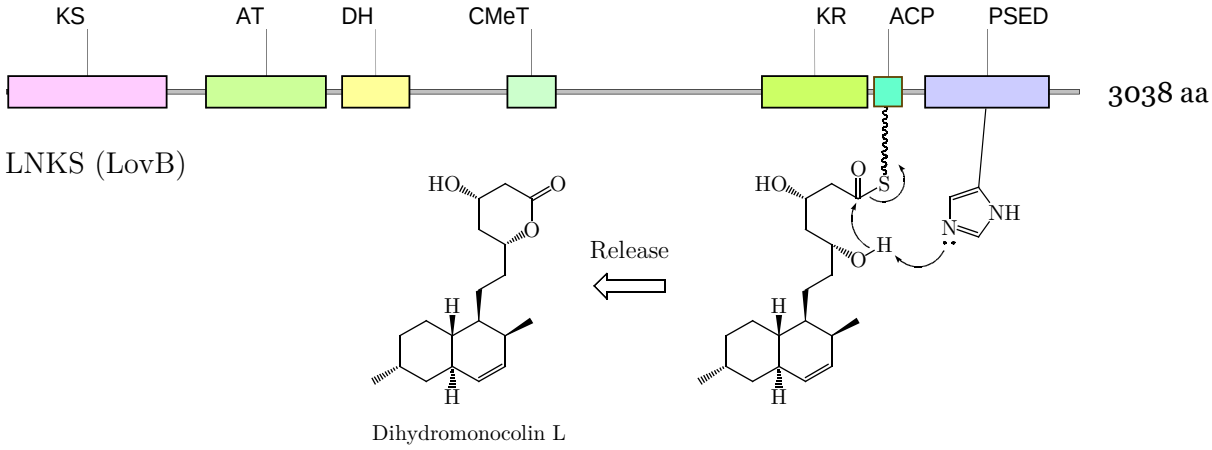


Figure I.45. Proposed role of a PSED-like domain of the *Aspergillus terreus* LNKS in lovastatin biosynthesis (adapted from Fujii *et al.*, 2004).

2.3.4. PKS involved in squalestatin S1 biosynthesis

Squalestatins, or zaragozic acids, are a group of fungal secondary metabolites known as potent and selective inhibitors of squalene synthase (Figure I.46) and therefore display significant potential for the treatment of diseases associated with elevated serum cholesterol concentrations (Bergstrom *et al.*, 1995). Squalestatin S1 was also shown to possess unprecedented curative properties towards prion infected neurons (Bate *et al.*, 2004).

The first isolation of squalestatins was reported in 1992 by scientists at Glaxo from a previously unknown ascomycete *Phoma sp.* C2932 (Baxter *et al.*, 1992; Dawson *et al.*, 1992; Sidebottom *et al.*, 1992). Almost simultaneously a group at Merck published the isolation of squalestatins under the name "zaragozic acids" from other ascomycete fungi *Sporormiella intermedia*, *Leptodontium elatius*, and *Amauroascus niger* (Bergstrom *et al.*, 1993; Dufresne *et al.*, 1992; Dufresne *et al.*, 1993; Hensens *et al.*, 1993; Wilson *et al.*, 1992). Screening programmes for squalene synthase inhibitors of the squalestatin family across major groups of fungi uncovered squalestatins in a variety of ascomycete fungi (Bergstrom *et al.*, 1995; Bills *et al.*, 1994; Dabrah *et al.*, 1997; Hosoya *et al.*, 1997; Tanimoto *et al.*, 1997).

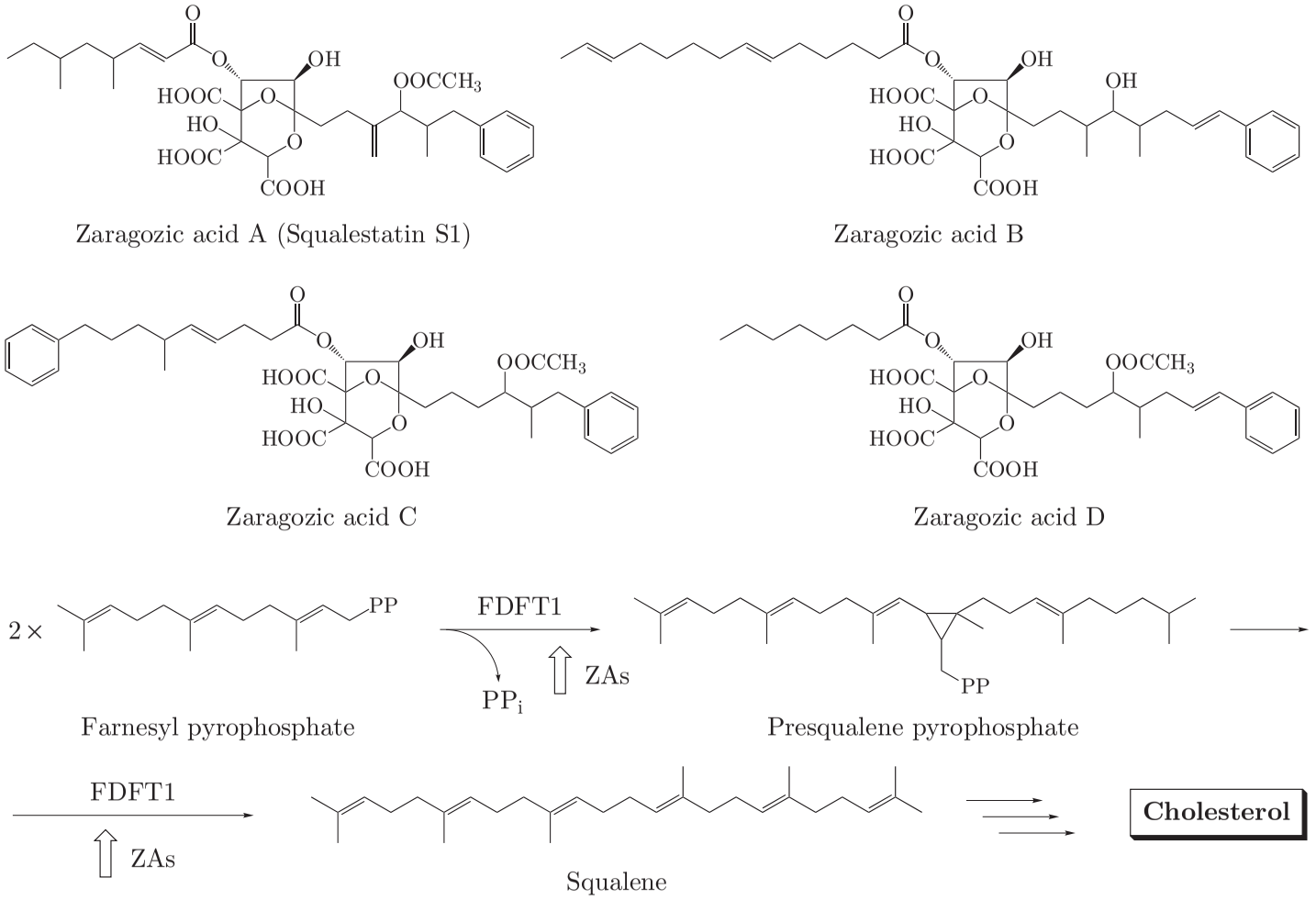


Figure I.46. Structures of selected zaragozic acids and proposed mechanism of squalene synthase inhibition. Squalene synthase (farnesyl-diphosphate farnesyltransferase, FDFT1) catalyzes a two-step reaction. Farnesyl pyrophosphate (F-PP) is converted to presqualene pyrophosphate (pS-PP) and then to squalene. The ZAs inhibit FDFT by effectively mimicking the binding of pS-PP to the enzyme. The ZAs are therefore competitive inhibitors of FDFT with respect to both F-PP and pS-PP as its substrates.

Structure elucidation of squalestatins revealed a characteristic 4,6,7-trihydroxy-2,8-dioxa-bicyclo[3.2.1]octane-3,4,5-tricarboxylic acid core with various side chains, termed the alkyl and the acyl side chains (Figure I.46). The biosynthetic origin of an archetypal member, squalestatin S1 (zaragozic acid A), has been established using stable and radioactive isotope feeding experiments (Figure I.47). These experiments revealed that the main alkyl chain of squalestatin S1 is a hexaketide, composed of a benzoate derived starter unit and five acetate-derived extender units, presumably in the activated form of malonyl-CoA. The acyl side chain is a tetraketide, with acetate as the starter unit. Four more carbons are derived from succinate, forming the distinctive 4,6,7-trihydroxy-2,8-dioxa-bicyclo[3.2.1]octane- 3,4,5-tricarboxylic acid core. The four branched-chain methyl/methylene groups on the alkyl and acyl side chains of squalestatin S1 are all derived from the methyl group of *S*-adenosyl-L-methionine (AdoMet) (Byrne *et al.*, 1993; Jones *et al.*, 1992).

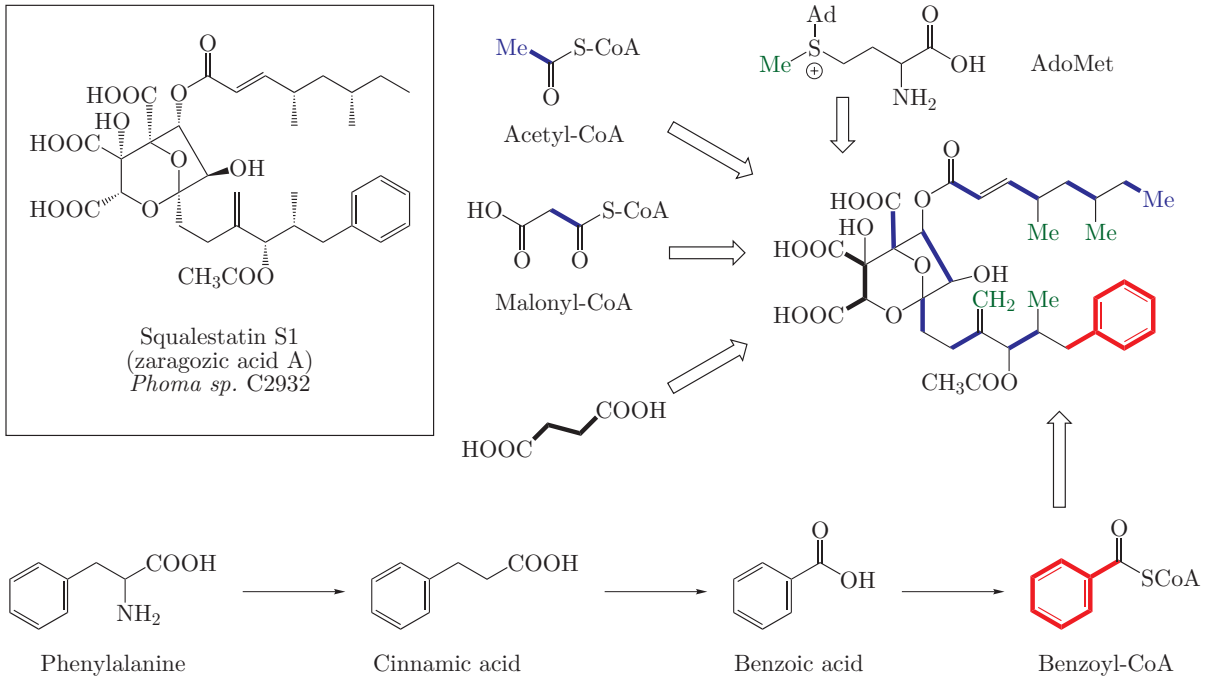


Figure I.47. Structure of squalestatin S1 and its biosynthetic origin. Me, methyl group –CH_3; Ad, adenosine; AdoMet, *S*-adenosyl-L-methionine.

The PKS gene involved in squalestatin S1 biosynthesis was accessed by screening the *Phoma sp.* C2932 cDNA library prepared from RNA isolated from cultures with increasing production of squalestatin S1 (Cox *et al.*, 2004). The probe used for screening was amplified from *Phoma sp.* cDNA using the degenerate primers for a *C*MeT domain (Nicholson *et al.*, 2001). Sequence analysis of the isolated clone revealed the presence of a 4.8-kb fragment of a Type I PKS gene, containing the polyadenylated 3′ end

of the sequence. The clone contained (5′→3′) a truncated fragment of a DH domain and complete *C*MeT, ER, KR and ACP domains. The remaining 5′ portion of the PKS gene was then obtained using rapid amplification of cDNA ends (RACE) to produce an overlapping clone, again of 4.8 kb containing KS, AT, DH and *C*MeT domains. A contiguous, full-length clone was then constructed by digestion and religation of the two cDNA fragments. The reconstructed PKS coding sequence was subcloned into the fungal expression vector pTAex3 (Figure II.2) using Gateway technology. The resulting plasmid was transformed into the heterologous fungal host *Aspergillus oryzae* M-2-3 (Figure I.64). The transformants obtained were capable of producing the tetraketide side chain of squalestatin S1 in detectable concentrations, establishing involvement of the *Phoma sp.* PKS1 in squalestatin S1 biosynthesis and indicating that all its domains are active under these conditions. The *Phoma sp.* PKS1 was therefore designated squalestatin tetraketide synthase (SQTKS). SQTKS exhibits a high degree of homology to the *Aspergillus terreus* LDKS (45% sequence identity, 63% sequence similarity). By analogy to LDKS, it possesses a functional ER domain, but SQTKS carries out three extensions in contrast to LDKS catalyzing only one cycle. Like LDKS, all modifying reactions – ketoreduction, dehydration, enoyl reduction and methylation – occur after the first round of extension. In both cases the stereochemistry of the branching methyl groups is the same. This is followed by two further rounds of extension and all modification reactions occur again after the first of them, but the ER and *C*MeT domains are not active after the final round of extension. SQTKS thus represents an example of a fungal PKS with so called *programming* of polyketide biosynthesis.

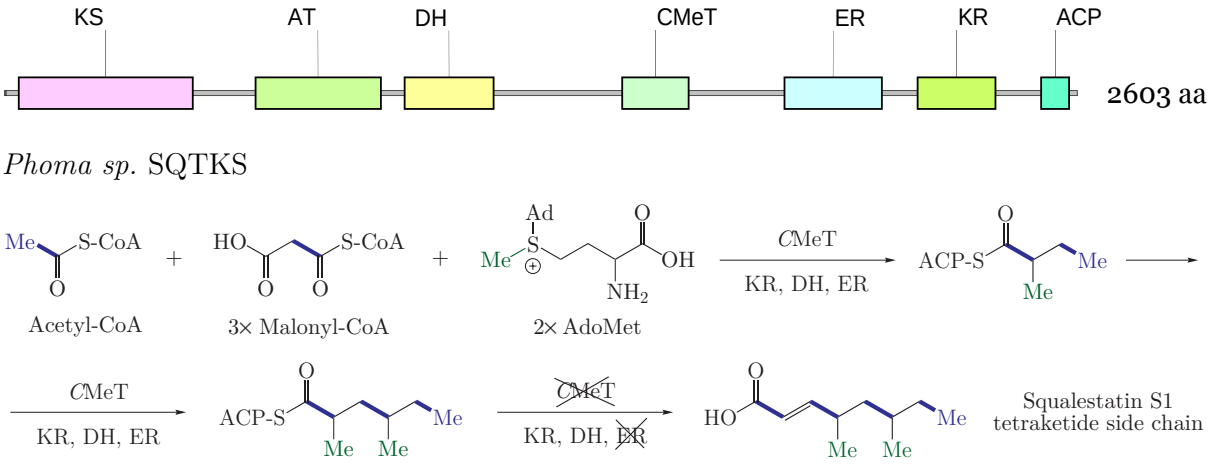


Figure I.48. Domain organization of *Phoma sp.* SQTKS (accession number AAO62426; Cox *et al.*, 2004) and proposed mechanism of squalestatin tetraketide biosynthesis catalyzed by this enzyme. KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein; Me, methyl group $-CH_3$; Ad, adenosine; AdoMet, *S*-adenosyl-L-methionine.

2.3.5. Highly reducing PKSs from *Alternaria solani*

Alternaria solani is a phytopathogenic fungus causing early blight disease of *Solanum* species, including potato and tomato. It produces various polyketide metabolites (Figure I.49), both aromatic and complex reduced-type compounds, including the phytotoxins alternaric acid (Brian *et al.*, 1949; Brian *et al.*, 1951; Brian *et al.*, 1952; Barteles-Keith, 1960) and solanopyrones (Ichihara *et al.*, 1983; Ichihara *et al.*, 1985; Oikawa *et al.*, 1989b). Feeding experiments with labeled acetate revealed the polyketide origin of these compounds (Oikawa *et al.*, 1989a; Ichihara, Oikawa, 1997; Oikawa *et al.*, 1998). Alternaric acid was shown to be derived from two polyketide chains via C–C bond formation rather than from a single chain (Ichihara, Oikawa, 1997). Biosynthetic studies on the solanapyrones strongly suggested that a enzymological Diels-Alder reaction is involved in formation of their decalin skeleton from a precursor octaketide α-pyrone, prosolanapyrone (Oikawa *et al.*, 1995; Oikawa *et al.*, 1998; Ichihara, Oikawa, 1999).

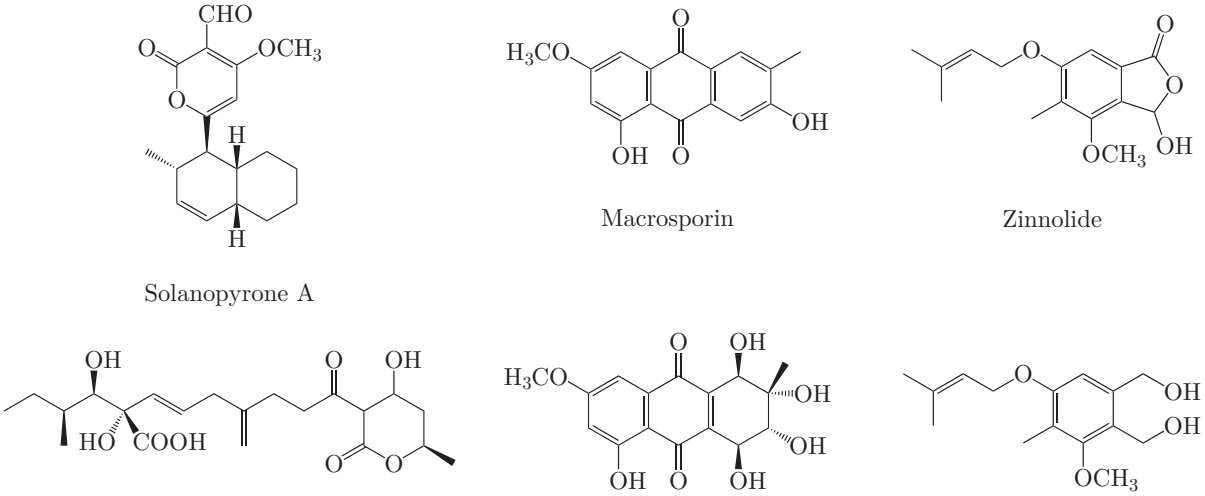


Figure I.49. Structures of polyketide metabolites isolated from *Alternaria solani*

To probe interesting features of polyketide biosynthesis in *A. solani*, Ebizuka's group carried out cloning of the highly reducing PKS genes from the solanapyrone producing strain of *A. solani*. A biosynthetic gene cluster containing five genes, *alt*1-5, was isolated from an *A. solani* genomic library using the fragment, amplified from *A. solani* genomic DNA with KS degenerate primers (Table I.3), as a probe (Fujii *et al.*, 2005). Homology searching showed that the *alt*1, 2, and 3 genes encoded cytochrome P450s and the *alt*4 gene encoded a FAD-dependent oxygenase/oxidase. The *alt*5 gene, interrupted by three introns, encoded a protein of 2 551 amino acids, named PKSN, that was found to

be an iterative type I complex reducing-type PKS with the following domain organization: KS→AT→DH→*C*MeT→ER→KR→ACP. To establish the function of PKSN, the *alt*5 gene was subcloned into the fungal expression plasmid pTAex3 under control of the *A. oryzae* α-amylase promoter (Figure II.2) and introduced into the fungal host *A. oryzae* M-2-3 (Figure I.64). The transformant produced a single polyketide compound in a good yield (approximately 15 $mg \cdot l^{-1}$), named alternapyrone. Feeding experiments with [1,2-$^{13}C_2$]acetate and [methyl-^{13}C]methionine confirmed that alternapyrone is a decaketide with octa-methylation. The interesting feature of the *A. solani* PKSN is its regio-specific methylation: it catalyzes transfer of a methyl group from AdoMet to every C2 unit of the polyketide chain except the third unit (Figure I.51).

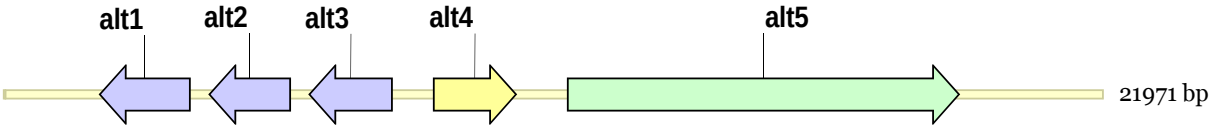


Figure I.50. The *Alternaria solani* gene cluster: the *alt*1, *alt*2, and *alt*3 genes encodes cytochrome P450s, the *alt*4 gene encodes a FAD-dependent oxygenase/oxidase, the *alt*5 gene encodes a PKS (accession number AB120221, Fujii *et al.*, 2005).

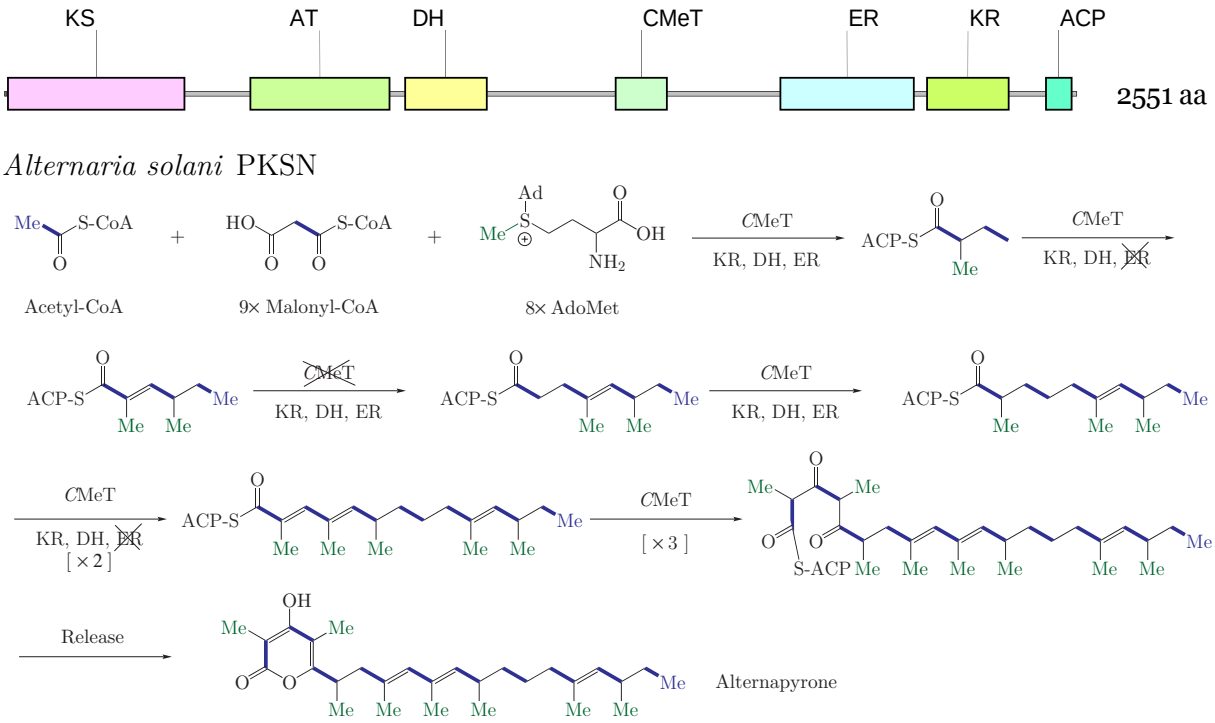


Figure I.51. Domain organization of *Alternaria solani* PKSN (accession number BAD83684; Fujii *et al.*, 2005) and proposed biosynthesis of alternapyrone catalyzed by this PKS. KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein.

A further searching for PKS genes in *A. solani* identified two more highly reducing PKSs, PKSF and PKSK encoded by *pks*F and *pks*K, respectively (KASAHARA *et al.*, 2006). Unlike PKSN, these synthases did not possess a *C*MeT domain, but were otherwise similar to other highly reducing PKSs and contained catalytic domains in the following order: KS→AT→DH→ER→KR→ACP (Figure I.53). Heterologous expression of PKSK in *A. oryzae* did not result in any products, but PKSF did produce the new compounds aslanipyrone and aslaniol (Figure I.52) although in low yields, together with nine other related minor compounds. The ER domain of PKSF was considered to be nonfunctional due to its defective NADPH binding motif which is consistent with the absence of reduction of the double bonds formed by the DH domain in the polyunsaturated PKSF products (Figures I.52, I.53). It is not yet known whether the *pks*F gene cluster possesses a *lov*C-like ER gene as the lovastatin and tenellin clusters do, but coexpression of PKSF with such a gene might form new polyketides.

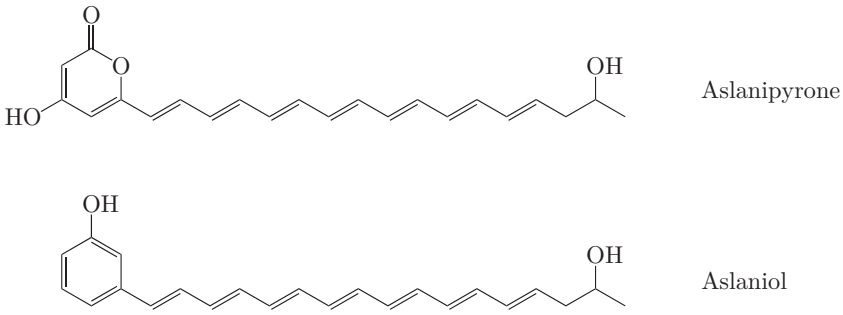


Figure I.52. Structures of aslanipyrone and aslaniol produced by the *Alternaria solani* PKSF

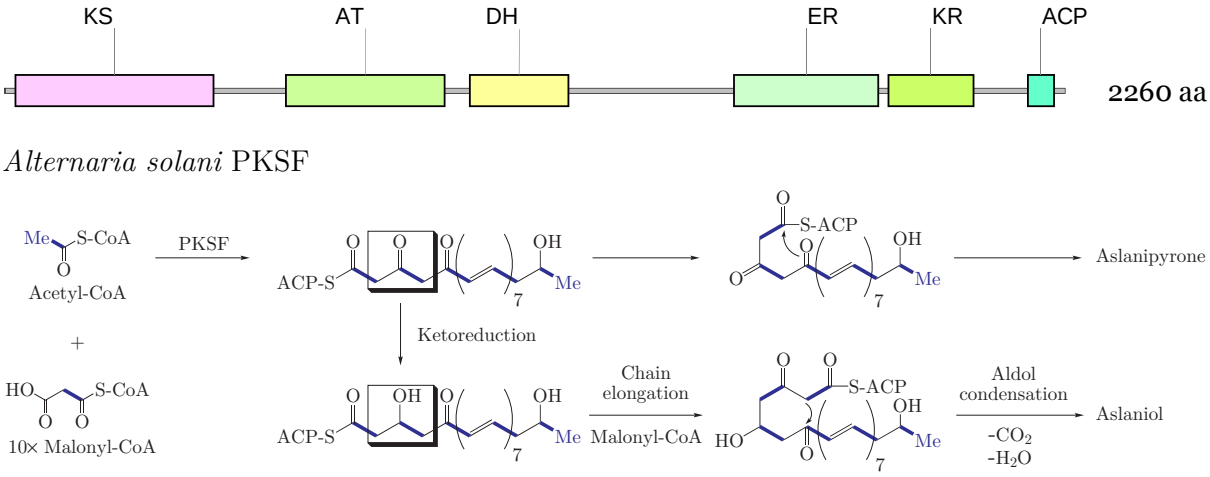


Figure I.53. Domain organization of *Alternaria solani* PKSF (accession number BAE80697; KASAHARA *et al.*, 2006) and proposed biosynthesis of aslanipyrone and aslaniol catalyzed by this PKS. KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein.

2.3.6. Highly reducing PKS-NRPSs

Fungi produce various bioactive metabolites derived from polyketide backbones fused to amino acids. This group of compounds includes fusarin C, equisetin, tenellin and the recently discovered aspyridones A and B (Figure I.54).

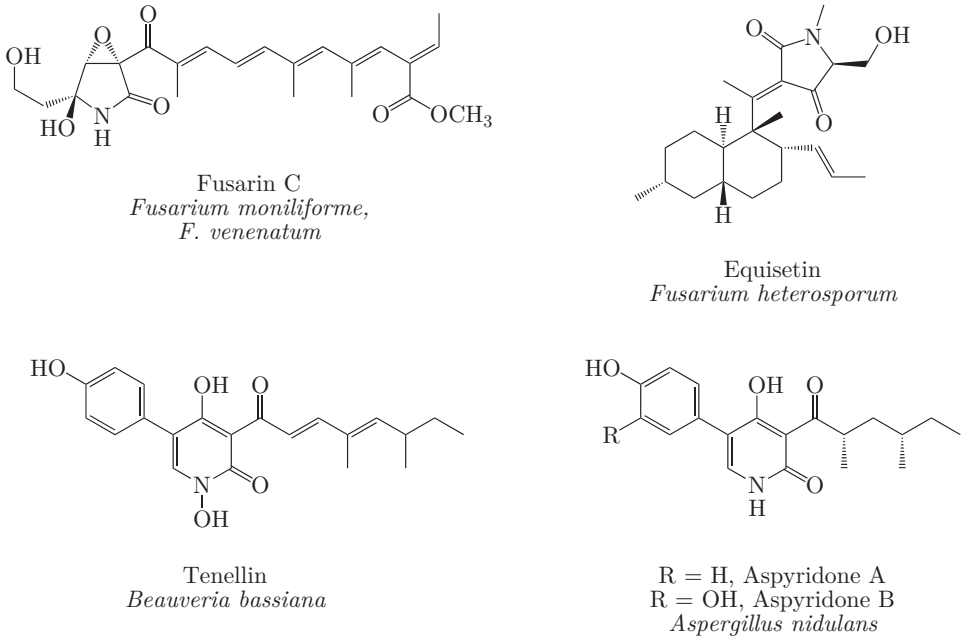


Figure I.54. Structures of selected polyketides produced by fungal highly reducing PKS-NRPSs

Fusarin C is a mycotoxin produced by strains of the plant pathogens *Fusarium moniliforme* (= *Gibberella fujikuroi*) and *F. venenatum*. This compound consists of a tetramethylated heptaketide fused to a modified homoserine moiety (Figure I.54). Genomic DNA libraries from fusarin C producers were used to isolate a gene cluster responsible for its biosynthesis using as a probe a fragment amplified from the *F. moniliforme* and *F. venenatum* cDNA with *C*MeT degenerate primers (Table I.3) (Song *et al.*, 2004). The cloned biosynthetic cluster included a 12-kb open reading frame encoding a protein of 3733 amino acids consisting of a highly reducing PKS fused to a non-ribosomal peptide synthetase (NRPS) module. The PKS region shows a high degree of homology to LNKS and possesses KS, AT and DH domains followed by *C*MeT, a defective ER, KR and ACP domains. Like in LNKS, the ACP domain lies upstream of an NRPS condensation (C) domain, but in this case, the NRPS module is complete, featuring downstream adenylation (A), thiolation (T) (peptidyl carrier protein, PCP) and *C*-terminal thiolester reductase (R) domains (Figure I.55).

Directed knockout of the PKS-NRPS gene resulted in strains incapable of producing fusarin proving its involvement in the biosynthesis of fusarin C. The encoded PKS-NRPS

was thus named fusarin synthetase (FUSS). The nonfunctional ER domain and the absence of any *lov*C homologue in the cluster is consistent with the polyunsaturated nature of the polyketide moiety of fusarin C.

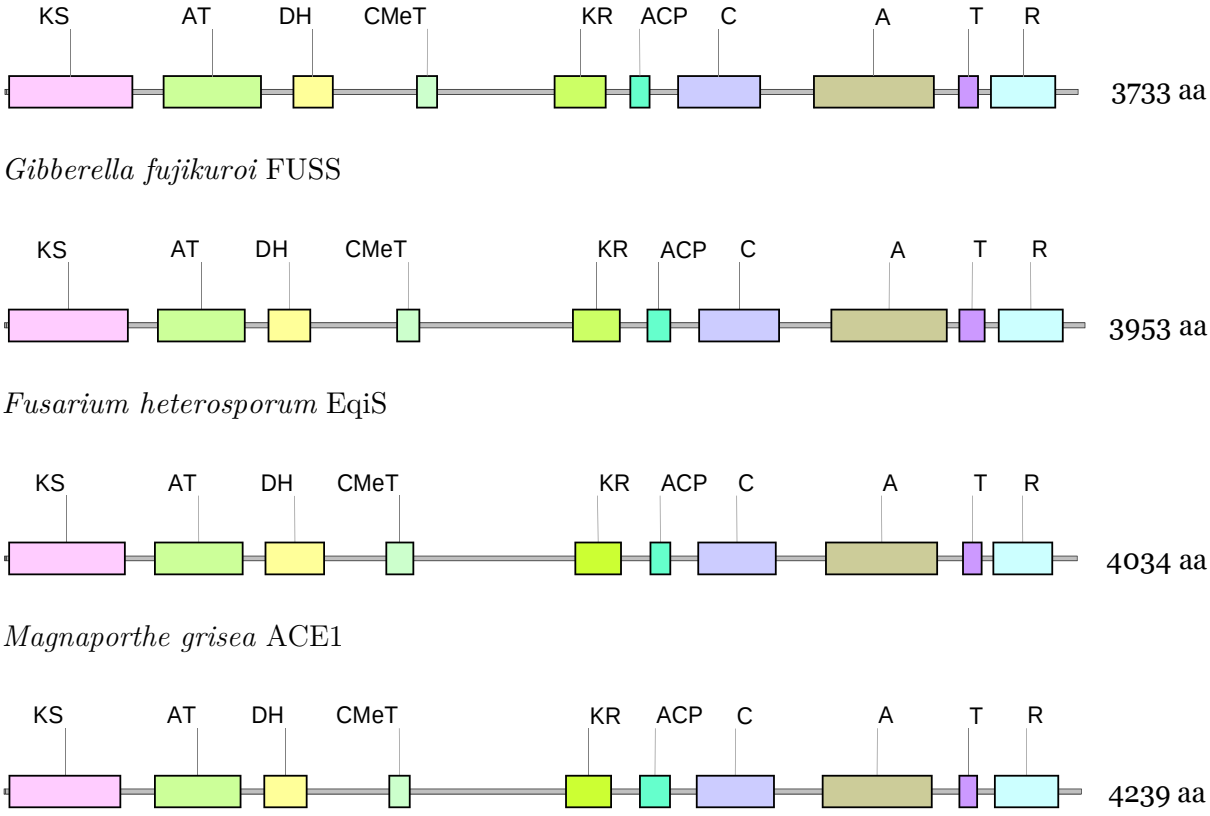


Figure I.55. Domain organization of selected fungal PKS-NRPSs: fusarin synthetase (FUSS) from *Gibberella fujikuroi* (accession number AAT28740; Song *et al.*, 2004), equisetin synthetase (EqiS) from *Fusarium heterosporum* (accession number AAV66106; Sims *et al.*, 2005), ACE1 from *Magnaporthe grisea* (accession number CAG28797; Böhnert *et al.*, 2004), tenellin synthetase (TENS) from *Cordyceps bassiana* (accession number CAL69597; Eley *et al.*, 2007). KT, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; KR, β-ketoreductase; ACP, acyl carrier protein; C, condensation domain; A, adenylation domain; T, thiolation domain (peptidyl carrier protein, PCP); R, thiolester reductase. Inactive ER domains are not shown.

It was proposed that FUSS assembles a tetramethylated heptaketide attached to the ACP remarkably resembling in its structure the heptaketide pyrone produced by LNKS in the absence of the LovC protein. In parallel, the A domain of the NRPS module appears to select, activate and load homoserine onto the thiolation domain. The C domain then utilizes the amine of homoserine to form an amide linkage with the ACP-bound polyketide, forming a covalently bound intermediate peptide. A similar reaction is probably catalyzed by the C domain of LNKS which may use a water nucleophile, or the C-11 hydroxyl of the putative enzyme-bound lovastatin precursor to release dihydromonacolin L (Figure I.45).

The final reaction governed by FUSS may be reductive release of the thiolester, forming a peptide aldehyde. A homologous thiolester reductase domain forms the *C*-terminal feature of fungal α-aminoadipate reductase and this catalyzes a single reduction of a thiolester to form an aldehyde. Finally, Knoevenagel cyclisation and release would give the putative pre-fusarin C. Other genes in the fusarin C biosynthetic gene cluster are presumably responsible for the required further transformation of pre-fusarin C itself: epoxidation, oxidation of a pendant methyl group to a carboxylate and its esterification (*O*-methyl transfer), and hydroxylation α to the nitrogen atom.

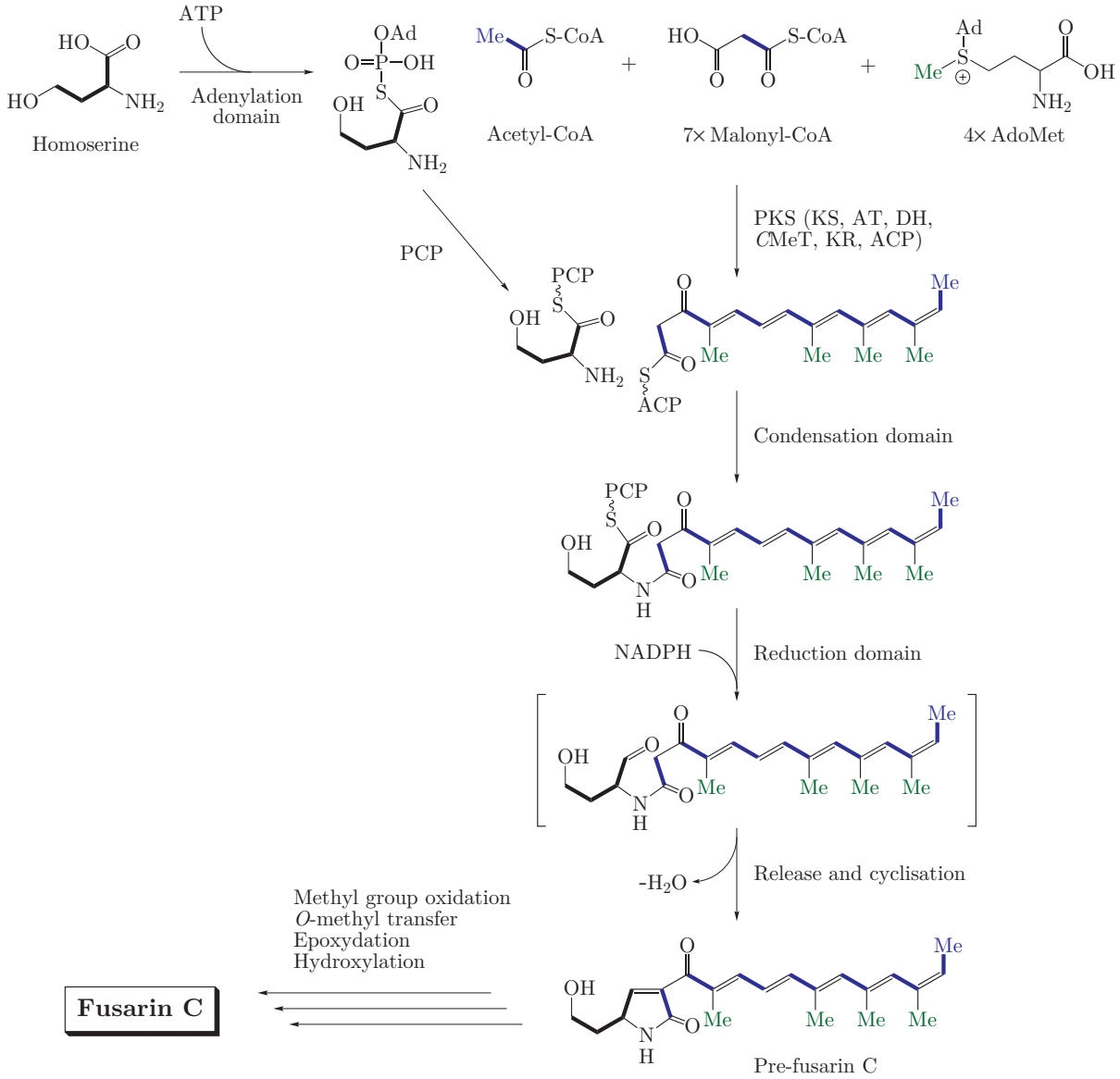


Figure I.56. The proposed biosynthesis of fusarin C via pre-fusarin C. ATP, adenosine 5′-triphosphate; Ad, adenosine; AdoMet, *S*-Adenosyl-L-methionine; KT, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; KR, β-ketoreductase; ACP, acyl carrier protein; PCP, peptidyl carrier protein. Adapted from Song *et al.*, 2004.

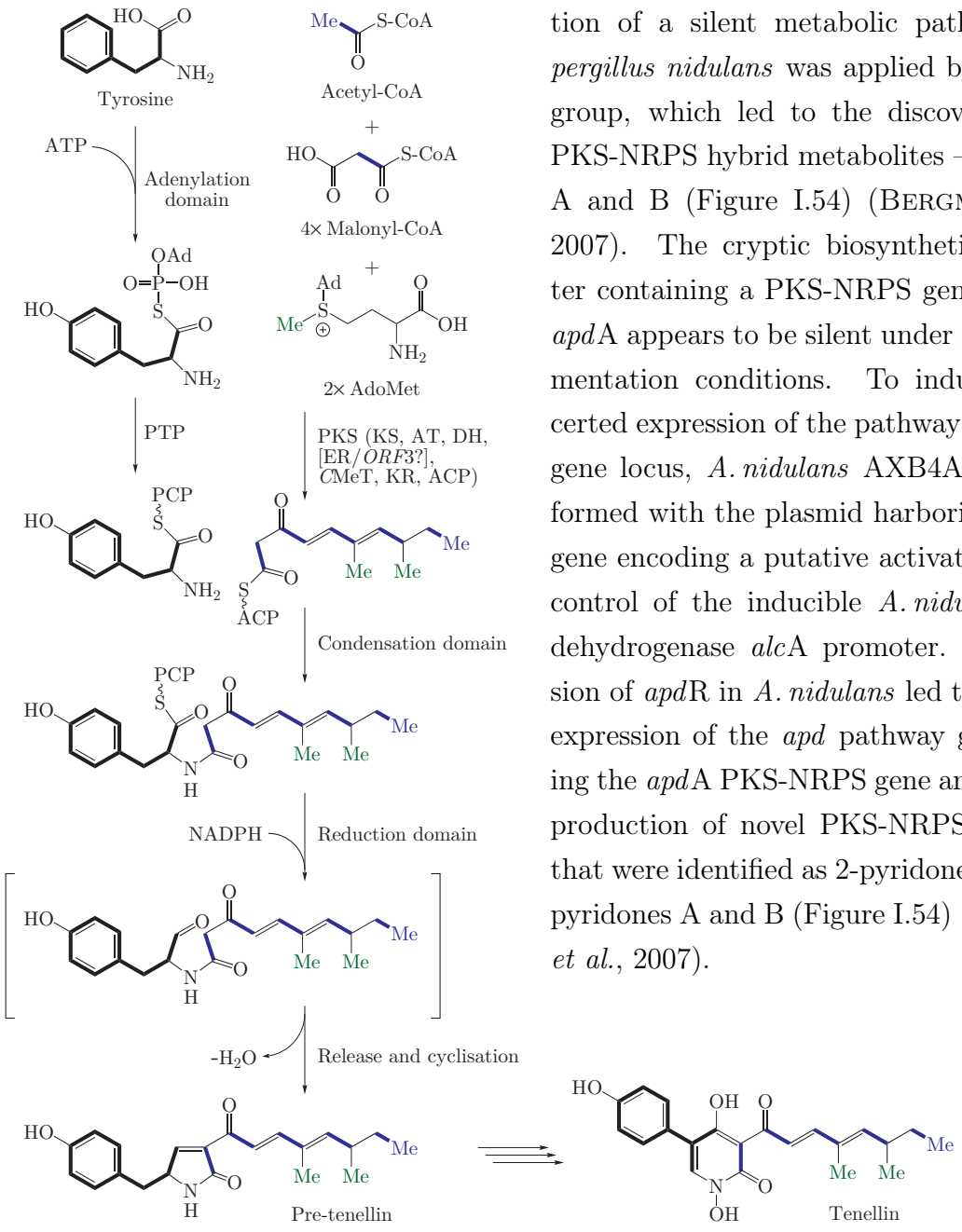
A highly homologous PKS-NRPS gene has been shown to be involved in the biosynthesis of equisetin in *Fusarium heterosporum* (Sims *et al.*, 2005). The encoded equisetin synthetase of 3953 amino acids, designated EqiS, possesses the same catalytic domains as FUSS, but the structure of equisetin suggests that the pyrollidinone carbon derived from the carboxylate of the amino acid (serine in this case) is not reduced, indicating either a reoxidation mechanism, or the fact that the R domain does not produce an aldehyde intermediate in this case.

PKS-NRPSs are also involved in the biosynthesis of yet unidentified compounds in the plant pathogen *Magnaporthe grisea* (Böhnert *et al.*, 2004). The Ace1 PKS-NRPS gene forms part of a biosynthetic gene cluster which is transiently expressed only during appressorium formation during the initial penetration event of pathogenesis. Other genes in the Ace1 cluster encode cytochrome P450s as well as LovC-type ER homologues. The compound produced by this cluster can be detected by specific genotypes of rice resistant to this pathogen and, if detected, this triggers a signalling cascade resulting in the effective mounting of resistance to further fungal penetration. Thus when Ace1 and its neighboring genes are expressed, *M. grisea* is rendered avirulent (Böhnert *et al.*, 2004).

PKS-NRPS proteins are also responsible for the production of fungal 2-pyridones. For example the yellow pigment tenellin from the insect pathogen *Beauveria bassiana* (= *Cordyceps bassiana*) (Figure I.57) is produced by a PKS-NRPS of 4239 amino acids, named TENS (Figure I.55), homologous to FUSS (Eley *et al.*, 2007). Like FUSS, the ER domain of TENS appears to be inactive. However, the tenellin biosynthetic gene cluster also contains a *lov*C homologue, ORF3, encoding a protein that presumably carries out the programmed enoyl reduction event required during the first round of polyketide biosynthesis. It appears that TENS selects tyrosine and fuses it to a doubly methylated pentaketide to form putative pre-tenellin after reductive release and Knoevenagel cyclisation (Figure I.58).

Figure I.57. Insects affected with *Beauveria bassiana*. Photos courtesy by Clive Shirley.

Two other genes in the tenellin biosynthetic gene cluster encode cytochrome P450s and they likely catalyze oxidative ring expansion to form the 2-pyridone, and *N*-hydroxylation (Eley *et al.*, 2007). *B. bassiana* is an effective insect pathogen and it has been hypothesized that tenellin may be involved in pathogenesis. However, directed knockout of the TENS gene did not significantly reduce the pathogenicity of *B. bassiana* towards wax moth (*Galleria mellonella*) larvae (Eley *et al.*, 2007).



A new approach for the successful induction of a silent metabolic pathway in *Aspergillus nidulans* was applied by Hertweck's group, which led to the discovery of novel PKS-NRPS hybrid metabolites – aspyridones A and B (Figure I.54) (Bergmann *et al.*, 2007). The cryptic biosynthetic gene cluster containing a PKS-NRPS gene designated *apd*A appears to be silent under standard fermentation conditions. To induce the concerted expression of the pathway genes in this gene locus, *A. nidulans* AXB4A2 was transformed with the plasmid harboring the *apd*R gene encoding a putative activator under the control of the inducible *A. nidulans* alcohol dehydrogenase *alc*A promoter. Overexpression of *apd*R in *A. nidulans* led to substantial expression of the *apd* pathway genes including the *apd*A PKS-NRPS gene and resulted in production of novel PKS-NRPS compounds that were identified as 2-pyridones, named aspyridones A and B (Figure I.54) (Bergmann *et al.*, 2007).

Figure I.58. The proposed biosynthesis of tenellin via pre-tenellin. Adapted from Eley *et al.*, 2007.

2.3.7. *C*-Methyltransferase domains in reducing fungal PKSs

Many polyketides isolated from fungi possess methyl group branches that are inconsistent with a molecular backbone derived from folding of a polyketide chain (Figures I.36, I.39, I.47, I.51, I.54, I.59). These methyl groups arise almost universally from the methyl group of L-methionine activated as *S*-adenosyl-L-methionine (O'Hagan, 1991).

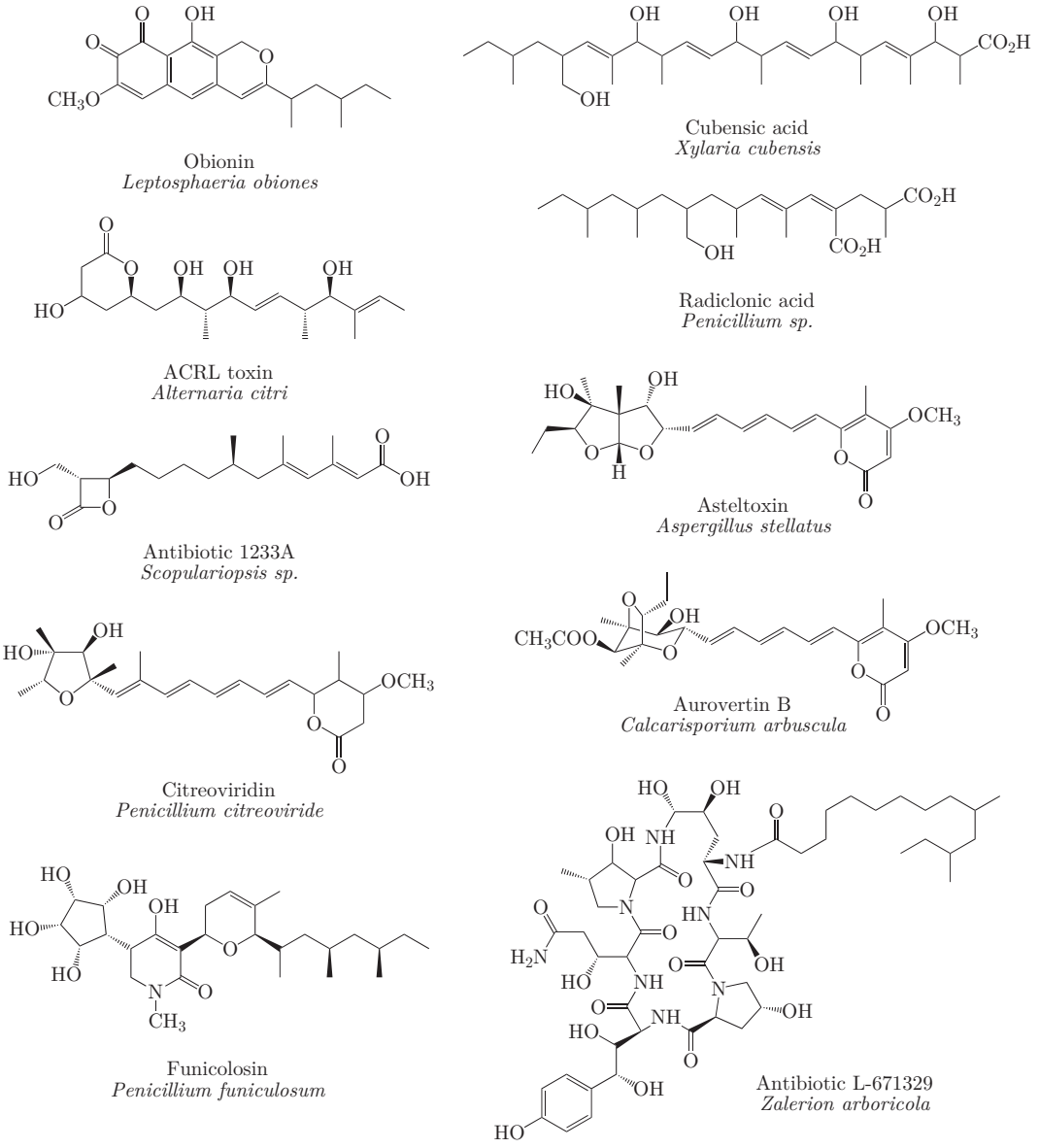


Figure I.59. Structures of selected fungal polyketides with pendant methyl groups derived from *S*-adenosyl-L-methionine: obionin from *Leptosphaeria obiones* (Poch, Gloer, 1989), cubensic acid from *Xylaria cubensis* (O'Hagan *et al.*, 1992), ACRL toxin from *Alternaria citri* (Gardner *et al.*, 1985), radiclonic acid from *Penicillium sp.* (Seto *et al.*, 1977), antibiotic 1233A from *Scopulariopsis sp.* (Kumagai *et al.*, 1992), asteltoxin from *Aspergillus stellatus* (Vleggaar, 1986), citreoviridin from *Penicillium citreoviride* (Vleggaar, 1986), aurovertin B from *Calcarisporium arbuscula* (Vleggaar, 1986), funiculosin from *Penicillium funiculosum* (Ando *et al.*, 1978), antibiotic L-671329 from *Zalerion arboricola* (Adefarati *et al.*, 1991).

S-Adenosyl-L-methionine (AdoMet or SAM) (Figure I.60) plays an essential role as a major methyl donor in various biochemical reactions[7] (Chiang *et al.*, 1996). AdoMet is synthesized by the transfer of an adenosyl group from ATP to the sulfur atom of methionine. The reaction is catalyzed by AdoMet synthetase (methionine adenosyltransferase) in a nonequilibrium manner. The AdoMet synthetase reaction is the only one known to date that requires hydrolysis of all three phosphates from ATP to drive the reaction. The resulting tripolyphosphate residue of ATP is hydrolyzed simultaneously to P_i and PP_i by the intrinsic tripolyphosphatase activity of the enzyme. The activated methyl group ($-CH_3$) in AdoMet allows donation of this group to an acceptor substrate in transmethylation reactions (Figure I.60). *S*-Adenosyl-L-homocysteine (AdoHcy or SAH) is formed after the transfer of a methyl group from AdoMet to an acceptor. AdoHcy is hydrolyzed to homocysteine and adenosine by *S*-adenosylhomocysteine hydrolase.

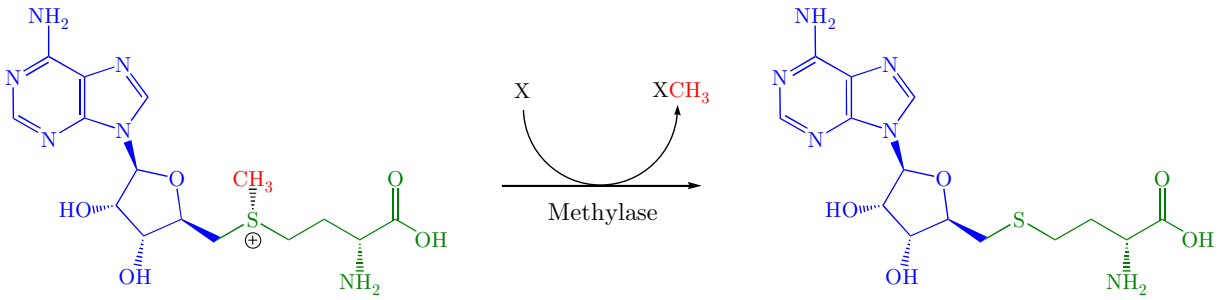


Figure I.60. Transmethylation. X represents a methyl acceptor, the methyl group is highlighted with read, the methionine and homocysteine moieties in AdoMet and AdoHcy respectively are highlighted with green, the adenosyl group is highlighted with blue. Adapted from Chiang *et al.*, 1996.

Homocysteine is then recycled back to methionine by transferring a methyl group from *N*-5-methyltetrahydrofolate by one of the two classes of methionine synthases: *N*-5-methyltetrahydrofolate-homocysteine *S*-methyltransferase (cobalamin-dependent methionine synthase), an enzyme that requires vitamin B_{12} as a cofactor, or 5-methyltetrahydropteroyltri-L-glutamate-homocysteine *S*-methyltransferase (cobalamin-independent methionine synthase). Alternatively, the remethylation of homocysteine to methionine can be catalyzed by betaine-homocysteine methyltransferase utilizing betaine as a methyl donor. This methionine can then be converted to AdoMet, completing the cycle (Figure I.61).

[7] Another major role of AdoMet is in polyamine biosynthesis. Here, AdoMet is decarboxylated by adenosylmethionine decarboxylase to form *S*-adenosyl-5′-deoxy-(5′)-3-methylthiopropylamine (decarboxylated AdoMet). This compound then donates its *n*-propylamine group in the biosynthesis of polyamines such as spermidine and spermine from putrescine.

AdoMet-dependent methyltransferases (MTases) are widespread in all living organisms and methylate a broad variety of molecular targets, such as DNA, RNA, proteins, lipids, polysaccharides, and various small molecules. According to the substrate classes MTases can be divided into families. According to target atoms (nitrogen, oxygen, sulfur or carbon), MTases are also referred to as *N*-, *O*-, *S*-, and *C*-methyltransferases, respectively. *C*-MTases are relatively rare, while *O*- and *N*-MTases are the most abundant (Fauman *et al.*, 1999).

In spite of the wide variety of methyl acceptor substrates and in general the low overall sequence homology to each other presumably due to their high divergence (Fauman *et al.*, 1999), AdoMet-dependent MTases contain several conserved motifs that are considered to be AdoMet binding regions. The first indication of sequence similarities between different types of AdoMet-dependent MTases were shown by Pósfai and coworkers in enzymes that

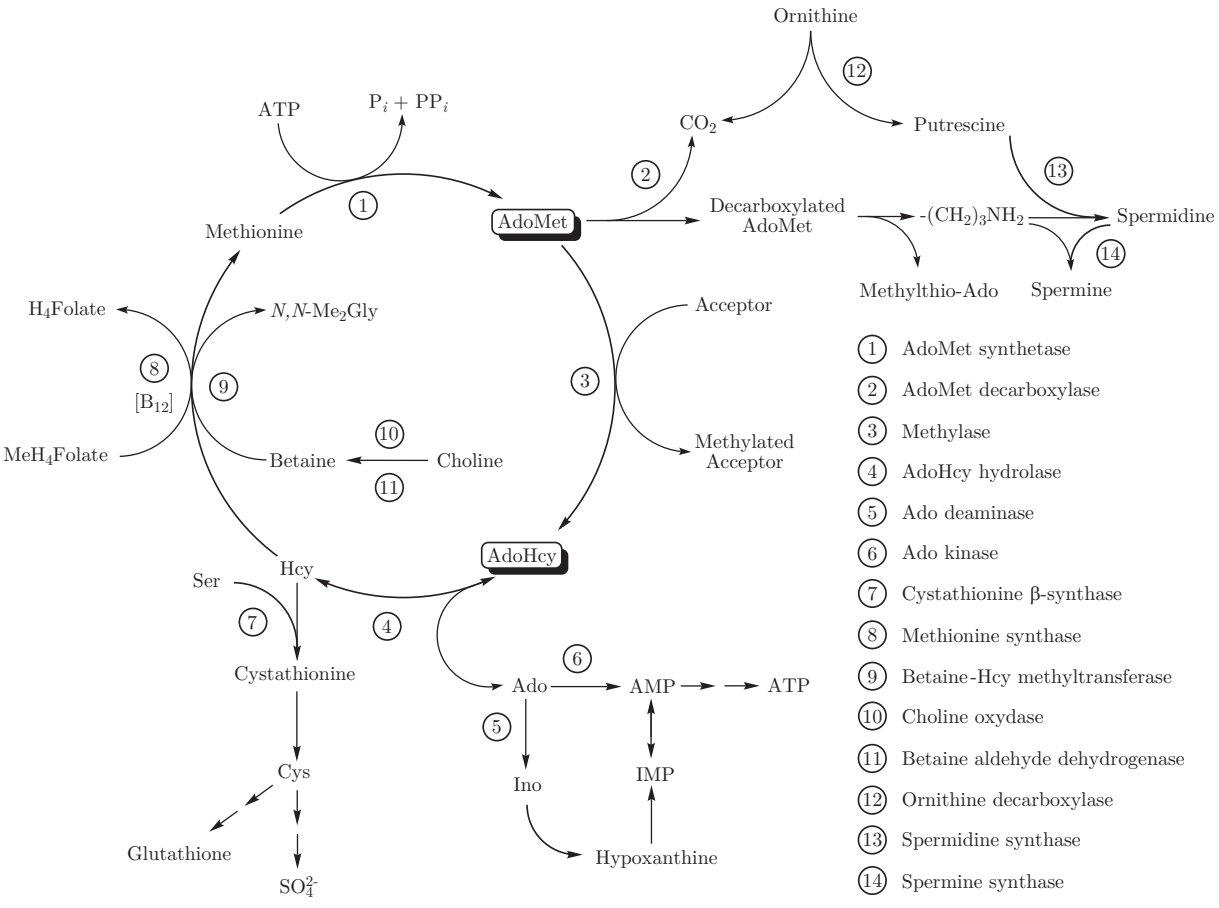


Figure I.61. Metabolic cycle of *S*-adenosylmethionine (adapted from Chiang *et al.*, 1996). AdoMet, *S*-adenosyl-L-methionine; AdoHcy, *S*-adenosyl-L-homocysteine; Ado, adenosine; Hcy, homocysteine; ATP, adenosine triphosphate; AMP, adenosine monophosphate; Ino, inosine; IMP, inosine monophosphate; Ser, serine; Cys, cysteine; MeH_4Folate, *N*-5-methyltetrahydrofolate; H_4Folate, tetrahydrofolate; *N,N*-Me_2Gly, *N,N*-dimethylglycine; B_{12}, vitamin B_{12}, cobalamin. *N*-5-methyltetrahydrofolate and betaine are known methyl donors besides AdoMet.

methylate DNA (PÓSFAI *et al.*, 1988, 1989). They defined 10 motifs with different degrees of conservation among 5-methylcytosine DNA MTases (m^5C MTases).

Kagan and Clarke have identified three widespread conserved sequence motifs for AdoMet-dependent MTases that do not act on DNA (KAGAN, CLARKE, 1994). Motif I can be identified in all MTases and has been previously described in diverse MTases of both prokaryotic and eukaryotic origin, as well as other AdoMet and AdoHcy utilizing enzymes, such as AdoMet synthases, AdoMet decarboxylases, and AdoHcy hydrolases. It comprises a 9-residue conserved region with the consensus sequence (V/I/L)(L/V)(D/E)(V/I)G (G/C)G(T/P)G (Figure I.62). The glycine (G) residues at positions 5, 7, and 9 are the most highly conserved. Motif I is followed by an aspartate (D) or a glutamate (E) residue 17–19 residues *C*-terminal to the motif. The consensus sequence for this region is *hh*X*h*(D/H), where *h* represents a hydrophobic residue. Motif II comprises an 8-residue conserved region found 57 ± 13 residues after the last glycine residue of the motif I and has the consensus sequence (P/G)(Q/T)(F/Y/A)DA(I/V/Y)(F/I)(C/V/L). The central aspartate residue is invariant. The distinctive feature of this motif is high frequency of the aromatic amino acids phenylalanine (F), tyrosine (Y), and tryptophan (W). Motif III comprises a 10-residue conserved region situated 22 ± 5 residues after the end of motif II and has the consensus sequence LL(R/K)PGG(R/I/L)(L/I)(L/F/I/V)(I/L). The central glycines of this region are highly conserved.

The presence of methyltransferase (MeT) domains in multifunctional enzymes such as type I PKSs and NRPSs is usually determined on the basis of conservation of these motifs and homology with known MTases or with MeT domains of enzymes known to produce methylated products, for example, lovastatin diketide and nonaketide synthases (LDKS and LNKS, respectively), both involved in biosynthesis of lovastatin (Figures I.59, I.62).

Fauman and coworkers redefined AdoMet binding motif nomenclature, identifying motifs I–VII. According to them, motifs I, IV and VI correspond to Kagan and Clarke motifs I, II and III, respectively (FAUMAN *et al.*, 1999). Mapping and molecular modeling of AdoMet binding sites in *N*-MeT domains of cyclosporin synthase from the ascomycete fungus *Tolypocladium inflatum* (= *Cordyceps subsessilis, Beauveria nivea*) supported the idea that the four noncontiguous motifs I–IV govern interaction with AdoMet (VELKOV, LAWEN, 2003).

Crystal structure elucidation of the MTase from *Thermotoga maritima* 0872 (TM0872) with both AdoMet and AdoHcy allowed Anderson and coworkers to identify four AdoMet binding motifs that have spatial and functional homology with other MTases (MILLER *et al.*, 2003). Anderson's motif I is a glycine-rich (G) region that forms the binding pocket surrounding the L-homocysteine portion of AdoHcy, and corresponds to motif I by Kagan

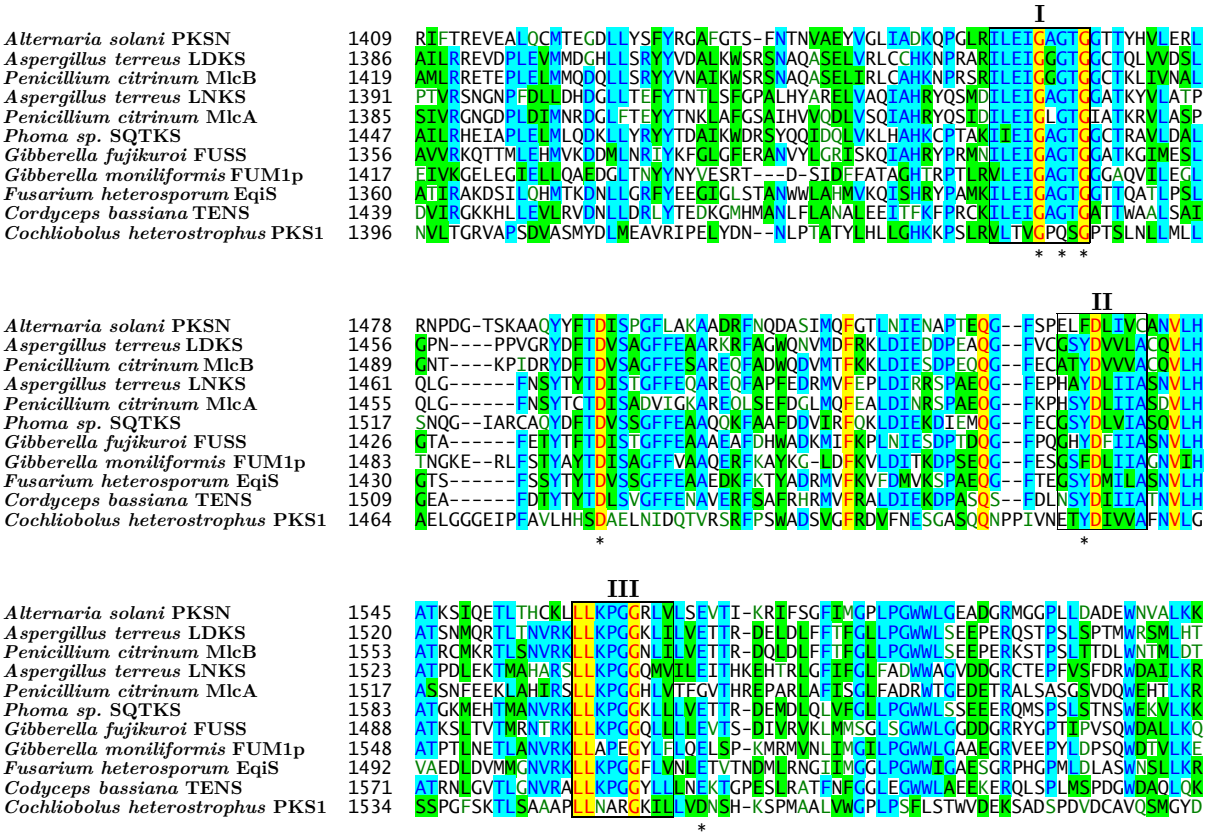


Figure I.62. Amino acid sequence alignment of *C*-methyltransferase domains in selected fungal iterative type I PKSs: PKSN for alternapyrone biosynthesis from *Alternaria solani* (accession number BAD83684; Fujii *et al.*, 2005), LDKS for lovastatin diketide biosynthesis from *Aspergillus terreus* (accession number AAD34559; Kennedy *et al.*, 1999), MlcB for compactin diketide biosynthesis from *Penicillium citrinum* (accession number BAC20566; Abe *et al.*, 2002), LNKS for lovastatin nonaketide biosynthesis from *A. terreus* (accession number Q9Y8A5; Hendrickson *et al.*, 1999), MlcA for compactin nonaketide biosynthesis from *P. citrinum* (accession number BAC20564; Abe *et al.*, 2002), SQTKS for squalestatin tetraketide biosynthesis from *Phoma sp.* (accession number AAO62426; Cox et al., 2004), FUSS for fusarin C biosynthesis from *Gibberella fujikuroi* (= *Fusarium moniliforme*) (accession number AAT28740; Song *et al.*, 2004), FUM1p for fumonisin biosynthesis from *G. moniliformis* (= *Fusarium verticillioides*) (accession number AAD43562; Proctor *et al.*, 1999), EqiS for equisetin biosynthesis from *F. heterosporum* (accession number AAV66106; Sims *et al.*, 2005), TENS for tenellin biosynthesis from *Cordyceps bassiana* (= *Beauveria bassiana*) (accession number CAL69597; Eley *et al.*, 2007), PKS1 for T-toxin biosynthesis from *Cochliobolus heterostrophus* (accession number AAB08104; Yang *et al.*, 1996).

The alignment was obtained using AlignX, a component of Vector NTI Advance v.10.3 Sequence Analysis Software (Invitrogen). The amino acid residues in the alignment are color-coded according to the following scheme: red on yellow, completely conserved; blue on cyan, consensus derived from block of similar residues; green on white, residue weakly similar to consensus residue; black on white, non-similar to consensus residue; black on green, consensus derived from majority residue. The conserved motifs I, II and III, identified by Kagan and Clarke, are shown in boxes. Key amino acid residues in the motifs identified by Anderson and coworkers are marked with an asterisk: the glycine (G) residues in motif I, the aspartic acid (D) residue in motif II, the tyrosine (Y) residue in motif III and the glutamic acid (E) residue in motif IV.

and Clarke, and Fauman and coworkers. The key amino acid residue in Anderson's motif II is aspartic acid (D) or glutamic acid (E), which interacts with the ribose 2′ and 3′ hydroxyls of AdoMet. Next to the acidic amino acid is a bulky aliphatic amino acid such as leucine (L), proline (P), or valine (V), which is in van der Waals contact with the AdoMet adenine ring. This motif II corresponds to that defined by Fauman and coworkers. An aromatic amino acid, tyrosine (Y) or phenylalanine (F), present in Anderson's motif III, which corresponds to the Kagan and Clarke motif II and motif IV of Fauman and coworkers, binds to the AdoMet adenine ring by π–π stacking interaction. Aspartic acid (D) or glutamic acid (E) in Anderson's motif IV, which corresponds to Kagan and Clarke's motif III, interacts with the amino group of AdoHcy. Conservation of these four motifs was also confirmed in caffeic acid/5-hydroxyferulic acid 3/5-*O*-methyltransferase from alfalfa (*Medicago sativa*), whose crystal structure has also been determined (Zubieta *et al.*, 2002).

In fungal iterative type I reducing PKSs, the AdoMet binding motifs are found in the central region of the PKS polypeptides. Interestingly, evidence for these motifs can also be observed in *Cochliobolus heterostrophus* PKS1 involved in T-toxin biosynthesis (Yang *et al.*, 1996) and *Penicillium citrinum* MlcA for compactin nonaketide biosynthesis (Abe *et al.*, 2002), although their products are apparently nonmethylated polyketides. Visual inspection of their *C*MeT domains reveals critical mutations in motifs I and IV for MlcA and in motif I for PKS1 possibly abrogating methylation by these PKSs.

In contrast to post-PKS *O*- and *N*-methylation reactions, which are catalyzed by discrete enzymes after polyketide assembly, *C*-methylations of the polyketide carbon backbone take place much earlier, during chain formation. It appears reasonable to conclude that they occur at a nucleophilic center such as a double bond or more likely active methylene on a poly-β-keto chain prior to ketoreduction by attacking the methyl group of AdoMet. It is not likely that methylation occurs on a saturated chain (O'Hagan, 1991). On the other hand, how regio-specific methylation of the polyketide intermediate is controlled during reducing PKS-catalyzed chain extension is not understood. Cubensic acid (Figure I.59), a metabolite of *Xylaria cubensis* (O'Hagan *et al.*, 1992), and alternapyrone (Figure I.50), a product of heterologous expression of *Alternaria solani* PKSN in *Aspergillus oryzae* (Fujii *et al.*, 2005), are probably the most extreme examples of methylated fungal polyketides. Both compounds possess eight methyl branches with strict regio- and stereospecificity.

Engineering of conserved MTase motifs in reducing PKSs may become a powerful tool to control *C*-methylations of carbon backbones in reduced polyketides and produce compounds with desirable numbers and regio-specificity of methyl groups (Fujii *et al.*, 2004).

3. Locating and cloning fungal PKS genes

The availability of numerous fungal genomes provides tremendous potential for accessing yet unexplored putative PKS gene clusters (Galagan *et al.*, 2003, 2005; Kroken *et al.*, 2003; Machida *et al.*, 2005; Nierman *et al.*, 2005; Pel *et al.*, 2007). In most cases isolation of a specific PKS gene from a particular producer involves creating genomic (cosmid, BAC), as well as cDNA (phage) libraries and screening them by hybridization with heterologous or homologous probes and/or by PCR using degenerate primers. Early efforts of targeting PKS genes in fungi exploited related gene fragments from bacterial genomes. However, this approach was not successful since bacterial heterologous probes did not hybridize well with fungal PKS genes due to their quite low sequence conservation.

The first isolation of a fungal PKS gene was achieved by Beck and coworkers in 1990 with the *Penicillium griseofulvum* MSAS gene by means of an alternative approach. An *E. coli* expression library of *P. griseofulvum* cDNA was screened with antibodies raised against the purified MSAS protein (Beck *et al.*, 1990; Spencer, Jordan, 1992). In turn, the cloned *P. griseofulvum* MSAS was then used as a probe to search for MSAS homologous genes in various other fungal genomes (Fujii *et al.*, 1996; Feng, Leonard, 1998).

This approach, however, failed in isolating the PKS genes involved in lovastatin biosynthesis in *Aspergillus terreus* as only MSAS homologues were cloned (Hendrickson *et al.*, 1999). The genes for lovastatin biosynthesis were identified by screening mutants of the lovastatin producing strain of *A. terreus*. Several mutants were found that were unable to produce either the nonaketide part of lovastatin (the LNKS mutant) or its methylbutyryl side chain. An *A. terreus* genomic cosmid library was prepared in an *Aspergillus*/*E. coli* shuttle cosmid (pLO9) and used to complement the LNKS mutant incapable of producing lovastatin. One lovastatin producing transformant was identified and the cosmid was recovered. In addition, an *A. terreus* cDNA expression library prepared from a lovastatin-producing culture was screened with antiserum raised against a polypeptide shown to represent a PKS associated with lovastatin production. One of the isolated clones that hybridized with the complementing cosmid insert contained a 3′-untranslated region and the *C*-terminus of a protein with an ACP domain. The cDNA clone was then used to isolate overlapping clones from the genomic library yielding an 11.6-kb region coding for the LNKS. Specific primers were then used to obtain the cDNA sequence by PCR (Hendrickson *et al.*, 1999).

The mutant complementation approach was also used for isolating the *Colletotrichum lagenarium* PKS1 gene for DHN-melanin biosynthesis (Takano *et al.,* 1995), the *Aspergillus nidulans w*A gene involved in the biosynthesis of green spore pigment (Mayorga,

Timberlake, 1990) and more recently a PKS gene involved in the formation of a red perithecial pigment of *Nectria haematococca* (Graziani *et al.*, 2004).

A gene encoding the PKS involved in T-toxin biosynthesis in *Cochliobolus heterostrophus* was accessed by tagged mutations using the restriction enzyme-mediated integration (REMI) procedure (see next chapter) resulting in a T-toxin negative phenotype. A PKS gene locus required for T-toxin production and virulence and flanking the tag was then identified and recovered (Lu *et al.*, 1994; Yang *et al.*, 1996).

In 1995 Keller and coworkers designed a set of degenerate primers to identify enzyme and regulatory genes for the sterigmatocystin pathway in *A. nidulans* (Keller *et al.*, 1995). The catalytic domains they targeted were KS and AT. At that time sequence information on fungal PKSs was limited only to the *P. griseofulvum* MSAS and *A. nidulans* *w*A genes. They therefore decided to include a series of published bacterial PKS and FAS sequences in their amino acid comparison. Based on two conserved amino acid regions in the KS domains of fungal FASs and type I fungal and bacterial PKS two sets of long (up to 33 bp) and highly degenerate (up to 256-fold) primers were designed (Figure I.63, Table I.3). In contrast, the sequences for the AT domain did not fall into specific PKS and FAS subgroups. One set of primers (PKS KS Reg. 1 and 2) was then used in PCR with *A. nidulans* genomic DNA as a template. The amplified PCR fragments were sequenced and probed for hybridization with cosmids known to carry genes involved in sterigmatocystin biosynthesis. The results strongly suggested that fragments of PKS genes associated with sterigmatocystin production had been cloned along with fragments of FAS genes.

The same set of primers was used to amplify the KS region of the PKS gene *fum*5 for fumonisin biosynthesis in *Gibberella fujikuroi* (Proctor *et al.*, 1999). PCR experiments using genomic DNA as a template resulted in seven unrelated PKS KS gene fragments but not the desired fumonisin PKS gene fragment. However, PCR with a cDNA template obtained from RNA isolated during fumonisin biosynthesis gave a single homologous probe corresponding to the desired FUM1p gene. It is therefore apparent that Keller's primers are more suitable for amplification of a great variety, but not the entirety, of putative fungal PKS KS domain fragments from genomic DNA templates.

A breakthrough in specifically targeting fungal PKS genes by PCR was achieved by Bingle and coworkers in 1999 using degenerate primers based on known fungal PKS KS domains rather than bacterial PKS or FAS sequences (Bingle *et al.*, 1999). The four fungal PKS genes available at the time were *w*A from *A. nidulans*, PKS1 encoding THNS *C. lagenarium*, and the MSAS genes from *P. griseofulvum* and *A. parasiticus* (Feng, Leonard, 1995). Two sets of degenerate primers (LC1 and LC2c, LC3 and

LC5c) were designed to amplify KS domain fragments from fungal PKS genes belonging to one of the two classes of non-reducing and partially reducing PKSs, WA-type and MSAS-type, respectively (Figure I.63, Table I.3). Sequence analysis of the fragments amplified by these primers showed that the products fell into two classes matching the WA- and MSAS-types, thus allowing selection for different classes of PKS.

In a subsequent study, additional sets of degenerate primers were developed to target conserved regions (about seven amino acids long) of particular fungal PKS domains, such as KS, KR and *C*MeT domains, allowing the selective and rapid cloning of specific fungal PKS genes (Figure I.63, Table I.3) (Nicholson *et al.*, 2001). It has been shown that these primers allow distinguishing between non-reducing PKS (NR PKS), partially reducing PKS (PR PKS), and highly reducing PKS (HR PKS) genes. Primers derived from the consensus sequences of the KS domain (DTACS) and the AT domain (HSSGEIA), as well as the above mentioned primers designed by Bingle and coworkers were successfully exploited for cloning the PKS gene involved in citrinin biosynthesis in *Monascus purpureus* (Shimizu *et al.*, 2005).

Primers derived from intrinsic conserved *C*MeT motifs proved to be a valuable tool for locating genes encoding highly reducing PKSs involved in production of methylated polyketides. They have been successfully used for cloning PKS genes involved in formation of squalestatin S1 in *Phoma sp.* C2932 (Cox *et al.*, 2004), tenellin in *Beauveria bassiana* (Eley *et al.*, 2007), and fusarin C in *Fusarium moniliforme* (Song *et al.*, 2004).

Other sets of KS primers were developed by Gibson and coworkers for accessing PKS genes in insect- and nematode associated fungi (Lee *et al.*, 2001). For this study the degenerate primers were designed for the KS domain using sequence information from known fungal, bacterial, and plant PKS as well as FAS genes (Figure I.63, Table I.3). A very similar approach was applied to study PKS genes in non-sporulating endophytic fungi associated with wild-type *Vaccinium macrocarpon* (Sauer *et al.*, 2002). A pair of degenerate primers was designed based on 12 known fungal PKS genes (Figure I.63, Table I.3) and used in PCR screening. The amplified fragments were sequenced and phylogenetic analysis of the deduced amino acid sequences of the KS domains gave rise to clusters similar to those obtained by Bingle and coworkers (Bingle *et al.*, 1999). One group of sequences was composed of THNS PKS homologues involved in melanin formation. Another group consisted of sequences clustered with aflatoxin encoding fungal PKSs. A third group included sequences that were not clustered with any of the known fungal PKSs. The same set of primers and approach were applied successfully to clone the gene for THNS as well as two more PKS genes in *Glarea lozoyensis* (Zhang *et al.*, 2003).

Degenerate PCR primers were also used to screen the *Penicillium citrinum* genome for its biosynthetic potential (Abe *et al.*, 2002). Three pairs of PCR primers, based on the conserved amino acid sequences found in the active site regions of KS and AT domains, successfully amplified four PKS gene fragments 1–2 kb long, subsequently used in Northern hybridizations as probes. One primer pair (designated comKS and comAT in Table I.3) generated a fragment corresponding to the PKS gene which expression profile appeared to correlate with the pattern of compactin production. This probe was then used successfully for screening the *P. citrinum* cosmid library and cloning a PKS gene cluster responsible for compactin biosynthesis.

Another viable approach to target fungal PKS genes based on nitrogen repression was used to isolate a PKS gene for biosynthesis of the red pigment bikaverin in *G. fujikuroi*. RNA isolated from mycelia grown under repressed and derepressed conditions was subjected to reverse transcriptase reaction and the obtained cDNAs were used as templates in differential display PCR (Linnemannstons *et al.*, 2002).

The PKS gene required for ochratoxin A (Figure I.18) biosynthesis in *Aspergillus ochraceus* was also identified by a suppression subtractive hybridization PCR-based approach (SSH-PCR) (O'Callaghan *et al.*, 2003).

Finally it is noteworthy that fungal researchers should be aware, when searching for particular PKS genes, that not all metabolites isolated from fungi need be fungal metabolites. A recent study revealed that *Rhizopus microsporus*, which does not seem to possess any PKS genes, harbors a bacterial endosymbiont, namely *Burkholderia rhizoxina*, for the production of the polyketide rhizoxin, the causative agent of rice seedling blight (Partida-Martinez, Hertweck, 2005).

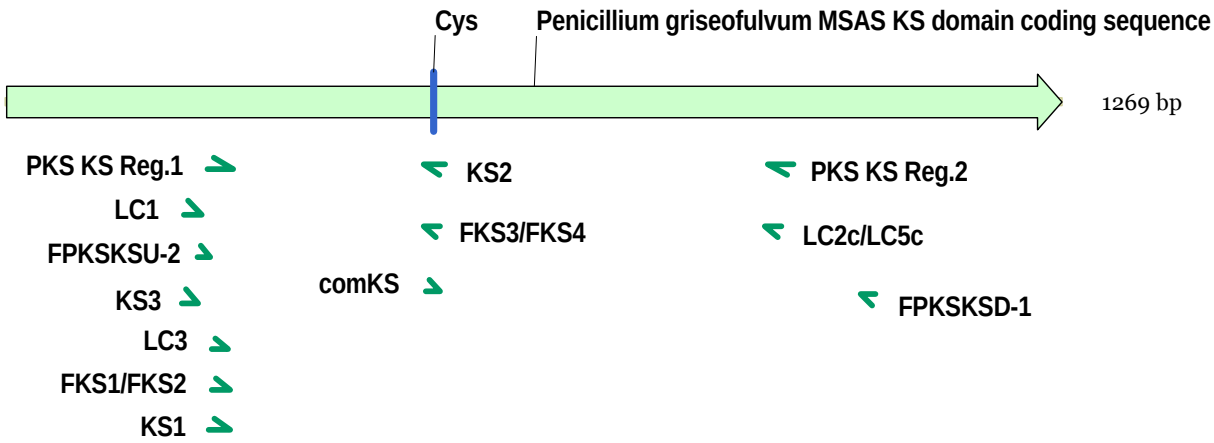


Figure I.63. Approximate positions of binding sites, and directions of some degenerate KS PCR primers relative to the *Penicillium griseofulvum* MSAS KS domain coding sequence. Cys, the active site cysteine codon.

Table I.3. Sequences of degenerate primers from different studies

Primer name	**Sequence** $5' \rightarrow 3'$[a]	**Domain**	**Degeneracy**	**Direction**	**Partner**
FAS KS Reg. 1[b]	GAR GGI GTI GAR ATG GCI TGG ATH ATG GG	FAS KS	24	Forward	FAS KS Reg. 3
FAS KS Reg. 2[b]	GAR WSI TTY ATY ATY ACI ATG WSI GCI TGG G	FAS KS	256	Forward	FAS KS Reg. 3
FAS KS Reg. 3[b]	CAT ACC IGC NCC IGC NGC ICC YTT IGG RTG TCC	FAS KS	64	Reverse	FAS KS Reg. 1/Reg. 2
PKS KS Reg. 1[b]	MGI GAR GCI CAR ATG GAY CCI CAR CAR MG	PKS KS	256	Forward	PKS KS Reg. 2
PKS KS Reg. 2[b]	GG RTI NCC IAR YTG IGT ICC IGT ICC RTG IGC	PKS KS	64	Reverse	PKS KS Reg. 1
PKS AT Reg. 1[b]	GGI CAI GGI GYI CAR TGG GYI GGI ATG G	PKS AT	8	Forward	PKS AT Reg. 2
PKS AT Reg. 2[b]	GCR ATY TCI CCI ARI SWR TGI CCD ATI ACI GC	PKS AT	96	Reverse	PKS AT Reg. 1
LC1[c]	GAY CCI MGI TTY TTY AAY ATG	NR PKS KS	32	Forward	LC2c
LC2c[c]	GTI CCI GTI CCR TGC ATY TC	NR PKS KS	4	Reverse	LC1
LC3[c]	GCI GAR CAR ATG GAY CCI CA	PR PKS KS	8	Forward	LC5c
LC5c[c]	GTI GAI GTI GCR TGI GCY TG	PR PKS KS	4	Reverse	LC3
KS1[d]	GAR GCN GCN GCN YTN GAY CCN CAR CA	PKS KS	4096	Forward	KS2
KS2[d]	AGN CCN CYN GAR CAN GCN GTR TCN AC	PKS KS	16384	Reverse	KS1
FKS1[d]	GCN BHN CAR ATG GAY CCN GCN CA	PKS KS	2058	Forward	FKS3/FKS4

Continued on the next page

Table I.3. (*Continued*)

Primer name	**Sequence** $5' \rightarrow 3'$	**Domain**	**Degeneracy**	**Direction**	**Partner**
FKS2[d]	GCN BHN CAR ATG GAY CCN CAR CA	PKS KS	1024	Forward	FKS3/FKS4
FKS3[d]	GAN GAR CAN GCN GTR TCN AC	PKS KS	1024	Reverse	FKS1/FKS2
FKS4[d]	GAN GAR CAN GCN GTR TCR TT	PKS KS	512	Reverse	FKS1/FKS2
KS3[e]	TTY GAY GCI GCI TTY TTY AA	HR PKS KS	16	Forward	KS4c
KS4c[e]	RTG RTT IGG CAT IGT IAT ICC	HR PKS KS	4	Reverse	KS3
KR1[e]	YTI ATI ACI GGI GGI YTI GGI	PR PKS KS	4	Forward	KR2c/KR3c
KR2c[e]	TAI SWI GCY TGI CCI GGR AA	PR PKS KS	16	Reverse	KR1
KR3c[e]	RTT IGC IGC RTA RTT IC	PR PKS KS	8	Reverse	KR1
CMeT1[e]	GAR ATI GGI GSI GGI ACI GG	PKS *C*MeT	4	Forward	CMeT2c/CMeT3c
CMeT2c[e]	ATI ARY TTI CCI CCI GGY TT	PKS *C*MeT	8	Reverse	CMeT1
CMeT3c[e]	ACC ATY TGI CCI CCI GGY TT	PKS *C*MeT	4	Reverse	CMeT1
comKS[f]	GAY ACI GCI TGY AST TC	PKS KS	8	Forward	comAT
comAT[f]	TCI CCI KIR CWG TCI CC	PKS AT	8	Reverse	comKS
FPKSKSU-2[g]	ATS TCK CCY MRR GAR GC	PKS KS	128	Forward	FPKSKSD-1
FPKSKSD-1[g]	CHM SRT GRC CCR AYT TKG	PKS KS	512	Reverse	FPKSKSU-2

Continued on the next page

Table I.3. (*Continued*)

Primer name	Sequence $5' \rightarrow 3'$	Domain	Degeneracy	Direction	Partner
KAF1[h]	GAR KSI CAY GGI ACI GGI AC	PKS KS	16	Forward	KAR1/KAR2
KAF2[h]	GAR GCI CAY GCI ACI TCI AC	PKS KS	4	Forward	KAR1/KAR2
KAR1[h]	CCA YTG IGC ICC RTG ICC IGA RAA	PKS AT	8	Reverse	KAF1/KAF2
KAR2[h]	CCA YTG IGC ICC YTG ICC IGT RAA	PKS AT	8	Reverse	KAF1/KAF2
XKS1[h]	TTY GAY GCI BCI TTY TTY RA	PKS KS	96	Forward	XKS2
XKS2[h]	CRT TIG YIC CIC YDA AIC CAA A	PKS KS	12	Reverse	XKS1
KSD-F[i]	GAR GCN CAY GGN ACN GGN AC	HR-PKS KS	1024	Forward	KSD-R
KSD-R[i]	CCA YTG NGC NCC YTG NCC	HR-PKS KS	256	Reverse	KSD-F

[a] B = C or G or T; D = A or G or T; H = A or C or T; I = inosine; K = G or T; M = A or C; N = A or C or G or T; R = A or G; S = C or G; W = A or T; Y = C or T.
[b] Keller *et al.*, 1995.
[c] Bingle *et al.*, 1999.
[d] Lee *et al.*, 2001.
[e] Nicholson *et al.*, 2001.
[f] Abe *et al.*, 2002.
[g] Sauer *et al.*, 2002.
[h] Amnuaykanjanasin *et al.*, 2005.
[i] Fujii *et al.*, 2005.

4. Functional analysis of fungal PKS genes

Determining the function of a particular fungal PKS gene can be a challenging and cumbersome task and typically involves either inactivation of a PKS gene of interest or heterologous expression of fungal PKS and related genes in a surrogate host and identifying the product(s) of this expression. As a prerequisite both procedures require introduction of DNA in the form of disruptive or expression constructs into a fungal strain (SCHÜMANN, HERTWECK, 2006).

The first genetic manipulation of filamentous fungi was reported in 1973 by Tatum and coworkers for an inositol-auxotrophic mutant of *Neurospora crassa* that was transformed to prototrophy although the integration of exogenous DNA was not confirmed by Southern hybridization analysis (MISHRA *et al.,* 1973; MISHRA, TATUM, 1973). This was followed by further development of transformation systems for *Neurospora crassa* (CASE *et al.*, 1979), for *Aspergillus nidulans* (BALLANCE *et al.*, 1983; JOHN, PEBERDY, 1984; TILBURN *et al.,* 1983; YELTON *et al.,* 1984) and for *A. niger* (BUXTON *et al.,* 1985). At present many filamentous fungi from all fungal taxa can be subjected to DNA-mediated transformation (HYNES, 1996).

Typical fungal transformation techniques require protoplasts obtained by cell wall degradation with lytic enzymes (β-glucanse, cellulase, protease, chitinase, etc.) under osmotically stabilized conditions and permeabilization of cell membranes with calcium chloride in the presence of polyethylene glycol (PEG) (FINCHAM, 1989; MACH, 2003; RUIZ-DÍEZ, 2002). Protocols for electroporation of germinating conidia have been reported for a large number of filamentous fungi such as *Neurospora crassa*, *Penicillium urticae*, *Leptosphaeria maculans*, *Aspergillus oryzae* (CHAKRABORTY, KAPOOR, 1990; CHAKRABORTY *et al.*, 1991), *Scedosporium prolificans* (RUIZ-DÍEZ, MARTÍNEZ-SUÁREZ, 1999), *Humicola grisea var. termoidea* (DANTAS-BARBOSA *et al.*, 1998), *Beauveria bassiana* (JIANG *et al.*, 2007), *A. niger* (OZEKI *et al.*, 1994), *A. nidulans* (SÁNCHEZ, AGUIRRE, 1996; SÁNCHEZ *et al.*, 1998), *A. fumigatus* and *Wangiella (Exophiala) dermatitidis* (KWONG-CHUNG *et al.*, 1998; PENG *et al.*, 1995; WEIDNER *et al.*, 1998), etc. A procedure involving transformation of germinating conidia with DNA in the presence of lithium acetate to increase permeability of their cell wall has been successfully applied to *N. crassa* (DHAWALE *et al.*, 1984), *Coprinus cinereus* (BINNINGER *et al.*, 1987), *Trichoderma harzianum* (GOLDMAN *et al.*, 1990) and *T. longibrachiatum* (SÁNCHEZ-TORRES *et al.*, 1994), *H. grisea var. termoidea* (DANTAS-BARBOSA *et al.*, 1998), and *Ustilago violacea* (BEJ, PERLIN, 1989). Biolistic transformation involving bombardment of fungal spores or hyphae with tungsten particles coated with heterologous DNA has also been used for transforming a number of filamentous fungi such as *N. crassa*, *Gliocladium virens*, *Magna-*

porthe grisea, *T. harzianum* (Lorito *et al.*, 1993; Riach, Kinghorn, 1996), *Botryotinia fuckeliana* (Hilber *et al.*, 1994), *A. nidulans* (Fungaro *et al.*, 1995; Barcellos *et al.*, 1998), *Trichoderma reesei* (Hazell *et al.*, 2000; Te'o *et al.*, 2002), *Fusarium oxysporum* (Aboul-Soud *et al.*, 2004). An alternative approach is the use of *Agrobacterium tumifaciens* that has proven successful in its ability to efficiently transfer DNA into both plants and fungi (de Groot *et al.*, 1998; for the review see Hooykaas, 2004).

In general, selection of fungal transformants can be achieved with either nutritional markers, which complement auxotrophic mutations (*arg*B, *pyr*G, *trp*C, *amd*S, *nia*A, etc.) or dominant antibiotic resistance markers against hygromycin B, bleomycin and phleomycin (Zeocin™), oligomycin, bialophos (BASTA), sulfonylurea, benomyl, and G418 (Geneticin™) (Fincham, 1989; Mach, 2003; Ruiz-Díez, 2002; Bowyer, 2001). Both types of markers have their advantages and disadvantages. The use of nutritional markers is generally more reliable and cheap, but requires a corresponding auxotrophic mutant strain as a host for transformation. Antibiotic resistance markers can be used for transforming virtually all fungal species including genetically uncharacterized ones. It is a feature that has resulted in the application of this markers to fungal species for which there are few or no auxotrophic mutants. One disadvantage of antibiotic resistance markers is that the resistance allele may not show significant dominance over the wild-type allele, resulting in selection difficulties. Other disadvantages of many of these selective markers are that the compounds used to select transformants are often very expensive and may require extensive optimization of transformation procedure such as the concentration of antibiotic, composition of a selective medium, type of osmotic stabilizer, and timing of selection – at once or after some period of time. All these factors hamper the use of antibiotic resistance markers as general selective markers and may be a reason for why they are not in wide use.

In some special cases, the expression of the transformed gene can be seen by eye on a transformation plate and it can be used for visual selection of transformants although this approach has not gained widespread use due to its species specificity and the absence of direct positive selection. The *Aspergillus nidulans* *y*A gene encoding a laccase that converts a yellow intermediate to the mature green spore pigment (O'Hara, Timberlake, 1989; Aramayo, Timberlake, 1990) is probably the most well-known example of markers based on visual selection of transformants: $w\mathrm{A}^{+}$ $y\mathrm{A}^{-}$ strains of *A. nidulans* producing yellow conidia can be transformed with DNA carrying the *y*A gene and transformants with green spores can be selected in this way. Another approach for the study of *Aspergillus* transformation was provided by Van Gorcom and coworkers by fusing the *Escherichia coli* *lac*Z (β-galactosidase) gene in frame with the coding region of the *A. nidulans* *trp*C gene. Colonies of *A. nidulans* transformants harboring this construction turned blue on medium

containing the chromogenic β-galactosidase substrate X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (Van Gorcom *et al.*, 1985).

A special type of selection marker is based on the suppression of an auxotrophic mutation by tRNA suppressor genes. This was demonstrated in a *Podospora anserina* strain carrying the nonsense mutation *leu*1$^-$, suppressed to leucine prototrophy by the tRNA suppressor genes *su*4–1 and *su*8–1 (Brygoo, Debuchy, 1985).

In spite of the various molecular tools for gene targeting now available to fungal researchers, a major problem for functional gene analyses is that the integration of transformed DNA can occur via either homologous or non-homologous (ectopic) recombination (Fincham, 1989). The frequency of homologous recombination can vary greatly between different organisms and depending on the length of homologous DNA. If two separate homologous sites lie within one DNA fragment a double homologous integration can be achieved, resulting in replacement of the target gene. Non-homologous, or ectopic, recombination occurs when the transformed DNA integrates at a random site in the fungal genome. This can also happen at high frequency if the DNA contains homologous regions. Restriction enzyme-mediated integration (REMI) transformation can be used to facilitate non-homologous integration of transforming DNA (Schiestl, Petes, 1991). This technique is based on the addition of a restriction enzyme with a single site in the transforming plasmid to the transformation reaction resulting in nicks in chromosomal DNA and facilitates recognition of the insertion sites (Brakhage, Langfelder, 2002; Brown *et al.*, 1998; Lu *et al.*, 1994; Bowyer, 2001).

On the other hand, abolishing non-homologous end-joining (NHEJ) activity facilitates the efficiency of homologous integration. The major break-through in gene targeting for filamentous fungi was obtained in *Neurospora crassa* by eliminating the NHEJ activity, thus achieving gene-targeting efficiencies of more than 90% (Ninomiya *et al.*, 2004). Similar results have been presented for five different *Aspergillus* species by permanent or transient disruption of the genes encoding the Ku70 component of the NHEJ machinery (Krappmann *et al.*, 2006; Meyer *et al.*, 2007; Nayak *et al.*, 2006; Nielsen *et al.*, 2007; Takahashi *et al.*, 2006). Another approach exploiting a PKS gene (*pks*12) from *Fusarium graminearum* as a visible marker has recently been used to develop a highly efficient gene targeting system for this fungus (Maier *et al.*, 2005).

Complementation of randomly generated mutants by transforming them with wild-type DNA is another viable method for analyzing the functions of targeted genes with recognizable phenotypes. Today several techniques are known to generate fungal mutants, such as REMI (Brown *et al.*, 1998), UV mutagenesis (Jahn *et al.*, 1997; Bowyer, 2001), chemical mutagenesis (Bowyer, 2001) or transposon mutagenesis (Kempken *et*

al., 1998; Kempken, Kuck, 2000; Brakhage, Langfelder, 2002).

Most of the gene targeting techniques resulting in gene inactivation have been applied to the functional analyses of fungal polyketide biosynthesis genes that are involved in the formation of pigments and/or pathogenicity factors. DHN-melanin, a widely distributed type of melanin among ascomycete fungi, is probably the most important and known example of both a pigment and a pathogenicity factor derived via the polyketide pathway. Function of many PKS genes involved in DHN-melanin biosynthesis in fungi was established by gene disruption/replacement experiments. Generated mutants displayed an albino phenotype and in some cases reduced pathogenicity (Feng *et al.*, 2001; Fulton *et al.*, 1999; Loppnau *et al.*, 2004; Moriwaki *et al.*, 2004; Tanguay *et al.*, 2006; Tsai *et al.*, 1998; Tsai *et al.*, 1999; Zhang *et al.*, 2003).

In addition to gene disruption/deletion, RNA mediated gene silencing represents another genetic tool that could be useful for exploring gene functions in fungi. This method represents a posttranscriptional gene-silencing phenomenon, in which double stranded RNA triggers degradation of sequence homologous mRNA. Successful RNA mediated silencing of specific genes in *Magnaporthe oryzae* and *C. lagenarium* was reported in 2005 by Nakayashiki and coworkers (Nakayashiki *et al.*, 2005). While being an attractive method to shut down gene expression relatively easily and quickly, RNA silencing may be partial and not necessarily lead to a null phenotype, like the gene knockout, enabling the analysis of essential genes. Furthermore, this method provides flexibility in gene inactivation as it is sequence specific, and not locus specific. Homologous genes have been shown to be silenced simultaneously (Allen *et al.*, 2004; Baulcombe, 1999; Moriwaki *et al.*, 2007; Nakayashiki *et al.*, 2005; Tanquay *et al.*, 2006a; Thierry, Vaucheret, 1996).

5. Heterologous expression of fungal PKS genes

Harvesting of fungal polyketide metabolites from their native producers by large scale fermentation may be hampered by the requirement for sophisticated growth conditions, inefficient production of the compounds of interest, and a high biosynthetic background, which interferes with analysis and isolation of metabolites. For these reasons, development of universal fungal PKS gene expression systems with good fermentation characteristics and a well characterized, if not low, endogenous chemical background, are highly desirable. They might not only facilitate the production of native polyketides, but also serve for functional analyses of fungal PKS genes and engineering of fungal polyketide biosynthesis. However, to date a limited number of fungal PKS genes have been expressed in heterologous hosts and only few fungal expression systems are available.

In contrast to the impressive progress in expressing bacterial PKS genes, heterologous expression of fungal PKS genes has been hampered by many factors. First of all, the presence of introns in fungal genes and differences in splicing may require the use of full-length cDNA, hard to obtain, or laborious removal of introns. Second, the ACP domains require posttranslational phosphopantetheinylation catalyzed by 4′-phosphopantetheinyl transferases (PPTase) (Figure I.8). These enzymes can be quite specific, which necessitates coexpression of a specific PPTase with the PKS in order to obtain a functioning *holo*-form of ACP (Kealey *et al.*, 1998). Third, the differences in regulatory signals may require some retrofit of genes before they are introduced into a heterologous host. In some cases, the large size of the gene cluster encoding the biosynthetic pathway may render expression constructs instable and result in deletions.

In general, a variety of heterologous expression hosts could be suitable for the production of polyketides and this has been reviewed elsewhere (Pfeifer, Khosla, 2001; Rude, Khosla, 2004). Fungal PKS genes have been successfully expressed in different hosts of both eukaryotic and prokaryotic origin.

Expression of fungal PKS genes in bacteria requires removal of introns and may require some modification of the 5′ ends of the PKS genes. To date it has been achieved for the *P. griseofulvum* MSAS gene expressed in both *E. coli* (Kealey *et al.*, 1998) and *S. coelicolor* CH999 (Bedford *et al.*, 1995) and for the *Aspergillus terreus lov*B and *Gibberella fujikuroi* PKS4 genes expressed in *E. coli* (Ma *et al.*, 2007; Ma, Tang, 2007). In all cases, expression of functional PKSs in *E. coli* was facilitated by the coexpression of a heterologous PPTase, specifically the surfactin phosphopantetheinyl transferase (*sfp*) from *Bacillus subtilis*, which posttranslationally modified the active site serines of the ACP domains of the fungal PKSs (Figure I.8). Heterologous expression of a MSAS gene alone in *S. coelicolor* CH999 resulted in high production levels of 6-MSA, indicating correct posttranslational phosphopantetheinylation of the ACP domain by this bacterium. *S. coelicolor* CH999 was originally developed for functional expression of bacterial PKS genes and does not produce any polyketides, because one of its own PKS genes, involved in actinorhodin biosynthesis, was removed via homologous recombination and another one, *whi*E, involved in spore pigment biosynthesis, is exclusively spore associated and is not expressed in liquid cultures (McDaniel *et al.*, 1993).

By analogy to bacteria, heterologous expression of fungal PKS genes in the yeast *Saccharomyces cerevisiae* necessitates removal of introns and coexpression of a heterologous PPTase. Production levels of 6-MSA two fold higher than in the native host were achieved in *S. cerevisiae* on expression of the *P. griseofulvum* MSAS gene, indicating that this host is capable of synthesizing polyketides at high levels (Kealey *et al.*, 1998).

Moreover, the *S. cerevisiae* genome contains no PKS genes and therefore it does not have any polyketide background that can interfere with identification and isolation of heterologously produced polyketides. Further optimization of heterologous production of 6-MSA in *S. cerevisiae* was achieved by exploiting the NpgA PPTase from *Aspergillus nidulans* (Wattanachaisaereekul *et al.*, 2006).

Filamentous fungi have proven suitable hosts for expression of PKS genes from their counterparts. A major advantage of such hosts is their ability to process and correctly splice introns, often found in fungal PKS genes. Similar transcriptional signals allow expression of the fungal PKS genes in heterologous fungal hosts under control of their native promoters. The utility of *A. nidulans* as a heterologous host has been successfully demonstrated in the context of lovastatin biosynthesis (Kennedy *et al.*, 1999). In other studies, *A. nidulans* was used as a heterologous host for successful expression of the *A. terreus* and *Glarea lozoyensis* MSAS genes, using either the fungal expression system pTAex3 (Figure II.2) or under control of the *trp*C promoter, respectively (Fujii *et al.*, 1996; Lu *et al.*, 2005). The pTAex3 system, exploited in this study, utilises the *Aspergillus oryzae* α-amylase (Taka-amylase A, *amy*B) promoter which is repressed by glucose and induced by starch (Fujii *et al.*, 1995). Thus, overexpression of the PKS gene, cloned downstream of the α-amylase promoter and introduced into a fungal host, can be achieved by culturing fungal transformants in medium containing starch. At present, this system is widely used for expression and functional analysis of fungal PKS genes together with *A. oryzae* M-2-3 (*arg*B$^-$) (Gomi *et al.*, 1987), that allows selection of transformants using the *A. nidulans arg*B gene in pTAex3 and ensures the optimal function of its own α-amylase promoter.

Figure I.64. The filamentous fungus *Aspergillus oryzae* M-2-3 (*arg*B$^-$), exploited in this study. Potato dextrose agar (PDA) supplemented with arginine, 30°C, 2 weeks.

To date the only case of fungal PKS gene expression in plants is the expression of the MSAS gene from *P. griseofulvum* in tobacco (Figure I.65) (Yalpani *el al.*, 2001). This example indicates that endogenous plant PPTases appear to be sufficiently flexible to phosphopantetheinylate ACP domains of fungal PKS and suggests that plants can serve as heterologous hosts for functional expression of other fungal PKS genes.

Figure I.65. Tobacco (*Nicotiana tabacum*) plants expressing the *P. griseofulvum* MSAS and effect of this expression on leaf phenotype. Leaves formed just before opening of the first inflorescence grow in a nonsymmetrical pattern. Foliage that formed subsequently appears normal. Photo taken from YALPANI *el al.*, 2001.

It is noteworthy that in some cases heterologous expression of fungal PKS genes resulted in production of compounds that have never been detected in the natural hosts (FUJII *et al.*, 2005; KASAHARA *et al.*, 2006; KENNEDY *et al.*, 1999). Although transcription of individual PKS genes on the mRNA level was confirmed by reverse transcriptase PCR in the natural hosts, their products are probably produced in very low, non-detectable yields and/or undergo extensive post-PKS modification by enzymes, such as cytochrome P450s and FAD-dependent oxygenases/oxidases. On the other hand, standalone heterologous expression of some PKS genes resulted in malfunctioning enzymes that were unable to coordinate all enzymatic activities properly. Only coexpression of accessory proteins complementing the lacking enzymatic activity restored the ability to produce original products. This phenomenon was observed for the *A. terreus lov*B gene involved in lovastatin biosynthesis as its heterologous expression in *A. nidulans* resulted in production of shunt pyrone products with shorter carbon chains and a lower degree of reduction (Figure I.41) than the original intermediate dihydromonacolin L and only the coexpression of *lov*B with the *lov*C gene encoding the lacking ER activity resulted in dihydromonacolin L production (Figure I.40), showing that these two enzymes work closely together in lovastatin biosynthesis (KENNEDY *et al.*, 1999). Finally, the absence of an obvious product release domain, such as an TE domain, may result in the mature product staying bound to the PKS enzyme and therefore blocking it. This possibility was proposed for *A. terreus* LDKS involved in lovastatin biosynthesis (KENNEDY *et al.*, 1999). The LDKS lacks a TE domain and requires a specialized acyl transferase/transesterase LovC for off-loading the diketide product. However, successful heterologous expression of some fungal PKSs without any TE domain suggests a possible dissociative mode of attachment of these enzymes to their products resulting in product release and accumulation. Therefore, one must be aware when expressing individual PKS components of a particular fungal PKS gene cluster that coexpression of accessory proteins may be required.

6. *Solorina crocea*

6.1. Taxonomy and systematics

S o l o r i n a c r o c e a (L.) Ach. 1808.

Synonyms: *Arthonia crocea* (L.) Ach. 1806, *Lichen croceus* L. 1753, *Parmelia crocea* (L.) Spreng. 1827, *Peltidea crocea* (L.) Ach. 1803, *Peltigera crocea* (L.) Hoffm. 1794.

Superkingdom	*Eukaryota*
Kingdom	*Fungi*
Phylum	*Ascomycota (Ascomycetes)*
Subphylum	*Pezizomycotina*
Class	*Lecanoromycetes*
Subclass	*Lecanoromycetidae*
Order	*Peltigerales* W. Watson 1929
Suborder	*Peltigerineae*
Family	*Peltigeraceae* Dumort. 1822
Genus	*Solorina* Ach. 1808 (Eriksson *et al.*, 2006)
Species	*Solorina crocea* (L.) Ach.

Placement of the genus *Solorina* into the separate monotypic family *Solorinaceae* Bayrh. 1951 (Eriksson, 1981; Hafellner, 1988; Hafellner *et al.*, 1994) is not feasible since molecular phylogenetic classification based on ribosomal RNA small and large subunits did not confirm a monophyletic origin of the genus *Solorina* (Miadlikowska, Lutzoni, 2004).

6.2. Morphological description

Solorina crocea is a terricolous foliose lichen with a large, distinctly dorsiventral heteromerous thallus, 1–10 cm broad. The lobes are irregular to rounded, 3–15 mm broad, their margins are entire, often slightly raised. The upper surface of the thallus is green when moist, becoming olive-brown or olive-gray to red-brown and rather scabrous when dry. The medulla and the lower surface of the thallus are bright orange. The lower surface apparently without a cortex, often with a reticulum of flat, brown veins and scattered rhizines. The apothecia are of the lecideine type, frequent, spread all over the surface of the thallus, oblong to round, plane, without a thalline margin, level with the thallus or even somewhat convex, not depressed into the thallus, brown to red-brown, 2–6 (up to 10) mm in diameter. The asci are 6–8-spored, clavate. The paraphyses are not anastomosing, simple, not apically thickened, adglutinated. The ascospores are

1-septate, 2-celled, oblong-ellipsoid, hyaline, becoming brown, 11.0 ± 1.3 μm (with a range $8.4-21.2$ μm) in width and 39.0 ± 4.8 μm (with a range $26.6-51.1$ μm) in length, with rounded free papillae ornamentation (Brodo *et al.*, 2001; Galloway, 1985; Garibova *et al.*, 1978; Martinez, Burgaz, 1998; Thompson, Thompson, 1984).

The medulla gives a positive reaction with KOH[10] (deep red-purple coloration) and a negative reaction with phenylenediamine[11]. Reactions wih hypochlorite[12] and KOH-hypochlorite[13] are considered to be negative although it is not correct since they are masked by pigments (anthraquinones) present in the lichen.

Figure I.66. *Solorina crocea* (L.) Ach., habitus. Huckleberry Mountain, Glacier National Park, Montana, USA. Note the dark brown apothecia and the bright orange lower surface of the thallus. Photo courtesy by Mike Tyler.

[10] Potassium hydroxide solution (about 10–20%), gives deep red coloration with quinones.

[11] *p*-Phenylenediamine ethanolic solution (2–5%), gives yellow to orange coloration with depsides and depsidones containing an aldehyde group.

[12] Calcium ($Ca(OCl)_2$) or sodium hypochlorite (NaOCl) solution, gives yellowish to red coloration with depsides and xanthones containing two free hydroxyl groups in *meta*-position. Some dibenzofuranes show a very characteristic green coloration with this reagent.

[13] Calcium or sodium hypochlorite solution with preliminary addition of KOH solution to facilitate hydrolysis of the depside ester groups.

S. crocea is a tripartite lichen consisting of one mycobiont (fungus) and two photobionts – a phycobiont (green alga) and a cyanobiont (cyanobacterium). Phycobionts are the primary photobionts in tripartite lichens, they occupy most of the lichen thallus and produce most of the total photosynthate. In contrast cyanobionts in tripartite lichens provide the mycobiont with very little or no fixed carbon. Their main function is nitrogen fixation and they are localized in specialized structures called cephalodia. The phycobiont in *S. crocea* is situated rather discontinuously below the upper cortex of the thallus and belongs to the species *Coccomyxa solorinae-crocae* Chod. (*Coccomyxaceae*[14], *Chlorophyta*), sometimes considered to be a variant of *C. solorinae* (Chod.) Jaag – *C. solorinae var. croceae*. The cyanobiont belongs to the genus *Nostoc* (*Nostocaceae, Cyanobacteria*) and forms internal ovoid cephalodia above the lower surface of the thallus (Galloway, 1985; Guiry, Guiry, 2007; Rai, Bergman, 2002).

6.3. Early molecular investigations of *S. crocea*

A *pyr*G gene, encoding orotidine 5′-monophosphate decarboxylase (OMPD), was cloned from a phage library of *S. crocea* genomic DNA (Sinnemann *et al.*, 2000). Both coding and upstream sequences of *S. crocea pyr*G exhibited features typical of fungal genes. A 132-bp intron intervening the *pyr*G coding region between nucleotides 157 and 288 was confirmed by reverse-transcriptase PCR and sequencing. Transformation of the heterologous host *Aspergillus nidulans* with *S. crocea pyr*G, under control of either its native promoter or the *A. nidulans trp*C promoter, resulted in uridine-independent strains that exhibited appreciable growth only at 24℃. The observed effect of temperature on growth of the *Aspergillus* transformants was ascribed to the source of the *pyr*G protein as the natural habitat of *S. crocea* lies in snow-bed communities in subalpine and alpine environments and its metabolism and enzymes may be adapted to lower temperatures. Southern analysis revealed multiple integrations of *S. crocea pyr*G into the genome of *A. nidulans*. This study demonstrated for the first time that *Aspergillus* could be successfully used as a host for heterologous expression and investigation of genes from lichens.

The *S. crocea* RNA polymerase II largest subunit genes (RPB1 and RPB2)[15] and 28S and 18S subunit ribosomal RNA genes[16] were sequenced and used in subsequent phylogenetic analyses and classification of the class *Lecanoromycetes*, particularly the order *Peltigerales* (Eriksson, Strand, 1995; Miadlikowska, 2006; Miadlikowska, Lutzoni, 2000; Miadlikowska, Lutzoni, 2004).

[14]Sometimes placed into the family *Chlorococcaceae*.
[15]Partial sequences.
[16]Partial and complete sequences.

6.4. Polyketide biosynthesis in *S. crocea*

Solorina crocea produces various polyketide metabolites belonging to both anthraquinones and depsides (Figure I.67) (Anderson *et al.*, 1966; Brodo *et al.*, 2001; Ebizuka *et al.*, 1970; Huneck, Yoshimura, 1996).

All anthraquinones isolated from *S. crocea* are colorful compounds (from yellow to red) and therefore responsible to a different extent for the orange color of the medulla and the lower surface of the thallus. In contrast to the anthraquinones, gyrophoric acid and methyl gyrophorate are colorless compounds.

An interesting feature of polyketide biosynthesis in *S. crocea* is the presence of norsolorinic acid and averantin, major precursors in the biosynthesis of aflatoxins and related sterigmatocystins, potent carcinogenic and mutagenic mycotoxins produced by several *Aspergillus* species.

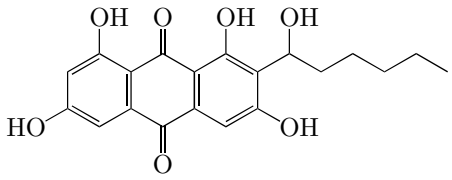
Similarity in structures of the anthraquinones and depsides isolated to date from *S. crocea* suggests that their biosynthesis requires only 2 different PKSs and their structural differences are due to the action of various post-PKS modifying enzymes. Judging from information on PKSs involved in aflatoxin and sterigmatocystin biosynthesis in *Aspergillus* species, biosynthesis of *S. crocea* anthraquinones most likely requires a non-reducing PKS with the architecture KS→AT→ACP→(ACP)→CYC (Chang *et al.,* 1995; Feng, Leonard, 1995; Yu, Leonard, 1995).

Both gyrophoric acid and methyl gyrophorate are derived from condensation of orsellinic acid. The PKS architecture necessary for its biosynthesis is KS→AT→ACP, e.g. in the AviM and CalO5 orsellinic acid synthases from *Streptomyces viridochromogenes* (Gaisser *et al.*, 1997) and *Micromonospora echinospora* (Ahlert *et al.,* 2002), respectively.

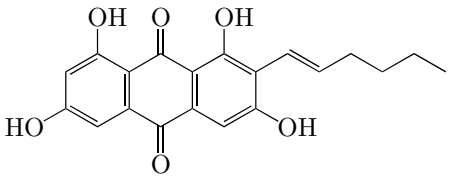
The *S. crocea* PKS explored in this study is not likely to be involved in biosynthesis of any polyketide isolated to date from this lichen since it belongs to the reducing PKSs with the following organization of domains: KS→AT→DH→*C*MeT→ER→KR→ACP. Other PKS enzymes with the same domain organization and identified products are PKSN for alternapyrone biosynthesis from *Alternaria solani* (accession number BAD83684, Fujii *et al.*, 2005), LDKS for lovastatin diketide biosynthesis from *Aspergillus terreus* (accession number AAD34559, Kennedy *et al.*, 1999), MlcB for compactin diketide biosynthesis from *Penicillium citrinum* (accession number BAC20566, Abe *et al.*, 2002), SQTKS for squalestatin tetraketide biosynthesis from *Phoma sp.* (accession number AAO62426, Cox *et al.*, 2004) and FUM1p involved in fumonisin biosynthesis from *Gibberella moniliformis* (= *Fusarium verticillioides*) (accession number AAD43562, Proctor *et al.*, 1999).

Results of a PCR-based study with DNA isolated from *S. crocea* specimens collected at five different locations in southern British Columbia (Canada) suggested a high degree

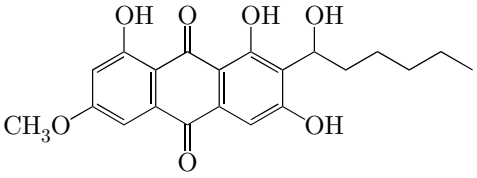
Anthraquinones



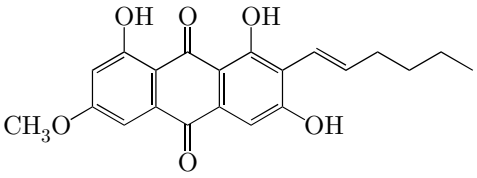
Averantin



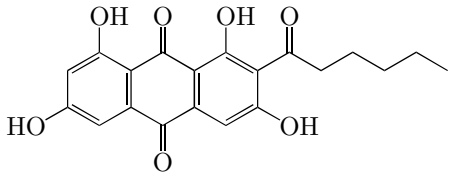
Averythrin



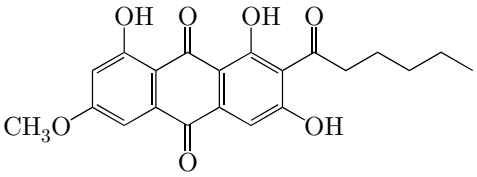
6-*O*-Methylaverantin



6-*O*-Methylaverythrin

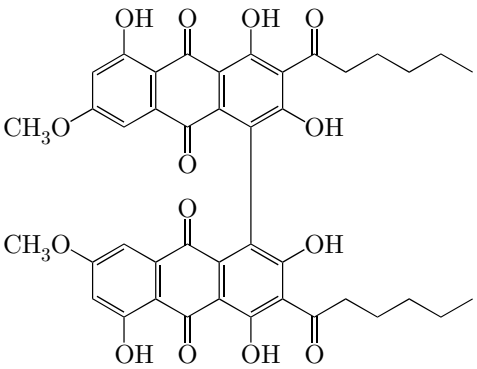


Norsolorinic acid

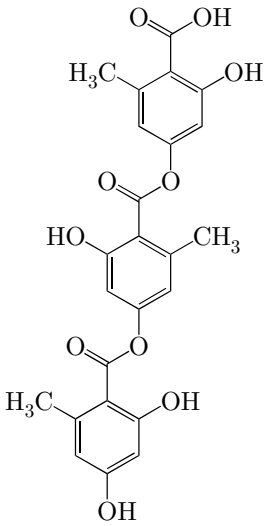


Solorinic acid

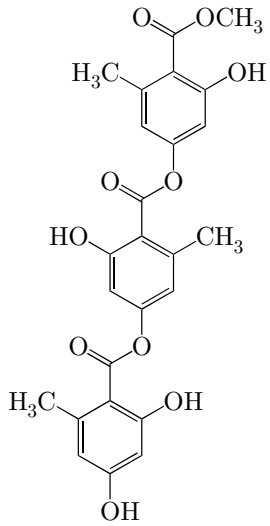
Depsides



4,4'-Disolorinic acid



Gyrophoric acid



Methyl gyrophorate

Figure I.67. Structures of polyketide metabolites isolated from *Solorina crocea*

of variability and diversity of PKS genes in the genomes of both the mycobiont and the cyanobiont (Table I.4) (Miao *et al.*, 2001). Five different tags were obtained from a reference specimen collected at site A1 using KS degenerate primers. Four tags represented the mycobiont's PKS genes since they carried typical fungal introns (EFA, E6B and E6H) or fell into the same cluster with other fungal KS domains in dendrograms based on alignments of deduced amino acid sequences (E6B1). The fifth tag represented a cyanobacterial PKS gene, because the corresponding clone isolated from the *S. crocea* genomic library was completely sequenced and found to be a bacterial type II PKS gene. The absence of amplified products for some PCR reactions suggests that the PKS genes that they were intended to amplify were either rather diverged from the corresponding genes in the reference specimen or were simply not present in the specimen (Miao *et al.*, 2001).

Table I.4. Occurrence of KS tags among populations of *Solorina crocea* in southern British Columbia (Miao *et al.*, 2001)

Site	**KS tag**					**Positive control (*pyr*G gene)**
	EFA	**P6A**	**E6B**	**E6B1**	**E6H**	
A1–1[a]	+[b]	+	+	+	+	+
A1–2	+	+	+	+	+	+
A2	+	–	–	–	+	+
A3	+	–	–	–	+	+
B	–	–	–	–	+	+
C	–	–	–	–	+	+
D	–	–	–	–	–	+
E	–	–	–	–	+	+

[a] Specimens were collected at three locations at site A. DNA extracted from two portions of the same sample collected at location A1 are distinguished as A1–1 and A1–2. Collection sites were distributed to the north and east of Vancouver, at distances from 50 km to 375 km from the city. The locations in site A were within ∼15 km of each other.

[b] "+" – a PCR product of the expected size was observed; "–" – no PCR product was observed.

II

Materials and methods

Bacterial and fungal strains

Escherichia coli DH10B T1 phage resistant (Invitrogen, Carlsbad, CA, USA) was used for bacterial transformations and routine plasmid propagation. *E. coli ccd*B Survival T1 phage resistant (Invitrogen) was used for propagation of the plasmid containing the *ccd*B gene (destination vector pTAex3GS) (Table II.1).

Aspergillus oryzae strain M-2-3 (*arg*B$^-$) (Gomi *et al.*, 1987) was used as a host for lichen PKS expression. It is capable of growing on minimal medium supplemented with arginine or citrulline but not with ornithine and has a spontaneous reversion rate to arginine prototrophy lower than 2×10^{-8}. The strain was obtained from Dr. Russell J. Cox (School of Chemistry, University of Bristol, Bristol, UK).

Table II.1. Genotype of bacterial strains used in this study

Strain	**Genotype**
DH10B T1^R	F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) *ϕ*80*lac*ZΔM15 Δ*lac*X74 *deo*R *rec*A1 *end*A1 *ara*Δ139 Δ(*ara, leu*)*7697 gal*U *gal*K *λ*$^-$ *rps*L (Str^R) *nup*G *ton*A
*ccd*B Survival T1^R	F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) *ϕ*80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*Δ139 Δ(*ara, leu*)*7697 gal*U *gal*K *rps*L (Str^R) *nup*G *ton*A::P$_{trc}$-*ccd*A

Plasmids

The plasmid pSocPKS1 (Andrésson, Davíðsson, *unpublished*) (Appendix A-1) was used for amplification of the full-length *S. crocea* PKS1 gene. pSocPKS1 contains approximately 8.4 kb encoding a polyketide synthase from the lichen *S. crocea* together with approximately 630 bp preceding and 360 bp trailing the coding sequence. Directional TOPO entry vector pENTR/D-TOPO (Figure II.1, Appendix A-2) was from Invitrogen. The expression vector pTAex3 (Fujii *et al.*, 1995) and its destination version pTAex3GS (Figure II.2) were obtained from Dr. Colin Lazarus (School of Biological Sciences, University of Bristol, Bristol, UK). The pTAex3 expression system utilises the α-amylase (Taka-amylase A, *amy*B) promoter which is repressed by glucose and induced by starch.

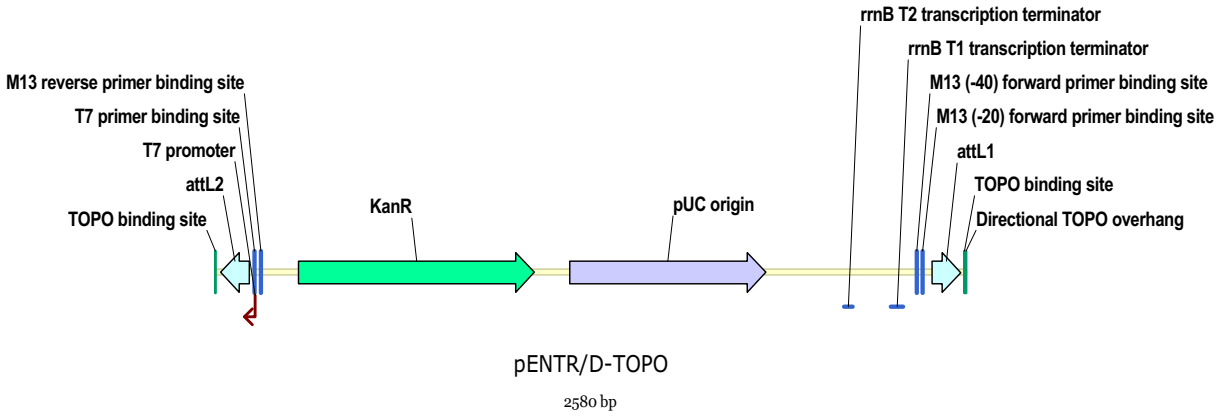


Figure II.1. Directional TOPO vector pENTR/D-TOPO

Chemicals and reagents

All chemicals used in this study were of analytical or molecular biology grade and were purchased from standard commercial sources such as Sigma-Aldrich (St. Louis, MO, USA) and Merck (Darmstadt, Germany) unless otherwise indicated. Glucanex (lysing enzymes from *Trichoderma harzianum*) was a gift from Novozymes (Bagsværd, Denmark). All PCR reagents, nucleases, Antarctic phosphatase, restriction enzymes, buffers and bovine serum albumin (BSA) were from New England Biolabs (Beverly, MA, USA) unless otherwise indicated. Synthetic oligonucleotides were purchased from TAG Copenhagen (Copenhagen, Denmark). Agarose was from Cambrex Bio Science Rockland (Rockland, ME, USA). Culture media components were from Difco Laboratories (Becton Dickinson, Sparks, MD, USA).

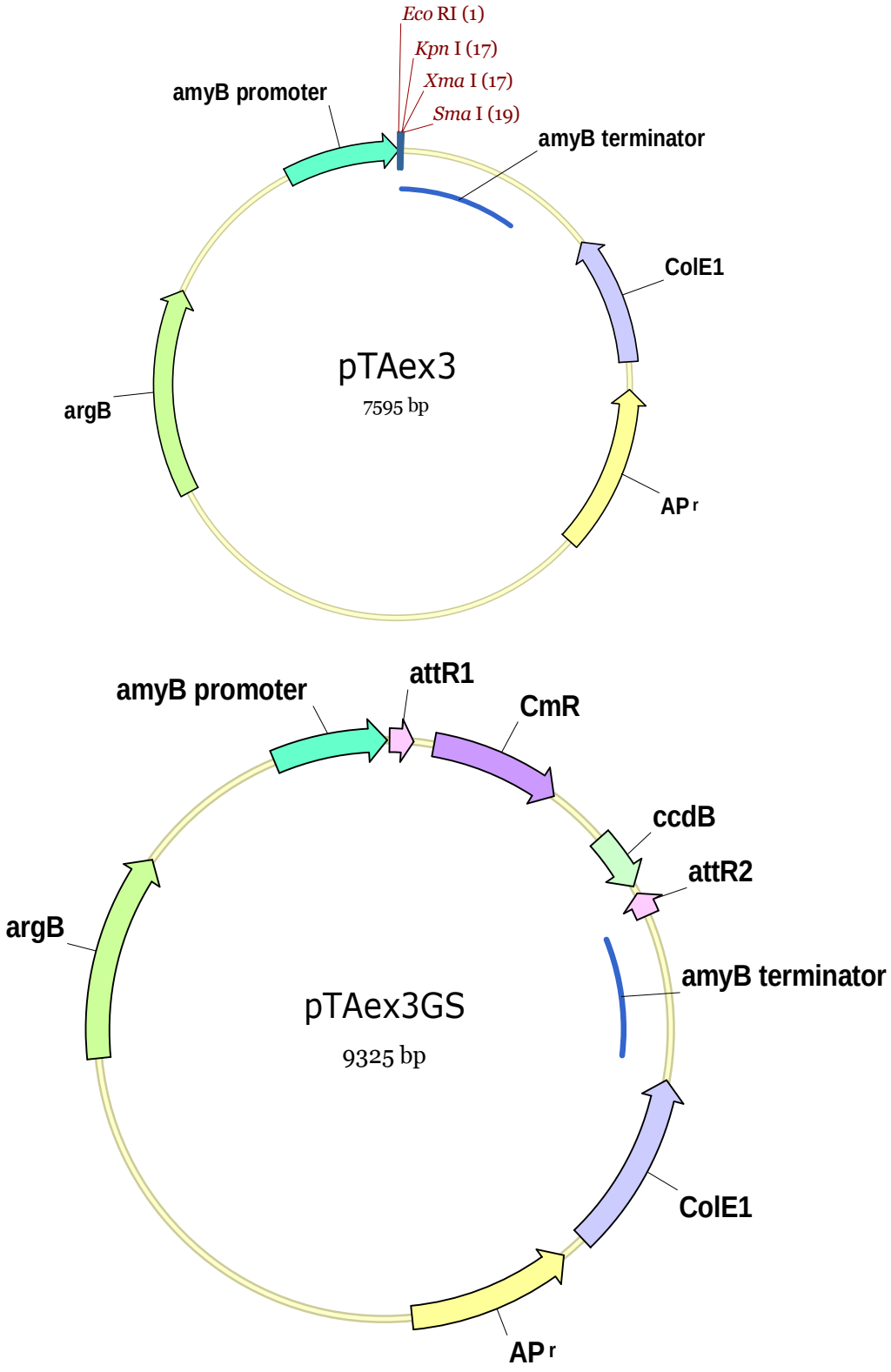


Figure II.2. Expression vector pTAex3 and its destination version pTAex3GS

Culture media

E. coli strains carrying plasmids were grown in Luria-Bertani (LB) medium (0.5% yeast extract, 1% tryptone, 1% NaCl, pH 7.0) and selected with appropriate antibiotics. For LB plates 1.5% agar was added. For recovery of *E. coli* cells after electroporation, SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was used.

Potato-dextrose agar (PDA) supplemented with arginine (potato infusion from 200 g of raw potatoes $\cdot$ liter^{-1} of water, 2% glucose, 1.5% agar, 4 mM arginine) was used for routine cultivation of *A. oryzae* M-2-3. Dextrin-peptone-yeast extract (DPY) medium (2% dextrin, 1% polypeptone, 0.1% yeast extract, 0.5% KH_2PO_4, 0.05% $MgSO_4 \cdot 7H_2O$) was used for mycelia preparations. Czapek-Dox medium (3% sucrose, 0.3% $NaNO_3$, 0.1% $K_2HPO_4 \cdot 3H_2O$, 0.05% KCl, 0.05% $MgSO_4 \cdot 7H_2O$, 0.001% $FeSO_4 \cdot 7H_2O$) and *Aspergillus* minimal medium (1% glucose, 0.6% $NaNO_3$, 0.152% $K_2HPO_4 \cdot 3H_2O$, 0.052% KCl, 0.052% $MgSO_4 \cdot 7H_2O$, 1× Hutner's trace element solution[1]; Barratt *et al.*, 1965) were used as minimal media.

Preparation of electrocompetent cells

Five ml of LB medium were inoculated with a single colony of *E. coli* DH10B $T1^R$ strain and grown overnight at 37°C with moderate shaking at 150 rpm. The overnight preculture was inoculated into 500 ml of LB medium in a 2-liter flask (1 : 100 dilution of the original culture volume) and incubated at 37°C with vigorous shaking at 200–300 rpm until A_{600} reached 0.5–0.6. The culture was immediately chilled in an ice-water bath for 10–15 min and then divided into two prechilled 250-ml centrifuge bottles. The cells were pelleted by centrifugation at 4200 rpm in a Sorvall GS-3 rotor for 20 min at 0–4°C. The supernatant was poured off and the bottles were inverted on paper towels at 4°C to drain off the rest of the supernatant. The cells were resuspended in a volume of ice-cold Milli-Q water equal to the original culture volume and pelleted by centrifugation at 4200 rpm in the same rotor for 20 min at 0–4°C. The supernatant was carefully decanted and the pellet was again resuspended in the same volume of ice-cold Milli-Q water. The cells were centrifuged under the same conditions and the supernatant was carefully decanted. The cells from two 250-ml centrifuge bottles were pooled together in a 50 ml centrifuge tube, resuspended in 40 ml of ice-cold 10% glycerol and centrifuged at 4200 rpm in a Sorvall SS-34 rotor for

[1]Hutner's trace element solution (1000× stock): 2.2% $ZnSO_4 \cdot 7H_2O$, 1.1% H_3BO_3, 0.5% $MnCl_2 \cdot 4H_2O$, 0.5% $FeSO_4 \cdot 7H_2O$, 0.16% $CoCl_2 \cdot 6H_2O$, 0.16% $CuSO_4 \cdot 5H_2O$, 0.11% $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, 5% EDTA, pH 6.5–6.8 (Hutner *et al.*, 1950).

10 min at 0–4℃. The supernatant was decanted and the pellet volume was estimated. An equal volume of ice-cold 10% glycerol was added and the cells were gently resuspended. Aliquots of 40 μl were transferred into prechilled 1.5 ml microcentrifuge tubes, frozen and stored at −80℃.

DNA sequencing

Sequencing with synthetic primers was carried out using an ABI PRISM 3100 Genetic analyzer (Applied Biosystems, Foster City, CA, USA/Hitachi High-Technologies, Japan) with a BigDye Terminator Cycle Sequencing kit v.1.1 (Applied Biosystems).

A 10 μl sequencing reaction contained 150 ng of plasmid DNA template, 1.6 pmole of a primer, 1 μl of TRR mix and 3 μl of 2.5× Dilution buffer. The cycle sequencing was performed in a DNA Engine Tetrad 2 Peltier Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, USA) with 35 cycles at 96℃ for 10 sec, 50℃ for 5 sec and 60℃ for 4 min.

After the cycle sequencing 50 μl of NaOAc-glycogen stock solution (0.1 $\mu g \cdot ml^{-1}$ glycogen (MBI Fermentas, Lithuania), 0.3 M sodium acetate) was added to the sequencing reaction and DNA was precipitated by the addition of 125 μl of cold 100% ethanol. The samples were spun down at 4000 rpm in a Beckman Coulter S5700 rotor for 30 min at 0–4℃. The supernatant was immediately poured off. The samples were wrapped with kimwipes and spun inverted at 300 rpm in the same rotor for 2 min at 0–4℃ to remove the residual ethanol. The DNA pellet was rinsed with 200 μl of cold 70% ethanol and the samples were spun 4000 rpm in the same rotor for 5 min at 0–4℃. The ethanol was immediately discarded and the tubes were spun again inverted and wrapped with kimwipes for 5 min at 300 rpm in the same rotor at 0–4℃. The samples were air dried in the dark for 15 min, dissolved in 10 μl of HiDi Formamide (Applied Biosystems) and loaded onto the analyzer.

For direct sequencing of PCR products 5 μl of PCR reaction was treated with 2.5 units of Antarctic phosphatase and 4 units of Exonuclease I in 1× Antarctic phosphatase reaction buffer (50 mM Bis-Tris-propane · HCl, 1mM $MgCl_2$, 0.1 mM $ZnCl_2$, pH 6.0) in a total reaction volume of 7 μl. The reaction was incubated at 38℃ for 35 min, heat-inactivated at 80℃ for 20 min and 0.8 μl of it was directly used for the sequencing reaction as described above.

The obtained raw sequence data were processed with DNA Sequencing Analysis Software v.5.2 (Applied Biosystems). Contig sequence assembly and alignment were performed using ContigExpress and AlignX, components of Vector NTI Advance v.10.3 Sequence Analysis Software (Invitrogen).

Plasmid construction

Entry clone construction

To improve specificity and yield of the subsequent long PCR reaction, 10 μg of the plasmid pSocPKS1 were digested with 10 units of *Pme*I and 20 units of *Bgl*I (Amersham Pharmacia Biotech, Piscataway, NJ, USA) at 37°C overnight. The reaction was carried out in 1× NEBuffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol) supplemented with 100 $\mu g \cdot ml^{-1}$ BSA in a total volume of 50 μl. The digestion mixture was heat-inactivated at 65°C for 20 min, combined with 6× Crystal violet loading dye (30% glycerol, 20 mM EDTA, 100 $\mu g \cdot ml^{-1}$ crystal violet) and loaded onto a 0.8% agarose gel with 1.6 $\mu g \cdot ml^{-1}$ crystal violet in 1× TAE buffer (50 mM Tris-acetate pH 8.0, 1 mM EDTA). Crystal violet is used to visualize DNA bands during gel electrophoresis in order to avoid damage to large DNA fragments or long PCR products by UV light using conventional ethidium bromide staining, which in turn improves cloning efficiency (Rand, 1996). The gel was run until the dye had migrated about a quarter of the gel and DNA bands had been sufficiently resolved. The DNA band of interest (9.5 kb) containing the whole *S. crocea* PKS1 gene was excised from the gel with a sterile scalpel under normal light, purified using a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol and DNA was finally eluted using 30 μl of prewarmed Tris buffer (10 mM Tris · HCl pH 8.5).

The full-length *S. crocea* PKS1 gene was amplified from the previously purified DNA fragment using Platinum *Pfx* DNA polymerase (Invitrogen) with primers forward GW-PKSF (5′ – C<u>AC CAT GG</u>G TTC GAT GCC GCT TGA – 3′) and reverse GW-PKSR (5′ – CTT CTC AAA T<u>CT A</u>CT TCT CTT CAG – 3′). At the 5′ end the forward primer contains the sequence CACC, which base pairs with the overhang sequence GTCC in the pENTR/D-TOPO vector (Figure II.1) and facilitates directional TOPO cloning. The Kozak translation initiation sequence with the ATG initiation codon (Kozak, 1984, 1987) for proper initiation of translation is shown boxed. The initiation codon in the forward primer and the TAG stop codon in the reverse primer are underlined. The PCR reaction was carried out in 1× *Pfx* Amplification buffer with 1 mM $MgSO_4$, 0.2 mM dNTPs, 0.3 μM of each primer, 120 ng of DNA template, and 5 units of Platinum *Pfx* DNA polymerase in a total volume of 50 μl. PCR was performed in a DNA Engine Tetrad 2 Peltier Thermal Cycler with the following programme: an initial denaturing step 94°C for 2 min, then 10 cycles of 94°C for 10 sec, 55°C for 30 sec, and 68°C for 8 min 30 sec. This was followed by 10 cycles of 94°C for 10 sec, 55°C for 30 sec, 68°C for 8 min 30 sec, increasing extension time by 10 sec each cycle. The final extension step was 30 min to

ensure that all PCR products were completely extended.

The amplified fragment of 8.4 kb was gel purified using crystal violet as described above and TOPO cloned into the pENTR/D-TOPO vector. The TOPO reaction in a final volume of 6 μl contained 4 μl of the gel purified PCR product (100–120 ng), 1 μl of TOPO vector (15–20 ng) and 1 μl of 4-fold diluted salt solution (300 mM NaCl, 15 mM $MgCl_2$ after dilution, supplied with TOPO vector). The reaction was incubated for 30 min at room temperature. Two μl of TOPO reaction were added to 40 μl of DH10B $T1^R$ electrocompetent cells and gently mixed by swirling with a pipette tip. The cells were kept on ice for 1 min and then transferred into a prechilled 0.1 cm electroporation cuvette (Bio-Rad Laboratories). Electroporation was performed using a Cellject Duo electroporator (Thermo Fisher Scientific, Waltham, MA, USA) under the following conditions: output voltage 1800 V, capacitance 25 μF, shunt resistance 335 Ω. Immediately after electroporation the cells were mixed with 1 ml of prewarmed SOC medium, transferred to a 1.5 ml microcentrifuge tube and incubated at 30℃ for 1.5 hours with shaking. One hundred μl aliquots of the electroporation mixture were plated onto LB plates containing 50 $\mu g \cdot ml^{-1}$ kanamycin. The plates were incubated at 30℃ overnight.

Twenty colonies were randomly picked from the plates and cultured in LB medium containing 50 $\mu g \cdot ml^{-1}$ kanamycin at 30℃ overnight with shaking at 200 rpm. Plasmid DNA was isolated using GenElute Plasmid Miniprep Kit (Sigma-Aldrich), according to the manufacturer's protocol, and analyzed by agarose gel electrophoresis. Clones of the appropriate size were subjected to further analysis by digestion with restriction enzymes and sequencing to confirm the presence and correct orientation of the insert. Sequencing was performed with M13 Forward (-20) (5′ – GTA AAA CGA CGG CCA G – 3′) and M13 Reverse (5′ – CAG GAA ACA GCT ATG AC – 3′) primers supplied with the TOPO vector. One of the positive clones, designated pENTR-SocPKS1-16 (Figure II.3), was sequenced again with the primers used for the original PKS1 gene sequencing (Appendix A-3). The primers were chosen in such a way to cover the whole cloned gene with 4× coverage on average. The obtained contig was compared with the original sequence. This was done to test for possible errors introduced by *in vitro* amplification using PCR.

Expression plasmid construction

The entry clone pENTR-SocPKS1-16 was then subjected to a LR recombination reaction with pTAex3GS destination vector (Figure II.2) using Gateway LR Clonase enzyme mix (Invitrogen). A 10 μl LR reaction contained 2 μl of 5× LR Clonase reaction buffer, 150 ng of both the entry clone and destination vector, 1 μl of Gateway LR Clonase enzyme mix and TE buffer (10 mM Tris · HCl pH 8.0, 1 mM EDTA) up to 10 μl. The reaction

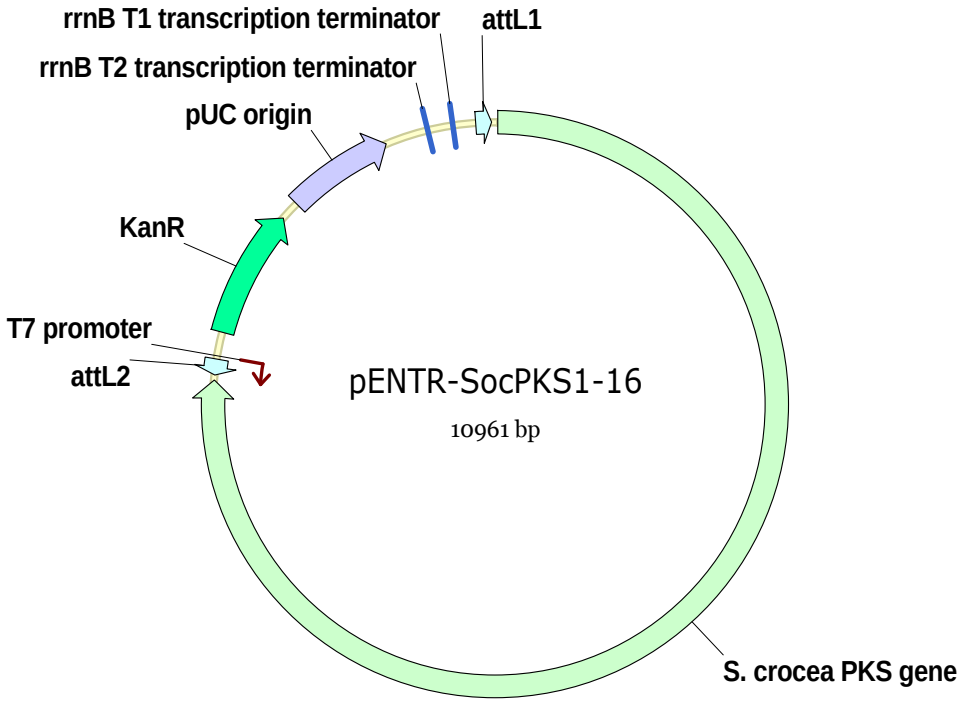


Figure II.3. Entry clone pENTR-SocPKS1-16

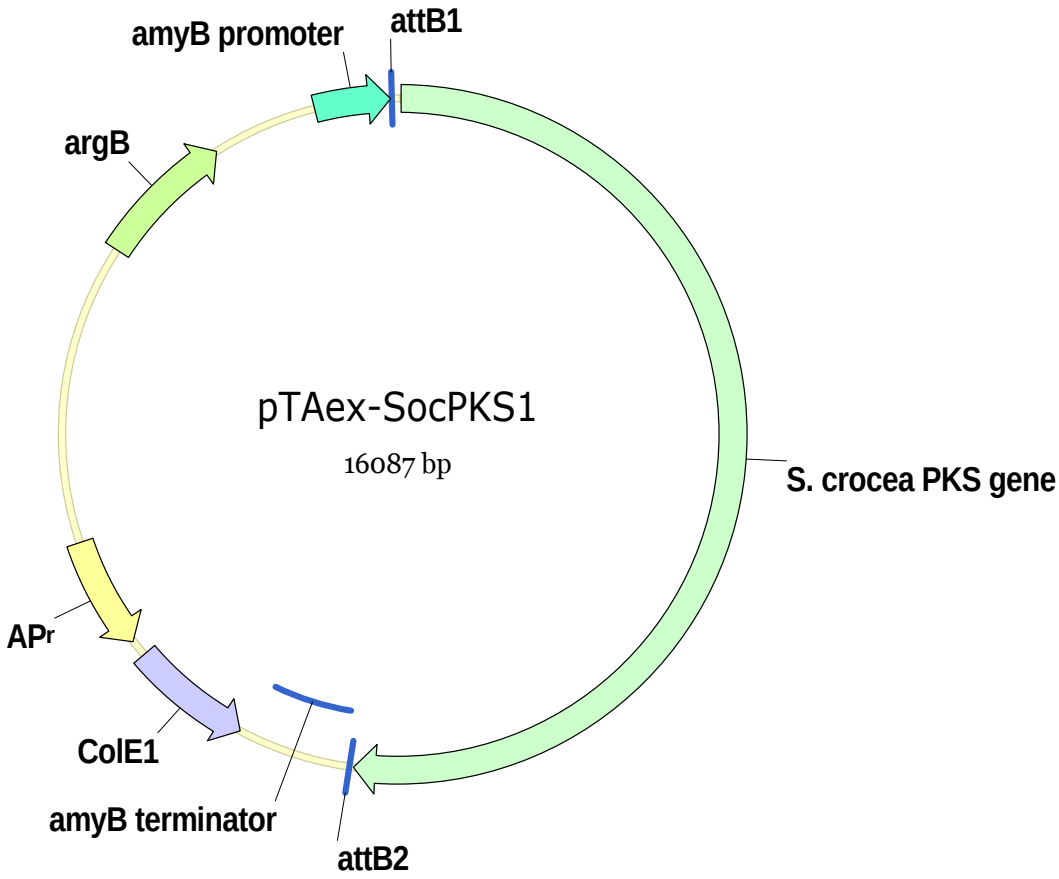


Figure II.4. Expression plasmid pTAex-PKS1

was incubated at 25℃ for 2 hours and terminated by adding 1 μl of Proteinase K solution (supplied with Gateway LR Clonase enzyme mix) and then incubating at 37℃ for 10 min. One μl of the LR reaction was then transformed into 40 μl of DH10B $T1^R$ electrocompetent cells by electroporation as described above. One hundred μl aliquots of the electroporation mixture were plated onto LB plates containing 100 $\mu g \cdot ml^{-1}$ ampicillin. The plates were incubated at 30℃ overnight.

Ten colonies were randomly picked from the plates and cultured in LB medium containing 100 $\mu g \cdot ml^{-1}$ ampicillin at 30℃ overnight with vigorous shaking at 200 rpm. Plasmid DNA was isolated using GenElute Plasmid Miniprep Kit and analyzed by digestion with restriction enzymes. The resulting expression plasmid pTAex-SocPKS1 (Figure II.4) was used to transform the fungal host *A. oryzae* M-2-3 by standard $CaCl_2$/PEG-mediated genetic transformation of protoplasts.

Protoplasts preparation

A. oryzae M-2-3 was grown on PDA plates at 30℃ until conidiation (usually about 2 weeks). Conidia from one well sporulating plate were inoculated into 200 ml of DPY medium. The cultures were incubated for 20–22 hrs at 25℃ with shaking at 150 rpm. Protoplasts were prepared from the cultures in which the fungi grew as fine pellets.

Mycelium was collected by vacuum filtration through sterile Whatman №1 filter paper and washed once with sterile distilled water and twice with 0.8 M NaCl. The pellet was then resuspended in 20 ml of filter-sterilized protoplasting solution (50 $mg \cdot ml^{-1}$ Glucanex, 0.8 M NaCl, 10 mM Na-phosphate buffer pH 6.0) and incubated at 25℃ for 2.5–3 hours with gentle agitation at 100 rpm.

Protoplasts were separated from the mycelium by filtration through sterile Miracloth (Calbiochem, La Jolla, CA, USA). The filtrate was centrifuged at 3500 rpm in a Beckman Coulter JS-5.3 rotor for 10 min. The pelleted protoplasts were washed twice with 0.8 M NaCl and once with transformation buffer (0.8 M NaCl, 10 mM $CaCl_2$, 10 mM Tris $\cdot$ HCl pH 7.5), counted in a Improved Neubauer hemacytometer (Bright-Line, Warner-Lambert Technologies, Buffalo, NY, USA) and resuspended in the same buffer at the concentration of 2.5×10^8 protoplasts $\cdot ml^{-1}$.

Fungal transformation

The protoplasts were divided into 200 μl aliquots, to which circular plasmid DNA (5–10 μg in a total volume of no more than 20 μl) was added. In addition one reaction was done with the plasmid pTAex3 and another one without DNA as a negative control. The protoplasts

were incubated on ice for 5–10 min. Then 1.2 ml of polyethylene glycol solution (60% PEG 4000 (*w/v*), 50 mM $CaCl_2$, 10 mM Tris · HCl pH 7.5) was added in three steps (0.2, 0.2 and 0.8 ml) with gentle agitation by hand after each addition. The protoplasts were incubated for 20 min at room temperature. After dilution with a 10-fold volume of transformation buffer the protoplasts were collected by centrifugation at 3500 rpm in a Beckman Coulter JS-5.3 rotor for 10 min. The supernatant was carefully decanted and the pellet was gently resuspended in a small volume of the same buffer. Aliquots of 500 μl were mixed with 5 ml of soft agar (Czapek-Dox medium with 0.5% agar and 1 M sorbitol, equilibrated to 45℃) and immediately poured onto prewarmed plates with 10 ml of solidified Czapek-Dox medium containing 1.5% agar and 1 M sorbitol. The plates were left to sit for 15–20 min at room temperature and then incubated inverted at 25℃ for 7–10 days. The obtained transformants were transferred onto new plates with minimal medium and then subcultured twice through conidia on the same medium at 30℃.

Fungal DNA isolation

A. oryzae transformants were grown on minimal media plates at 30℃ until conidiation. Conidia from the sporulating plates were inoculated into 200 ml of DPY medium. The cultures were incubated for 2 days at 30℃ with shaking at 150 rpm. Mycelia were collected by vacuum filtration through sterile Miracloth and washed several times with chilled sterile distilled water, followed by continuous vacuum until dry. Residual water was removed by blotting the mycelial mats between sterile paper towels. The mycelia were then used fresh or stored frozen at −80℃ until ready for use. For DNA isolation, the mycelia were frozen in liquid nitrogen and ground to fine powder using a pre-cooled mortar and pestle. The mycelial powder (approximately 100–150 mg) were gently resuspended in 750 μl of cold spermidine-SDS buffer (4 mM spermidine, 10 mM EDTA, 0.1 M NaCl, 0.5% SDS, 10 mM β-mercaptoethanol, 40 mM Tris · HCl pH 8.0; Borges *et al.*, 1990) in a 2-ml microcentrifuge tube. The mixture was then immediately extracted with an equal volume of phenol[2]/chloroform/isoamyl alcohol (ΦOH/Chl/IAA) (25:24:1), gently mixed by inversion and spun at maximum speed in a tabletop microcentrifuge for 15 min. The aqueous phase was removed to a new tube and re-extracted once with an equal volume of ΦOH/Chl/IAA and once with an equal volume of Chl/IAA (24:1), with spinning between extractions at maximum speed in a tabletop microcentrifuge for 10 min. The DNA was then precipitated with 0.1 volume of 3 M sodium acetate, pH 5.2, followed by addition of one volume of 100% isopropanol, and pelleted by spinning at maximum speed in a

[2]Tris-equilibrated phenol, pH 8.0.

tabletop microcentrifuge for 10 min. The pellet was washed twice with 70% ethanol, air dried for 15 min and dissolved in 100 μl of TE buffer (10 mM Tris $\cdot$ HCl pH 8.0, 1 mM EDTA) with 25 $\mu g \cdot ml^{-1}$ RNase A (Sigma-Aldrich). The samples were incubated at 37℃ for 1 hour. The DNA was again precipitated with 0.1 volume of 3 M sodium acetate, pH 5.2, followed by addition of 2.5 volumes of cold 100% ethanol, and pelleted by spinning at maximum speed in a tabletop microcentrifuge for 10 min. The pellet was washed twice with 70% ethanol, air dried for 15 min and dissolved in 100 μl of TE buffer. Purity and average fragment size were checked by agarose gel electrophoresis with $\lambda/$*Hin*dIII molecular weight marker. The isolated DNA was stored at −20℃ for later analysis.

PCR screening of the transformants

To verify the integration of the *S. crocea* PKS1 gene into the fungal genome, the isolated genomic DNA was used as a template in PCR reactions with 4 different pairs of specific primers chosen in such a way as to span 1–1.2 kb fragments from different sections of the PKS gene (Table II.2, Figure II.5). PCR reactions were carried out in 1× ThermoPol reaction buffer (20 mM Tris·HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100) with 0.2 mM dNTPs, 0.3 μM of each primer, 100–500 ng of DNA template, and 2 units of *Taq* DNA polymerase in a total volume of 20 μl. PCR was performed in a DNA Engine Tetrad 2 Peltier Thermal Cycler with the following programme: an initial denaturing step 94℃ for 2 min, then 30 cycles of 94℃ for 10 sec, 55℃ for 30 sec, and 72℃ for 1 min, followed by a final extension step of 7 min.

Table II.2. Primer pairs used for PCR screening of the transformants

№	Primer pair[a]	The size of the amplicon (bp)
1	B8R1 – PKSR1	1074
2	BF1685 – BR2760	1113
3	BF3417 – PKSR2	1295
4	Ex3-6 – K34nt1	1261

[a] For primer sequences see Appendix A-3.

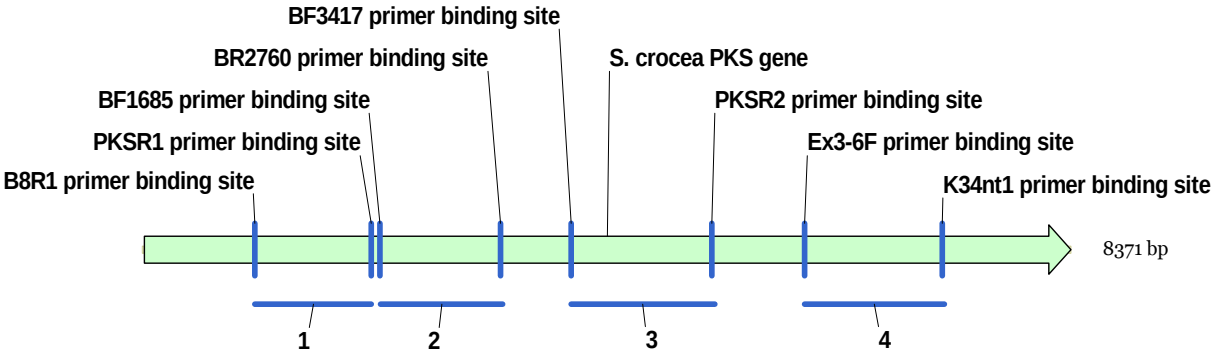


Figure II.5. The location of the primer binding sites used for PCR screening of the transformants and the fragments (indicated with numbers 1–4) flanked by these sites.

PCR with internal transcribed spacer (ITS) region primers ITS1 (5′ – TCC GTA GGT GAA CCT GCG G – 3′) and ITS4 (5′ – TCC TCC GCT TAT TGA TAT GC – 3′) (White *et al.*, 1990) was performed in parallel as a control to evaluate the quality of isolated DNA. The primers span ITS1 and ITS2 regions flanking the DNA sequence encoding the 5.8S rRNA (Figure II.6). According to the sequenced genome of *Aspergillus oryzae* RIB40 (accession number AP007172; Machida *et al.*, 2005), the size of the fragment amplified with the primers ITS1 and ITS4 should be 595 bp. The PCR was performed as described above in this section, except the extension time was 30 sec.

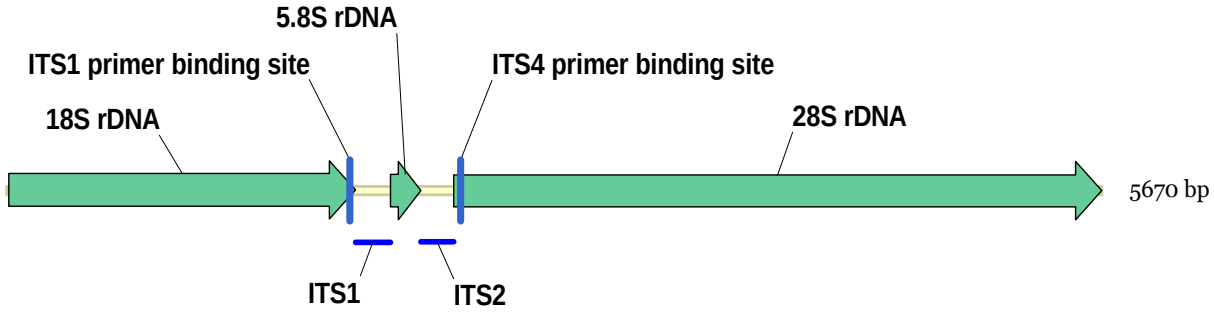


Figure II.6. Region of the ribosomal DNA of *Aspergillus oryzae*. The map represents the genes encoding 18S, 5.8S, and 28S ribosomal RNA (rRNA), intervened by two internal transcribed spacers ITS1 and ITS2.

The PCR products were analyzed on 1.5% agarose gels containing 0.5 $\mu g \cdot ml^{-1}$ ethidium bromide in 0.5× TBE buffer (45 mM Tris-borate, 1 mM EDTA).

Expression of the *S. crocea* PKS gene

To verify the expression of the *S. crocea* PKS1 gene, 10 pTAex-SocPKS1 transformants were grown in liquid Czapek-Dox medium with 2% glucose as a sole carbon source at 25°C and 200 rpm for 4 days. Five pTAex3 transformants and *A. oryzae* M-2-3 (wild type) were

grown in parallel as a control. Mycelia were then collected by vacuum filtration through sterile Miracloth, washed with sterile distilled water and transferred into Czapek-Dox medium with 2% starch as a sole carbon source. The cultures were grown under the same conditions for further 1–2 days depending on application.

RNA isolation

Total RNA was prepared from the fungal cultures after 1 day induction with starch. Mycelia were collected by vacuum filtration through sterile Miracloth, flash frozen in liquid nitrogen and ground to fine powder using a pre-cooled mortar and pestle. Approximately 100 mg of the mycelial powder were resuspended in 1 ml of TRIzol reagent (Life Technologies, Rockville, MD, USA) and the samples were processed according to the manufacturer's protocol. The resulting RNA was dissolved in 100 μl of DEPC-treated water[3] and stored at -80℃ for later analysis.

Prior to Reverse-Transcriptase PCR, the isolated RNA was treated with DNase I (RNase-free) to get rid of residual genomic DNA. A 10 μl reaction contained 1 μg of RNA, 1 unit of DNase I, and 10 units of RiboLock Ribonuclease Inhibitor (MBI Fermentas) in 1× DNase I reaction buffer (10 mM Tris $\cdot$ HCl, 2.5 mM $MgCl_2$, 0.1 mM $CaCl_2$). The reaction was incubated at 37℃ for 30 min and terminated by adding 1 μl 25 mM EDTA and then incubating at 65℃ for 10 min. The prepared RNA was directly used as a template for the subsequent reverse transcriptase reaction.

Reverse-Transcriptase PCR

First strand cDNA was synthesized from 1 μg of total RNA treated with DNase I in a final volume of 20 μl using RevertAid First Strand cDNA Synthesis Kit (Fermentas) with random hexamer primers according to the manufacturer's protocol and used directly as a template in PCR with 8 pairs of *S. crocea* PKS1 gene specific primers to cover the entire cDNA. PCR reactions were carried out in 1× ThermoPol reaction buffer with 0.2 mM dNTPs, 0.3 μM of each primer, 1 μl of cDNA template from reverse transcriptase reaction, and 2 units of *Taq* DNA polymerase in a total volume of 20 μl. PCR was performed in a DNA Engine Tetrad 2 Peltier Thermal Cycler with the following programme: an initial denaturing step 94℃ for 2 min, then 35 cycles of 94℃ for 10 sec, 55℃ for 30 sec, and 72℃ for 1–1.2 min[4], followed by a final extension step of 7 min.

[3] 0.1% Diethudiethylpyrocarbonate (DEPC) solution, incubated at 37℃ overnight and then autoclaved for 15 minutes at 121℃ on liquid cycle to get rid of residual DEPC.

[4] Depending on the size of the amplified fragment (Table III.6)

Table II.3. Primer pairs used for Reverse-Transcriptase PCR screening of the transformants

№	Primer pair[a]	Spans putative intron(s)	The size of the amplicon (bp)[b]
1	GW-PKSF – B8AAF	Yes	1421
2	B8R1 – PKSR1	Yes	1074
3	B8AAR – Y13R	No	1349
4	B8FCY – 15-9FF	No	1460
5	Ex2-3F – PKSR2	No	1192
6	PKSF2 – BR5524	No	1234
7	Ex3-6F – K34nt1	Yes	1261
8	K34R1 – GW-PKSR	Yes	1256

[a] For primer sequences see Appendix A-3.
[b] Deduced from genomic DNA sequence.

PCR with total RNA treated with DNase I but not subjected to reverse transcriptase reaction was carried out in parallel as a negative control with actin gene specific primers forward AspActF (5′ – CCC AAG TCC AAC CGT GAG AA – 3′) and reverse AspActR (5′ – GGC CTG GAT GGA GAC GTA GA – 3′).

The PCR products were analyzed on 1.5% agarose gels containing 0.5 $\mu g \cdot ml^{-1}$ ethidium bromide in 0.5× TBE buffer. Eight cDNA fragments obtained from the positive transformants were directly sequenced with the primers used to amplify the fragments. Exon-intron boundaries of the *S. crocea* PKS1 gene were established by aligning the cDNA sequences with the genomic sequence.

III

Results and discussion

Slow growth and low metabolic activity of lichens and their mycobionts prohibit studies of lichen polyketide biosynthesis in its natural hosts. However, molecular genetic techniques using genomic library construction, PCR and heterologous expression in surrogate hosts provide an alternative approach to access lichen PKS genes and examine features of polyketide biosynthesis in this largely unexplored group of organisms.

The objectives of this study were expression of a polyketide synthase gene from the lichen *Solorina crocea* in the filamentous fungus *Aspergillus oryzae*, analysis of the products of its expression and determination of the gene structure: exon–intron boundaries, 5′- and 3′-prime ends, the native promoter region and the domain organization of the deduced protein sequence.

Plasmid construction

To construct the plasmid for expression of the *S. crocea* PKS1 gene in a heterologous host, the gene was transferred from the plasmid pSocPKS1 (Appendix A-1) into the pTAex3 expression system (Figure II.2) using Gateway cloning technology (Invitrogen). The conventional cloning techniques using restriction enzymes and ligases are time-consuming and tedious and subcloning of such large sequences is hampered by the lack of suitable restriction sites.

The entire 8.4 kb *S. crocea* PKS1 gene was amplified using a proofreading polymerase with flanking primers GW-PKSF and GW-PKSR, containing translation initiation and stop codons respectively. The choice of the ATG translation initiation codon of the *S. crocea* PKS1 coding sequence for design of the primer GW-PKSF was not confirmed by experiments involving the rapid amplification of cDNA ends (RACE) and was based on consensus sequences surrounding the translation initiation site identified for eukaryotes and filamentous fungi in particular (Figure III.7, see the section "mRNA processing and

translation") (Ballance, 1986; Gurr *et al.*, 1987; Kozak, 1984; Turner *et al.*, 1993; Unkles, 1992). The amplified fragment was TOPO cloned into the pENTR/D-TOPO vector (Figure II.1) to obtain the entry clone.

The sequencing of the positive entry clone, designated pENTR-SocPKS1-16 (Figure II.3), revealed the presence and correct orientation of the insert and its *error-free* sequence. The absence of errors after *in vitro* amplification of the insert using PCR was achieved by the use of a proofreading polymerase, reduced total number of cycles (20), relatively high DNA template concentration ($\sim$2.4 ng $\cdot$ μl^{-1}) and lower concentration of dNTPs (200 μM, instead of 300 μM suggested by the manufacturer).

The expression plasmid pTAex-SocPKS1 was constructed by a LR reaction between the entry clone pENTR-SocPKS1-16 and the vector pTAex3GS, a destination version of the plasmid pTAex3 (Figure III.1). pTAex-SocPKS1 contains the *S. crocea* PKS1 gene under control of the inducible α-amylase (*amy*B) promoter and terminator and the whole *Aspergillus nidulans arg*B gene as an auxotrophic marker for selection of transformants. The resulting plasmid was introduced into the host fungus *A. oryzae* M-2-3 by standard $CaCl_2$/PEG-mediated genetic transformation of protoplasts.

The plasmid pTAex-SocPKS1 was not linearized with restriction enzymes prior to transformation and its integration into the host chromosomal DNA after entering the cell could occur by means of both non-homologous (ectopic) and homologous recombination (Fincham, 1989). Ectopic integration at any site can lead to disruption of the transformed gene of interest and therefore to its nonfunctionality. Homologous integration should occur at a single site via crossover at *arg*B with the transformed PKS gene lying between the 5′ end of the mutated *A. oryzae arg*B region (*arg*B$^-$) and the 5′ end of the selection marker gene (*arg*B$^+$) of *A. nidulans*.

Homologous integration was not checked by PCR or Southern hybridization analysis with DNA isolated from the transformants but limited homology of the *A. nidulans arg*B gene with its equivalent in *A. oryzae* (73% identity of coding sequences) suggests that integration of the plasmid DNA into the host genome most likely occurred by means of non-homologous rather than homologous recombination.

Fungal transformation

Auxotrophic *Aspergillus oryzae* M-2-3 was transformed to prototrophy with the expression plasmid pTAex-SocPKS1 and selected on arginine deficient medium. Transformation with the plasmid pTAex3 was carried out in parallel to obtain "blank" *arg*B$^+$ transformants containing no PKS gene insert and for use as a control in later analysis.

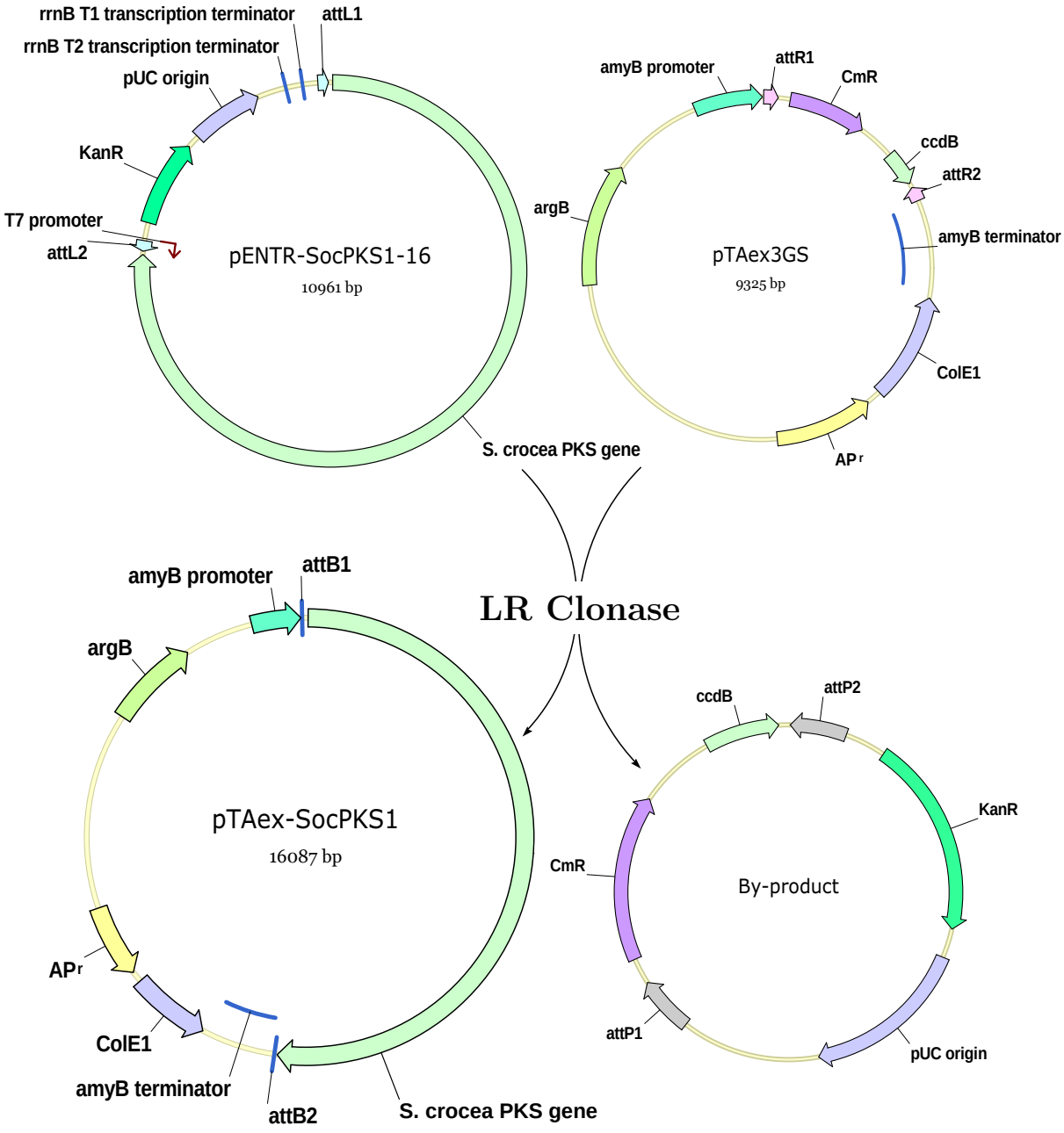


Figure III.1. Gateway LR reaction. Under the action of LR Clonase enzyme, site-specific recombination between *att*L and *att*R sites (*att*R × *att*L → *att*B × *att*P) generates an expression plasmid and a by-product. Selection of the expression plasmid is achived by transformation of the LR reaction mixture into *E. coli* strains that do not contain an F′ episome. Only plasmids without the *ccd*B gene and with ampicillin resistance (AP^r) yield colonies on plates with ampicillin.

After the incubation of the transformation plates a few large colonies were observed together with several small ones (Figure III.2). No colonies appeared on minimal medium in the experiment with no plasmid DNA. Most of the small colonies showed stunted growth and were thought to be so called *abortive* transformants. The small colonies failed to grow when they were subcultured onto plates with minimal medium, whereas the large colonies exhibited vigorous growth. Similar observations have been made previously on the transformation of other fungal species (Buxton *et al.,* 1985; Gomi *et al.,* 1987; Tilburn *et al.,* 1983; Yelton *et al.,* 1984). This phenomenon still remains unclear and abortive transformants are thought to be derived from protoplasts that have failed to integrate plasmid DNA into the genome. Some growth can take place due to transient expression of the *arg*B gene on the unintegrated plasmid but it is rapidly lost in the absence of integration and the colony stops growing (Buxton *et al.*, 1985).

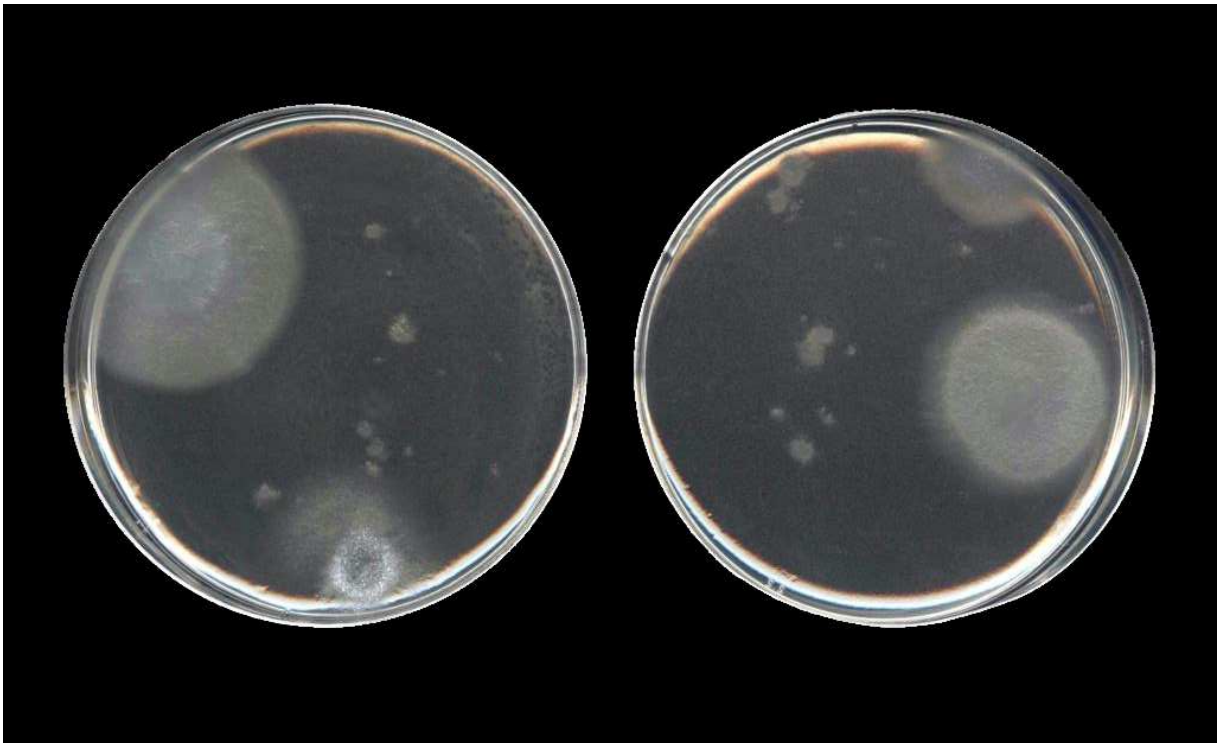


Figure III.2. Stable and abortive transformants. Petri dishes containing minimal selective medium show examples of large and vigorously growing stable transformants and very small abortive transformants. Czapek-Dox medium with 1 M sorbitol, 25℃, 1 week.

Counting the large colonies as real transformants, the average transformation frequency was approximately 2 transformants per μg of plasmid DNA. The frequency in this study was much lower than that in general obtained for *A. nidulans* but comparable with results from the transformation experiments with *A. niger* and *A. oryzae* using plasmids containing the *A. nidulans arg*B gene (Table III.1).

Table III.1. Transformation frequencies (number of transformants $\cdot \mu g^{-1}$ of plasmid DNA) obtained with *Aspergillus* species using different nutritional selective markers[a]

Species	Selection marker	Transformation frequency	Reference
A. nidulans	*amd*S	25	Tilburn *et al.,* 1983
A. nidulans	*trp*C	20–300	Yelton *et al.,* 1984
A. nidulans	*arg*B	50	John, Peberdy, 1984
A. nidulans	*arg*B	500	Johnstone *et al.,* 1985
A. nidulans	*pyr*4	50–150	Ballance, Turner, 1985
A. nidulans	*amd*S	300–400	Wernars *et al.,* 1985
A. niger	*arg*B	0.6–4	Buxton *et al.*, 1985
A. niger	*amd*S	3.75	Kelly, Hynes, 1985
A. niger	*pyr*G	50	Goosen *et al.*, 1987
A. oryzae	*arg*B	0.7	Gomi *et al.,* 1987
A. oryzae	*pyr*G	16	Mattern *et al.,* 1987
A. oryzae	*amd*S	5–30	Christensen *et al.,* 1988
A. oryzae	*arg*B	5–10 (90% abortive)	Hahm, Batt, 1988
A. oryzae	*arg*B	2	This study

[a] Only results of integrative transformations with *non*-autonomously replicating plasmids are listed.

Forty pTAex-SocPKS1 and five pTAex3 transformants exhibiting vigorous growth were transferred onto new plates with minimal medium and then subcultured twice through conidia on the same medium. This was done to ensure the mitotic stability of the transformants, since *A. oryzae* conidia are multinucleate (generally 4 nuclei per conidium) (Ishitani, Sakaguchi, 1955) and protoplasts derived from them can give rise to heterokaryotic transformants, in which nuclei containing the integrated plasmid and nuclei not having such sequences may coexist. During mitotic divisions and conidiogenesis homokaryotic conidia with a stable transformed phenotype are formed.

No remarkable morphological differences were observed comparing the obtained pTAex-SocPKS1 transformants with the original *A. oryzae* M-2-3 strain except for an irregular or wavy colony front of some transformants. There were also slight differences among the

transformants in the speed of growth and the timing of sporulation. These differences might be due to random integrative inactivation of different chromosomal genes.

Subculturing the transformants on *Aspergillus* minimal medium instead of Czapek-Dox medium allowed their more rapid and abundant sporulation, which may be due to the presence of essential microelements and a lower C : N ratio in this medium.

PCR screening of the transformants

In order to evaluate the quality of DNA isolated for PCR analysis it was used as a template in PCR with ITS primers flanking internal transcribed spacers in the ribosomal DNA region (Figure II.6). The PCR fragment of the expected size was obtained with a good yield for all samples tested (Figure III.3). Thus, the quality of isolated DNA was suitable for subsequent PCR analysis with *S. crocea* PKS1 gene specific primers.

PCR with 4 different primer pairs specific for the *S. crocea* PKS1 gene (Table II.2, Figure II.5) yielded fragments of the expected sizes for all 40 pTAex-SocPKS1 transformants screened. No PCR products were observed for pTAex3 transformants and *A. oryzae* M-2-3 (wild type) (Figure III.4). These results confirm integration of the plasmid pTAex-SocPKS1 into the genome of the transformants.

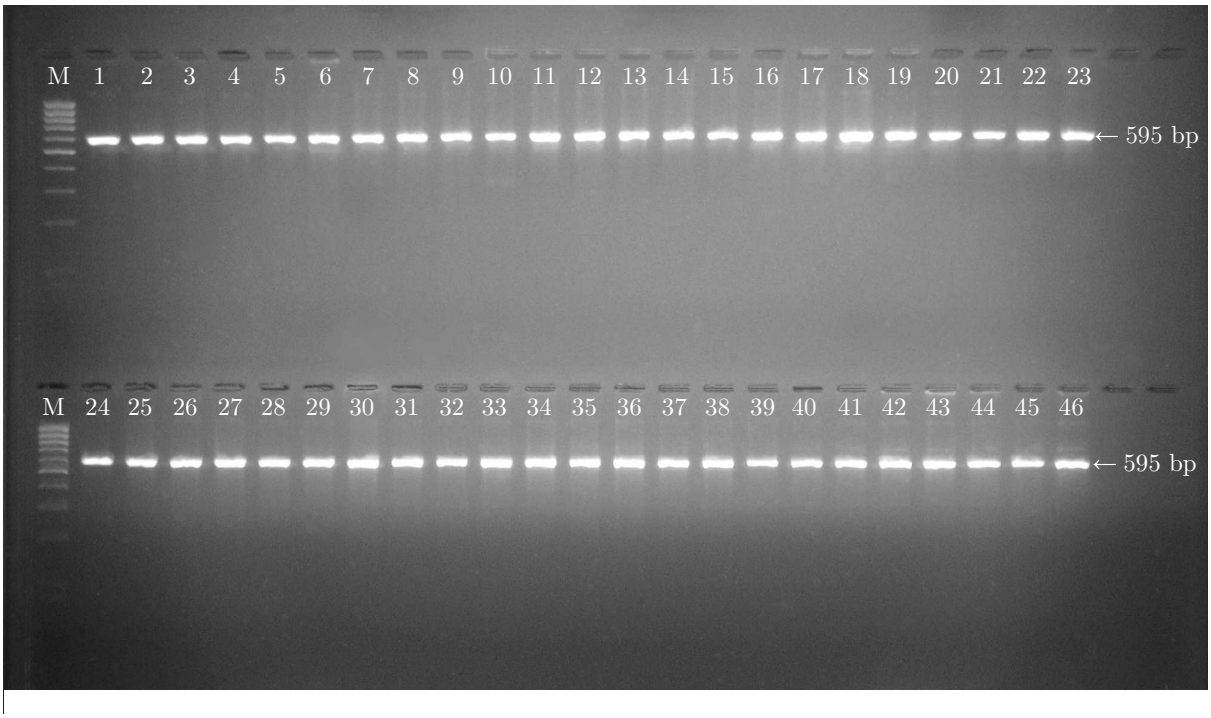
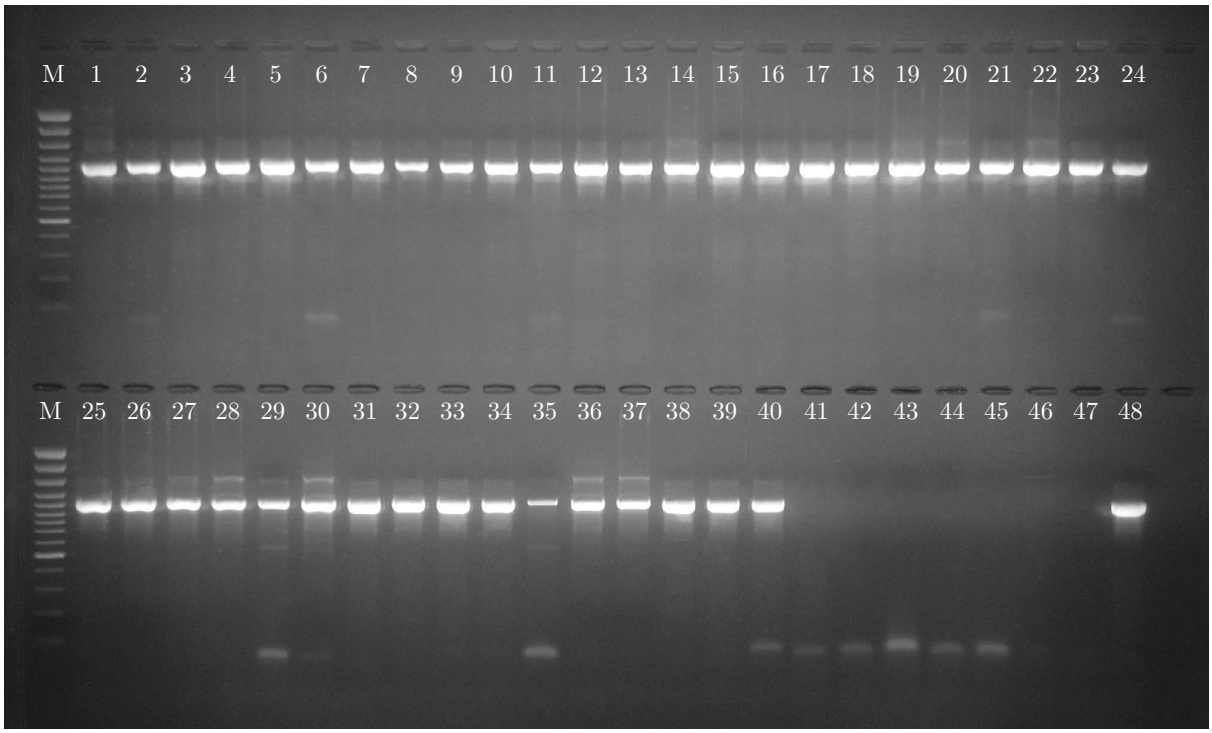
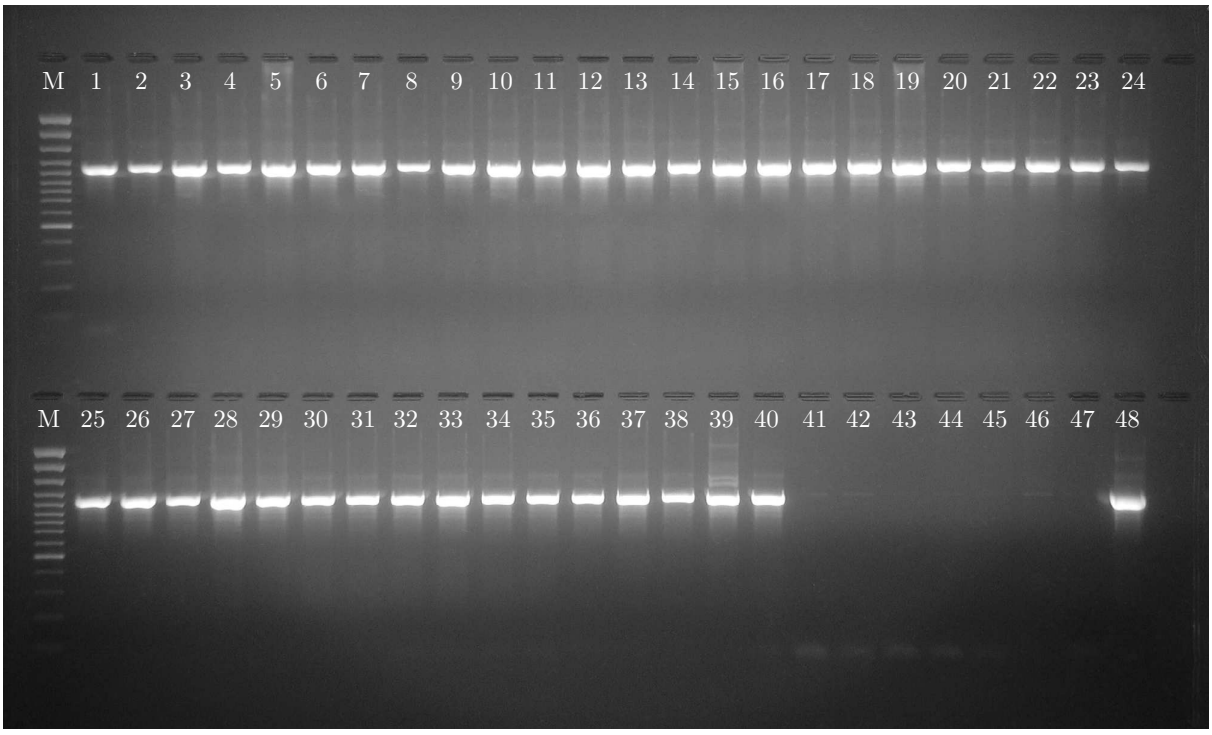


Figure III.3. Agarose gel electrophoresis of PCR products using genomic DNA as a template and ITS primers. The expected size of the amplified fragment is 595 bp. Lanes 1–40 – pTAex-SocPKS1 transformants, lanes 41–45 – "blank" pTAex3 transformants, lane 46 – *A. oryzae* M-2-3 (wild type), M – MassRuler DNA Ladder, Low Range (MBI Fermentas).

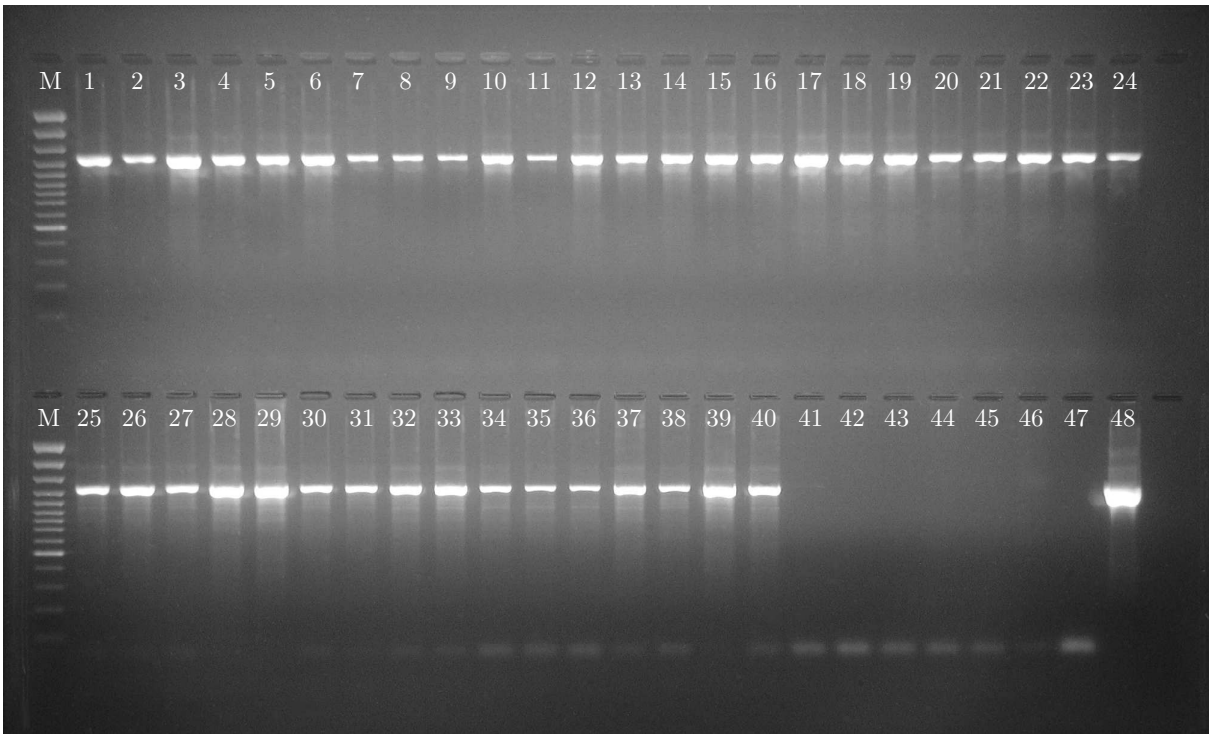


(a)

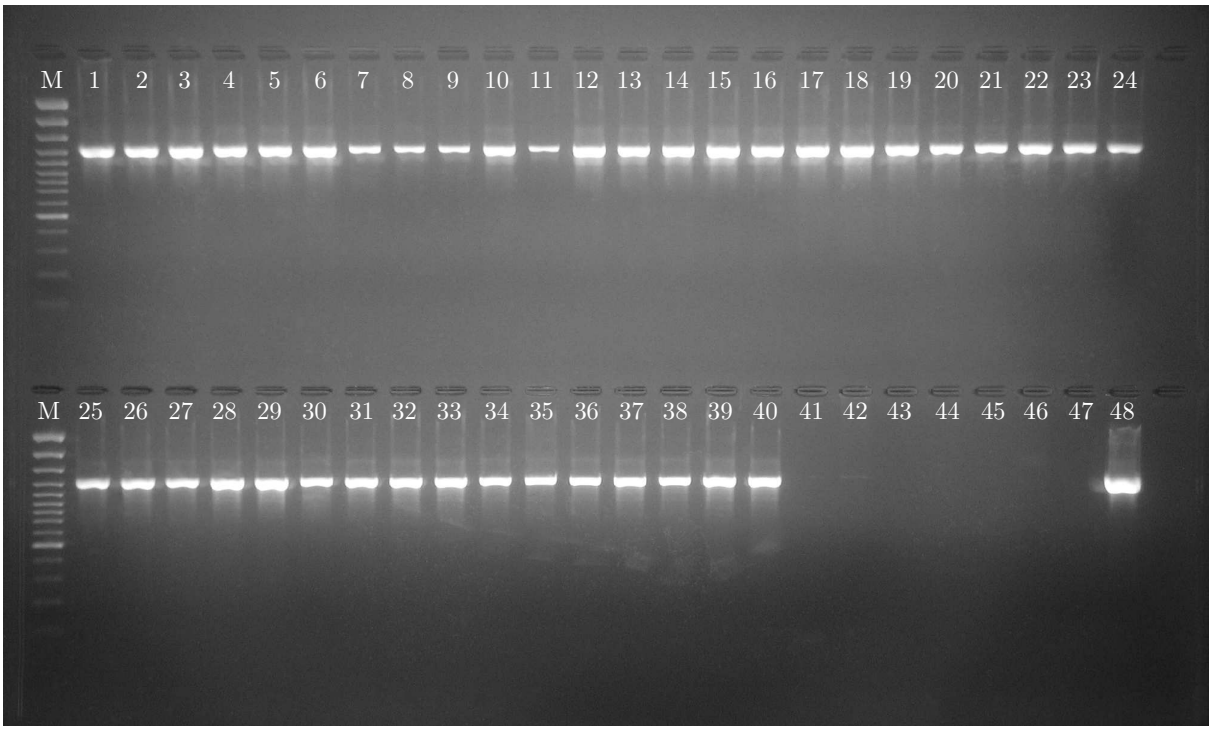


(b)

Figure III.4. Agarose gel electrophoresis of PCR products using genomic DNA as a template and *S. crocea* PKS1 gene specific primers: (a) B8R1 – PKSR1; (b) BF1685 – BR2760; (c) BF3417 – PKSR2; (d) Ex3-6 – K34nt1. Lanes 1–40 – pTAex-SocPKS1 transformants, lanes 41–45 – "blank" pTAex3 transformants, lane 46 – *A. oryzae* M-2-3 (wild type), lane 47 – negative control (*no* DNA), lane 48 – positive control (plasmid pTAex-SocPKS1), M – GeneRuler 100 bp DNA Ladder Plus (MBI Fermentas).



(c)



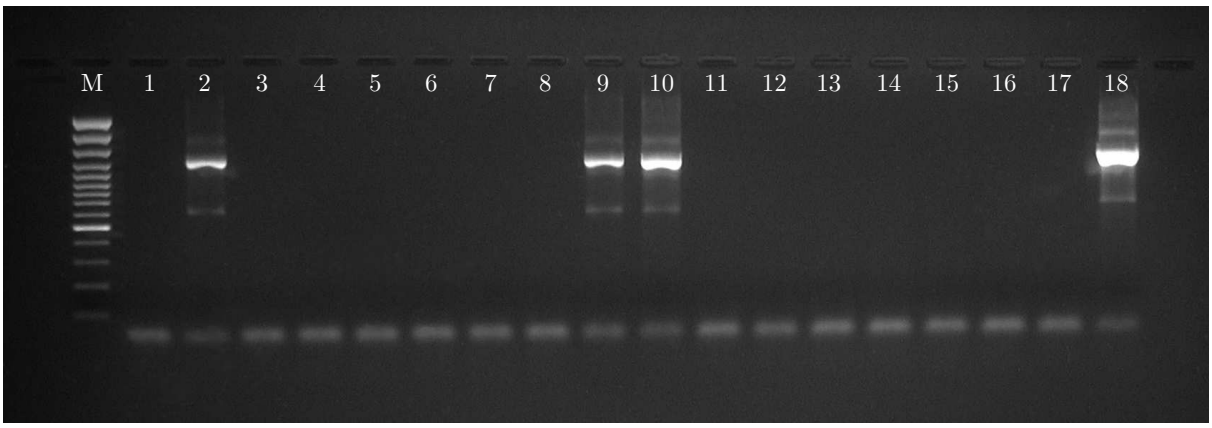
(d)

Figure III.4. *Continued*

Reverse-Transcriptase PCR

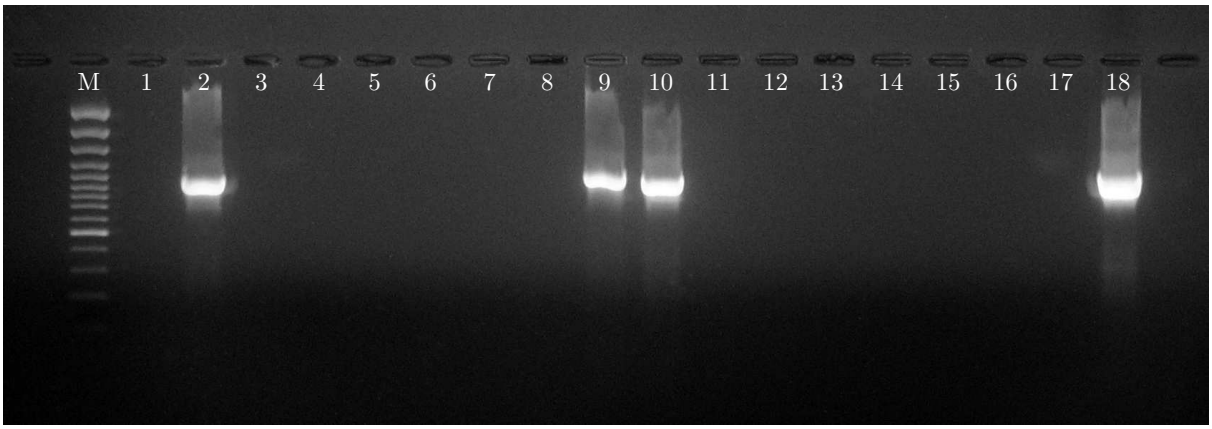
The pTAex3 expression system exploited in this study utilizes the α-amylase (Taka-amylase A, *amy*B) promoter which is repressed by glucose and induce by starch (Fujii *et al.*, 1995). Thus, 10 pTAex-SocPKS1 transformants were grown in liquid Czapek-Dox medium with glucose as a sole carbon source for 4 days and then subjected to the same medium but containing starch as a sole carbon source to induce the α-amylase promoter. Five pTAex3 transformants and *A. oryzae* M-2-3 (wild type) were grown in parallel as a control. Total RNA isolated from the mycelia of the transformants after 1 day induction with starch was treated with DNase I to get rid of residual genomic DNA and then subjected to reverse-transcriptase reaction. Synthesized first strand cDNA was directly used as a template in subsequent PCR with *S. crocea* PKS1 gene specific primers spanning putative introns.

Reverse transcriptase PCR revealed expression of the *S. crocea* PKS1 gene in 3 transformants – №№2, 9 and 10 – with all 4 pairs of primers tested. Faint bands were also detected for the transformant №8 with 2 primer pairs but it was not processed for the subsequent analysis. No products were observed for minus Reverse Transcriptase control samples in PCR with actin primers ensuring that all genomic DNA contamination was eliminated in isolated RNA prior to Reverse Transcriptase PCR (Figures III.5, III.6). Positive samples №№2, 9 and 10 were then subjected to PCR with 4 additional pair of primers to cover the entire cDNA.

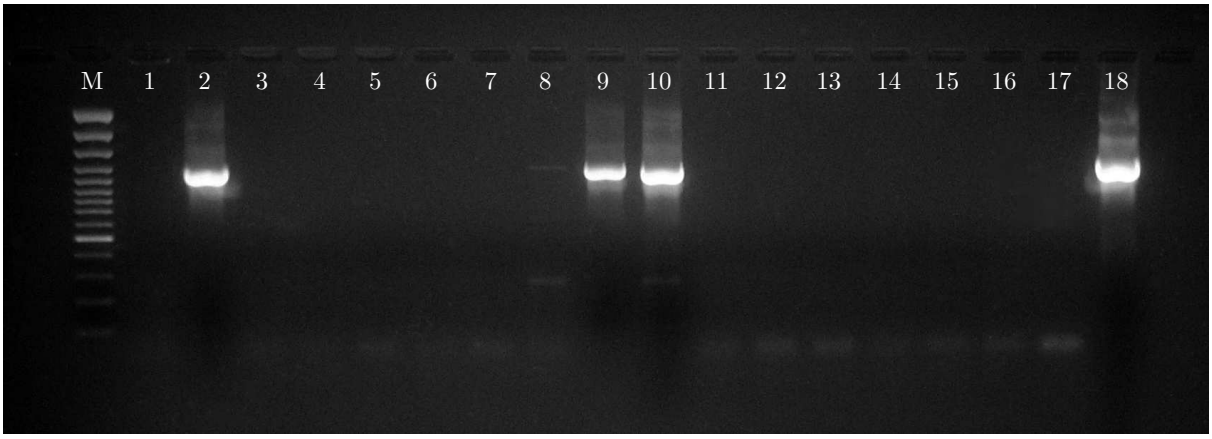


(e)

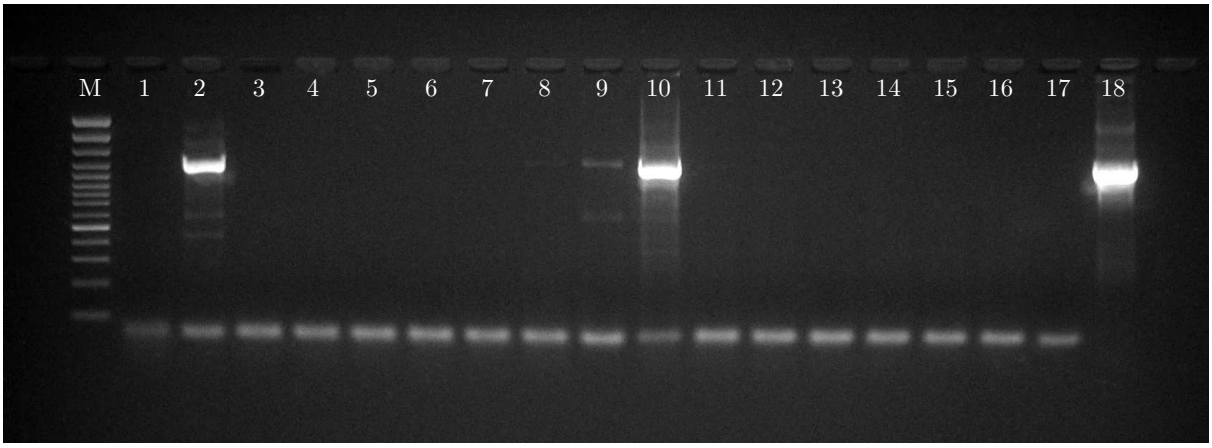
Figure III.5. Agarose gel electrophoresis of reverse transcriptase PCR products using cDNA as a template and *S. crocea* PKS1 gene specific primers: (a) GW-PKSF – B8AAF; (b) B8R1 – PKSR1; (c) Ex3-6F – K34nt1; (d) K34R1 – GW-PKSR. Lanes 1–10 – pTAex-SocPKS1 transformants, lanes 11–15 – "blank" pTAex3 transformants, lane 16 – *A. oryzae* M-2-3 (wild type), lane 17 – negative control (*no* DNA), lane 18 – positive control (plasmid pTAex-SocPKS1), M – GeneRuler 100 bp DNA Ladder Plus (MBI Fermentas).



(f)



(g)



(h)

Figure III.5. *Continued*

The results of all reverse transcriptase PCR reactions are summarized in Table III.2.

Eight overlapping cDNA fragments in three replicates obtained from the positive transformants (№№2, 9 and 10) were directly sequenced from both ends with the primers used

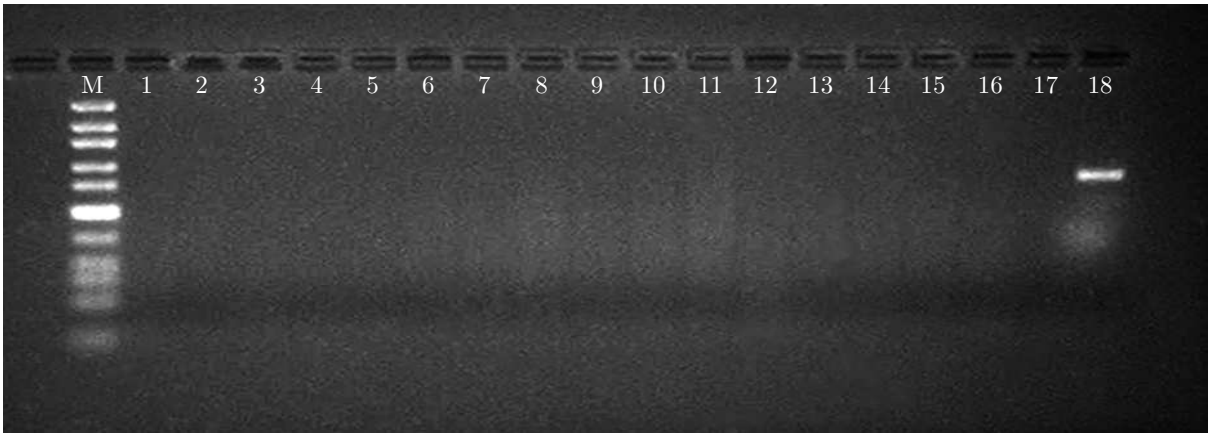


Figure III.6. Agarose gel electrophoresis of minus reverse transcriptase PCR products using RNA treated with DNase I and actin gene specific primers. The expected size of the amplified fragment is 81 bp. Lanes 1–10 – pTAex-SocPKS1 transformants, lanes 11–15 – "blank" pTAex3 transformants, lane 16 – *A. oryzae* M-2-3 (wild type), lane 17 – negative control (*no* RNA), lane 18 – positive control (*A. oryzae* M-2-3 genomic DNA), M – GeneRuler DNA Ladder, Ultra Low Range (MBI Fermentas).

to amplify the fragments. Aligning the cDNA sequences with the genomic sequence established exon-intron boundaries of the *S. crocea* PKS1 gene and confirmed its expression and successful splicing of pre-mRNA in the heterologous *Aspergillus oryzae* host.

Promoter region

Promoters of eukaryotes are regulatory regions of DNA located upstream of genes and consist of so called *core* promoter elements, such as CAAT, TATA and in fungi the CT-rich box, that serve as specific binding sites for general transcription factors determining the start site and efficiency of transcription. Unfortunately, although core elements may be seen in the promoter region of the *S. crocea* PKS1 gene, it may be impossible to evaluate their relative importance for transcription in the natural host. However, comments can be made about their presence, position and possible role by analogy to better characterized fungal systems. Because the transcription start site has not been determined, positions of the *S. crocea* PKS1 gene promoter elements are given in base pairs from the A(+1) of the putative translation start site.

GC-rich boxes. GC-rich boxes have been found in many eukaryotic promoters and are generally recognized and bound by Sp1-like transcription factors. The Sp1 binding regions contain one or more GC-boxes with the consensus sequence $^{G}_{T}$GGGCGG$^{GGC}_{AAT}$, which may be present in either orientation with respect to transcription (Dynan *et al.*, 1986;

Table III.2. Results of Reverse Transcriptase PCR screening with different pair of *S. crocea* PKS1 gene specific primers

№	Primer pair	pTAex-SocPKS1 transformants										pTAex3 transformants					WT[a]	Negative[b] control	Positive[c] control
		1	2	3	4	5	6	7	8	9	10	1	2	3	4	5			
1	GW-PKSF – B8AAF	–[d]	+[e]	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	+
2	B8R1 – PKSR1	–	+	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	+
3	B8AAR – Y13R	nt[f]	+	nt	nt	nt	nt	nt	nt	+	+	nt	nt	nt	nt	nt	nt	–	+
4	B8FCY – 15-9FF	nt	+	nt	nt	nt	nt	nt	nt	+	+	nt	nt	nt	nt	nt	nt	–	+
5	Ex2-3F – PKSR2	nt	+	nt	nt	nt	nt	nt	nt	+	+	nt	nt	nt	nt	nt	nt	–	+
6	PKSF2 – BR5524	nt	+	nt	nt	nt	nt	nt	nt	+	+	nt	nt	nt	nt	nt	nt	–	+
7	Ex3-6F – K34nt1	–	+	–	–	–	–	–	+	+	+	–	–	–	–	–	–	–	+
8	K34R1 – GW-PKSR	–	+	–	–	–	–	–	+	+	+	–	–	–	–	–	–	–	+

[a]Wild type, *A. oryzae* M-2-3.
[b]No DNA.
[c]Plasmid pTAex-SocPKS1.
[d]PCR product was not observed.
[e]PCR product was observed.
[f]Not tested.

Dynan, Tjian, 1985).

The promoter region of the *S. crocea* PKS1 gene contains two GC-rich boxes: the proximal GC-rich box is located 254 bp upstream of the putative translation start site and has the sequence GGGGGGCG, the distal one is located 336 bp upstream of the translation start site and has the sequence CCCGCCCCGC, which has complementarity to the proximal GC-rich box (Figure III.7). Similar observations have been made for the *amy*1 gene for α-amylase from *A. oryzae* (Wirsel *et al.,* 1989).

CAAT box. A CAAT box with the conserved sequence GG$^{\mathrm{C}}_{\mathrm{T}}$CAATCT is often found in the promoter region of many eukaryotic genes at around 70–90 bp upstream of the transcription start site but in yeast this element is rare. In filamentous fungal genes, sequences similar to the CAAT consensus are common but occur at far greater distances from the transcription start site than in mammalian promoters. Most of them lie 100–200 bp from the transcription start site (Benoist *et al.*, 1980; Gurr *et al.*, 1987; Nussinov *et al.*, 1986; Unkles, 1992).

A putative CAAT box is located 216 bp upstream of the putative translation initiation site of the *S. crocea* PKS1 gene (Figure III.7).

TATA box. The conserved TATA box is an AT-rich region common in eukaryotic promoters, recognized and bound by the general transcription factor TFIID. This element with the consensus TATAAA is found approximately 25–30 bp upstream of the major transcription start site of many animal and human genes. Most fungal TATA boxes are found 50–150 bp from the translation initiation site and 40–100 bp from the transcription start site. Some fungal genes do not possess a clear TATA box and may simply have an AT-rich region 30–100 bp prior to the transcription start site (Ballance, 1986; Gurr *et al.*, 1987; Nussinov *et al.*, 1986; Smale, Kadonaga, 2003; Turner *et al.*, 1993; Unkles, 1992)

A putative TATA box with the sequence TATTAT is found in the natural promoter region of the *S. crocea* PKS gene 147 bp upstream of the putative translation start site (Figure III.7).

CT-rich box. CT boxes or pyrimidine-rich tracts have been found in the promoter regions of yeast and filamentous fungal genes often immediately before the major transcription start site. In *Saccharomyces cerevisiae*, the presence of a pyrimidine-rich sequence preceding the transcription initiation site may markedly affect the level of transcription. The importance of the CT-rich box for determining the position of transcription initiation was confirmed by deletion analysis of this region in the *A. nidulans gpd*A and *oli*C

```
-522  CCCGAGTAAG GTTGCAGTAT GCTGCAATCC TGCTGTAACG GCTAAAGCAA TCGACAATCC TCCTGGTGTT
-452  TCGCAACAGG AGACGGATGA CGATAGGTTG ATGTTCTCAG GTATCGGCCC GTAGAATCCA CTGGCAGGCT
-382  CCGCTCCGCA TTAATGCTAC CTGTTGAGCT GGTATCAGCA CTAACACCCG CCCCGCAATC GGTCCGCTGT
-312  TTAGGTTCAA CCATGACTAG GCGGAACATA GCAGAGCAGA GTTGCAAGTT TAGCAATTGG GGGGCGAGCA
-242  CAAAGGTTAT CATTACAGTA GCATTACAAT GCACCTCTGG TGGACTGAAA TGGCTAGCAT GCCTTGGATG
-172  GGTTTTGCTA CATCTACAAC AGAAATATTA TCCGCTGCAT ACGACCTGAG CAAGATATTC GAACCGGTTT
-102  CTTCGTTGGA ATGCTCAATG CAGGGTGGTC GATTCAGATG TTATAGACAT ATCAGCGCCT TCCAAGTCAA
                                                GW-PKSF →
 -32  GTTGTCGAAG AACTCCGGAA AGTCTAAGCA AGATGGGTTC GATGCCGCTT GACGATGGCA TCCTGGGTTC
   1                                          M  G  S  M  P  L  D  D  G  I  L  G  S·
```

Figure III.7. Native promoter region of the *S. crocea* PKS1 gene. The putative GC-rich, CAAT, CT-rich boxes and ATG translation initiation codon is marked with grey, yellow, turquoise, and green, respectively. The putative TATA element is boxed. The primer GW-PKSF binding site is indicated with dotted underlining.

genes (Punt *et al.*, 1990; Turner *et al.*, 1989). CT-rich motifs ranging from 10 up to 60 bp (average 20 bp) in length can be found in many fungal genes. They are especially noticeable in genes lacking TATA and CAAT motifs and in highly expressed genes. These observations therefore suggest the role of CT motifs as possible promoter elements. CT boxes have a variable position relative to other promoter elements in different genes, occurring immediately upstream from the transcription start site, between the transcription and translation start sites, between the CAAT and TATA boxes or even further upstream (Gurr *et al.*, 1987; Unkles, 1992).

A 10 bp sequence reminiscent of this motif is located 105 bp upstream of the putative translation start site of the *S. crocea* PKS1 gene (Figure III.7). Based on the observations made for filamentous fungi, the unclear character of this motif suggests that the *S. crocea* PKS gene is expressed at a low level in its natural host. In contrast, the generally highly expressed *pyr*G gene from the same lichen has a remarkable 18 bp CT motif TTTTTTTCTCTCGTTCTT located 158 bp upstream of the transcription start site (Sinnemann *et al.*, 2000).

Terminator region

In eukaryotes, terminators are nucleotide sequences that act as signals for termination of transcription. They are recognized by protein factors, notably CstF (cleavage stimulation factor F) and CPSF (cleavage and polyadenylation specificity factor), that co-transcriptionally cleave the nascent RNA at a polyadenylation site, halting further elongation of the transcript by RNA polymerase. Following cleavage of the primary transcript,

```
8285  GAGATTCAAG GCACCATGAG TGTCACAGAT CTCTCCAGGA AGATTGCAAG CGTGAGTAGC GCGGTTCAGA
2664   E  I  Q  G  T  M  S  V  T  D  L  S  R  K  I  A  S  V  S  S  A  V  Q  I·
                  ← GW-PKSR
8355  TCACTGAAGA GAAGTAGATT TGAGAAGTAG ATTTTTTTTC TTGTCCGATA GAAATTGGAA CCAGAGGTGG
2668  ·  T  E  E  K  *
9425  TCTCTTTGAG ATCGCGACGA GAAGACGAAT GACAGATGGA GTGTTCTACT CGAATTTATG TGGCAAAGCA
9495  ACTGGCCTTC AACAGCATAA GACAACGACT GATCTGACAA TACCAGCTTT ACCTGCTATG AGTGCAGTCA
9565  AAAAAAGAGC CTTTCGATCA TGTGATGTAA CTTCGCTTGT GTAATTGATA CCCGATATTA CGTCGCGTGT
9635  ATCTTTGGGT AATTTCGCCG GGTTATTCTG CCCTGAGAAT AACAATACTT GGGCAGATCG GAGAGGGACC
```

Figure III.8. Native terminator region of the *S. crocea* PKS1 gene. The sites presumably involved in polyadenylation and termination are underlined. The putative stop codon TAG is marked with pink and an asterisk. The primer GW-PKSR binding site is indicated with dotted underline.

a poly(A) tail is added to most eukaryotic mRNAs by poly(A) polymerase stabilizing and allowing them to be exported out of the nucleus. In vertebrate genes, the highly conserved sequence AATAAA upstream of the cleavage site is recognized as the polyadenylation site. A second element associated with polyadenylation and cleavage lies just downstream of these sites and is composed of a less conserved T- or GT-rich sequence (Alberts et al., 2002; Proudfoot et al., 2002).

Filamentous fungal genes typically have several polyadenylation sites and, judging from 3′ end heterogeneity observed in the sequences of cDNA clones, at least some of these sites are real rather than artefactual. The conserved polyadenylation signal AATAAA, found in mammalian genes, is rarely present in yeast and filamentous fungal genes. A motif, reminiscent of the conserved AATAAA sequence, can be found in the 3′ end region of some fungal genes at approximately the expected position, but since the 3′ non-coding region is often AT-rich, its functional significance is rather uncertain. In any case, many fungal genes have nothing remotely similar to this sequence and the lack of identifiable polyadenylation and cleavage signals in most fungal genes may suggest differences in processing machinery (Ballance, 1986; Gurr *et al.*, 1987; Unkles, 1992).

The 3′ non-coding region of the *S. crocea* PKS1 gene was not determined by 3′ RACE experiments because of the difficulties in obtaining the corresponding cDNA from the natural host but several potential polyadenylation sites can be found in the terminator region of the *S. crocea* PKS1 gene, with the most notable blocks, ATTTTTTTT and AAAAAAA, being complementary to each other and located 14 and 193 bp, respectively, downstream of the putative stop codon (Figure III.8). The region surrounding the first block exhibits overall AT-richness (Figure III.9). These observations correspond to the results obtained by Nussinov on distribution analysis of $(T)_{4-5}$ and $(A)_{4-5}$ homooligomers

in the 3′ ends of mammalian genes based on information about 3′ poly(A) ends of their mRNAs. It was found that the frequency of $(A)_{4-5}$ was lower prior to position −50 (with respect to a site of poly(A) tail addition) than following it. The frequency of $(A)_{4-5}$ was also higher prior to the transcription initiation site than following it. Thus, untranscribed regions have a higher $(A)_{4-5}$ frequency than transcribed ones. In contrast to $(A)_{4-5}$, $(T)_{4-5}$ homooligomer exhibited the opposite distribution rising to its maximum near the site of poly(A) tail addition and then falling off (Nussinov, 1986).

The consensus sequence YGTGTTYY, where Y represents a pyrimidine base, has been found in many mammalian genes downstream from the polyadenylation site and is considered to be a termination signal (McClauchlan *et al.*, 1985). A similar block, AGTGTTCT, differing from the consensus only in the first base, can be observed 93 bp downstream of the putative stop codon, but again this sequence may not be significant.

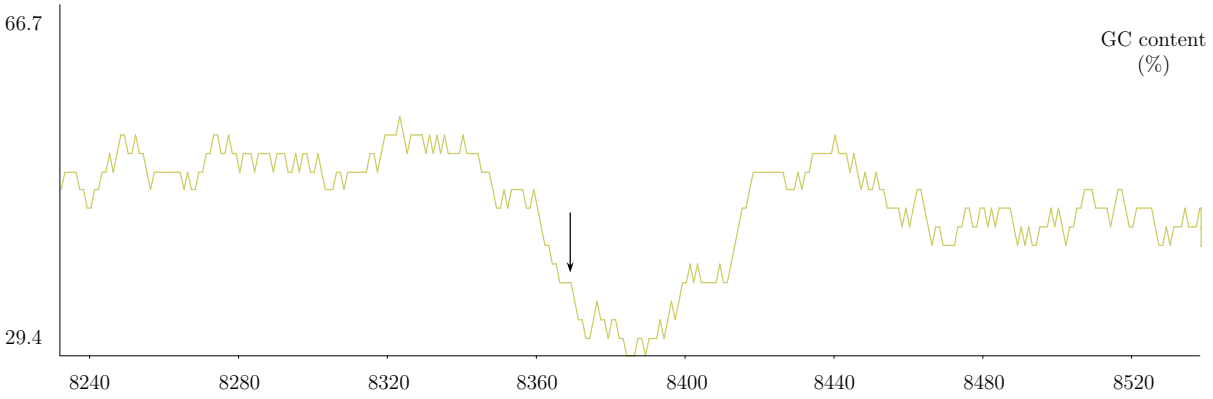


Figure III.9. GC content of the 3′ end region of the *S. crocea* PKS1 gene. The numbers on the X axis represents positions in base pairs from the A(+1) of the putative translation start site of the *S. crocea* PKS1 gene. The GC content (%) on the Y axis represents the ratio of the number of scored occurrences of G and C bases in any window divided by the total number of bases in that window (50). The position of the putative stop codon is indicated by an arrow. Analysis was performed using BioAnnotator, a component of Vector NTI Advance v.10.3 Sequence Analysis Software (Invitrogen).

mRNA processing and translation

Translation start site

The putative ATG translation initiation codon of the *S. crocea* PKS1 coding region (Figure III.7) was deduced based on:

1. Conservation of some amino acid residues at the *N*-terminus of fungal reducing PKSs;

2. Corresponding open reading frame shift, with the position of the first putative intron between codons (Table III.4);

3. Consensus sequences surrounding the translation initiation site identified for eukaryotes and filamentous fungi in particular (Ballance, 1986; Gurr *et al.*, 1987; Kozak, 1984; Turner *et al.*, 1993; Unkles, 1992).

A consensus sequence around ATG at which translation begins has been identified as CC $^{A}_{G}$CC ATG G based on analysis of mainly vertebrate genes (Table III.3) (Kozak, 1984, 1987). The most highly conserved nucleotides are a purine, most often represented by an adenine, at position −3 (three nucleotides upstream from the ATG codon) and a guanine at position +4, following the ATG codon. Approximately 80% of fungal genes examined show a clear preference for a purine at position −3, and approximately 60% for an adenine (Unkles, 1992).

There are two potential initiator ATG codons belonging to the same open reading frame and lying six residues apart (Figure III.7), but the sequence around the more 5′ ATG, GC AAG ATG G, matches the Kozak consensus sequence, CC $^{A}_{G}$CC ATG G, and a consensus for fungal genes, TC A$^{CA}_{AC}$ ATG $^{G}_{T}$C (Ballance, 1986), and is therefore considered as the most likely translation start.

Table III.3. Conservation of ATG sequence environment

Organisms	**Consensus sequence**	**Reference**
Eukaryotes (n^a = 211)	CC $^{A}_{G}$CC ATG G	Kozak, 1984
Vertebrates (n = 699)	GCC GCC $^{A}_{G}$CC ATG G	Kozak, 1987
Fungi (n = 31)	TC A$^{CA}_{AC}$ ATG $^{G}_{T}$C	Ballance, 1986
Neurospora spp. (n = 77)	CN NNC A$^{A}_{C}$"T" ATG GC[b]	Bruchez *et al.*, 1993
Neurospora spp. (n = 88)	C A$^{CA}_{AC}$ ATG GCT	Edelmann, Staben, 1993

[a] n represents a number of genes analyzed.

[b] "T" indicates the conserved absence of that particular nucleotide at this position.

Translation stop site

Like other eukaryotic and, in particular, fungal genes that exhibit a preference (70%) for a purine at the position following the stop codon, the putative stop codon of the *S. crocea* PKS1 coding region is followed by an adenine (Cavener, Ray, 1991; Cridge *et al.*, 2006; Unkles, 1992).

Introns

Sequencing of the cDNA fragments obtained by reverse transcriptase PCR confirmed the presence of 5 putative introns and allowed to identify their splice sites (Figure III.10, Table III.4). In contrast to mammalian and plant introns, and like typical fungal introns (Unkles, 1992), they are relatively short, with the average length being 71 bp and a range of 47 bp to 151 bp.

All identified 5′ and 3′ splice sites match the consensus sequences, GTRNGY and YAG, respectively, determined for fungal introns, where R is a purine, N is any base and Y is a pyrimidine (Ballance, 1986; Gurr *et al.*, 1987; Unkles, 1992). No alternative splicing was observed.

Putative internal branch point sites for lariat formation were identified based on the consensus sequences YGCTAACN (Gurr *et al.*, 1987) and NNCTRAY (Unkles, 1992) found for filamentous fungal introns (Appendix A-4). Like in most fungal introns, the CT at positions 3 and 4 was found in all sites identified. Four sites out of five also exhibit an adenine at position 6.

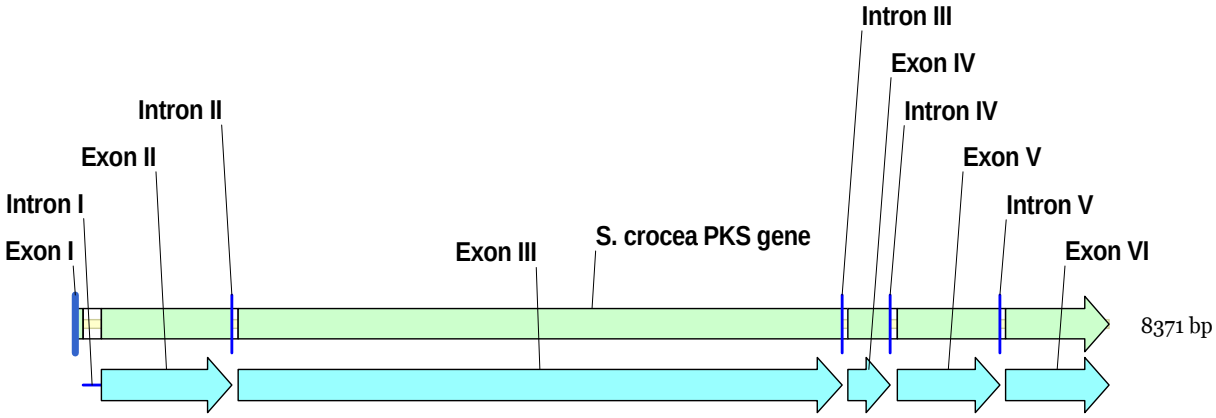


Figure III.10. Structure of the *S. crocea* PKS1 gene

Table III.4. Location and characteristics of exons and introns in the *S. crocea* PKS1 gene

Exon				Intron[a]				Codon phase[b]	Amino acid[c]
№	Starts	Ends	Length	№	5′ splice site	Length	3′ splice site		
1	1	63	63	1	CCGATGgtaagc	151	tcacagGCCATC	0	Met^{21}/Ala^{22}
2	215	1269	1055	2	TGCTGGgtaagt	51	ttgcagAGTCAT	2	Gly^{373}
3	1321	6206	4886	3	CTGCAGgtatgt	50	tctcagGTATTG	1	Gly^{2002}
4	6257	6606	344	4	CGAGCAgtaagt	56	ctttagCTTGTC	0	Ala^{2116}/Leu^{2117}
5	6657	7484	828	5	TTGAAGgtgagt	47	gtttagGATACA	0	Lys^{2392}/Asp^{2393}
6	7532	8371	840	–	–	–	–	–	–

[a]The introns are indicated in lower case.
[b]The position of introns between codons is indicated by phase 0, interruption after the first base is indicated by phase 1, and interruption after the second base is indicated by phase 2.
[c]The amino acid(s) encoded at the splice site.

Codon usage

Codon usage in the *S. crocea* PKS1 gene does not exhibit any significant bias and is rather even (Table III.5). It was observed previously that highly expressed genes in filamentous fungi have a preferred subset of codons, probably reflecting the abundance of their corresponding tRNA molecules. In contrast, filamentous fungal genes which are expressed at low levels show much less codon bias (Gurr *et al.*, 1987; Unkles, 1992). Judging from these observations, it can be assumed that the *S. crocea* PKS1 gene is not highly expressed (or expressed at a low level) in its natural host.

Deduced amino acid sequence

The open reading frame of the 8 371 bp *S. crocea* PKS1 coding region encodes a deduced protein of 2 671 amino acids (Figures III.10, III.12). By analogy to other known fungal PKSs, the protein was named PKS1. The molecular weight of the *S. crocea* PKS1 was predicted to be 294.84 kDa.

Seven catalytic domains were identified in the *S. crocea* PKS1 based on the presence of their corresponding conserved motifs. These domains are β-ketoacyl synthase (KS), acyltransferase (AT), dehydratase (DH), *C*-methyltransferase (*C*MeT), enoyl reductase (ER), β-ketoreductase (KR), and acyl carrier protein (ACP), with the order being KS→AT→DH→*C*MeT→ER→KR→ACP. Other fungal PKS enzymes with the same domain organization and identified products are PKSN for alternapyrone biosynthesis from *Alternaria solani* (accession number BAD83684, Fujii *et al.*, 2005), LDKS for lovastatin diketide biosynthesis from *Aspergillus terreus* (accession number AAD34559, Kennedy *et al.*, 1999), MlcB for compactin diketide biosynthesis from *Penicillium citrinum* (accession number BAC20566, Abe *et al.*, 2002), SQTKS for squalestatin tetraketide biosynthesis from *Phoma sp.* (accession number AAO62426, Cox *et al.*, 2004) and FUM1p involved in fumonisin biosynthesis from *Gibberella moniliformis* (= *Fusarium verticillioides*) (accession number AAD43562, Proctor *et al.*, 1999). Alignments of *S. crocea* PKS1 catalytical domains with those of the listed fungal PKSs are shown in Figure III.11.

Conserved amino acid residues and motifs important for catalysis can be identified in all seven domains of the *S. crocea* PKS1: the cysteine (C) residue necessary for thioester formation in the KS domain, the pantetheine binding serine (S) residue in the AT domain, the histidine (H) residue required for dehydration in the DH domain, the NADPH binding motifs in the ER and KR domains, the phosphopantetheine binding serine in the ACP domain and all motifs in the *C*MeT domain identified by Anderson and coworkers: the glycine (G) residues in motif I, the aspartic acid (D) residue in motif II, the tyrosine (Y)

Table III.5. Codon usage in the *S. crocea* PKS1 gene

Codon	Amino acid[a]	n[b]	%[c]	Ratio[d]	Codon	Amino acid	n	%	Ratio	Codon	Amino acid	n	%	Ratio	Codon	Amino acid	n	%	Ratio
UUU	Phe (F)	37	13.85	0.39	UCU	Ser (S)	39	14.60	0.18	UAU	Tyr (Y)	34	12.72	0.52	UGU	Cys (C)	10	3.74	0.34
UUC	Phe (F)	59	22.08	0.61	UCC	Ser (S)	33	12.35	0.15	UAC	Tyr (Y)	31	11.60	0.48	UGC	Cys (C)	19	7.11	0.66
UUA	Leu (L)	17	6.36	0.07	UCA	Ser (S)	41	15.34	0.19	UAA	STOP	0	0.00	0.00	UGA	STOP	0	0.00	0.00
UUG	Leu (L)	49	18.34	0.19	UCG	Ser (S)	29	10.85	0.13	UAG	STOP	1	0.37	1.00	UGG	Trp (V)	34	12.72	1.00
CUU	Leu (L)	56	20.96	0.22	CCU	Pro (P)	35	13.10	0.27	CAU	His (H)	41	15.34	0.54	CGU	Arg (R)	26	9.73	0.18
CUC	Leu (L)	63	23.58	0.24	CCC	Pro (P)	21	7.86	0.16	CAC	His (H)	35	13.10	0.46	CGC	Arg (R)	19	7.11	0.13
CUA	Leu (L)	17	6.36	0.07	CCA	Pro (P)	49	18.34	0.37	CAA	Gln (Q)	49	18.34	0.59	CGA	Arg (R)	39	14.60	0.27
CUG	Leu (L)	58	21.71	0.22	CCG	Pro (P)	26	9.73	0.20	CAG	Gln (Q)	34	12.72	0.41	CGG	Arg (R)	15	5.61	0.10
AUU	Ile (I)	43	16.09	0.30	ACU	Thr (T)	31	11.60	0.21	AAU	Asn (H)	48	17.96	0.50	AGU	Ser (S)	27	10.10	0.12
AUC	Ile (I)	79	29.57	0.55	ACC	Thr (T)	34	12.72	0.23	AAC	Asn (H)	48	17.96	0.50	AGC	Ser (S)	49	18.34	0.22
AUA	Ile (I)	21	7.86	0.15	ACA	Thr (T)	54	20.21	0.36	AAA	Lys (K)	70	26.20	0.51	AGA	Arg (R)	28	10.48	0.19
AUG	Met (M)	81	30.31	1.00	ACG	Thr (T)	30	11.23	0.20	AAG	Lys (K)	66	24.70	0.49	AGG	Arg (R)	18	6.74	0.12
GUU	Val (V)	54	20.21	0.28	GCU	Ala (A)	73	27.32	0.33	GAU	Asp (D)	85	31.81	0.57	GGU	Gly (G)	54	20.21	0.28
GUC	Val (V)	68	25.45	0.35	GCC	Ala (A)	50	18.71	0.23	GAC	Asp (D)	64	23.95	0.43	GGC	Gly (G)	68	25.45	0.36
GUA	Val (V)	19	7.11	0.10	GCA	Ala (A)	60	22.46	0.28	GAA	Glu (E)	94	35.18	0.53	GGA	Gly (G)	44	16.47	0.23
GUG	Val (V)	55	20.58	0.28	GCG	Ala (A)	35	13.10	0.16	GAG	Glu (E)	82	30.69	0.47	GGG	Gly (G)	24	8.98	0.13

[a]The letter in parenthesis represents the one-letter code for the amino acid.
[b]*n* represents the total number of times this codon is observed in the sequence.
[c]% represents the average frequency this codon is used per 1000 codons.
[d]Ration represents the abundance of this codon relative to all of the codons for that particular amino acid.

β-Ketoacyl synthase (KS)

```
Alternaria solani PKSN          132  NLQARDNEGPSIYHATGMSASLASNRLSWFYNLRGPSLTVDTACSSSLTAFHLACQSIRTGEAEMSVVGGANLMFGPDMS
Aspergillus terreus LDKS        140  GSFQRQPEALPRYFITGNAGTMLANRVSHFYDLRGPSVSIDTACSTTLTALHLAIQSLRAGESDMAIVAGANLLLNPDVF
Penicillium citrinum MlcB       159  DSLNRDPETVPRYFITGNSGTMLSNRISHFYDLRGPSVTVDTACSTTLTALHLACQSLRTGESDTAIVIGANLLLNPDVF
Phoma sp. SQTKS                 159  DSLVRDPDLVPRFLLTGNGAAMSSNRVSHFYDLRGASMTVDTGCSTTLTALHLACQGLRNRESKTSIVTGANVILNPDMF
Gibberella moniliformis FUM1p   158  DLHSKDLYDSGTYRVSGGHDFAISNRISYEYDLKGPSFTIKAGCSSSLIALHEAVRAIRAGDCDGAIVAGTNLVFSPTMS
Solorina crocea PKS1            147  ANLNADPESTMKYKPTGTTNSLLSNRVSWFFDLKAPSVTLDTACSSSMVALHLGCQSLRTRECDMSVISGVTMISFPTDI
                                                                    *
```

Acyltransferase (AT)

```
Alternaria solani PKSN          619  GCPWSLAEELSRSKAEYKLRETDYSQPACTAIQVALVDLLSGFDVRPVAVLGHSSGEIAAAYAAGFIDQEAAIKIAWLRG
Aspergillus terreus LDKS        582  GSTWSIIEELSRPETESRVDQAEFSLPLSTALQIALVRLLWSWNIQPVAVTSHSSGEAAAAYAIGALTARSAICGISYIRG
Penicillium citrinum MlcB       605  GASWNFMEELHRDELTTRVNDAEYSLPLSTAIQIALVRLLWSWGIRPTGITSHSSGEAAAAYAAGALSARSAIGITYIRG
Phoma sp. SQTKS                 645  GCEWSIIDELSRDAENSNVNDMTLSPPLCTAVQISLVQLLESWGIVPTAVTGHSSGEIAAAYAAGALDFKSAMAVTYFRG
Gibberella moniliformis FUM1p   648  PPTWDLIRELSRPAESSNVGRAEFSQPLVCAVQVALVNLLRSWGITPAAVVGHSSGEMAAAYAAGAISSEEAITIAYYRG
Solorina crocea PKS1            661  GSPWSLWDELLRNKSDTKVNEPHLAHPACAALQMALVDLLRSWQITPSRVIGHSSGEIAAAYCAGKLTRESAWKVAYFRG
                                                                    *
```

Dehydratase (DH)

```
Alternaria solani PKSN          936  AWNHKP-YWDESHLSRQHRLRKSPRTDLVGYR--LLGTPEHTWRNFLRCNENPWIREHKVQGDILYPGAGMLVMAIEAAH
Aspergillus terreus LDKS        916  PWNHETRYWKEPRISQSARQRKGPVHDLIGLQEPLNLPLARSWHNVLRVSDLPWLRDHVVGSHIVFPGAGFVCMAVMGIS
Penicillium citrinum MlcB       941  PWNHQTRYWREPRGSRESKQRTHPPHTLIGSRESLSPQFAPKWKHVLRLSDIPWIRDHVVGSSIIFPGAGFISMAIEGFS
Phoma sp. SQTKS                 975  PWNHDNSHWWEPRLNKAHRQRVHPPHDLLGSLIVGRDLREPTWRHFIRVQDIPWIRDHVVQSALVYPGAGFICMAMEAMV
Gibberella moniliformis FUM1p   955  PWNHDREYWSESRVSKDWRFRKFPNHELLGSRTLESSSLQPEWRNLIRLDGIPWLRDHQVLNDVVFPCAGYLAMAVEAVR
Solorina crocea PKS1           1003  TFNHSEKLIYESRLTKNYRLRKQPRHDLFGAPVPDWNSETPRWRNILRLAEQPWLRDHIVTNQVVLPGVGYLIAAIEASR
                                                                          *
```

C-Methyltransferase (*C*MeT)

```
Alternaria solani PKSN         1409  RIFTREVEALQCMT-EGDLLYSFYRGAFGTSFN-TNVAEYVGLIADKQPGLRILEIGAGTGGTTYHVLERLRNPDGT---
Aspergillus terreus LDKS       1386  AILRREVDPLEVMM-DGHLLSRYYVDALKWSRSNAQASELVRLCCHKNPRARILEIGGGTGGCTQLVVDSLGPN------
Penicillium citrinum MlcB      1419  AMLRRETEPLELMM-QDQLLSRYYVNAIKWSRSNAQASELIRLCAHKNPRSRILEIGGGTGGCTKLIVNALGNT------
Phoma sp. SQTKS                1447  AILRHEIAPLELML-QDKLLYRYYTDAIKWDRSYQQIDQLVKLHAHKCPTAKIIEIGAGTGGCTRAVLDALSNQG-I---
Gibberella moniliformis FUM1p  1412  VDHCEEIVKGELEGIELLQAEDGLTNYYNYVESRTDSIDFFATAGHTRPTLRVLEIGAGTGGGAQVILEGLTNGK-----
Solorina crocea PKS1           1488  KVLRKEEDPLHLMFGVDDILDRVYAQVAHLGNLPALQAALLKIVADNSTNLRVLEIGAGTGGSTVGMLEGLISTSDDGVF
                                                                          * * *

Alternaria solani PKSN         1484  -SKAAQYYFTDISPGFLAKAADRFNQDASIMQFGTLNIENAPTEQGFSPELFDLIVCANVLHATKSIQETLTHCKLLLKP
Aspergillus terreus LDKS       1459  -PPVGRYDFTDVSAGFFEAARKRFAGWQNVMDFRKLDIEDDPEAQGFVCGSYDVVLACQVLHATSNMQRTLTNVRKLLKP
Penicillium citrinum MlcB      1492  -KPIDRYDFTDVSAGFFESAREQFADWQDVMTFKKLDIESDPEQQGFECATYDVVVACQVLHATRCMKRTLSNVRKLLKP
Phoma sp. SQTKS                1522  -ARCAQYDFTDVSSGFFEAAQQKFAAFDDVIRFQKLDIEKDIEMQGFECGSYDLVIASQVLHATGKMEHTMANVRKLLKP
Gibberella moniliformis FUM1p  1487  ERLFSTYAYTDISAGFFVAAQERFKAYKG-LDFKVLDITKDPSEQGFESGSFDLIIAGNVIHATPTLNETLANVRKLLAP
Solorina crocea PKS1           1568  SSSVSTYTFTDISLAFFAKAKEKFKNYRDIMEYKVLDIEKEPDEEVFELGSYDIVVAQNVVHATTNLRRTLSHIRSLLKP
                                           *                                 *
```

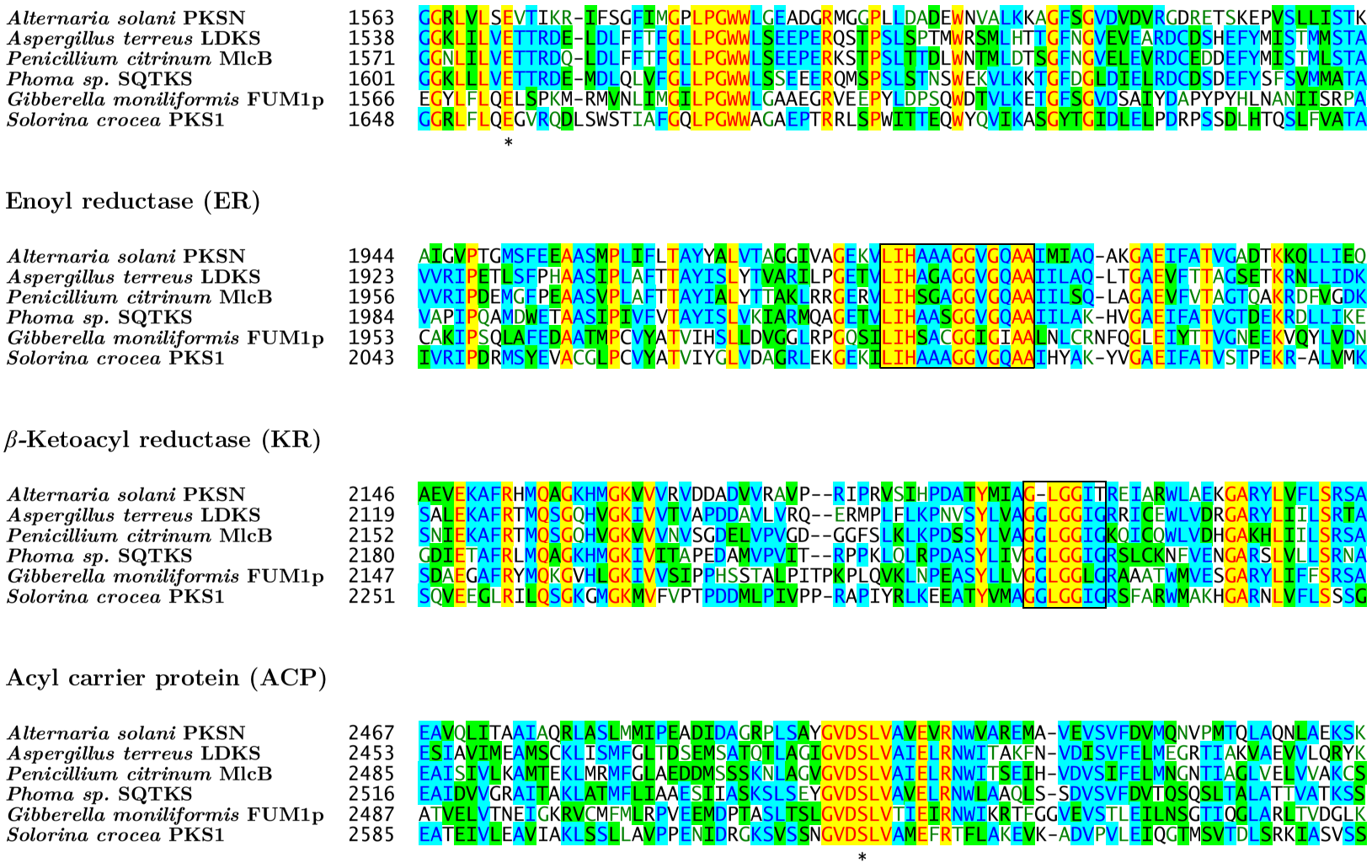


Figure III.11. Amino acid sequence alignment of the catalytical domains of the *S. crocea* PKS1 and selected fungal iterative type I PKSs with the same domain organization. Conserved key amino acid residues and motifs at catalytic sites are marked with an asterisk and shown boxed, respectively. The alignment was obtained using AlignX, a component of Vector NTI Advance v.10.3 Sequence Analysis Software (Invitrogen). The amino acid residues in the alignment are color-coded according to the following scheme: red on yellow, completely conserved; blue on cyan, consensus derived from block of similar residues; green on white, residue weakly similar to consensus residue; black on white, non-similar to consensus residue; black on green, consensus derived from majority residue.

residue in motif III and the glutamic acid (E) residue in motif IV (Miller *et al.*, 2003).

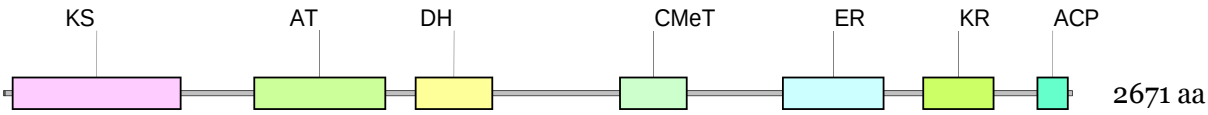


Figure III.12. Domain organization of the *S. crocea* PKS1. KS, β-ketoacyl synthase; AT, acyltransferase; DH, dehydratase; *C*MeT, *C*-methyltransferase; ER, enoyl reductase; KR, β-ketoreductase; ACP, acyl carrier protein.

Table III.6. Approximate domain boundaries[a] of the *S. crocea* PKS1

Domain	**Starts**	**Ends**
β-Ketoacyl synthase (KS)	19	439
Acyltransferase (AT)	624	951
Dehydratase (DH)	1028	1219
C-Methyltransferase (*C*MeT)	1540	1706
Enoyl reductase (ER)	1949	2199
β-Ketoreductase (KR)	2299	2474
Acyl carrier protein (ACP)	2585	2660

[a] Approximate domain boundaries were determined using the SearchPKS program (*http://www.nii.res.in/searchpks.html*; Yadav *et al.*, 2003) and by visual examination of the amino acid sequence alignment with other fungal PKSs.

According to a protein BLAST query[1] the *S. crocea* PKS1 shares the highest end-to-end homology with:

1. The polyketide synthase ChPKS5 from *Cochliobolus heterostrophus* (53% sequence identity, 71% sequence similarity) (accession number AAR90260, Kroken *et al.*, 2003),

2. A putative PKS from *Aspergillus clavatus* NRRL 1 (40% sequence identity, 59% sequence similarity) (accession number XP_001276282, Nierman, *unpublished*),

[1] http://www.ncbi.nlm.nih.gov/BLAST

3. A hypothetical protein from *Aspergillus terreus* NIH2624 (40% sequence identity, 59% sequence similarity) (accession number XP_001212624, Birren *et al.*, *unpublished*).

Unfortunately, the products of these proteins have not been identified yet. The nearest PKS relatives with known products are:

1. PKSN from *Alternaria solani* involved in biosynthesis of the decaketide alternapyrone when expressed in *A. oryzae* (36% sequence identity, 55% sequence similarity) (accession number BAD83684, Fujii *et al.*, 2003),

2. SQTKS from *Phoma sp.* involved in biosynthesis of the squalestatin tetraketide side chain (35% sequence identity, 54% sequence similarity) (accession number AAO62426, Cox *et al.*, 2004),

3. MlcB from *Penicillium citrinum* involved in biosynthesis of the compactin diketide side chain (34% sequence identity, 53% sequence homology) (accession number BAC20566, Abe *et al.*, 2002).

According to the classification based on phylogenomic analysis of type I PKSs proposed by Kroken and coworkers (Kroken *et al.*, 2003), the *S. crocea* PKS1 falls into the clade I of fungal reducing PKSs with the general domain organization KS→AT→DH→(*C*MeT)[2]→ER→KR→ACP.

[2]A *C*MeT domain is variable in its presence/absence within this clade.

IV

Conclusion

To our knowledge, the *Solorina crocea* PKS1 gene is the first lichen polyketide synthase gene that was cloned, sequenced, characterized and expressed in a heterologous host. After the pioneering work of Sinneman and coworkers on cloning and expression of the *S. crocea pyr*G (orotidine 5′-monophosphate decarboxylase) gene in *Aspergillus nidulans* (Sinneman *et al.*, 2000), this is only the second gene from lichen mycobionts, apart from the genes involved in RNA biosynthesis, that has been fully sequenced and characterized.

Our future study on the *S. crocea* PKS1 gene will concentrate on its functional analysis involving isolation and identification of product(s) of this PKS. The results presented in this work open new horizons for exploring and harnessing the untapped polyketide biosynthetic potential of lichens.

References

Abe, Y., Suzuki, T., Ono, C., Iwamoto, K., Hosobuchi, M., Yoshikawa, H. 2002. Molecular cloning and characterization of an ML-236B (compactin) biosynthetic gene cluster in *Penicillium citrinum*. *Molecular Genetics and Genomics* **267**: 636–646.

Aboul-Soud, M.A., Yun, B.W., Harrier, L.A., Loake, G.J. 2004. Transformation of *Fusarium oxysporum* by particle bombardment and characterisation of the resulting transformants expressing a GFP transgene. *Mycopathologia* **158**: 475–482.

Adefarati, A.A., Giacobbe, R.A., Hensens, O.D., Tkacz, J.S. 1991. Biosynthesis of L-671329, an echinocandin-type antibiotic produced by *Zalerion arboricola*: origins of some of the unusual amino acids and the dimethylmyristic acid side chain. *Journal of the American Chemical Society* **113**: 3542–3545.

Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B.O., Huang, K., Fonstein, L., Czisny, A., Whitwam, R.E., Farnet, C.M., Thorson, J.S. 2002. The calicheamicin gene cluster and its iterative type I enediyne PKS. *Science* **297**: 1173–1176.

Al-Rawi, J.M.A., Elvidge, J.A., Jaiswal, D.K., Jones, J.R., Thomas, R. 1974. Use of tritium nuclear magnetic resonance for the direct location of ^{3}H in biosynthetically-labeled penicillic acid. *Journal of the Chemical Society. Chemical communications*: 220–221.

Alberts, A.W., Chen, J., Kuron, G., Hunt, V., Huff, J., Hoffman, C., Rothrock, J., Lopez, M., Joshua, H., Harris, E., Patchett, A., Monaghan, R., Currie, S., Stapley, E., Albers-Schonberg, G., Hensens, O., Hirshfield, J., Hoogsteen, K., Liesch, J., Springer, J. 1980. Mevinolin: a highly potent competitive inhibitor of hydroxymethylglutaryl-coenzyme A reductase and a cholesterol-lowering agent. *Proceedings of the National Academy of Sciences of the USA* **77**: 3957–3961.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. 2002. Molecular Biology of the Cell. Fourth Edition, Garland Science Publishing, New York.

Aldridge, D.C., Davies, A.B., Jackson, M.R., Turner, W.B. 1974. Pentaketide metabolites of the fungus *Phialophora lagerbergii*. *Journal of the Chemical Society, Perkin Transactions 1*: 1540–1541.

Allen, R.S., Millgate, A.G., Chitty, J.A., Thisleton, J., Miller, J.A., Fist, A.J., Gerlach, W.L., Larkin, P.J. 2004. RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nature Biotechnology* **22**: 1559–1566.

Allport, D.C., Bu'Lock, J.D. 1960. Biosynthetic pathways in *Daldinia concentrica*. *Journal of the Chemical Society*: 654–662.

Amnuaykanjanasin, A., Punya, J., Paungmoung, P., Rungrod, A., Tachaleat, A., Pongpattanakitshote, S., Cheevadhanarak, S., Tanticharoen, M. 2005. Diversity of type I polyketide synthase genes in the wood-decay fungus *Xylaria sp.* BCC 1067. *FEMS Microbiology Letters* **251**: 125–136.

Anderson, H.A., Thomson, R.H., Wells, J.W. 1966. Naturally occurring quinones. Part VIII. Solorinic acid and norsolorinic acid. *Journal of the Chemical Society* **C**: 1727–1729.

Ando, K., Matsuura, I., Nawata, Y., Endo, H., Sasaki, H., Okytomi, T., Saehi, T., Tamura, G. 1978. Funiculosin, a new antibiotic. II. Structure elucidation and antifungal activity. *The Journal of Antibiotics* **31**: 533–538.

Aramayo, R., Timberlake, W.E. Sequence and molecular structure of the *Aspergillus nidulans yA* (laccase I) gene. *Nucleic Acids Research* **18**: 3415.

Axberg, K., Gatenbeck, S. 1975. Intermediates in the penicillic acid biosynthesis in *Penicillium cyclopium*. *Acta Chemica Scandinavica. Series B: Organic Chemistry and Biochemistry* **29**: 749–751.

Axberg, K., Gatenbeck, S. 1975a. The enzymic formation of penicillic acid. *FEBS Letters* **54**: 18–20.

Baker, S.E., Kroken, S., Inderbitzin, P., Asvarak, T., Li, B.Y., Shi, L., Yoder, O.C., Turgeon, B.G. 2006. Two polyketide synthase-encoding genes are required for biosynthesis of the polyketide virulence factor, T-toxin, by *Cochliobolus heterostrophus*. *Molecular Plant-Microbe Interactions* **19**: 139–149.

Ballance, D.J. 1986. Sequences important for gene expression in filamentous fungi. *Yeast* **2**: 229–236.

Ballance, D.J., Buxton, F.P., Turner, G. 1983. Transformation of *Aspergillus nidulans* by the orotidine-5′-phosphate decarboxylase gene of *Neurospora crassa*. *Biochemical and Biophysical Research Communications* **112**: 284–289.

Ballance, D.J., Turner, G. 1985. Development of a high-frequency transforming vector for *Aspergillus nidulans*. *Gene* **36**: 321–331.

Barber, J., Staunton, J. 1980. New insights into polyketide metabolism; the use of protium as a tracer in the biosynthesis of citrinin by *Penicillium citrinum*. *Journal of the Chemical Society, Perkin Transactions 1*: 2244–2248.

Barcellos, F.G., Fungaro, M.H., Furlaneto, M.C., Lejeune, B., Pizzirani-Kleiner, A.A., de Azevedo, J.L. 1998. Genetic analysis of *Aspergillus nidulans* unstable transformants obtained by the biolistic process. *Canadian Journal of Microbiology* **44**: 1137–1141.

Barratt, R.W., Johnson, G.B., Ogata, W.N. 1965. Wild-type and mutant stocks of *Aspergillus nidulans*. *Genetics* **52**: 233–246.

Barteles-Keith, J.R. 1960. Alternaric acid. Part III. Structure. *Journal of the Chemical Society*: 1662–1665.

Baulcombe, D.C. 1999. Gene silencing: RNA makes RNA makes no protein. *Current Biology* **22**: 1559–1566.

Baxter, A., Fitzgerald, B.J., Hutson, J.L., McCarthy, A.D., Motteram, J.M., Ross, B.C., Sapra, M., Snowden, M.A., Watson, N.S., Williams, R.J., Wright, C. 1992. Squalestatin 1, a potent inhibitor of squalene synthase, which lowers serum cholesterol *in vivo*. *The Journal of Biological Chemistry* **267**: 11705-11708.

Beck, J., Ripka, S., Siegner, A., Schiltz, E., Schweizer, E. 1990. The multifunctional 6-methylsalicylic acid synthase gene of *Penicillium patulum*. Its gene structure relative to that of other polyketide synthases. *European Journal of Biochemistry* **192**: 487–498.

Bedford, D.J., Schweizer, E., Hopwood, D.A., Khosla, C. 1995. Expression of a functional polyketide synthase in the bacterium *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* **177**: 4544–4548.

Bej, A.K., Perlin, M.H. 1989. A high efficiency transformation system for the basidiomycete *Ustilago violacea* employing hygromycin resistance and lithium-acetate treatment. *Gene* **80**: 171–176.

Bell, A.A., Wheeler, M.H. 1986. Biosynthesis and functions of fungal melanins. *Annual Review of Phytopathology* **24**: 411–451.

Bennett, J.W., Klich, M. 2003. Mycotoxins. *Clinical Microbiology Reviews* **16**: 497–516.

Benoist, C., O'Hare, K., Breathnach, R., Chambon, P. 1980. The ovalbumin gene-sequence of putative control regions. *Nucleic Acids Research* **8**: 127–142.

Bentley, R., Keil, J.G. 1962, Tetronic acid biosynthesis in molds. II. Formation of penicillic acid in *Penicillium cyclopium*. *The Journal of Biological Chemistry* **237**: 867–872.

Bentley, R., Zwitkowits, P.M. 1967. Biosynthesis of tropolones in *Penicillium stipitatum*. VII. The formation of polyketide lactones and other nontropolone compounds as a result of ethionine inhibition. *Journal of the American Chemical Society* **89**: 676–680.

Bergmann, S., Schümann, J., Scherlach, K., Lange, C., Brakhage, A.A., Hertweck, C. 2007. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nature Chemical Biology* **3**: 213–217.

Bergstrom, J.D., Dufresne, C., Bills, G.F., Nallin-Omstead, M., Byrne, K. 1995. Discovery, biosynthesis, and mechanism of action of the zaragozic acids: potent inhibitors of squalene synthase. *Annual Review of Microbiology* **49**: 607-639.

Bergstrom, J.D., Kurtz, M.M., Rew, D.J., Amend, A.M., Karkas, J.D., Bostedor, R.G., Bansal, V.S., Dufresne, C., VanMiddlesworth, F.L., Hensens, O.D., Liesch, J.M., Zink, D.L., Wilson, K.E., Onishi, J., Milligan, J.A., Bills, G., Kaplan, L., Omstead, M.N., Jenkins, R.G., Huang, L., Meinz, M.S., Quinn, L., Burgh, R.W., Kong, Y.L., Mochales, S., Mojena, M., Martin, I., Pelaez, F., Diez, M.T., Alberts A.W. 1993. Zaragozic acids: a family of fungal metabolites that are picomolar competitive inhibitors of squalene synthase. *Proceed-*

ings of the National Academy of Sciences of the USA **90**: 80–84.

Bevitt, D.J., Cortes, J., Haydock, S.F., Leadlay, P.F. 1992. 6-Deoxyerythronolide-B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *European Journal of Biochemistry* **204**: 39–49.

Bhatnagar, D., Ehrlich, K.C., Cleveland, T.E. 2003. Molecular genetic analysis and regulation of aflatoxin biosynthesis. *Applied Microbiology and Biotechnology* **61**: 83-93.

Bills, G.F., Peldez, F., Polishook, J.D., Diez-Matas, M.T., Harris, G.H., Clapp, W.H., Dufresne, C., Byrne, K.M., Nallin-Omstead, M., Jenkins, R.G., Mojena, M., Huang, L., Bergstrom, J.D. 1994. Distribution of zaragozic acids (squalestatins) among filamentous ascomycetes. *Mycological Research* **98**: 733–739.

Bingle, L.E., Simpson T.J., Lazarus C.M. 1999. Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. *Fungal Genetics and Biology* **26**: 209–223.

Binninger, D.M., Le Chevanton, L., Skrzynia, C., Shubkin, C.D., Pukkila, P.J. 1991. Targeted transformation in *Coprinus cinereus*. *Molecular and General Genetics* **227**: 245–251.

Birren, B., Lander, E., Galagan, J., Devon, K., Nusbaum, C., Jaffe, D., Butler, J., Alvarez, P., Gnerre, S., Grabherr, M., Kleber, M., Mauceli, E., Brockman, W., Rounsley, S., Young, S., LaButti, K., Pushparaj, V., DeCaprio, D., Crawford, M., Koehrsen, M., Engels, R., Montgomery, P., Pearson, M., Howarth, C., Kodira, C., Zeng, Q., Yandava, C., Alvarado, L., Oleary, S., Oliver, R.O., Solomon, P. Annotation of the *Phaeosphaeria nodorum* SN15 genome. *Unpublished*.

Birren, B., Lander, E., Galagan, J., Nusbaum, C., Devon, K., Henn, M., Ma, L.-J., Jaffe, D., Butler, J., Alvarez, P., Gnerre, S., Grabherr, M., Kleber, M., Mauceli, E., Brockman, W., Rounsley, S., Young, S., LaButti, K., Pushparaj, V., DeCaprio, D., Crawford, M., Koehrsen, M., Engels, R., Montgomery, P., Pearson, M., Howarth, C., Larson, L., Luoma, S., White, J., Alvarado, L., Kodira,C., Zeng, Q., Oleary, S., Yandava, C., Denning, D., Nierman, B., Milne,T., Madden, K. Annotation of the *Aspergillus terreus* NIH2624 genome. *Unpublished*.

Birren, B., Lander, E., Galagan, J., Devon, K., Nusbaum, C., Jaffe, D., Butler, J., Alvarez, P., Gnerre, S., Grabherr, M., Kleber, M., Mauceli, E., Brockman, W., Rounsley, S., Young, S., LaButti, K., Pushparaj, V., DeCaprio, D., Crawford, M., Koehrsen, M., Engels, R., Montgomery, P., Pearson, M., Howarth, C., Kodira, C., Zeng, Q., Yandava, C., Alvarado, L., Oleary, S., Oliver, R.O., Solomon, P. Annotation of the *Phaeosphaeria nodorum* SN15 genome. *Unpublished*.

Blackwell, B.A., Edwards, O.E., Fruchier, A., ApSimon, J.W., Miller, J.D. 1996. NMR structural studies of fumonisin B_1 and related compounds from *Fusarium moniliforme*. In: Jackson, L.S., DeVries, J.W., Bullerman, L.B. [*eds.*] *Fumonisins in Food*, pp. 75–91. Plenum, New York.

Blackwell, B.A., Miller, J.D., Savard, M.E. 1994. Production of carbon 14-labeled fumonisin in liquid culture. *Journal of AOAC International* **77**: 506–511.

Bloomer, J.L., Caggiano, T.J., Smith, C.A. 1982. Demonstration of an alternative chain folding in the pyranonapthalene fonsecin. *Tetrahedron Letters* **23**: 5103–5106.

Böhnert, H.U., Fudal, I., Dioh, W., Tharreau, D., Notteghem, J.L., Lebrun, M.H. 2004. A putative polyketide synthase/peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. *Plant Cell* **16**: 2499–2513.

Borges, M.I., Azevedo, M.O., Bonatelli, R.Jr., Felipe, M.S.S., Astolfi-Filho, S. 1990. A practical method for the preparation of total DNA from filamentous fungi. *Fungal Genetics Newsletter* **37**: 10.

Bowyer, P. 2001. DNA-mediated transformation of fungi. In: Talbot, N. [*ed.*] *Molecular and Cellular Biology of Filamentous Fungi*, pp. 33–46. Oxford University Press, Oxford.

Bradshaw, R.E., Bhatnagar, D., Ganley, R.J., Gillman, C.J., Monahan, B.J., Seconi, J.M. 2002. *Dothistroma pini*, a forest pathogen, contains homologs of aflatoxin biosynthetic pathway genes. *Applied and Environmental Microbiology* **68**: 2885–2892.

Bradshaw, R.E., Jin, H., Morgan, B.S., Schwelm, A., Teddy, O.R., Young, C.A., Zhang, S. 2006. A polyketide synthase gene required for biosynthesis of the aflatoxin-like toxin, dothistromin. *Mycopathologia* **161**: 283–294.

Bradshaw, R.E., Zhang, S. 2006. Biosynthesis of dothistromin. *Mycopathologia* **162**: 201–213.

Brakhage, A.A., Langfelder, K. 2002. Menacing mold: the molecular biology of *Aspergillus fumigatus*. *Annual Review of Microbiology* **56**: 433–455.

Branham, B.E., Plattner, R.D. 1993. Alanine is a precursor in the biosynthesis of fumonisin B_1 by *Fusarium moniliforme*. *Mycopathologia* **124**: 99–104.

Brian, P.W., Curtis, P.J., Hemming, H.G., Unwin, C.H., Wright, J.M. 1949. Alternaric acid, a biologically active metabolic product of the fungus *Alternaria solani*. *Nature* **164**: 534.

Brian, P.W., Curtis, P.J., Hemming, H.G., Jefferys, E.G., Unwin, C.H., Wright, J.M. 1951. Alternaric acid – active metabolic product of *Alternaria solani* – production, and antifungal properties. *Journal of General Microbiology* **1**: 619–632.

Brian, P.W., Elson, G.W., Hemming, H.G., Wright, J.M. 1952. Phytotoxic properties of alternaric acid in relation to the etiology plant diseases caused by *Alternaria solani*. *Annals of Applied Biology* **39**: 308–321.

Brodo, M.I., Sharnoff, S.D., Sharnoff, S. 2001. Lichens of North America. Yale University Press.

Brown, D.W., Adams, T.H., Keller, N.P. 1996. *Aspergillus* has distinct fatty acid synthases for primary and secondary metabolism. *Proceedings of the National Academy of Sciences of the USA* **93**: 14873–14877.

Brown, D.W., Hauser, F.M., Tommasi, R., Corlett, S., Salvo, J.J. 1993. Structural elucidation

of a putative conidial pigment intermediate in *Aspergillus parasiticus*. *Tetrahedron Letters* **34**: 419–422.

Brown, D.W., Salvo, J.J. 1994. Isolation and characterization of sexual spore pigments from *Aspergillus nidulans*. *Applied and Environmental Microbiology* **60**: 979–983.

Brown, J.S., Aufauvre-Brown, A., Holden, D.W. 1998. Insertional mutagenesis of *Aspergillus fumigatus*. *Molecular and General Genetics* **259**: 327–335.

Bruchez, J.J.P., Eberle, J., Russo, V.E.A. 1993. Regulatory sequences involved in the translation of *Neurospora crassa* mRNA: Kozak sequences and stop codons. *Fungal Genetics Newsletters* **40**: 85–88.

Brygoo, Y. Debuchy, R. 1985. Transformation by integration in Podospora anserina: methodology and phenomenology. *Molecular and General Genetics* **200**: 128–131

Bu'Lock, J.D., Hamilton, D., Hulme, M.A., Powell, A.J., Smalley, H.M., Shepherd, D., Smith, G.N. 1965. Metabolic development and secondary biosynthesis in *Penicillium urticae*. *Canadian Journal of Microbiology* **11**: 765–778.

Bu'Lock, J.D., Hulme, M.A., Shepherd, D. 1966. Inhibition and possible feedback by 6-methylsalicylic acid. *Nature* **211**: 1090–1091.

Bu'Lock, J.D., Ryan, A.J. 1958. The biogenesis of patulin. *Proceedings of the Chemical Society* **B**: 222–223.

Bu'Lock, J.D., Shepherd, D., Winstanley, D.J. 1969. Regulation of 6-methylsalicylate and patulin synthesis in *Penicillium urticae*. *Canadian Journal of Microbiology* **15**: 279–285.

Burkart, M.D. 2003. Metabolic engineering – a genetic toolbox for small molecule organic synthesis. *Organic and Biomolecular Chemistry* **1**: 1–4.

Butler, M.J., Day, A.W. 1998. Fungal melanins: a review. *Canadian Journal of Microbiology* **44**: 1115–1136.

Buxton, F.P., Gwynne, D.I., Davies, R.W. 1985. Transformation of *Aspergillus niger* using the *arg*B gene of *Aspergillus nidulans*. Gene **37**: 207–214.

Byrne, K.M., Arison, B.H., Nallin-Omstead, M., Kaplan, L. 1993. Biosynthesis of the zaragozic acids. I. Zaragozic acid A. *Journal of Organic Chemistry* **58**: 1019–1024.

Case, M.E., Schweizer, M., Kushner, S.R., Giles, N.H. 1979. Efficient transformation of *Neurospora crassa* by utilizing hybrid plasmid DNA. *Proceedings of the National Academy of Sciences of the USA* **76**: 5259–5263.

Caldas, E.D., Sadilkova, K., Ward, B.L., Jones, A.D., Winter, C.K., Gilchrist, D.G. 1998. Biosynthetic studies of fumonisin B_1 and AAL toxins. *Journal of Agricultural and Food Chemistry* **46**: 4734–4743.

Cavener, D.R., Ray, S.C. 1991. Eukaryotic start and stop translation sites. *Nucleic Acids Research* **19**: 3185–3192.

Chakraborty, B.N., Kapoor, M. 1990. Transformation of filamentous fungi by electroporation. *Nucleic Acids Research* **18**: 6737.

Chakraborty, B.N., Patterson, N.A., Kapoor, M. 1991. An electroporation-based system for high-efficiency transformation of germinated conidia of filamentous fungi. *Canadian Journal of Microbiology* **37**: 858–863.

Chakravarti, R., Sahai, V. Compactin – a review. *Applied Microbiology and Biotechnology* **64**: 618–624.

Chan, L.K., Moore, R.N., Nakashima, T., Vederas, J.S. 1983. Biosynthesis of mevinolin, spectral assignment by double quantum coherence NMR after high carbon-13 incorporation. *Journal of the American Chemical Society* **105**: 3334–3336.

Chang, P.K., Cary, J.W., Yu, J., Bhatnagar, D., Cleveland, T.E. 1995. The *Aspergillus parasiticus* polyketide synthase gene *pks*A, a homolog of *Aspergillus nidulans w*A, is required for aflatoxin B1 biosynthesis. *Molecular and General Genetics* **248**: 270–277.

Cheng, Q., Kinney, K.A., Whitman, C.P., Szaniszlo, P.J. 2004. Characterization of two polyketide synthase genes in *Exophiala lecanii-corni*, a melanized fungus with bioremediation potential. *Bioorganic Chemistry* **32**: 92–108.

Chiang, P.K., Gordon, R.K., Tal, J., Zeng, G.C., Doctor, B.P., Pardhasaradhi, K., McCann, P.P. 1996. *S*-Adenosylmethionine and methylation. *The FASEB Journal* **10**: 471–480.

Christensen, T., Woeldike, H., Boel, E., Mortensen, S.B., Hjortshoej, K., Thim, L., Hansen, M.T. 1988. High level expression of recombinant genes in *Aspergillus oryzae*. *Nature Biotechnology* **6**: 1419–1422.

Chu, F.S. 1991. Mycotoxins: food contamination, mechanism, carcinogenic potential and preventive measures. *Mutation Research* **259**: 291Ũ306.

Ciegler, A., Detroy, R.W., Lillejoj, E.B. 1971. Patulin, penicillic acid and other carcinogenic lactones. In: Ciegler, A., Kadis, S., Ajl, S.J. [*eds.*] *Microbial Toxins. Vol. VI. Fungal Toxins*, pp. 409–434. Academic Press, New York.

Clutterbuck, A.J. 1972. Absence of laccase from yellow-spored mutants of *Aspergillus nidulans*. *Journal of General Microbiology* **70**: 423–435.

Colombo, L., Gennari, C., Potenza, D., Scolastico, C. 1981. Biosynthesis of citrinin and synthesis of its biogenetic precursors. *Journal of the Chemical Society, Perkin Transactions 1*: 2594–2597.

Cox, R.J. 2007. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Organic and Biomolecular Chemistry* **5**: 2010–2026.

Cox, R.J., Glod, F., Hurley, D., Lazarus, C.M., Nicholson, T.P., Rudd, B.A., Simpson, T.J., Wilkinson, B., Zhang, Y. 2004. Rapid cloning and expression of a fungal polyketide synthase gene involved in squalestatin biosynthesis. *Chemical Communications* **20**: 2260–2261.

Crawford, J.M., Dancy, B.C., Hill, E.A., Udwary, D.W., Townsend, C.A. 2006. Identification

of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proceedings of the National Academy of Sciences of the USA* **103**: 16728-16733.

Cridge, A.G., Major, L.L., Mahagaonkar, A.A., Poole, E.S., Isaksson, L.A., Tate, W.P. 2006. Comparison of characteristics and function of translation termination signals between and within prokaryotic and eukaryotic organisms. *Nucleic Acids Research* **34**: 1959–1973.

Dabrah, T.T., Harwood, H.J., Huang, L.H., Jankovich, N.D., Kaneko, T., Li, J.C., Lindsey, S., Moshier, P.M., Subashi, T.A., Therrien, M., Watts, P.C. 1997. CP-225,917 and CP-263,114, novel Ras farnesylation inhibitors from an unidentified fungus. I. Taxonomy, fermentation, isolation, and biochemical properties. *The Journal of Antibiotics* **50**: 1–7.

Danko, S.J., Kono, Y., Daly, J.M., Suzuki, Y., Takeuchi, S., McCrery, D.A. 1984. Structure and biological activity of a host-specific toxin produced by the fungal corn pathogen *Phyllosticta maydis*. *Biochemistry* **23**: 759–766.

Dantas-Barbosa, C., Araújo, E.F., Moraes, L.M., Vainstein, M.H., Azevedo, M.O. 1998. Genetic transformation of germinated conidia of the thermophilic fungus *Humicola grisea var. thermoidea* to hygromycin B resistance. *FEMS Microbiol Letters* **169**: 185–190.

Dao, H.P., Mathieu, F., Lebrihi, A. Molecular cloning and characterization of the 6-methylsalicylic acid synthase (6MSAS) gene in *Aspergillus ochraceus* NRRL 3174. *Unpublished*.

Davis, N.D., Diener, U.L. 1987. Mycotoxins. In: Beuchat, L.R. [*ed.*] *Food and Beverage Mycology*, 2nd edition, pp. 517–570. Van Nostrand Reinhold, New York.

Dawson, M.J., Farthing, J.E., Marshall, P.S., Middleton, R.F., O'Neill, M.J., Shuttleworth, A., Stylli, C., Tait, R.M., Taylor, P.M., Wildman, H.G., Buss, A.D., Langley, D., Hayes, M.V. 1992. The squalestatins, novel inhibitors of squalene synthase produced by a species of *Phoma*. I. Taxonomy, fermentation, isolation, physico-chemical properties and biological activity. *The Journal of Antibiotics* **45**: 639–647.

Dayan, F.E., Romagni, J.G. 2001. Lichens as a potential source of pesticides. *Pesticide Outlook* **12**: 229–232.

de Crécy-Lagard, V., Marliere, P., Saurin, W. 1995. Multienzymatic non-ribosomal peptide biosynthesis: identification of the functional domains catalysing peptide elongation and epimerisation. *Comptes Rendus de l'Académie des Sciences. Série III, Sciences de la Vie* **318**: 927–936.

de Groot, M.J., Bundock, P., Hooykaas, P.J., Beijersbergen, A.G. 1998. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nature Biotechnology* **16**: 839–842.

Dhawale, S.S., Paietta, J.V., Marzluf, G.A. 1984. A new, rapid and efficient transformation procedure for *Neurospora*. *Current Genetics* **8**: 77–79.

Dimroth, P., Walter, H., Lynen, F. 1970. Biosynthese von 6-Methylsalicylsäure. *European Journal of Biochemistry* **13**: 98–110.

Domsch, K.H., Gams, W., Anderson T.-H. 1980. Compendium of soil fungi. Vol. 1. Academic Press,

London.

Dufresne, C., Wilson, K.E., Singh, S.B., Zink, D.L., Bergstrom, J.D., Rew, D., Polishook, J.D., Meinz, M., Huang, L., Silverman, K.C., Lingham, R.B. 1993. Zaragozic acids D and D2: potent inhibitors of squalene synthase and of Ras farnesyl-protein transferase. *Journal of Natural Products* **56**: 1923–1929.

Dufresne, C., Wilson, K.E., Zink, D., Smith, J., Bergstrom, J.D., Kurtz, M., Rew, D., Nallin, M., Jenkins, R., Bartizal, K., Trainor, C., Bills, G., Meinz, M., Huang, L., Onishi, J., Milligan, J., Mojena, M., Pelaez, F. 1992. The isolation and structure elucidation of zaragozic acid C, a novel potent squalene synthase inhibitor. *Tetrahedron* **48**: 10221–10226.

Dynan, W.S., Sazer, S., Tjian, R., Schimke, R.T. 1986. Transcription factor Sp1 recognizes a DNA sequence in the mouse dihydrofolate reductase promoter. *Nature* **319**: 246–248.

Dynan, W.S., Tjian, R. 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. *Nature* **316**: 774–778.

Ebizuka, Y., Sankawa, U., Shibata, S. 1970. The constituents of *Solorina crocea*: averythrin 6-monomethyl ether and methyl gyrophorate. *Phytochemistry* **9**: 2061–2063.

Edelmann, S.E., Staben, C. 1994. A statistical analysis of sequence features within genes from *Neurospora crassa*. *Experimental Mycology* **18**: 70–81.

Eley, K.L., Halo, L.M., Song, Z., Powles, H., Cox, R.J., Bailey, A.M., Lazarus, C.M., Simpson, T.J. 2007. Biosynthesis of the 2-pyridone tenellin in the insect pathogenic fungus *Beauveria bassiana*. *ChemBioChem* **8**: 289–297.

Elvidge, J.A., Jaiswal, D.K., Jones, J.R., Thomas, R. 1977. Tritium nuclear magnetic resonance spectroscopy. Part 7. New information from the tritium distribution in biosynthetically labelled penicillic acid. *Journal of the Chemical Society, Perkin Transactions 1*: 1080–1083.

Endo, A. 1979. Monacolin K, a new hypocholesterolemic agent produced by a *Monascus* species. *The Journal of Antibiotics* **32**: 852–854.

Endo, A., Hasumi, K. 1993. HMG-CoA reductase inhibitors. *Natural Product Reports* **10**: 541–550.

Endo, A., Kuroda, M., Tanzawa, K. 2004. Competitive inhibition of 3-hydroxy-3-methylglutaryl coenzyme A reductase by ML-236A and ML-236B fungal metabolites, having hypocholesterolemic activity. 1976. *Atherosclerosis. Supplements.* **5**: 39–42.

Eriksson, O.E. 1981. The families of bitunicate ascomycetes. *Opera Botanica* **60**: 1–209.

Eriksson, O.E. [*ed.*] 2006. Outline of *Ascomycota* – 2006. *Myconet* **12**: 1–82.

Eriksson, O.E., Strand, A. 1995. Relationships of the genera *Nephroma*, *Peltigera* and *Solorina* (*Peltigerales, Ascomycota*) inferred from 18S rDNA sequences. *Systema Ascomycetum* **14**: 33–39.

Fauman, E.B., Blumenthal, R.M., Cheng, X. 1999. Structure and evolution of AdoMet-dependent methyltransferases. In: Cheng, X., Blumenthal, R.M. [*eds.*] *S-Adenosylmethionine-Dependent Methyltransferases: Structure and Functions*, pp. 1–38. World Scientific Publishing, Singapore.

Feng, B., Wang, X., Hauser, M., Kaufmann, S., Jentsch, S., Haase, G., Becker, J.M., Szaniszlo, P.J. 2001. Molecular cloning and characterization of WdPKS1, a gene involved in dihydroxynaphthalene melanin biosynthesis and virulence in *Wangiella (Exophiala) dermatitidis*. *Infection and Immunity* **69**: 1781–1794.

Feng, G.H., Leonard, T.J. 1995. Characterization of the polyketide synthase gene (*pks*L1) required for aflatoxin biosynthesis in *Aspergillus parasiticus*. *Journal of Bacteriology* **177**: 6246–6254.

Feng, G.H., Leonard, T.J. 1998. Culture conditions control expression of the genes for aflatoxin and sterigmatocystin biosynthesis in *Aspergillus parasiticus* and *A. nidulans*. *Applied and Environmental Microbiology* **64**: 2275–2277.

Fincham J.R.S. 1989. Transformation in fungi. *Microbiological Reviews* **53**: 148–170.

Fungaro, M.H., Rech, E., Muhlen, G.S., Vainstein, M.H., Pascon, R.C., de Queiroz, M.V., Pizzirani-Kleiner, A.A., de Azevedo, J.L. 1995. Transformation of *Aspergillus nidulans* by microprojectile bombardment on intact conidia. *FEMS Microbiology Letters* **125**: 293–297.

Fujii, I. 1999. Polyketide biosynthesis in filamentous fungi. In: Sankawa, U. [*ed.*] *Comprehensive Natural Products Chemistry*, Vol. 1, pp. 409–441. Elsevier, Oxford.

Fujii, I., Mori, Y., Watanabe, A., Kubo, Y., Tsuji, G., Ebizuka, Y. 1999. Heterologous expression and product identification of *Colletotrichum lagenarium* polyketide synthase encoded by the *PKS*1 gene involved in melanin biosynthesis. *Bioscience, Biotechnology, and Biochemistry* **63**: 1445–1452.

Fujii, I., Ono, Y., Tada, H., Gomi, K., Ebizuka, Y., Sankawa, U. 1996. Cloning of the polyketide synthase gene *at*X from *Aspergillus terreus* and its identification as the 6-methylsalicylic acid synthase gene by heterologous expression. *Molecular and General Genetics* **253**: 1–10.

Fujii, I., Watanabe, A., Ebizuka, Y. 2004. More functions for multifunctional polyketide synthases. In: Tkacz, J.S., Lange, L. [*eds.*] *Advances in Fungal Biotechnology for Industry, Agriculture, and Medicine*, pp. 97–125. Kluwer Academic/Plenum Publishers, New York.

Fujii, I., Watanabe, A., Mori, Y., Ebizuka, Y. 1998. Structures and functional analyses of fungal polyketide synthase genes. *Actinomycetologica* **12**: 1–14.

Fujii, I., Watanabe, A., Sankawa, U., Ebizuka, Y. 2001. Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of *Aspergillus nidulans*. *Chemistry and Biology* **8**: 189–197.

Fujii, I., Yasuoka, Y., Tsai, H.F., Chang, Y.C., Kwon-Chung, K.J., Ebizuka, Y. 2004. Hydrolytic polyketide shortening by Ayg1p, a novel enzyme involved in fungal melanin biosynthesis. *The Journal of Biological Chemistry* **279**: 44613–44620.

Fujii, I., Yoshida, N., Shimomaki, S., Oikawa, H., Ebizuka, Y. 2005. An iterative type I polyketide synthase PKSN catalyzes synthesis of the decaketide alternapyrone with region-specific octamethylation. *Chemistry and Biology* **12**: 1301–1309.

Fujii, T., Yamaoka, H., Gomi, K., Kitamoto, K., Kumagai, C. 1995. Cloning and nucleotide se-

quence of the ribonuclease T1 gene (*rnt*A) from *Aspergillus oryzae* and its expression in *Saccharomyces cerevisiae* and *Aspergillus oryzae*. *Bioscience, Biotechnology and Biochemistry* **59**: 1869–1874.

Fulton, T.R., Ibrahim, N., Losada, M.C., Grzegorski, D., Tkacz, J.S. 1999. A melanin polyketide synthase (PKS) gene from *Nodulisporium sp.* that shows homology to the *pks*1 gene of *Colletotrichum lagenarium*. *Molecular and General Genetics* **262**: 714–720.

Gaisser, S., Trefzer, A., Stockert, S., Kirschning, A., Bechthold, A. 1997. Cloning of an avilamycin biosynthetic gene cluster from *Streptomyces viridochromogenes* Tü57. *Journal of Bacteriology* 179: 6271–6278.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.J., Smirnov, S., Purcell, S., Rehman, B., Elkins, T., Engels, R., Wang, S., Nielsen, C.B., Butler, J., Endrizzi, M., Qui, D., Ianakiev, P., Bell-Pedersen, D., Nelson, M.A., Werner-Washburne, M., Selitrennikoff, C.P., Kinsey, J.A., Braun, E.L., Zelter, A., Schulte, U., Kothe, G.O., Jedd, G., Mewes, W., Staben, C., Marcotte, E., Greenberg, D., Roy, A., Foley, K., Naylor, J., Stange-Thomann, N., Barrett, R., Gnerre, S., Kamal, M., Kamvysselis, M., Mauceli, E., Bielke, C., Rudd, S., Frishman, D., Krystofova, S., Rasmussen, C., Metzenberg, R.L., Perkins, D.D., Kroken, S., Cogoni, C., Macino, G., Catcheside, D., Li, W., Pratt, R.J., Osmani, S.A., DeSouza, C.P., Glass, L., Orbach, M.J., Berglund, J.A., Voelker, R., Yarden, O., Plamann, M., Seiler, S., Dunlap, J., Radford, A., Aramayo, R., Natvig, D.O., Alex, L.A., Mannhaupt, G., Ebbole, D.J., Freitag, M., Paulsen, I., Sachs, M.S., Lander, E.S., Nusbaum, C., Birren, B. 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422**: 859–868.

Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.J., Wortman, J.R., Batzoglou, S., Lee, S.I., Basturkmen, M., Spevak, C.C., Clutterbuck, J., Kapitonov, V., Jurka, J., Scazzocchio, C., Farman, M., Butler, J., Purcell, S., Harris, S., Braus, G.H., Draht, O., Busch, S., D'Enfert, C., Bouchier, C., Goldman, G.H., Bell-Pedersen, D., Griffiths-Jones, S., Doonan, J.H., Yu, J., Vienken, K., Pain, A., Freitag, M., Selker, E.U., Archer, D.B., Penalva, M.A., Oakley, B.R., Momany, M., Tanaka, T., Kumagai, T., Asai, K., Machida, M., Nierman, W.C., Denning, D.W., Caddick, M., Hynes, M., Paoletti, M., Fischer, R., Miller, B., Dyer, P., Sachs, M.S., Osmani, S.A., Birren, B.W. 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* **438**: 1105–1115.

Galloway, D.J. 1985. Flora of New Zealand. Lichens. P.D. Hasselberg, Government Printer, Wellington, New Zealand.

Galmarini, O.L., Stodola, F.H. 1965. Fonsecin, a pigment of an *Aspergillus fonsecaeus* mutant. *Journal of Organic Chemistry* **30**: 112–115.

Galmarini, O.L., Stodola, F.H., Raper, K.B., Fennell, D.I. 1962. Fonsecin, a naphthopyrone pigment from a mutant of *Aspergillus fonsecaeus*. *Nature* **195**: 502–503.

Gardner, J.M., Kono, Y., Tatum, J.H., Suzuki, Y., Takeuchi, S. 1985. Structure of major component of ACRL toxins, host-specific phytotoxic compound produced by *Alternaria citri*. *Agricultural and Biological Chemistry* **49**: 1235–1238.

Garibova, L.V., Dundin, Yu.K., Koptyaeva, T.F., Filin, V.P. 1978. Algae, lichens and bryophytes of the USSR. (*in Russian*). Misl, Moscow.

Garson, M.J., Staunton, J. 1976. Biosynthesis of sclerin, a metabolite of *Sclerotinia sclerotiorum*: incorporation of [1-^{13}C]- and [1,2-$^{13}C_2$]-acetates. *Journal of the Chemical Society. Chemical Communications*: 928–929.

Garson, M.J., Staunton, J. 1978. Biosynthesis of sclerin: incorporation studies with advanced precursors. *Journal of the Chemical Society. Chemical Communications*: 158–159.

Gaucher, G.M., Shepherd, M.G. 1968. Isolation of orsellinic acid synthase. *Biochemical and Biophysical Research Communications* **32**: 664–671.

Goldman, G.H., van Montagu, M. Herrera-Estrella, A. 1990. Transformation of *Trichoderma harzianum* by high-voltage electric pulse. *Current Genetics* **17**: 169–174.

Gomi, K., Iimura, Y., Hara, S. 1987. Integrative transformation of *Aspergillus oryzae* with a plasmid containing the *Aspergillus nidulans arg*B gene. *Agricultural and Biological Chemistry* **51**: 2549–2555.

Goosen, T., Bloemheuvel, G., Gysler, C., de Bie, D.A., van den Broek, H.W., Swart, K. 1987. Transformation of *Aspergillus niger* using the homologous orotidine-5′-phosphate-decarboxylase gene. *Current Genetics* **11**: 499–503.

Graziani, S., Vasnier, C., Daboussi, M.J. 2004. Novel polyketide synthase from *Nectria haematococca*. *Applied and Environmental Microbiology* **70**: 2984–2988.

Guiry, M.D., Guiry, G.M. 2007. AlgaeBase version 4.2. World-wide electronic publication, National University of Ireland, Galway. *http://www.algaebase.org*; searched on 05 May 2007.

Gunde-Cimerman, N., Cimerman, A. 1995. *Pleurotus* fruiting bodies contain the inhibitor of 3-hydroxy-3-methylglutaryl-coenzyme A reductase – lovastatin. *Experimental Mycology* **19**: 1–6.

Gunde-Cimerman, N., Plemenitas, A., Cimerman, A. 1993. *Pleurotus* fungi produce mevinolin, an inhibitor of HMG-CoA reductase. *FEMS Microbiology Letters* **113**: 333–337.

Gurr, S.J., Unkles, S.E., Kinghorn, J.R. 1987. The structure and organization of nuclear genes of filamentous fungi. In: Kinghorn, J.R. [*ed.*] *Gene structure in eukaryotic microbes*, pp. 93–139. IRL Press, Oxford.

Hacking, A., Rosser, W.R. 1981. Patulin production by *Paecilomyces* species isolated from silage in the United Kingdom. *Journal of the Science of Food and Agriculture* **32**: 620–623.

Hafellner, J. 1988. Principles of classification and main taxonomic groups. In: Galun, M. [*ed.*] *CRC Handbook of Lichenology*, vol. III, pp. 41–52. CRC Press, Boca Raton, Florida.

Hafellner, J., Hertel, H., Rambold, G., Timdal, E. 1994. Discussion 4. *Lecanorales.* In: Hawksworth, D.L. [*ed.*] *Ascomycete Systematics: Problems and Perspectives in the Nineties*,

pp. 379–387. Plenum Press, New York.

Hajjaj, H., Klaebe, A., Loret, M.O., Goma, G., Blanc, P.J., Francois, J. 1999. Biosynthetic pathway of citrinin in the filamentous fungus *Monascus ruber* as revealed by ^{13}C nuclear magnetic resonance. *Applied and Environmental Microbiology* **65**: 311–314.

Hazell, B.W., Te'o, V.S., Bradner, J.R., Bergquist, P.L., Nevalainen, K.M. 2000. Rapid transformation of high cellulase producing mutant strains of *Trichoderma reesei* by microprojectile bombardment. *Letters in Applied Microbiology* **30**: 282–286.

Hendrickson, L., Davis, C.R., Roach, C., Nguyen, D.K., Aldrich, T., McAda, P.C., Reeves, C.D. 1999. Lovastatin biosynthesis in *Aspergillus terreus*: characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chemistry and Biology* **6**: 429–439.

Hensens, O.D., Dufresne, C., Liesch, J.M., Zink, D.L., Reamer, R.A., VanMiddlesworth, F. 1993. The zaragozic acids: structure elucidation of a new class of squalene synthase inhibitors. *Tetrahedron Letters* **34**: 399–402.

Henson, J.M., Butler, M.J., Day, A.W. 1999. The dark side of the mycelium: melanins of phytopathogenic fungi. *Annual Review of Phytopathology* **37**: 447–471.

Hilber, U.W., Bodmer, M., Smith, F.D., Köller, W. 1994. Biolistic transformation of conidia of *Botryotinia fuckeliana*. *Current Genetics* **25**: 124–127.

Hoffman, W.F., Alberts, A.W., Anderson, P.S., Chen, J.S., Smith, R.L., Willard, A.K. 1986. 3-Hydroxy-3-methylglutaryl-coenzyme A reductase inhibitors. 4. Side chain ester derivatives of mevinolin. *Journal of Medicinal Chemistry* **29**: 849–852.

Hooykaas, P.J.J. 2004. Transformation mediated by *Agrobacterium tumefaciens*. In: Tkacz, J.S., Lange, L. [*eds.*] *Advances in Fungal Biotechnology for Industry, Agriculture, and Medicine*, pp. 41–68. Kluwer Academic/Plenum Publishers, New York.

Hosoya, T., Tanimoto, T., Onodera, K., Kurihara, Y., Takamatsu, Y., Tsujita, Y. 1997. Zaragozic acids production from discomycetes. *Mycoscience* **38**: 305–311.

Howard, P.C., Eppley, R.M., Stack, M.E., Warbritton, A., Voss, K.A., Lorentzen, R.J., Kovach, R.M., Bucci, T.J. 2001. Fumonisin B_1 carcinogenicity in a two-year feeding study using F344 rats and B6C3F_1 mice. *Environmental Health Perspectives* **109** (Supplement 2): 277–282.

Huneck, S. 1999. The significance of lichens and their metabolites. *Naturwissenschaften* **86**: 559–570.

Huneck, S., Yoshimura, I. 1996. Identification of lichen substances. Springer-Verlag, Berlin Heidelberg.

Hutner, S., Provasoli, L., Schatz, A., Haskins, C.P. 1950. Some approaches to the study of the role of metals in the metabolism of microorganisms. *Proceedings of The American Philosophical Society* **94**: 152–170.

Hynes, M.J. 1996. Genetic transformation of filamentous fungi. *Journal of Genetics* **73**: 297–311.

Ichihara, A., Miki, M., Sakamura, S. 1985. Absolute configuration of (−)-solanapyrone A. *Tetrahe-*

dron Letters **26**: 2453–2454.

Ichihara, A., Oikawa, H. 1997. Biosynthesis of phytotoxins from *Alternaria solani*. *Bioscience, Biotechnology, and Biochemistry* **61**: 12–18.

Ichihara, A., Oikawa, H. 1999. The Diels-Alder reaction in biosynthesis of polyketide phytotoxins. In: Sankawa, U. [*ed.*] *Comprehensive Natural Products Chenmistry*, Vol. 1, pp. 367–408. Elsevier, Oxford.

Ichihara, A., Tazaki, H., Sakamura, S. 1983. Solanapyrones A, B and C, phytotoxic metabolites from the fungus *Alternaria solani*. *Tetrahedron Letters* **24**: 5373–5376.

Igel, M., Sudhop, T., von Bergmann, K. 2001. Metabolism and drug interactions of 3-hydroxy-3-methylglutaryl coenzyme A-reductase inhibitors (statins). *European Journal of Clinical Pharmacology* **57**: 357–364.

Ishitani, C., Sakaguchi, K. 1955. Hereditary variation and genetic recombination in *Koji*-molds (*Aspergillus oryzae* and *Asp. sojae*). IV. Cytological studies. *Journal of General and Applied Microbiology* **1**: 283–297.

Jacobson, E.S. 2000. Pathogenic roles for fungal melanins. *Clinical Microbiology Reviews* **13**: 708–717.

Jahn, B., Koch, A., Schmidt, A., Wanner, G., Gehringer, H., Bhakdi, S., Brakhage, A.A. 1997. Isolation and characterization of a pigmentless-conidium mutant of *Aspergillus fumigatus* with altered conidial surface and reduced virulence. *Infection and Immunity* **65**: 5110–5117.

Jiang, Q., Ying, S.H., Feng, M.G. 2007. Enhanced frequency of *Beauveria bassiana* blastospore transformation by restriction enzyme-mediated integration and electroporation. *Journal of Microbiological Methods* **69**: 512–517.

John, M.A., Peberdy, J.F. 1984. Transformation of *Aspergillus nidulans* using the *arg*B gene. *Enzyme and Microbial Technology* **6**: 386–389.

Jones, C.A., Sidebottom, P.J., Cannell, R.J.P., Noble, D., Rudd, B.A.M. 1992. The squalestatins, novel inhibitors of squalene synthase produced by a species of *Phoma*. III. Biosynthesis. *The Journal of Antibiotics* **45**: 1492–1498.

Johnstone, I.L., Hughes, S.G., Clutterbuck, A.J. 1985. Cloning an *Aspergillus nidulans* developmental gene by transformation. *The EMBO Journal* **4**: 1307–1311.

Jordan, P.M., Spencer, J.B. 1990. Stereospecific manipulation of hydrogen atoms with opposite absolute orientations during the biosynthesis of the polyketide 6-methylsalicylic acid from chiral malonates in *Penicillium patulum*. *Journal of the Chemical Society. Chemical communications*: 238–242.

Jordan, P.M., Spencer, J.B. 1991. Mechanistic and stereochemical investigation of fatty acid and polyketide biosynthesis using chiral malonates. *Tetrahedon* **47**: 6015–6028.

Jordan, P.M., Spencer, J.B. 1993. The biosynthesis of tetraketides: enzymology, mechanism and molecular programming. *Biochemical Society Transaction* **21**: 222–228.

Julien, B., Shah, S., Ziermann, R., Goldman, R., Katz, L., Khosla, C. 2000. Isolation and

characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*. *Gene* **249**: 153–160.

Kagan, R.M., Clarke, S. 1994. Widespread occurrence of three sequence motifs in diverse *S*-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Archives of Biochemistry and Biophysics* **310**: 417–427.

Karolewiez, A., Geisen, R. 2005. Cloning a part of the ochratoxin A biosynthetic gene cluster of *Penicillium nordicum* and characterization of the ochratoxin polyketide synthase gene. *Systematic and Applied Microbiology* **28**: 588–595.

Kasahara, K., Fujii, I., Oikawa, H., Ebizuka, Y. 2006. Expression of *Alternaria solani* PKSF generates a set of complex reduced-type polyketides with different carbon-lengths and cyclization. *ChemBioChem* **7**: 920–924.

Kealey, J.T., Liu, L., Santi, D.V., Betlach, M.C., Barr, P.J. 1998. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proceedings of the National Academy of Sciences of the USA* **95**: 505–509.

Keatinge-Clay, A.T., Shelat, A.A., Savage, D.F., Tsai, S.C., Miercke, L.J., O'Connell, J.D. III, Khosla, C., Stroud, R.M. 2003. Catalysis, specificity, and ACP docking site of *Streptomyces coelicolor* malonyl-CoA:ACP transacylase. *Structure* **11**: 147–154.

Keller, N.P., Brown, D., Butchko, R.A.E., Fernandes, M., Kelkar, H.S., Nesbitt, T.C., Segner, S., Bhatnagar, D. 1995. A conserved polyketide mycotoxin gene cluster in *Aspergillus nidulans*. In: Eklund, M., Richard, R.L., Mise, K. [*eds.*] *Molecular Approaches to Food Safety Issues Involving Toxic Microorganisms*, pp. 263–277. Alaken, Fort Collins.

Kelly, J.M., Hynes, M.J. 1985. Transformation of *Aspergillus niger* by the *amd*S gene of *Aspergillus nidulans*. *EMBO Journal* **4**: 475–479.

Kempken, F., Jacobsen, S., Kück, U. 1998. Distribution of the fungal transposon Restless: full-length and truncated copies in closely related strains. *Fungal Genetics and Biology* **25**: 110–118.

Kempken, F., Kuck, U. 2000. Tagging of a nitrogen pathway-specific regulator gene in *Tolypocladium inflatum* by the transposon Restless. *Molecular and General Genetics* **263**: 302–308.

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

Kodama, M., Rose, M.S., Yang, G., Yun, S.H., Yoder, O.C., Turgeon, B.G. 1999. The translocation-associated *Tox*1 locus of *Cochliobolus heterostrophus* is two genetic elements on two different chromosomes. *Genetics* **151**: 585–96.

Kono, Y., Danko, S.J., Suzuki, Y., Takeuchi, S., Daly, J.M. 1983. Structure of the host-specific pathotoxins produced by *Phyllosticta maydis*. *Tetrahedron Letters* **24**: 3803–3806.

Kono, Y., Daly, J.M. 1979. Characterization of the host-specific pathotoxin produced by *Helminthospo-*

rium maydis, race T, affecting corn with Texas male sterile cytoplasm. *Bioorganic Chemistry* **8**: 391–397.

Kono, Y., Takeuchi, S., Kawarada, A., Daly, J.M., Knoche, H.W. 1980. Structure of the host-specific pathotoxins produced by *Helminthosporium maydis*, race T. *Tetrahedron Letters* **21**: 1537–1540.

Kono, Y., Takeuchi, S., Kawarada, A., Daly, J.M., Knoche, H.W. 1981. Studies on the host-specific pathotoxins produced in minor amounts by *Helminthosporium maydis*, race T. *Bioorganic Chemistry* **10**: 206–218.

Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acid Research* **12**: 857–872.

Kozak, M. 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research* **15**: 8125–8148.

Krappmann, S., Sasse, C., Braus, G. 2006. Gene targeting in *Aspergillus fumigatus* by homologous recombination is facilitated in a nonhomologous end-joining-deficient genetic background. *Eukaryotic Cell* **5**: 212–215.

Kroken, S., Glass, N.L., Taylor, J.W., Yoder, O.C., Turgeon, B.G. 2003. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proceedings of the National Academy of Sciences of the USA* **100**: 15670–15675.

Kubo, Y., Nakamura, H., Kobayashi, K., Okuno, T., Furusawa, T. 1991. Cloning of a melanin biosynthetic gene essential for appressorial penetration of *Colletotrichum lagenarium*. *Molecular Plant-Microbe Interactions* **4**: 440–445.

Kumagai, H., Tomoda, H., Omura, S. 1992. Biosynthesis of antibiotic 1233A (F-244) and preparation of [^{14}C]1233A. *Journal of Antibiotics* **45**: 563–567.

Kwon-Chung, K.J., Goldman, W.E., Klein, B., Szaniszlo, P.J. 1998. Fate of transforming DNA in pathogenic fungi. *Medical Mycology* **36** (Supplement 1): 38–44.

Lam, K.S., Neway, J.O., Gaucher, G.M. 1988. In vitro stabilization of 6-methylsalicylic acid synthetase from *Penicillium urticae*. *Canadian Journal of Microbiology* **34**: 30–37.

Langfelder, K., Streibel, M., Jahn, B., Haase, G., Brakhage, A.A. 2003. Biosynthesis of fungal melanins and their importance for human pathogenic fungi. *Fungal Genetics and Biology* **38**: 143–158.

Lee, T., Yun, S.H., Hodge, K.T., Humber, R.A., Krasnoff, S.B., Turgeon, G.B., Yoder, O.C., Gibson, D.M. 2001. Polyketide synthase genes in insect- and nematode-associated fungi. *Applied Microbiology and Biotechnology* **56**: 181–187.

Leeper, F.J., Staunton, J. 1984. The biosynthesis of rubrofusarin, a polyketide naphthopyrone from *Fusarium culmorum*: ^{13}C n.m.r. assignments and incorporation of ^{13}C- and ^{2}H-labelled acetates. *Journal of the Chemical Society, Perkin Transactions 1*: 2919–2925.

Light, R.J. 1967. The biosynthesis of 6-methylsalicylic acid. *The Journal of Biological Chemistry* **242**: 1880–1886.

Linnemannstons, P., Schulte, J., del Mar Prado, M., Proctor, R.H., Avalos, J., Tudzynski, B. 2002. The polyketide synthase gene *pks*4 from *Gibberella fujikuroi* encodes a key enzyme in the biosynthesis of the red pigment bikaverin. *Fungal Genetics and Biology* **37**: 134-48.

Loppnau, P., Tanguay, P., Breuil, C. 2004. Isolation and disruption of the melanin pathway polyketide synthase gene of the softwood deep stain fungus *Ceratocystis resinifera*. *Fungal Genetics and Biology* **41**: 33–41.

Lorito, M., Hayes, C.K., Di Pietro, A., Harman, G.E. 1993. Biolistic transformation of *Trichoderma harzianum* and *Gliocladium virens* using plasmid and genomic DNA. *Current Genetics* **24**: 349–356.

Lu, S., Lyngholm, L., Yang, G., Bronson, C., Yoder, O.C., Turgeon, B.G. 1994. Tagged mutations at the *Tox*1 locus of *Cochliobolus heterostrophus* by restriction enzyme-mediated integration. *Proceedings of the National Academy of Sciences of the USA* **91**: 12649–12653.

Lu, P., Zhang, A., Dennis ,L.M., Dahl-Roshak, A.M., Xia, Y.-Q., Arison, B., An, Z., Tkacz, J.S. 2005. A gene (*pks*2) encoding a putative 6-methylsalicylic acid synthase from *Glarea lozoyensis*. *Molecular Genetics and Genomics* **273**: 207–216.

Lynen, F., Tada, M. 1961. Die biochemischen Grundlagen der "Polyacetat-Regel". *Angewandte Chemie* **73**: 513–519.

Ma, S.M., Tang, Y. 2007. Biochemical characterization of the minimal polyketide synthase domains in the lovastatin nonaketide synthase LovB. *FEBS Journal* **274**: 2854–2864.

Ma, S.M., Zhan, J., Watanabe, K., Xie, X., Zhang, W., Wang, C.C., Tang, Y. 2007. Enzymatic synthesis of aromatic polyketides using PKS4 from *Gibberella fujikuroi*. *Journal of the American Chemical Society* **129**: 10642–10643.

Mach, R.L. 2004. Transformation and gene manupulation in filamentous fungi: an overview. In: Arora, D.K. [*ed.*] *Handbook of Fungal Biotechnology*, 2nd edition, pp. 109–120. CRC Press.

Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., Kusumoto, K., Arima, T., Akita, O., Kashiwagi, Y., Abe, K., Gomi, K., Horiuchi, H., Kitamoto, K., Kobayashi, T., Takeuchi, M., Denning, D.W., Galagan, J.E., Nierman, W.C., Yu, J., Archer, D.B., Bennett, J.W., Bhatnagar, D., Cleveland, T.E., Fedorova, N.D., Gotoh, O., Horikawa, H., Hosoyama, A., Ichinomiya, M., Igarashi, R., Iwashita, K., Juvvadi, P.R., Kato, M., Kato, Y., Kin, T., Kokubun, A., Maeda, H., Maeyama, N., Maruyama, J., Nagasaki, H., Nakajima, T., Oda, K., Okada, K., Paulsen, I., Sakamoto, K., Sawano, T., Takahashi, M., Takase, K., Terabayashi, Y., Wortman, J.R., Yamada, O., Yamagata, Y., Anazawa, H., Hata, Y., Koide, Y., Komori, T., Koyama, Y., Minetoki, T., Suharnan, S., Tanaka, A., Isono, K., Kuhara, S., Ogasawara, N., Kikuchi, H. 2005. Genome sequencing

and analysis of *Aspergillus oryzae*. *Nature* **438**: 1157–1161.

Mahanti, N., Bhatnagar, D., Cary, J.W., Joubran, J., Linz, J.E. 1996. Structure and function of *fas*-1A, a gene encoding a putative fatty acid synthetase directly involved in aflatoxin biosynthesis in *Aspergillus parasiticus*. *Applied and Environmental Microbiology* **62**: 191–195.

Maier, F.J., Malz, S., Losch, A.P., Lacour, T., Schafer, W. 2005. Development of a highly efficient gene targeting system for *Fusarium graminearum* using the disruption of a polyketide synthase gene as a visible marker. *FEMS Yeast Research* **5**: 653–662.

Mann, J. 1987. Secondary metabolism. Oxford University Press, Oxford.

Marahiel, M.A., Stachelhaus, T., Mootz, H.D. 1997. Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chemical Reviews* **97**: 2651–2674.

Marasas, W.F.O. 1996. Fumonisins: history, world-wide occurrence and impact. In: Jackson, L.S., DeVries, J.W., Bullerman, L.B. [*eds.*] *Fumonisins in Food*, pp. 1–17. Plenum, New York.

Martínez, I., Burgaz, A.R. 1998. Revision of the genus *Solorina* (*Lichenes*) in Europe based on spore size variation. *Annales Botanici Fennici* **35**: 137–142.

Mattern, I.E., Unkles, S., Kinghorn, J.R., Pouwels, P.H., van den Hondel, C.A.M.J.J. 1987. Transformation of *Aspergillus oryzae* using the *A. niger pyr*G gene. *Molecular and General Genetics* **210**: 460–461.

Mauro, V.F., MacDonald, J.L. 1991. Simvastatin: a review of its pharmacology and clinical use. *DICP: the Annals of Pharmacotherapy* **25**: 257–264.

Mayorga, M.E., Timberlake, W.E. 1990. Isolation and molecular characterization of the *Aspergillus nidulans w*A gene. *Genetics* **126**: 73–79.

Mayorga, M.E., Timberlake, W.E. 1992. The developmentally regulated *Aspergillus nidulans w*A gene encodes a polypeptide homologous to polyketide and fatty acid synthases. *Molecular and General Genetics* **235**: 205–212.

McDaniel, R., Ebert-Khosla, S., Hopwood, D.A., Khosla, C. 1993. Engineered biosynthesis of novel polyketides. *Science* **262**: 1546–1550.

McLauchlan, J., Gaffney, D., Whitton, J.L., Clements, J.B. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3′ termini. *Nucleic Acids Research* **13**: 1347–1368.

McNaught, A.D., Wilkinson A. [*comps.*] 1997. Compendium of Chemical Terminology - IUPAC Recommendations (IUPAC Chemical Nomenclature Series). 2nd edition. Blackwell Science.

McTavish, D., Sorkin, E.M. 1991. Pravastatin. A review of its pharmacological properties and therapeutic potential in hypercholesterolaemia. *Drugs* **42**: 65–89.

Merrill, A.H. Jr., Schmelz, E.M., Dillehay, D.L., Spiegel, S., Shayman, J.A., Schroeder, J.J., Riley, R.T., Voss, K.A., Wang, E. 1997. Sphingolipids – the enigmatic lipid class: biochemistry, physiology, and pathophysiology. *Toxicology and Applied Pharmacology* **142**: 208–225.

Meyer, V., Arentshorst, M., El-Ghezal, A., Drews, A.C., Kooistra, R., van den Hondel, C.A., Ram, A.F. 2007. Highly efficient gene targeting in the *Aspergillus niger kus*A mutant. *Journal of Biotechnology* **128**: 770Ũ-775.

Miadlikowska, J., Kauff, F., Hofstetter, V., Fraker, E., Grube, M., Hafellner, J., Reeb, V., Hodkinson, B.P., Kukwa, M., Lucking, R., Hestmark, G., Otalora, M.G., Rahut, A., Budel, B., Scheidegger, C., Timdal, E., Stenroos, S., Brodo, I., Perlmutter, G.B., Ertz, D., Diederich, P., Lendemar, J.C., May, P., Schoch, C.L., Arnold, A.E., Gueidan, C., Tripp, E., Yahr, R., Robertson, C., Lutzoni, F. 2006. New insights into classification and evolution of the *Lecanoromycetes* (*Pezizomycotina, Ascomycota*) from phylogenetic analyses of three ribosomal RNA- and two protein-coding genes. *Mycologia* **98**: 1088–1103.

Miadlikowska, J.M., Lutzoni, F.M. 2000. Phylogenetic revision of the genus *Peltigera* (lichen-forming ascomycota) based on morphological, chemical and large subunit nuclear ribosomal DNA data. *International Journal of Plant Science* **161**: 925–958.

Miadlikowska, J., Lutzoni, F. 2004. Phylogenetic classification of peltigeralean fungi (*Peltigerales, Ascomycota*) based on ribosomal RNA small and large subunits. *American Journal of Botany* **91**: 449-464.

Miao, V., Coëffet-LeGal, M.F., Brown, D., Sinnemann, S., Donaldson, G., Davies, J. 2001. Genetic approaches to harvesting lichen products. *Trends in Biotechnology* **19**: 349–355.

Miller, D.J., Ouellette, N., Evdokimova, E., Savchenko, A., Edwards, A., Anderson, W.F. 2003. Crystal complexes of a predicted *S*-adenosylmethionine-dependent methyltransferase reveal a typical AdoMet binding domain and substrate recognition domain. *Protein Science* **12**: 1432–1442.

Mishra, N. C., Szabo, G., Tatum, E.L. 1973. Nucleic acid induced genetic changes in *Neurospora*. In: Niu, M.C., Segal, S.G. [*eds.*] *The Role of RNA in Reproduction and Development.*, pp. 259–268. Elsevier/North Holland Publishing, Amsterdam.

Mishra, N.C., Tatum, E.L. 1973. Non-Mendelian inheritance of DNA-induced inositol independence in *Neurospora*. *Proceedings of the National Academy of Sciences of the USA* **70**: 3875–3879.

Molnár, I., Aparicio, J.F., Haydock, S.F., Khaw, L.E., Schwecke, T., Konig, A., Staunton, J., Leadlay, P.F. 1996. Organisation of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of genes flanking the polyketide synthase. *Gene* **169**: 1–7.

Molnár, I., Schupp, T., Ono, M., Zirkle, R., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratmann, A., Cyr, D.D., Gorlach, J., Mayo, J.M., Hu, A., Goff, S., Schmid, J., Ligon, J.M. 2000. The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. *Chemistry and Biology* **7**: 97–109.

Moore, R.N., Bigam, G., Chan, J.K., Hogg, A.M., Nakasima, T.T., Vederas J.C. 1985. Biosynthesis of the hypocholesterolemic agent mevinolin by *Aspergillus terreus*. Determination of the origin of carbon, hydrogen and oxygen atoms by ^{13}C NMR and mass spectrometry. *Journal of the American*

Chemical Society **107**: 3694–3701.

Moriwaki, A., Kihara, J., Kobayashi, T., Tokunaga, T., Arase, S., Honda, Y. 2004. Insertional mutagenesis and characterization of a polyketide synthase gene (PKS1) required for melanin biosynthesis in *Bipolaris oryzae*. *FEMS Microbiology Letters* **238**: 1–8.

Moriwaki, A., Ueno, M., Arase, S., Kihara, J. 2007. RNA-mediated gene silencing in the phytopathogenic fungus *Bipolaris oryzae*. *FEMS Microbiology Letters* **269**: 85–89.

Mosbach, K. 1959. Presence of orsellinic acid in *Chaetomium cochliodes*. *Zeitschrift für Naturforschung. Teil B: Chemie, Biochemie, Biophysik, Biologie* **14**: 69–70.

Mosbach, K. 1960. Die Biosynthese der Orsellinsäure und Penicillinsäure. *Acta Chemica Scandinavica* **14**: 457–464.

Müller, K. 2001. Pharmaceutically relevant metabolites from lichens. *Applied Microbiology and Biotechnology* **56**: 9–16.

Müller, R. 2004. Don't classify polyketide synthases. *Chemistry and Biology* **11**: 4–6.

Nakamura, T., Komagata, D., Murakawa, S., Sakai, K., Endo, A. 1990. Isolation and biosynthesis of 3α-hydroxy-3,5-dihydromonacolin. *Journal of Antibiotics* **43**: 1597–1600.

Nakayashiki, H., Hanada, S., Nguyen, B.Q., Kadotani, N., Tosa, Y., Mayama, S. 2005. RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genetics and Biology* **42**: 275–283.

Nayak, T., Szewczyk, E., Oakley, C.E., Osmani, A., Ukil, L., Murray, S.L., Hynes, M.J., Osmani, S.A., Oakley, B.R. 2006. A versatile and efficient gene targeting system for *Aspergillus nidulans*. *Genetics* **172**: 1557–1566.

Nelson, P.E., Desjardins, A.E., Platter, R.D. 1993. Fumonisins, mycotoxins produced by *Fusarium* species: biology, chemistry, and significance. *Annual Review of Phytopathology* **31**: 233–252.

Nicholson, T.P., Rudd, B.A., Dawson, M., Lazarus, C.M., Simpson, T.J., Cox, R.J. 2001. Design and utility of oligonucleotide gene probes for fungal polyketide synthases. *Chemistry and Biology* **8**: 157–178.

Nicolaisen, M., Frisvad, J.C., Rossen, L. 1997. A *Penicillium freii* gene that is highly similar to the β-ketoacyl synthase domain of polyketide synthase genes from other fungi. *Letters in Applied Microbiology* **25**: 197–201.

Nielsen, J.B., Nielsen, M.L., Mortensen, U.H. 2007. Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus *Aspergillus nidulans*. *Fungal Genetics and Biology*, doi:10.1016/j.fgb.2007.07.003.

Nierman, W.C. 2006. Direct submission to the GenBank/EMBL database.

Nierman, W.C., Pain, A., Anderson, M.J., Wortman, J.R., Kim, H.S., Arroyo, J., Berriman, M., Abe, K., Archer, D.B., Bermejo, C., Bennett, J., Bowyer, P., Chen, D., Collins, M., Coulsen, R., Davies, R., Dyer, P.S., Farman, M., Fedorova, N., Fedorova,

N., Feldblyum, T.V., Fischer, R., Fosker, N., Fraser, A., Garcia, J.L., Garcia, M.J., Goble, A., Goldman, G.H., Gomi, K., Griffith-Jones, S., Gwilliam, R., Haas, B., Haas, H., Harris, D., Horiuchi, H., Huang, J., Humphray, S., Jimenez, J., Keller, N., Khouri, H., Kitamoto, K., Kobayashi, T., Konzack, S., Kulkarni, R., Kumagai, T., Lafon, A., Latge, J.P., Li, W., Lord, A., Lu, C., Majoros, W.H., May, G.S., Miller, B.L., Mohamoud, Y., Molina, M., Monod, M., Mouyna, I., Mulligan, S., Murphy, L., O'Neil, S., Paulsen, I., Penalva, M.A., Pertea, M., Price, C., Pritchard, B.L., Quail, M.A., Rabbinowitsch, E., Rawlins, N., Rajandream, M.A., Reichard, U., Renauld, H., Robson, G.D., Rodriguez de Cordoba, S., Rodriguez-Pena, J.M., Ronning, C.M., Rutter, S., Salzberg, S.L., Sanchez, M., Sanchez-Ferrero, J.C., Saunders, D., Seeger, K., Squares, R., Squares, S., Takeuchi, M., Tekaia, F., Turner, G., Vazquez de Aldana, C.R., Weidman, J., White, O., Woodward, J., Yu, J.H., Fraser, C., Galagan, J.E., Asai, K., Machida, M., Hall, N., Barrell, B., Denning, D.W. 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* **438**: 1151–1156.

Ninomiya, Y., Suzuki, K., Ishii, C., Inoue, H. 2004. Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous endjoining. *Proceedings of the National Academy of Sciences of the USA* **101**: 12248–12253.

Nixon, J.E., Putz, G.R., Porter, J.W. 1968. Synthesis of triacetic acid lactone by the pigeon liver fatty acid synthetase complex. *The Jornal of Biolpgical Chemistry* **243**: 5471–5478.

Nussinov, R. 1986. Sequence signals which may be required for efficient formation of mRNA 3′ termini. *Nucleic Acids Research* **14**: 3557–3571.

Nussinov, R., Owens, J., Maizel, J.V. Jr. 1986. Sequence signals in eukaryotic upstream regions. *Biochimica et Biophysica Acta* 866: 109–119.

O'Callaghan, J., Caddick, M.X., Dobson, A.D. 2003. A polyketide synthase gene required for ochratoxin A biosynthesis in *Aspergillus ochraceus*. *Microbiology* **149**: 3485–3491.

O'Hagan, D. 1991. The polyketide metabolites. Ellis Horwood, Chichester, United Kingdom.

O'Hagan, D. 1992. Biosynthesis of polyketide metabolites. *Natural Product Reports* **9**: 447–479.

O'Hagan, D. 1993. Biosynthesis of fatty acid and polyketide metabolites. *Natural Product Reports* **10**: 593–624.

O'Hagan, D. 1995. Biosynthesis of fatty acid and polyketide metabolites. *Natural Product Reports* **12**: 1–32.

O'Hagan, D., Rogers, S.V., Duffin, G.R., Edwards, R.L. 1992. Biosynthesis of the fungal polyketide cubensic acid from *Xylaria cubensis*. *Tetrahedron Letters* **33**: 5585–5588.

O'Hara, E.B., Timberlake, W.E. 1989. Molecular characterization of the *Aspergillus nidulans yA* locus. *Genetics* **121**: 249–254.

Oikawa, H., Katayama, K., Suzuki, Y., Ichihara, A. 1995. Enzymatic activity catalyzing exo-

selective Diels-Alder reaction in solanapyrone biosynthesis. *Journal of the Chemical Society. Chemical communications*: 1321–1322.

Oikawa, H., Yokota, T., Abe, T., Ichihara, A., Sakamura, S., Yoshizawa, Y., Vederas, J.C. 1989a. Biosynthesis of solanapyrone A, a phytotoxin of *Alternaria solani*. *Journal of the Chemical Society. Chemical communications*: 1282–1284.

Oikawa, H., Yokota, T., Ichihara, A., Sakamura, S. 1989b. Structure and absolute configuration of solanapyrone D: a new clue to the occurrence of biological Diels-Alder reactions. *Journal of the Chemical Society. Chemical communications*: 1284–1285.

Oikawa, H., Yokota, T., Sakano, C., Suzuki, Y., Naya, A., Ichihara, A. 1998. Solanapyrones, phytotoxins produced by *Alternaria solani*: biosynthesis and isolation of minor components. *Bioscience, Biotechnology, and Biochemistry* **62**: 2016–2022.

Ozeki, K., Kyoya, F., Hizume, K., Kanda, A., Hamachi, M., Nunokawa, Y. Transformation of intact *Aspergillus niger* by electroporation. *Bioscience, Biotechnology, and Biochemistry* **58**: 2224–2227.

Partida-Martinez L.P., Hertweck, C. 2005. Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* **437**: 884–888.

Pazoutova, S., Linka, M., Storkova, S., Schwab, H. 1997. Polyketide synthase gene *pks*M from *Aspergillus terreus* expressed during growth phase. *Folia Microbiologica (Praha)* **42**: 419-430.

Pel, H.J., de Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J., Turner, G., de Vries, R.P., Albang, R., Albermann, K., Andersen, M.R., Bendtsen, J.D., Benen, J.A., van den Berg, M., Breestraat, S., Caddick, M.X., Contreras, R., Cornell, M., Coutinho, P.M., Danchin, E.G., Debets, A.J., Dekker, P., van Dijck, P.W., van Dijk, A., Dijkhuizen, L., Driessen, A.J., d'Enfert, C., Geysens, S., Goosen, C., Groot, G.S., de Groot, P.W., Guillemette, T., Henrissat, B., Herweijer, M., van den Hombergh, J.P., van den Hondel, C.A., van der Heijden, R.T., van der Kaaij, R.M., Klis, F.M., Kools, H.J., Kubicek, C.P., van Kuyk, P.A., Lauber, J., Lu, X., van der Maarel, M.J., Meulenberg, R., Menke, H., Mortimer, M.A., Nielsen, J., Oliver, S.G., Olsthoorn, M., Pal, K., van Peij, N.N., Ram, A.F., Rinas, U., Roubos, J.A., Sagt, C.M., Schmoll, M., Sun, J., Ussery, D., Varga, J., Vervecken, W., van de Vondervoort, P.J., Wedler, H., Wosten, H.A., Zeng, A.P., van Ooyen, A.J., Visser, J., Stam, H. 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nature Biotechnology* **25**: 221–231.

Peng, M., Cooper, C.R. Jr., Szaniszlo, P.J. 1995. Genetic transformation of the pathogenic fungus *Wangiella dermatitidis*. *Applied Microbiology and Biotechnology* **44**: 444–450.

Pfeifer, B.A., Khosla, C. 2001. Biosynthesis of polyketides in heterologous hosts. *Microbiology and Molecular Biology* **65**: 106-118.

Plattner, R.D., Shackelford, D.D. 1992. Biosynthesis of labeled fumonisins in liquid cultures of *Fusarium moniliforme*. *Mycopathologia* **117**: 17–22.

Poch, G.K., Gloer, J.B. 1989. Obionin A: a new polyketide metabolite from the marine fungus *Leptosphaeria obiones*. *Tetrahedron Letters* **30**: 3483–3486.

Pontecorvo, G., Roper, J.A., Hemmons, L.M., Macdonald, K.D., Bufton, A.W. 1953. The genetics of *Aspergillus nidulans*. *Advances in Genetics* **5**: 141–238.

Pósfai, J., Bhagwat, A.S., Pósfai, G., Roberts, R.J. 1989. Predictive motifs derived from cytosine methyltransferases. *Nucleic Acid Research* **17**: 2421–2435.

Pósfai, J., Bhagwat, A.S., Roberts, R.J. 1988. Sequence motifs specific for cytosine methyltransferases. *Gene* **74**: 261–265.

Proctor, R.H., Brown, D.W., Plattner, R.D., Desjardins, A.E. 2003. Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in *Gibberella moniliformis*. *Fungal Genetics and Biology* **38**: 237–249.

Proctor, R.H., Desjardins, A.E., Plattner, R.D., Hohn, T.M. 1999. A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in *Gibberella fujikuroi* mating population A. *Fungal Genetics and Biology* **27**: 100–112.

Proudfoot, N.J., Furger, A., Dye, M.J. 2002. Integrating mRNA processing with transcription. *Cell* **108**: 501–512.

Puel, O., Tadrist, S., Delaforge, M., Oswald, I.P., Lebrihi, A. 2007. The inability of *Byssochlamys fulva* to produce patulin is related to absence of 6-methylsalicylic acid synthase and isoepoxydon dehydrogenase genes. *International Journal of Food Microbiology* **115**: 131–139.

Punt, P.J., Dingemanse, M.A., Kuyvenhoven, A., Soede, R.D., Pouwels,P.H., van den Hondel, C.A.M.J.J. 1990. Functional elements in the promoter region of the *Aspergillus nidulans gpd*A gene encoding glyceraldehyde-3-phosphate dehydrogenase. *Gene* **93**: 101–109.

Rai, A.N., Bergman, B. 2002. Cyanolichens. *Biology and Environment: Proceedings of The Royal Irish Academy* **102**: 19–22.

Rand, N.K. 1996. Crystal violet can be used to visualize DNA bands during gel electrophoresis and to improve cloning efficiency. *Elsevier Trends Journals Technical Tips Online*, T40022.

Rawlings, B.J. 1997. Biosynthesis of polyketides. *Natural Products Reports* **14**: 523–556.

Rawlings, B.J. 1999. Biosynthesis of polyketides (other than actinomycete macrolides). *Natural Product Reports* **16**: 425–484.

Reeves, C.D. 2003. The enzymology of combinatorial biosynthesis. *Critical Reviews in Biotechnology* **23**: 95–147.

Reio, L. 1958. A method for the paper-chromatographic separation and identification of phenol derivatives, mould metabolites and related compounds of biochemical interest, using a references system. *Journal of Chromatography* **1**: 338–373.

Riach, M.B.R., Kinghorn, J.R. 1996. Genetic transformation and vector developments in filamentous fungi. In: Bos, C.J. [*ed.*] *Fungal Genetics: Principles and Practice*, pp. 209–233. Marcel Dekker Inc., New York.

Rice, S.L., Beuchat, L.R., Worthington, R.E. 1977. Patulin production by *Byssochlamys spp.* in fruit juices. *Applied and Environmental Microbiology* **34**: 791–796.

Rose, M.S., Yun, S.H., Asvarak, T., Lu, S.W., Yoder, O.C., Turgeon, B.G. 2002. A decarboxylase encoded at the *Cochliobolus heterostrophus* translocation-associated *Tox*1B locus is required for polyketide (T-toxin) biosynthesis and high virulence on T-cytoplasm maize. *Molecular Plant-Microbe Interactions* **15**: 883–893.

Rude, M.A., Khosla, C. 2004. Engineered biosynthesis of polyketides in heterologous hosts. *Chemical Engineering Science* **59**: 4693–4701.

Ruiz-Díez, B. Strategies for the transformation of filamentous fungi. *Journal of Applied Microbiology* **92**: 189–195.

Ruiz-Díez, B., Martínez-Suárez, J.V. 1999. Electrotransformation of the human pathogenic fungus *Scedosporium prolificans* mediated by repetitive rDNA sequences. *FEMS Immunology and Medical Microbiology* **25**: 275–282.

Sánchez, O. Aguirre, J. 1996. Efficient transformation of *Aspergillus nidulans* by electroporation of germinated conidia. *Fungal Genetics Newsletters* **43**: 48–51.

Sánchez, O., Navarro, R.E., Aguirre, J. 1998. Increased transformation frequency and tagging of developmental genes in *Aspergillus nidulans* by restriction enzyme-mediated integration (REMI). *Molecular and General Genetics* **258**: 89–94.

Sánchez-Torres, P., González, R., Pérez-González, J.A., González-Candelas, L., Ramón, D. 1994. Development of a transformation system for *Trichoderma longibrachiatum* and its use for constructing multicopy transformants for the *egl*1 gene. *Applied Microbiology and Biotechnology* **41**: 440–446.

Sankawa, U., Ebizuka, Y., Noguchi, H., Isikawa, Y., Kitaghawa, S., Yamamoto, Y., Kobayashi, T., Iitak, Y. 1983. Biosynthesis of citrinin in *Aspergillus terreus*. *Tetrahedron* **39**: 3583–3591.

Sankawa, U., Shimada, H., Sato, T., Kinoshita, T., Yamasaki, K. 1977. Biosynthesis of scytalone. *Tetrahedron Letters* **18**: 483–486.

Sauer, M., Lu, P., Sangari, R., Kennedy, S., Polishook, J., Bills, G., An, Z. 2002. Estimating polyketide metabolic potential among nonsporulating fungal endophytes of *Vaccinium macrocarpon*. *Mycological Research* **106**: 460–470.

Schiestl, R.H., Petes, T.D. 1991. Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the USA* **88**: 7585–7589.

Schorr, R., Mittag, M., Müller, G., Schweizer, E. 1994. Differential activities and intramolecular location of fatty acid synthase and 6-methylsalicylic synthase component enzymes. *Journal of Plant*

Physiology **143**: 407–415.

Seo, J.A., Proctor, R.H., Plattner, R.D. 2001. Characterization of four clustered and coregulated genes associated with fumonisin biosynthesis in *Fusarium verticillioides*. *Fungal Genetics and Biology* **34**: 155–165.

Serizawa, N., Nakagawa, K., Hamano, K., Tsujita, Y., Terahara, A., Kuwano, H. 1983. Microbial hydroxylation of ML-236B (compactin) and monacolin K (MB-530B). *The Journal of Antibiotics* **36**: 604–607.

Serre, L., Verbree, E.C., Dauter, Z., Stuitje, A.R., Derewenda, Z.S. 1995. The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5- resolution. Crystal structure of a fatty acid synthase component. *The Journal of Biological Chemistry* **270**: 12961–12964.

Seto, H., Sasaki, T., Yonehara, H. 1977. Studies on the biosynthesis of radiclonic acid. *Tetrahedron Letters* **18**: 4083–4084.

Shimizu, T., Kinoshita, H., Ishihara, S., Sakai, K., Nagai, S., Nihira, T. 2005. Polyketide synthase gene responsible for citrinin biosynthesis in *Monascus purpureus*. *Applied and Environmental Microbiology* **71**: 3453–3457.

Schümann, J., Hertweck, C. 2006. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *Journal of Biotechnology* **124**: 690–703.

Sidebottom, P.J., Highcock, R.M., Lane, S.J., Procopiou, P.A., Watson, N.S. 1992. The squalestatins, novel inhibitors of squalene synthase produced by a species of *Phoma*. II. Structure elucidation. *The Journal of Antibiotics* **45**: 648–658.

Sims, J.W., Fillmore, J.P., Warner, D.D., Schmidt, E.W. 2005. Equisetin biosynthesis in *Fusarium heterosporum*. *Chemical Communications* **14**: 186–188.

Singh, N., Wakil, S.J., Stoops, J.K. 1984. On the question of half- or full-site reactivity of animal fatty acid synthetase. *The Journal of Biological Chemistry* **259**: 3605–3611.

Sinneman, S.J., Andrésson, Ó.S., Brown, D.W., Miao, V.P.W. 2000. Cloning and heterologous expression of *Solorina crocea pyr*G. *Current Genetics* **37**: 333–338.

Smale, S.T., Kadonaga, J.T. 2003. The RNA polymerase II core promoter. *Annual Review of Biochemistry* **72**: 449–479.

Sorensen, J.L., Auclair, K., Kennedy, J., Hutchinson, C.R., Vederas, J.C. 2003. Transformations of cyclic nonaketides by *Aspergillus terreus* mutants blocked for lovastatin biosynthesis at the *lov*A and *lov*C genes. *Organic and Biomolecular Chemistry* **1**: 50–59.

Spencer, J.B., Jordan, P.M. 1990. Use of chiral malonates to determine the absolute configuration of the hydrogen atoms eliminated during the formation of 6-methylsalicylic acid by 6-methylsalicylic acid synthase from *Penicillium patulum*. *Journal of the Chemical Society. Chemical communications*: 1704–1706.

Spencer, J.B., Jordan, P.M. 1992. Investigation of the mechanism and steric course of the reaction

catalyzed by 6-methylsalicylic acid synthase from *Penicillium patulum* using (*R*)-[1-^{13}C;2-^{2}H] and (*S*)-[1-^{13}C;2-^{2}H]malonates. *Biochemistry* **31**:9107–9116.

Spencer, J.B., Jordan, P.M. 1992. Purification and properties of 6-methylsalicylic acid synthase from *Penicillium patulum*. *The Biochemical Journal* **288**: 839–846.

Staunton, J., Weissman, K.J. 2001. Polyketide biosynthesis: a millenium review. *Natural Product Reports* **18**: 380–416.

Stout, G.H., Dreyer, D.L., Jensen, L.H. 1961. Structure of rubrofusarin. *Chemistry and Industry*: 289.

Su, Y.C., Wang, J.J, Lin, T.T, Pan, T.M. 2003. Production of the secondary metabolites gamma-aminobutyric acid and monacolin K by *Monascus*. *Journal of Industrial Microbiology and Biotechnology* **30**: 41–46.

Sutherland, A., Auclair, K., Vederas, J.C. 2001. Recent advances in the biosynthetic studies of lovastatin. *Current Opinion in Drug Discovery and Development* **4**: 229–236.

Sweeney, M.J., Dobson, A.D. 1999. Molecular biology of mycotoxin biosynthesis. *FEMS Microbiology Letters* **175**: 149–163.

Takahashi, T., Chang, P.K., Matsushima, K., Yu, J., Abe, K., Bhatnagar, D., Cleveland, T.E., Koyama, Y. 2002. Nonfunctionality of *Aspergillus sojae afl*R in a strain of *Aspergillus parasiticus* with a disrupted *afl*R gene. *Applied and Environmental Microbiology* **68**: 3737–3743.

Takahashi, T., Masuda, T., Koyama, Y. 2006. Enhanced gene targeting frequency in *ku*70 and *ku*80 disruption mutants of *Aspergillus sojae* and *Aspergillus oryzae*. *Molecular Genetics and Genomics* **275**: 460–470.

Takano, Y., Kubo, Y., Shimizu, K., Mise, K., Okuno, T., Furusawa, I. 1995. Structural analysis of *PKS*1, a polyketide synthase gene involved in melanin biosynthesis in *Colletotrichum lagenarium*. *Molecular and General Genetics* **249**: 162–167.

Tanaka, H., Tamura, T. 1961. The chemical constitution of rubrofusarin. *Tetrahedron Letters* **4**: 151–155

Tanenbaum, S.W., Bassett, E.W. 1959. The biosynthesis of patulin. III. Evidence for a molecular rearrangement of the aromatic ring. *The Journal of Biological Chemistry* **234**: 1861–1866.

Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., Julien, B. 2000. Cloning and heterologous expression of the epothilone gene cluster. *Science* **287**: 640–642.

Tanguay, P., Bozza, S., Breuil, C. 2006a. Assessing RNAi frequency and efficiency in *Ophiostoma floccosum* and *O. piceae*. *Fungal Genetics and Biology* **43**: 804–812.

Tanguay, P., Loppnau, P., Morin, C., Bernier, L., Breuil, C. 2006b. A spontaneous albino mutant of *Ceratocystis resinifera* results from a point mutation in the polyketide synthase gene, PKS1. *Canadian Journal of Microbiology* **52**: 501–507.

Tanguay, P., Tangen, K., Breuil, C. 2007. Identifying pigmentation-related genes in *Ophiostoma*

piceae using *Agrobacterium*-mediated integration. *Phytopathology. In press.*

Tanimoto, T., Hamano, K., Onodera, K., Hosoya, T., Kakusaka, M., Hirayama, T., Shimada, Y., Koga, T., Tsujita, Y. 1997. Biological activities of novel zaragozic acids, the potent inhibitors of squalene synthase, produced by the fungus, *Mollisia sp.* SANK 10294. *The Journal of Antibiotics* **50**: 390–394.

Tatusova, T.A., Madden, T.L. 1999. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiology Letters* **174**: 247–250.

Te'o, V.S., Bergquist, P.L., Nevalainen, K.M. 2002. Biolistic transformation of *Trichoderma reesei* using the Bio-Rad seven barrels Hepta Adaptor system. *Journal of Microbiological Methods* **51**: 393–399.

Tilburn, J., Scazzocchio, C., Taylor, G.G., Zabicky-Zissman, J.H., Lockington, R.A., Davies, R.W. 1983. Transformation by integration in *Aspergillus nidulans*. *Gene* **26**: 205–221.

Thierry, D., Vaucheret, H. 1996. Sequence homology requirements for transcriptional silencing of 35S transgenes and posttranscriptional silencing of nitrite reductase (trans)genes by the tobacco 271 locus. *Plant Molecular Biology* **32**: 1075–1083.

Thompson, N.F., Thompson, J.W. 1984. Spore ornamentation in the lichen genus *Solorina*. *Bryologist* **87**: 151–153.

Tominaga, M., Lee, Y.H., Hayashi, R., Suzuki, Y., Yamada, O., Sakamoto, K., Gotoh, K., Akita, O. 2006. Molecular analysis of an inactive aflatoxin biosynthesis gene cluster in *Aspergillus oryzae* RIB strains. *Applied and Environmental Microbiology* **72**: 484–490.

Townsend, C.A., Christensen, S.B., Trautwein, K. 1984. Hexanoate as a starter unit in polyketide biosynthesis. *Journal of the American Chemical Society* **106**: 3868–3869.

Treiber, L.R., Reamer, R.A., Rooney, C.S., Ramjit, H.G. 1989. Origin of monacolin L from *Aspergillus terreus* cultures. *Journal of Antibiotics* **42**: 30–36.

Tsai, H.F., Chang, Y.C., Washburn, R.G., Wheeler, M.H., Kwon-Chung, K.J. 1998. The developmentally regulated *alb*1 gene of *Aspergillus fumigatus*: its role in modulation of conidial morphology and virulence. *Journal of Bacteriology* **180**: 3031–3038.

Tsai, H.F., Fujii, I., Watanabe, A., Wheeler, M.H., Chang, Y.C., Yasuoka, Y., Ebizuka, Y., Kwon-Chung, K.J. 2001. Pentaketide melanin biosynthesis in *Aspergillus fumigatus* requires chain-length shortening of a heptaketide precursor. *The Journal of Biological Chemistry* **276**: 29292–29298.

Tsai, H.F., Wheeler, M.H., Chang, Y.C., Kwon-Chung, K.J. 1999. A developmentally regulated gene cluster involved in conidial pigment biosynthesis in *Aspergillus fumigatus*. *Journal of Bacteriology* **181**: 6469–6477.

Turgeon, B.G., Lu, S.-W. 2000. Evolution of host specific virulence in *Cochliobolus heterostrophus*. In: Kronstad, J.W. [*ed.*] *Fungal Pathology*, pp. 93–126. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Turner, G. 1993. Gene organisation in filamentous fungi. In: Broda, P., Oliver, S.G., Sims, P.F.G. [*eds.*] *The Eukaryotic Genome: Organisation and Regulation. 50th Symposium of the Society for General Microbiology*, pp. 107–125. Cambridge University Press, Cambridge.

Turner, G., Brown, J., Kerrey-Williams, S., Bailey, A.M., Ward, M., Punt, P.J., van den Hondel, C.A.M.J.J. 1989. Analysis of the *oli*C promoter of *Aspergillus nidulans*. In: Nevalainen, H., Pentilla, M. [*eds.*] *Proceedings of the EMBO-ALKO Workshop on Molecular Biology of Filamentous Fungi. Foundation for biotechnical and industrial fermentation research*, Vol. 6, pp. 101–110. Espoo, Finland.

Turner, W.B. 1971. Fungal Metabolites. Academic Press, London.

Turner, W.B., Aldridge, D.C. 1983. Fungal Metabolites II. Academic Press, London.

Tzeng, T., Chang, H., Bronson, C.R. 1989. Electrophoretic confirmation of chromosome rearrangements in *Cochliobolus heterostrophus*: translocation at *Tox*1. *Phytopathology* **79**:1204.

Tzeng, T.H., Lyngholm, L.K., Ford, C.F., Bronson, C.R. 1992. A restriction fragment length polymorphism map and electrophoretic karyotype of the fungal maize pathogen *Cochliobolus heterostrophus*. *Genetics* **130**: 81–96.

Udwary, D.W., Merski, M., Townsend, C.A. 2002. A method for prediction of the locations of linker regions within large multifunctional proteins, and application to a type I polyketide synthase. *Journal of Molecular Biology* **323**: 585–598.

Unkles, S.E. 1992. Gene organization in industrial filamentous fungi. In: Kinghorn, J.R., Turner, G. [*eds.*] *Applied Molecular Genetics of Filamentous Fungi*, pp. 28–53. Cambridge University Press, Cambridge.

van Gorcom, R.F., Pouwels, P.H., Goosen, T., Visser, J., van den Broek, H.W., Hamer, J.E., Timberlake, W.E., van den Hondel, C.A. 1985. Expression of an *Escherichia coli* β-galactosidase fusion gene in *Aspergillus nidulans*. *Gene* **40**: 99–106.

Venkatasubbaiah, P., Chilton, W.S. 1992. An epoxydon-derived ester from a *Phoma sp.* pathogenic to rhubarb. *Journal of Natural Products* **55**: 639–643.

Venkatasubbaiah, P., Sutton, T.B., Chilton, W.S. 1995. The structure and biological properties of secondary metabolites produced by *Peltaster fructicola*, a fungus associated with apple sooty blotch disease. *Plant disease* **79**: 1157–1160.

Vleggaar, R. 1986. Biosynthetic studies on some polyene mycotoxins. *Pure and Applied Chemistry* **58**: 239–256.

von Dören, H., Kleinkauf, H. 1997. Enzymology of peptide synthases. In: Strohl, W.R. [*ed.*] *Biotechnology of Antibiotics*, pp. 217–240. Mercel Dekker, New York.

Wagschal, K., Yoshizawa, Y., Witter, D.J., Liu, Y., Vederas, J.C. 1996. Biosynthesis of ML-236C and the hypocholesterolemic agents compactin by *Penicillium aurantiogriseus* and lovastatin by *Aspergillus terreus*: determination of the origin of carbon, hydrogen and oxygen atoms by ^{13}C NMR

spectroscopy and observation of unusual labeling of acetate-derived oxygens by $^{18}O_2$. *Journal of the Chemical Society, Perkin Transactions 1*: 2357–2363.

Wang, E., Norred, W.P., Bacon, C.W., Riley, R.T., Merrill, A.H. Jr. 1991. Inhibition of sphingolipid biosynthesis by fumonisins. Implications for diseases associated with *Fusarium moniliforme*. *The Journal of Biological Chemistry* **5**: 14486–14490.

Wang, H., Breuil, C. Polyketide synthase gene is involved in melanin production in *Ophiostoma floccosum*. *Unpublished.*

Wang, I.K., Reeves, C., Gaucher, G.M. 1991. Isolation and sequencing of a genomic DNA clone containing the 3′ terminus of the 6-methylsalicylic acid polyketide synthetase gene of *Penicillium urticae*. *Canadian Jouranl of Microbiology* **37**: 86–95.

Watanabe, A., Fujii, I., Tsai, H., Chang, Y.C., Kwon-Chung, K.J., Ebizuka, Y. 2000. *Aspergillus fumigatus alb*1 encodes naphthopyrone synthase when expressed in *Aspergillus oryzae*. *FEMS Microbiology Letters* **192**: 39–44.

Watanabe, A., Fujii, I., Sankawa, U., Mayorga, M.E., Timberlake, W.E., Ebizuka, Y. 1999. Re-identification of *Aspergillus nidulans w*A gene to code for a polyketide synthase of naphthopyrone. *Tetrahedron Letters* **40**: 91–94.

Watanabe, A., Ono, Y., Fujii, I., Sankawa, U., Mayorga, M.E., Timberlake, W.E., Ebizuka, Y. 1998. Product identification of polyketide synthase coded by *Aspergillus nidulans w*A gene. *Tetrahedron Letters* **39**: 7733–7736.

Watson, A.J., Fuller, L.J., Jeenes, D.J., Archer, D.B. 1999. Homologs of aflatoxin biosynthesis genes and sequence of *afl*R in *Aspergillus oryzae* and *Aspergillus sojae*. *Applied and Environmental Microbiology* **65**: 307–310.

Wattanachaisaereekul, S., Lantz, A.E., Nielsen, M.L., Andrésson, Ó.S., Nielsen, J. 2007. Optimization of heterologous production of the polyketide 6-MSA in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering* **97**: 893–900.

Weidner, G., d'Enfert, C., Koch, A., Mol, P., Brakhage, A.A. 1998. Development of a homologous transformation system for the human pathogenic fungus *Aspergillus fumigatus* based on the *pyr*G gene encoding orotidine 5′-monophosphate decarboxylase. *Current Genetics* **33**: 378–385.

Weitnauer, G., Mühlenweg, A., Trefzer, A., Hoffmeister, D., Süssmuth, R.D., Jung, G., Welzel, K., Vente, A., Girreser, U., Bechthold, A. 2001. Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the avi biosynthetic gene cluster of *Streptomyces viridochromogenes* Tü57 and production of new antibiotics. *Chemistry and Biology* **8**: 569–581.

Wernars, K., Goosen, T., Wennekes, L.M.J., Visser, J., Bos, C.J., van der Broek, H.W.J., van Gorcom, R.F.M., van den Hondel, C.A.M.J.J., Pouwels, P.H. 1985. Gene amplification in *Aspergillus nidulans* by transformation with vectors containing the *amd*S gene. *Current Genetics*

9: 361–368.

Wheeler, M.H., Bell, A.A. 1988. Melanins and their importance in pathogenic fungi. In: McGinnis, M.R. [*ed.*] *Current Topics in Medical Mycology*, Vol. 2, pp. 338–387. Springer-Verlag, New York.

White, S., O'Callaghan, J., Dobson, A.D.W. 2006. Cloning and molecular characterization of *Penicillium expansum* genes upregulated under conditions permissive for patulin biosynthesis. *FEMS Microbiology Letters* **255**: 17–26.

White, T. J., Bruns, T., Lee, S., Taylor, J.W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. [*eds.*] *PCR Protocols: A Guide to Methods and Applications,* pp. 315–322. Academic Press, New York.

Williams, J.H., Phillips, T.D., Jolly, P.E., Stiles, J.K., Jolly, C.M., Aggarwal, D. 2004. Human aflatoxicosis in developing countries: a review of toxicology, exposure, potential health consequences, and interventions. *The American Journal of Clinical Nutrition* **80**: 1106–1122.

Wilson, K.E., Burk, R.M., Biftu, T., Ball, R.G., Hoogsteen, K. 1992. Zaragozic acid A, a potent inhibitor of squalene synthase: initial chemistry and absolute stereochemistry. *The Journal of Organic Chemistry* **57**: 7151–7158.

Wirsel, S., Lachmund, A., Wildhardt, G., Ruttkowski, E. 1989. Three alpha-amylase genes of *Aspergillus oryzae* exhibit identical intron-exon organization. *Molecular Microbiology* **3**: 3–14.

Witter, D.J., Vederas, J.S. 1996. Putative Diels-Alder-catalyzed cyclization during the biosynthesis of lovastatin. *Journal of Organic Chemistry* **61**: 2613–2623

Woo, E.-R., Fujii, I., Ebizuka, Y., Sankawa, U., Kawaguchi, A., Huang, S., Beale, J.M., Shibuya, M., Mocek, U., Floss, H.G. 1989. Nonstereospecific proton removal in the enzymatic formation of orsellinic acid from chiral malonate. *Journal of the American Chemical Society* **111**: 5498–5500.

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

Wright, P.J., Pateman, J.A. 1970. Ultraviolet-light sensitive mutants of *Aspergillus nidulans*. *Mutation Research* **9**: 579–587.

Xie, X., Watanabe, K., Wojcicki, W.A., Wang, C.C.C., Tang, Y. 2006. Biosynthesis of lovastatin analogs with a broadly specific acyltransferase. *Chemistry and Biology* **13**: 1161–1169.

Yadav, G., Gokhale, R.S., Mohanty D. 2003. SearchPKS: a program for detection and analysis of polyketide synthase domains. *Nucleic Acids Research* **31**: 3654–3658.

Yalpani, M., Willecke, K., Lynen, F. 1969. Triacetic acid lactone, a derailment product of fatty acid biosynthesis. *European Journal of Biochemistry* **8**: 495–502.

Yalpani, N., Altier, D.J., Barbour, E., Cigan, A.L., Scelonge, C.J. 2001. Production of 6-methylsalicylic acid by expression of a fungal polyketide synthase activates disease resistance in

tobacco. *Plant Cell* **13**: 1401–1409.

Yang, G., Rose, M.R., Turgeon, B.G., Yoder, O.C. 1996. A polyketide synthase is required for fungal virulence and production of the polyketide T-toxin. *Plant Cell* **8**: 2139–2150.

Yelton, M.M., Hamer, J.E., Timberlake, W.E. 1984. Transformation of *Aspergillus nidulans* by using a *trp*C plasmid. *Proceedings of the National Academy of Sciences of the USA* **81**: 1470–1474.

Yoder, O.C. 1980. Toxins in pathogenesis. *Annual Review of Phytopathology* **18**: 103–129.

Yoder, O.C., Gracen, V.E. 1975. Segregation of pathogenicity types and host-specific toxin production in progenies of crosses between races T and O of *Helminthosporium maydis* (*Cochliobolus heterostrophus*). *Phytopathology* **65**: 273–276.

Yoder, O.C., Macko, V., Wolpert, T.J., Turgeon, B.G. 1997. *Cochliobolus spp.* and their host-specific toxins. In: Carroll, G., Tudzynski, P. [*eds.*] *The Mycota. Vol. 5: Plant Relationships, Part A.,* pp. 145-166. Springer-Verlag, Berlin.

Yoshizawa, Y., Witter, D.J., Liu, Y., Vederas, J.C. 1994. Revision of the biosynthetic origin of oxygens in mevinolin (lovastatin), a hypocholesterolemic drug from *Aspergillus terreus* MF 4845. *Journal of the American Chemical Society* **116**: 2693–2694

Yu, J., Bhatnagar, D., Cleveland, T.E. 2004a. Completed sequence of aflatoxin pathway gene cluster in *Aspergillus parasiticus*. *FEBS Letters* **564**: 126–130.

Yu, J., Chang, P.K., Ehrlich, K.C., Cary, J.W., Bhatnagar, D., Cleveland, T.E., Payne, G.A., Linz, J.E., Woloshuk, C.P., Bennett, J.W. 2004b. Clustered pathway genes in aflatoxin biosynthesis. *Applied and Environmental Microbiology* **70**: 1253–1262.

Yu, J.H., Leonard, T.J. 1995. Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type I polyketide synthase. *Journal of Bacteriology* **177**: 4792–4800.

Zhang, A., Lu, P., Dahl-Roshak, A.M., Paress, P.S., Kennedy, S., Tkacz, J.S., An, Z. 2003. Efficient disruption of a polyketide synthase gene (*pks*1) required for melanin synthesis through *Agrobacterium*-mediated transformation of *Glarea lozoyensis*. *Molecular Genetics and Genomics* **268**: 645–655.

Zhu, X., Yu, F., Bojja, R.S., Zaleta-Rivera, K., Du, L. 2006. Functional replacement of the ketosynthase domain of *FUM*1 for the biosynthesis of fumonisins, a group of fungal reduced polyketides. *Journal of Industrial Microbiology and Biotechnology* **33**: 859–868.

Zhu, X., Yu, F., Li, X.C., Du, L. 2007. Production of dihydroisocoumarins in *Fusarium verticillioides* by swapping ketosynthase domain of the fungal iterative polyketide synthase Fum1p with that of lovastatin diketide synthase. *Journal of the American Chemical Society* **129**: 36–37.

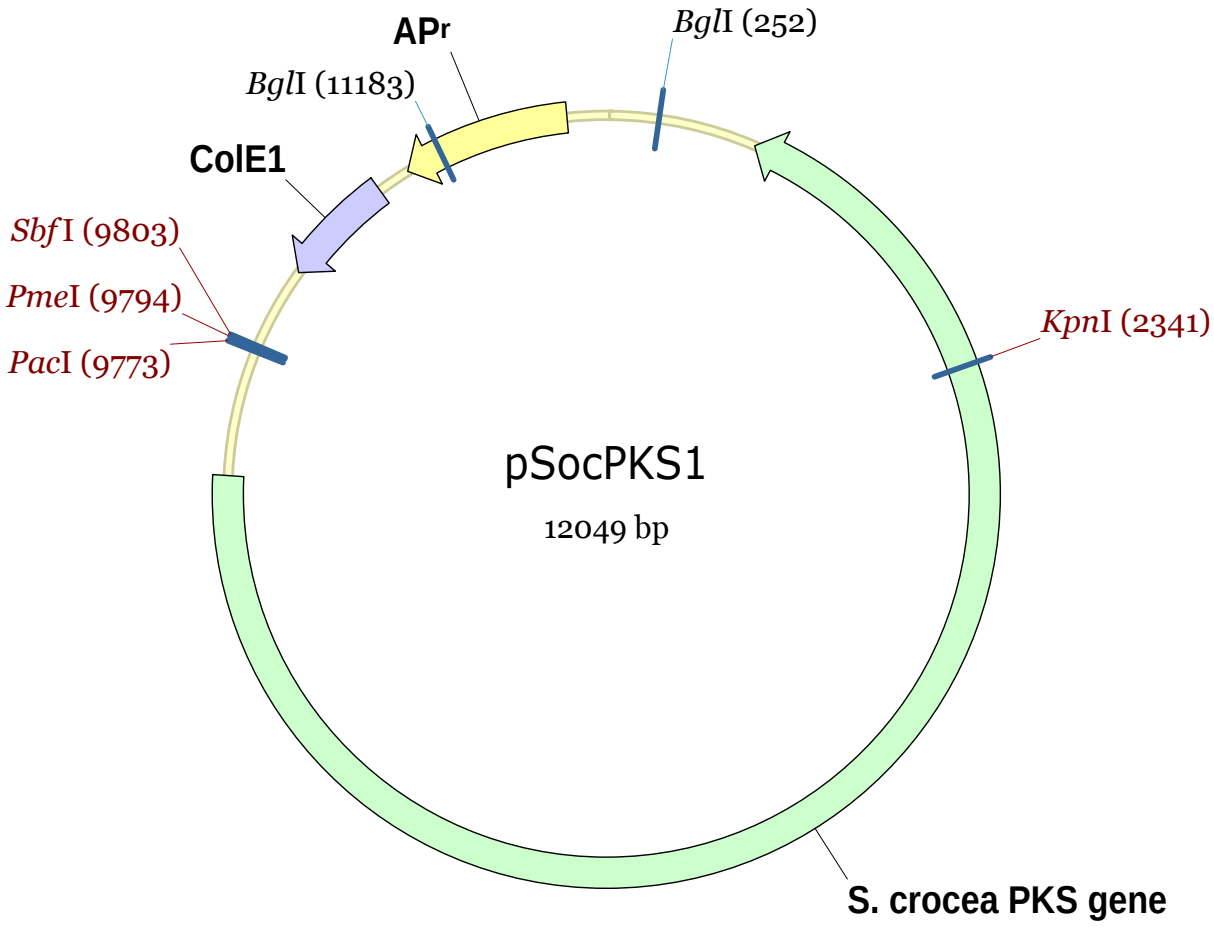
Zuber, P., Marahiel, M. 1997. Structure, function, and regulation of genes encoding multidomain peptide synthetase. In: Strohl, W.R. [*ed.*] *Biotechnology of Antibiotics*, pp. 187–216. Mercel Dekker, New York.

Zubieta, C., Kota, P., Ferrer, J.-L., Dixon, R.A., Noel, J.P. 2002. Structural basis for the mod-

ulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-*O*-methyltransferase. *Plant Cell* **14**: 1265–1277.

Appendix

A-1 Plasmid pSocPKS1 and its construction



The plasmid pSocPKS1 contains a gene of approximately 8.4 kb encoding a polyketide synthase from the lichen *Solorina crocea* together with approximately 630 bp preceding and 360 bp trailing the coding sequence. Resistance to ampicillin is conferred by β-lactamase gene (AP^r). ColE1 origin of replication allows high-copy replication and maintenance of the plasmid in *E. coli*.

The plasmid pSocPKS1 was constructed by:

1. Cleaving plasmid pK33 with restriction enzymes *Mlu*I, site in SocPKS1 sequence approximately 630 bp ahead of PKS coding sequence, and *Asc*I, site in pNEB193 vector (New England Biolabs), removing small fragment and religating. This was done in order to remove unwanted *Eco*RI site. Resulting plasmid is pAM.

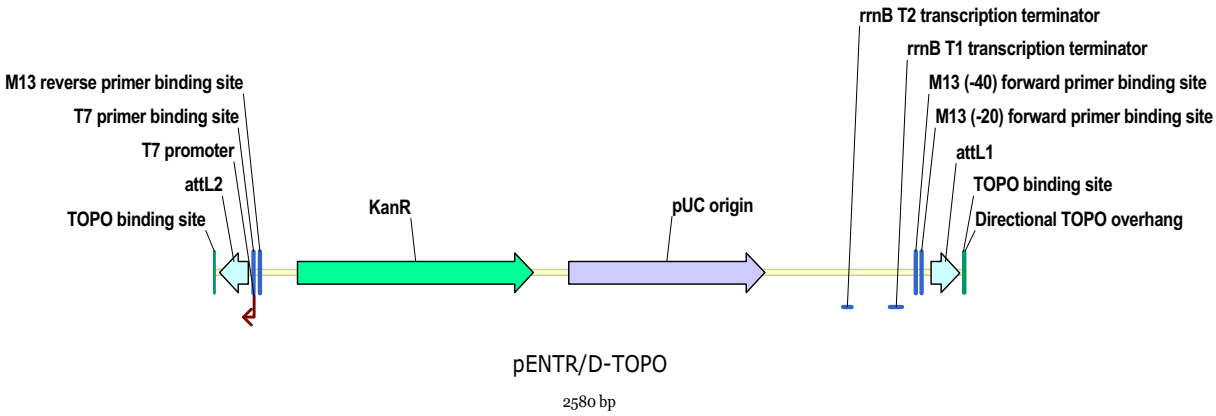
2. Generating a PCR copy from plasmid pK34 carrying *Kpn*I site internal to PKS gene and extending approximately 360 bp beyond coding sequence. Downstream oligo contained *Mfe*I site and PCR product was cloned and sequenced in vector LITMUS 38i (New England Biolabs).

3. Cleaving out *Kpn*I–*Mfe*I fragment from LITMUS 38i vector and ligating into *Kpn*I–*Eco*RI cleaved pAM. Resulting plasmid is pSocPKS1.

Origin of pK33 and pK34 plasmids

The plasmids pK33 and pK34 have a pNEB193 backbone and contain *Kpn*I generated fragments derived from a Lambda FIX II/*Xho* I Partial Fill-in vector (Stratagene) containing approximately 14 kb of genetic material derived from *S. crocea* genomic DNA (Terragen №345).

A-2 Directional TOPO vector pENTR/D-TOPO[1]



Feature	**Benefit**
*rrn*B T1 and T2 transcription termination sequences	Reduces potential toxicity in *E. coli* by preventing basal expression of the cloned PCR product
M13 (-20) forward and reverse primers binding sites	Allows sequencing of the insert
*att*L1 and *att*L2 sites	Bacteriophage λ-derived recombination sequences that allow recombinational cloning of a gene of interest in the entry construct with a Gateway destination vector
TOPO Cloning site (directional)	Allows directional cloning of PCR products
T7 promoter	Allows *in vitro* transcription of the insert
T7 primer binding site	Allows sequencing of the insert
Kanamycin resistance gene (KanR)	Allows selection of the plasmid in *E. coli*
pUC origin of replication	Allows high-copy replication and maintenance in *E. coli*

[1]pENTR Directional TOPO cloning Kits. User manual. Invitrogen, 2006.

A-3 List of oligonucleotides used in this study

Primer name	**Sequence 5′ → 3′**	**Direction**
	***S. crocea* PKS gene specific primers**	
B11R3	CGA CGG TTT GCA TAA TGC A	Forward
B11FF	GCG AGC ACA AAG GTT ATC AT	Forward
B11BT	CGA ACC CAT CTT GCT TAG A	Reverse
GW-PKSF	CAC CAT GGG TTC GAT GCC GCT TGA	Forward
GSP3	CGG AAA CAG TCT CAA GTG GCA AA	Reverse
GSP2	CGT TGT GCC TGT CGG TTT GTA CTT	Reverse
GSP1	GGG TAA CGC TCG GAG CTT TCA A	Reverse
B8R1	GTC AAT CAA GAT GGC CGG ACA	Forward
Ex1-2F	GCT TCG AAG ACA CGA TGA TGG T	Forward
BF843	CTG AAT GCA CAG TCT GTC A	Forward
B8AAF	GCA GGT TCC ACT TCT TCA	Reverse
Ex1-2R	AGC TGT GGG CGT CAT CAA GTA T	Reverse
PKSF1	GAC GGA TGC AAT CAA AGC GAT G	Forward
B8AAR	CGT CAG TCG TGC GAA GAG CGA	Forward
PKSR1	GAA CCC AAT CTT TGG CAC GTT C	Reverse
BF1685	CTT ACC CTG TGT TTG AGA GA	Forward
B8FCY	CGG TAA CGC AAA TGT GCT T	Forward
Y13R	CAA CGG TGT TCA ATA TTG TGA	Reverse
B8R2702	GAC GTG GTC TTG CTT GTG GTG A	Reverse
BR2760	CGT AGA TGA GCT TCT CAG AA	Reverse
B8R3008	GCT TCG ATG GCA GCA ATC AGG T	Reverse

Continued on the next page →

(*Continued*)

Primer name	**Sequence** $5' \rightarrow 3'$	**Direction**
Y13F	CTC TCA TCA TAC CTG ACA ACA	Forward
B8R3336	CGT ATA CAT CGT CGA TTT CGA T	Reverse
BF3417	CTC TCA GAG GTG AAA GGT A	Forward
Ex2-3F	ATC TCA TTC ACC CGT CCA CCA T	Forward
B8YB	GTA CAG ATC CAG AAC CCA TT	Reverse
B815-9R	GAG TTT GCG TAC AGA TCC AGA	Reverse
15-9FF	CAT TTC AAG GTA CGG AGT CCA	Reverse
BF3839	TCA GCT CTC TCA AAT GTT CA	Forward
Ex2-3R	TCA AAC TCC GCA AGA CGC ATA C	Reverse
PKSF2	CTT GGG AAA CCT TCC TGC ACT G	Forward
PKSR2	CTG CCA CCT GGT TTC AGT AAG C	Reverse
B8YA	GAA ACC AGG TGG CAG ACT	Forward
B8EEY	GAT ACC ATT GTT CGG TTG TGA	Reverse
BEE8R	TCG CAG ACA GAT ATG CAG ACA	Reverse
BR5524	CTT TGC TCG TGA GAA CTG A	Reverse
Ex3-6F	TCA GTT CTC ACG AGC AAA GCC A	Forward
B8R5777	CCT GAG AAA GTT AGT GGG GAT G	Reverse
B8ct1	GGC AAG CTG CTA TAC ACT A	Reverse
Ex3-6R	CCA TAG TGA CAT TGC GCA GGA A	Reverse
K34R1	GGA ATG GGT AAG ATG GTC TTT GTG	Forward
K34nt1	CCA GCC ATC ACA TAC GTA	Reverse
K34R2	GTT CCT TTC CAG TTC TGG CA	Forward
K34R3	CTG AAA ACC GAG ACC GTG CTA	Forward

Continued on the next page →

(*Continued*)

Primer name	**Sequence** $5' \rightarrow 3'$	**Direction**
K34R4	CGC CAA CTT ACT TGA GCT T	Forward
GW-PKSR	CTT CTC AAA TCT ACT TCT CTT CAG	Reverse
K34FR	AGA CCA CCT CTG GTT CCA A	Reverse
	Miscellaneous	
AspActF	CCC AAG TCC AAC CGT GAG AA	Forward
AspActR	GGC CTG GAT GGA GAC GTA GA	Reverse
ITS1	TCC GTA GGT GAA CCT GCG G	Forward
ITS4	TCC TCC GCT TAT TGA TAT GC	Reverse
M13 Forward (-20)	GTA AAA CGA CGG CCA G	–
M13 Reverse	CAG GAA ACA GCT ATG AC	–

A-4 Nucleotide and deduced amino acid sequence of the *Solorina crocea* PKS1 gene

Changes in translation frames are marked with green numbers. The introns are indicated in lower case. 5′ and 3′ splice sites are marked in bold. Putative internal branch point sites for lariat formation are underlined. The stop codon is indicated with an asterisk.

Nucleotide sequence of the *S. crocea* PKS gene has been deposited in the GenBank/EMBL/DDBJ database[2] under accession number EF554834.

```
     1  ATGGGTTCGA TGCCGCTTGA CGATGGCATC CTGGGTTCTA CTGGAGATAG CAATGAGCCG ATGgtaagct     Intron I
+1   1   M  G  S  M   P  L  D   D  G  I   L  G  S  T   G  D  S   N  E  P   M

    71  ctccgtcctt atatgcgtgt attggtgtgc ttggacatct ttcgcgtgag ccagcgatgt ctggtcagct

   141  cgtccttcgc aattcttgtc tttcagcacg agggcttgac ttgagcagaa acatgcttac tagatgggtc

   211  acagGCCATC ATCGGTCTGG CAACACGGTT TCCGCAAGAT GCAAAAAGCA CAGAGGAGCT GTGGAAGTTT
+2  22       A  I   I  G  L  A   T  R  F   P  Q  D   A  K  S  T   E  E  L   W  K  F

   281  CTTCTGAGAG GTAGATCTGC TCACACGCCC TTCCCAGAAG ATCGAATTGG CAAAGGGCAT TATCATCCAG
    44   L  L  R  G   R  S  A   H  T  P   F  P  E  D   R  I  G   K  G  H   Y  H  P  D·

   351  ACCCAGAACA CGGCGGCTCC TTCGCGGTCA AGGGCGGCCA TTTCTTATCT GAAGATCCTG CCTATTTTGA
    68   ·  P  E  H   G  G  S   F  A  V  K   G  G  H   F  L  S   E  D  P  A   Y  F  D·

   421  TGCCCCTTTC TTCAGCATTA CGAAAGGCGA GGTCATGGCC CTGGATCCTC AACAACGAGT GGTGCTCGAA
    91   ·  A  P  F   F  S  I   T  K  G  E   V  M  A   L  D  P  Q   Q  R  V   V  L  E

   491  AGCGTTTACC ACGCTCTGGA AAACGCTGGT TTGCCACTTG AGACTGTTTC CGGTTCAAAT ACCTCGGTCT
   114   S  V  Y  H   A  L  E   N  A  G   L  P  L  E   T  V  S   G  S  N   T  S  V  F·

   561  TTGTCTCGGG CTTCAATCAT GATCACCTTG CAAACCTCAA TGCCGATCCA GAAAGTACGA TGAAGTACAA
   138   ·  V  S  G   F  N  H   D  H  L  A   N  L  N   A  D  P   E  S  T  M   K  Y  K·

   631  ACCGACAGGC ACAACGAACT CATTACTGAG TAATCGAGTC TCCTGGTTTT TCGACTTGAA AGCTCCGAGC
   161   ·  P  T  G   T  T  N   S  L  L  S   N  R  V   S  W  F  F   D  L  K   A  P  S

   701  GTTACCCTTG ACACTGCGTG TTCCAGCAGC ATGGTGGCCC TGCATCTGGG CTGTCAAAGC CTCCGGACAC
   184   V  T  L  D   T  A  C   S  S  S   M  V  A  L   H  L  G   C  Q  S   L  R  T  R·

   771  GGGAGTGCGA TATGTCTGTG ATCAGTGGTG TTACGATGAT ATCATTTCCC ACAGATATCA TGAGTATGGG
   208   ·  E  C  D   M  S  V   I  S  G  V   T  M  I   S  F  P   T  D  I  M   S  M  G·

   841  CCATCATGGG TTCCTGTCGC CCGAAGGCAA ATGCTTCTCT TTTGATCACC GAGCAGATGG CTATGCACGA
   231   ·  H  H  G   F  L  S   P  E  G  K   C  F  S   F  D  H  R   A  D  G   Y  A  R

   911  GGCGAGGGCG TTGGAAGCTT AGTCGTCAAG CGTCTCTCGG ATGCTATTAG GGATGGAGAC ACAATTCGGT
   254   G  E  G  V   G  S  L   V  V  K   R  L  S  D   A  I  R   D  G  D   T  I  R  S·

   981  CTGTCATCCG GGGAATAGGT GTCAATCAAG ATGGCCGGAC ACCTGGAGTT TCCATGCCAA GCAGTGCGGC
   278   ·  V  I  R   G  I  G   V  N  Q  D   G  R  T   P  G  V   S  M  P  S   S  A  A·

  1051  TCAGGAAAGC TTGATGAGGA AAGTCTATGC TTCGGCAGGG TTGGGCTTCG AAGACACGAT GATGGTTGAA
   301   ·  Q  E  S   L  M  R  K   V  Y  A   S  A  G   L  G  F  E   D  T  M   M  V  E

  1121  GCTCATGGTA CAGGAACAGC TGCAGGCGAT CCGATTGAAG CAAACGCTAT CGCAAGGACC TTCGCATCGC
   324   A  H  G  T   G  T  A   A  G  D   P  I  E  A   N  A  I   A  R  T   F  A  S  R·

  1191  GAAAATCAAA GATCCCACTC TATGTCGGTG CGATCAAGTC AGGAGTGGGT CATCTTGAGG GAGGCGCAGG
   348   ·  K  S  K   I  P  L   Y  V  G  A   I  K  S   G  V  G   H  L  E  G   G  A  G·

  1261  AGTTGCTGGg taagtaatat gctccctgaa tgcacagtct gtcacttata atcattgcag AGTCATTAAA     Intron II
+2  371  ·  V  A  G·                                                        ·  V  I  K
```

[2] *http://www.ncbi.nlm.nih.gov/*

```
1331  GGTACACTGA TCCTCGAAAG TGGCATCATC CCACCAAACG TCAATTTCGA GAAAGTGAAT CCAAAAATCC
 371   G  T  L  I  L  E  S  G  I  I  P  P  N  V  N  F  E  K  V  N  P  K  I  P·

1401  CACTGAAGAA GTGGAACCTG CAGTTTCCAG TGAAGCCTGT CCCTTGGCCA ACCTCAGATG TTCGTCGCAT
 401  ·  L  K  K  W  N  L  Q  F  P  V  K  P  V  P  W  P  T  S  D  V  R  R  I·

1471  CTCGATCAAT TCCTTTGGAG TCGGGGGCAC CAATGCACAC GTTATACTTG ATGACGCCCA CAGCTTTCTA
 424  ·  S  I  N  S  F  G  V  G  G  T  N  A  H  V  I  L  D  D  A  H  S  F  L

1541  TCAAACCGCG GGATCGATGC CTCTCACAAC ACAAGGAGGT CTGTTCCGTC ACAGGATGAA GTCGATAGAC
 447   S  N  R  G  I  D  A  S  H  N  T  R  R  S  V  P  S  Q  D  E  V  D  R  L·

1611  TAGTGCAAAG TCTAGAGGCA GATGCTATTG ACGACAGCTC AGAGACGGAT GCAATCAAAG CGATGAGTAG
 471  ·  V  Q  S  L  E  A  D  A  I  D  D  S  S  E  T  D  A  I  K  A  M  S  S·

1681  CACGGATCAC GCCAACACCG TGAAAATGAA TGGTCATACG ACTAGCCTCA CGAACGGCGT CAGTCGTGCG
 494  ·  T  D  H  A  N  T  V  K  M  N  G  H  T  T  S  L  T  N  G  V  S  R  A

1751  AAGAGCGAAA CAAATGGAAT GACCAATGTG ACATTTGCTG ATGAGAACAT ATCGCCCGTA CCTCGATTGA
 517   K  S  E  T  N  G  M  T  N  V  T  F  A  D  E  N  I  S  P  V  P  R  L  I·

1821  TCACTTTGGC GGCTTTTGAT GAATCAGGGG TACAAAGGAT TGCAAGTGCG CAAGCAGAGT ACCTGCAAAC
 541  ·  T  L  A  A  F  D  E  S  G  V  Q  R  I  A  S  A  Q  A  E  Y  L  Q  T·

1891  CATTGAGCGG TCGCATCAAT TCAACGAGGA AGACTTCCTT GATGATTTTG CCTATACCAT TTCGAAACGC
 564  ·  I  E  R  S  H  Q  F  N  E  E  D  F  L  D  D  F  A  Y  T  I  S  K  R

1961  TCGCCATTCC CCTGGAAGAG CTTCGCATTG GCTAGTAGCT TCAATGAGGT CATTCACTCT CTGCAGGCTC
 587   S  P  F  P  W  K  S  F  A  L  A  S  S  F  N  E  V  I  H  S  L  Q  A  L·

2031  TCCCGAAGCC CATAAGAGCT CGGAACGTGC CAAAGATTGG GTTCATATTC ACAGGCCAGG GTGCTCAGTG
 611  ·  P  K  P  I  R  A  R  N  V  P  K  I  G  F  I  F  T  G  Q  G  A  Q  W·

2101  GTATGCAATG GGTCGCGAAC TTATGTCTTA CCCTGTGTTT GAGAGAAGCC TTGCAGATGC ATCCTCATAC
 634  ·  Y  A  M  G  R  E  L  M  S  Y  P  V  F  E  R  S  L  A  D  A  S  S  Y

2171  ATGCAGTCGA TTGGGTCGCC CTGGTCCTTG TGGGACGAAT TGCTGCGAAA CAAATCTGAT ACCAAAGTCA
 657   M  Q  S  I  G  S  P  W  S  L  W  D  E  L  L  R  N  K  S  D  T  K  V  N·

2241  ACGAACCGCA CCTGGCTCAT CCTGCCTGTG CAGCACTGCA GATGGCTCTG GTCGACTTAC TCAGATCATG
 681  ·  E  P  H  L  A  H  P  A  C  A  A  L  Q  M  A  L  V  D  L  L  R  S  W·

2311  GCAAATAACG CCTTCGCGAG TTATTGGTCA CTCTTCTGGA GAAATTGCTG CCGCCTATTG CGCTGGCAAG
 704  ·  Q  I  T  P  S  R  V  I  G  H  S  S  G  E  I  A  A  A  Y  C  A  G  K

2381  CTGACGCGAG AATCGGCATG GAAGGTGGCT TATTTTCGAG GCTTCGTTTC GAACCAACAG CTGAGCGCAA
 727   L  T  R  E  S  A  W  K  V  A  Y  F  R  G  F  V  S  N  Q  Q  L  S  A  K·

2451  AAGGATCAAT GATGGCTGTC GGACTTTCGG AGCAAGAACT TCAACCATAT CTGAATAAAG TCAACGAAGA
 751  ·  G  S  M  M  A  V  G  L  S  E  Q  E  L  Q  P  Y  L  N  K  V  N  E  E·

2521  GAACACGGGG GAGCTCATTG TGGCCTGTTT TAACAGCCCA AAGAATCTCA CAGTTTCAGG AGACCAGCAT
 774  ·  N  T  G  E  L  I  V  A  C  F  N  S  P  K  N  L  T  V  S  G  D  Q  H

2591  AAGGTTGATG CTCTCAAAGA AATCCTGGAC GAAGCGGAAA TCTTTGCCCG AAAACTCCAA GTGAAGAATG
 797   K  V  D  A  L  K  E  I  L  D  E  A  E  I  F  A  R  K  L  Q  V  K  N  A·

2661  CTTACCACTC CGCGCACATG AAAGAAGTGG CAGACGAATA TCTTAGCTTG CTCGGCAGCT TGCCTGAGCC
 821  ·  Y  H  S  A  H  M  K  E  V  A  D  E  Y  L  S  L  L  G  S  L  P  E  P·

2731  ACGTGTTGAA TCATCTGAGC CTGTGCATAT GTTCTCATCC GTGACGGGCG GCAGGATAGA AAGTGGACAA
 844  ·  R  V  E  S  S  E  P  V  H  M  F  S  S  V  T  G  G  R  I  E  S  G  Q

2801  CTTGGCGCTC AGTACTGGGT GGACAATCTG GTATCGCCTG TACGGTTTGC CGATGCGGTA ACGCAAATGT
 867   L  G  A  Q  Y  W  V  D  N  L  V  S  P  V  R  F  A  D  A  V  T  Q  M  C·

2871  GCTTCTCTCG TGTGGACAAA GGCCAAGCTT CAATACAAAT GAATGAAAAT ACCGGCAGTG TCTTCTCGGA
 891  ·  F  S  R  V  D  K  G  Q  A  S  I  Q  M  N  E  N  T  G  S  V  F  S  D·

2941  TATCATCATG GAGATTGGAC CCCATGCGGC ACTCCAAAGC GCGGTCAAGG AGATCCTAGC GAGAACACCA
 914  ·  I  I  M  E  I  G  P  H  A  A  L  Q  S  A  V  K  E  I  L  A  R  T  P

3011  TTTGCTGCGG CGATCACATC GCTGGCGGTA CTGAATCGTT CAGCTCCTGG CCTTTTCACA ATATTGAACA
 937   F  A  A  A  I  T  S  L  A  V  L  N  R  S  A  P  G  L  F  T  I  L  N  T·

3081  CCGTTGGAAG TCTCTATGCC AGAGGCTACC CAATTGATGC CACAGAGATC AATCAAAGCT CCAAGCGAAA
 961  ·  V  G  S  L  Y  A  R  G  Y  P  I  D  A  T  E  I  N  Q  S  S  K  R  K·
```

```
3151  GAGCGTGAAA TTCTCACCAC AAGCAAGACC ACGTCGTCTG GTGAATCTTC CAGGGTACAC CTTCAACCAT
 984  · S  V  K   F  S  P  Q   A  R  P   R  R  L   V  N  L  P   G  Y  T   F  N  H

3221  TCTGAGAAGC TCATCTACGA GAGTCGGTTG ACTAAGAACT ATCGCTTGAG GAAGCAACCA CGACATGATC
1007   S  E  K  L   I  Y  E   S  R  L   T  K  N  Y   R  L  R   K  Q  P   R  H  D  L·

3291  TTTTCGGTGC ACCTGTGCCC GACTGGAACT CAGAGACGCC CAGATGGAGG AATATACTCA GATTAGCCGA
1031  · F  G  A   P  V  P   D  W  N  S   E  T  P   R  W  R   N  I  L  R   L  A  E·

3361  ACAGCCATGG CTACGAGATC ATATTGTCAC CAATCAGGTT GTACTACCTG GCGTCGGCTA CCTGATTGCT
1054  · Q  P  W   L  R  D  H   I  V  T   N  Q  V   V  L  P  G   V  G  Y   L  I  A

3431  GCCATCGAAG CTTCTCGGCA AATGGCAGAC CCTGATAGGA AACTAGTAGC CTTCAGACTG CGCGATATTT
1077   A  I  E  A   S  R  Q   M  A  D   P  D  R  K   L  V  A   F  R  L   R  D  I  S·

3501  CCTTAAAAAG GGCTCTCATC ATACCTGACA ACAAGGATGG CGTTGAGACA ATGCTGAGTC TTACGCGAAT
1101  · L  K  R   A  L  I   I  P  D  N   K  D  G   V  E  T   M  L  S  L   T  R  M·

3571  GGATGAATCC AGCTTACAAG CATCCAAGGT GTGGCACAAG TTTGCTATGT CATCCTACAA CCCTCTCGGA
1124  · D  E  S   S  L  Q  A   S  K  V   W  H  K   F  A  M  S   S  Y  N   P  L  G

3641  GAAGACTGGA TCGAACACTG CACTGGGTAT ATAACCCTTG ACTATGAAGC ACCGCAAAAT CCGATCGACA
1147   E  D  W  I   E  H  C   T  G  Y   I  T  L  D   Y  E  A   P  Q  N   P  I  D  N·

3711  ATGGACTTGA AGCTAGAGAG GAATCCCTGA TATGGGAAAG ACAGTTGCAT CATGTGAAAG AGAAATGTGT
1171  · G  L  E   A  R  E   E  S  L  I   W  E  R   Q  L  H   H  V  K  E   K  C  V·

3781  AGTGCCAATC GAAATCGACG ATGTATACGA TAACCTGGTG ACCACCGGTC TTGAATTTGG CCCGCTGTTC
1194  · V  P  I   E  I  D  D   V  Y  D   N  L  V   T  T  G  L   E  F  G   P  L  F

3851  AAAAACCTCT CAGAGGTGAA AGGTACGGTC TCACGTTCTG GAGAGATGTG CGGTATCGTG ACAGTTCCGA
1217   K  N  L  S   E  V  K   G  T  V   S  R  S  G   E  M  C   G  I  V   T  V  P  N·

3921  ATGTCGCAGC TGCTATGCCC ATGCAAGCAA TTCACCCACA TCTCATTCAC CCGTCCACCA TGGATAGCTT
1241  · V  A  A   A  M  P   M  Q  A  I   H  P  H   L  I  H   P  S  T  M   D  S  F·

3991  CATGCATTTC TTCTTGGCCT CGGTCATGGA TGCAACTGGG AAGAAGACGC TCGATCGACC TATGGTGCCG
1264  · M  H  F   F  L  A  S   V  M  D   A  T  G   K  K  T  L   D  R  P   M  V  P

4061  ACTTTTATCA AAGATGTCTG GGTGTCAGCC GAGCTGAACT CTGCACCGGG TCATCAATTT TGTGGCTATG
1287   T  F  I  K   D  V  W   V  S  A   E  L  N  S   A  P  G   H  Q  F   C  G  Y  G·

4131  GCAAGAGCAA CTTATTAGCC TACGACAAGT ACGAGAGTGA CATCAGAGTG TGGGACATCA CCACCCATGA
1311  · K  S  N   L  L  A   Y  D  K  Y   E  S  D   I  R  V   W  D  I  T   T  H  E·

4201  AGCCCGGGTT GCTATTGAAG GTGTCCGTGC AACTCCTTTG GAAGCCACGG ACAATGGTTC TGGATCTGTA
1334  · A  R  V   A  I  E  G   V  R  A   T  P  L   E  A  T  D   N  G  S   G  S  V

4271  CGCAAACTCA ACCACGAGTT GAAATGGACT CCGTACCTTG AAATGCTCGA CAAGTCAGCT CTCTCAAATG
1357   R  K  L  N   H  E  L   K  W  T   P  Y  L  E   M  L  D   K  S  A   L  S  N  V·

4341  TTCAGTTGAC CTCCTTACAA GACAACGAAG ACTATAGAGG CTGGATCACT AAGTTCCAGA TTGCTACACA
1381  · Q  L  T   S  L  Q   D  N  E  D   Y  R  G   W  I  T   K  F  Q  I   A  T  H·

4411  TCTCCTGGTC AACGATGCTC TTGACGATTT GGGCGGTCAA GCGCCCCAAG GGTTGGACGG ACATCTGCTA
1404  · L  L  V   N  D  A  L   D  D  L   G  G  Q   A  P  Q  G   L  D  G   H  L  L

4481  CGATATTTCG AATGGATGAA ACACGCTTCA AAATGCTTGA GAAATGACGA GGTTTCCGGT ATGCGTCTTG
1427   R  Y  F  E   W  M  K   H  A  S   K  C  L  R   N  D  E   V  S  G   M  R  L  A·

4551  CGGAGTTTGA AAGATATCAA AGCAGTCCCG AGATGAAAAA GCAGCTATTG GAAGAAGTGG AAGAACATAA
1451  · E  F  E   R  Y  Q   S  S  P  E   M  K  K   Q  L  L   E  E  V  E   E  H  N·

4621  CGCAGATGGT GCTCTGGCTC TTCGAATGGG TAAAAGCATC GTCAAAGTCC TTCGCAAAGA AGAAGATCCA
1474  · A  D  G   A  L  A  L   R  M  G   K  S  I   V  K  V  L   R  K  E   E  D  P

4691  CTACATCTGA TGTTCGGCGT GGACGACATT CTTGATCGTG TTTATGCTCA GGTCGCTCAC TTGGGAAACC
1497   L  H  L  M   F  G  V   D  D  I   L  D  R  V   Y  A  Q   V  A  H   L  G  N  L·

4761  TTCCTGCACT GCAAGCCGCG CTACTGAAGA TTGTAGCCGA CAATTCCACG AACCTGAGAG TGCTCGAGAT
1521  · P  A  L   Q  A  A   L  L  K  I   V  A  D   N  S  T   N  L  R  V   L  E  I·

4831  CGGTGCAGGC ACAGGTGGTT CGACAGTTGG AATGCTAGAA GGCTTGATCA GCACGTCTGA CGATGGAGTC
1544  · G  A  G   T  G  G  S   T  V  G   M  L  E   G  L  I  S   T  S  D   D  G  V

4901  TTTTCTTCCT CTGTCAGTAC GTACACCTTT ACTGACATCT CCTTGGCGTT TTTCGCGAAG GCCAAAGAGA
1567   F  S  S  S   V  S  T   Y  T  F   T  D  I  S   L  A  F   F  A  K   A  K  E  K·
```

```
4971  AGTTCAAGAA TTACCGTGAT ATCATGGAGT ACAAAGTCCT CGATATTGAG AAGGAGCCCG ACGAAGAAGT
1591  · F  K  N  Y  R  D  I  M  E  Y  K  V  L  D  I  E  K  E  P  D  E  E  V·

5041  CTTCGAGCTT GGAAGCTACG ACATCGTCGT AGCCCAGAAC GTGGTTCACG CGACCACTAA TCTCAGGAGG
1614  · F  E  L  G  S  Y  D  I  V  V  A  Q  N  V  V  H  A  T  T  N  L  R  R

5111  ACTTTGAGTC ATATCCGAAG CTTACTGAAA CCAGGTGGCA GACTTTTCCT CCAGGAAGGT GTGAGACAAG
1637   T  L  S  H  I  R  S  L  L  K  P  G  G  R  L  F  L  Q  E  G  V  R  Q  D·

5181  ATCTCTCTTG GTCAACAATT GCATTTGGCC AGCTTCCGGG TTGGTGGGCT GGCGCTGAAC CCACTCGTCG
1661  · L  S  W  S  T  I  A  F  G  Q  L  P  G  W  W  A  G  A  E  P  T  R  R·

5251  TTTGAGCCCA TGGATCACAA CCGAACAATG GTATCAAGTC ATCAAAGCCT CGGGTTATAC AGGCATCGAT
1684  · L  S  P  W  I  T  T  E  Q  W  Y  Q  V  I  K  A  S  G  Y  T  G  I  D

5321  CTGGAATTGC CAGATAGACC TTCGTCTGAT CTGCATACAC AAAGCCTGTT CGTCGCGACA GCCTTGCCCG
1707   L  E  L  P  D  R  P  S  S  D  L  H  T  Q  S  L  F  V  A  T  A  L  P  E·

5391  AAGTCCGTCA CCATGAGCAC CACGGGCGTG AGATCGCGAT TGTGACTAAC ACTCCTAGAG GAGAGGGCGA
1731  · V  R  H  H  E  H  H  G  R  E  I  A  I  V  T  N  T  P  R  G  E  G  E·

5461  AAGTGAGCTC TGCAAGGCAC TCAAATCTTG CCTTCAGCAG GAATTGAAGA TGTCCCGGGT GTCTATCATC
1754  · S  E  L  C  K  A  L  K  S  C  L  Q  Q  E  L  K  M  S  R  V  S  I  I

5531  CACGTCCATG ATCTTGGCCA CAAAAAAATT GACAAAACTG TCTGCATATC TGTCTGCGAA CTCGAACGAC
1777   H  V  H  D  L  G  H  K  K  I  D  K  T  V  C  I  S  V  C  E  L  E  R  P·

5601  CATTGCTATC AACACTGACA AAGACCGAGT TTGAGAGTGT CAAGCAATTG CTGTGCACCT GCAAGGGCAT
1801  · L  L  S  T  L  T  K  T  E  F  E  S  V  K  Q  L  L  C  T  C  K  G  I·

5671  TTTGTGGGTT ACTGGTGACA CCACAAAATA CCCAGAGCTC GGCATGATCA CAGGACTTAT GAGAACCGTT
1824  · L  W  V  T  G  D  T  T  K  Y  P  E  L  G  M  I  T  G  L  M  R  T  V

5741  CGTTGGGAGC GAGACATAGA TGAGGCGAAT TTGGTGACAC TTGCCGTTGC TGATCCGACA CCATCTCACA
1847   R  W  E  R  D  I  D  E  A  N  L  V  T  L  A  V  A  D  P  T  P  S  H  S·

5811  GCCTACTTGT CGAGAACATT ACCACTTTGT ACAAACAACA ATTCTGCGAA CCTCTGTCCA TCGACAAATA
1871  · L  L  V  E  N  I  T  T  L  Y  K  Q  Q  F  C  E  P  L  S  I  D  K  Y·

5881  TCAAGGAGAG TTCCTCCTGC AAAATGGATC GTTCCTCACA AGCCGACTGG TTGAAGCCAA AGCTGCTGAC
1894  · Q  G  E  F  L  L  Q  N  G  S  F  L  T  S  R  L  V  E  A  K  A  A  D

5951  GCATACCTTA GCTCTCAGTT CTCACGAGCA AAGCCAGCCA TGACACCTTT CAAAGATGCT GGGCGCCCAA
1917   A  Y  L  S  S  Q  F  S  R  A  K  P  A  M  T  P  F  K  D  A  G  R  P  I·

6021  TCAAGCTTGC GACGTCTGCT CCTGGGTTGC TCGACAAATT GGAATGGGTG ACAGATGAGA TCTACGACGA
1941  · K  L  A  T  S  A  P  G  L  L  D  K  L  E  W  V  T  D  E  I  Y  D  E·

6091  GCCACTCGCA CCTACCCACG TGGAAATTGA CATCAAAGCT GTTGCAATGA ATTTTCGTGA TTTGATGATC
1964  · P  L  A  P  T  H  V  E  I  D  I  K  A  V  A  M  N  F  R  D  L  M  I

6161  GCCATTGGGG AGCATATGGC TTACAGCATG GGCAACGAGG CTGCAGgtat gtcgttgcta aaacgtgtct    Intron III
1987   A  I  G  E  H  M  A  Y  S  M  G  N  E  A  A  G·

6231  ttgcccatcc ccactaactt tctcagGTAT TGTTTCTCGT GTCGGGAAAG AAGTGAAAGA CCTCAAAGTT
+1 2003                          · I  V  S  R  V  G  K  E  V  K  D  L  K  V

6301  GGAGATCGCG TTGTATATCT TTGCGGTATC GAGAGCACAG GCTGTTTTCA TACATTCGGC CGTGTTGATC
2017   G  D  R  V  V  Y  L  C  G  I  E  S  T  G  C  F  H  T  F  G  R  V  D  Q·

6371  AGAATGTCAT CGTGAGGATT CCCGACCGTA TGAGTTATGA GGTGGCTTGT GGGCTCCCAT GTGTTTATGC
2041  · N  V  I  V  R  I  P  D  R  M  S  Y  E  V  A  C  G  L  P  C  V  Y  A·

6441  CACAGTCATA TATGGTCTTG TCGATGCCGG AAGGCTGGAG AAGGGCGAGA AAATCCTTAT TCATGCTGCG
2064  · T  V  I  Y  G  L  V  D  A  G  R  L  E  K  G  E  K  I  L  I  H  A  A

6511  GCAGGTGGTG TCGGGCAAGC TGCTATACAC TATGCGAAAT ACGTTGGTGC AGAAATCTTC GCAACTGTGT
2087   A  G  G  V  G  Q  A  A  I  H  Y  A  K  Y  V  G  A  E  I  F  A  T  V  S·

6581  CAACGCCGGA GAAGCGAGCA gtaagtagcc cgctcaattt caactcgtca gcatgctaac ctgagacttc    Intron IV
2111  · T  P  E  K  R  A

6651  ctttagCTTG TCATGAAGCT CGGAGTTGCG GAAGACCACA TCTTTTCAAG CAGAGATCTT ACCTTCGCCA
+3 2117        L  V  M  K  L  G  V  A  E  D  H  I  F  S  S  R  D  L  T  F  A  K·

6721  AAGGCATCAT GCGTGCAACT GACGGCAAAG GGGTTGACTT AGTCTTGAAC TCACTGTCGG GAGAGGTACC
2139  · G  I  M  R  A  T  D  G  K  G  V  D  L  V  L  N  S  L  S  G  E  V  P·
```

```
6791  CATGTCCTAT CCAGATCTAA TCTTTTGTCG ACAGGCCCTC CGCCGATCAT GGGACTTGCT TGCCCCTTTC
2162  · M  S  Y   P  D  L  I   F  C  R   Q  A  L   R  R  S  W   D  L  L   A  P  F

6861  GGTCGATTCA TCGAAATTGG TAAGAAGGAC GCTCAAAACA ATGGGAAGGT TGAGTTACGT CCGTTCCTGC
2185   G  R  F  I   E  I  G   K  K  D   A  Q  N  N   G  K  V   E  L  R   P  F  L  R·

6931  GCAATGTCAC TATGGCCAGT GTCGAATTGC CAACTATGAT GCGACATCGG CCTTCCCTGA TCAAACGCTT
2209  ·  N  V  T   M  A  S   V  E  L  P   T  M  M   R  H  R   P  S  L  I   K  R  L·

7001  GACAGAAGAC ACTGTACGCC TTTGGACAGA AGGCCATATC AAAGAGGCCG ATCCGATGAC AATCATGAAC
2232  ·  T  E  D   T  V  R  L   W  T  E   G  H  I   K  E  A  D   P  M  T   I  M  N

7071  TACTCCCAGG TTGAGGAAGG CCTTCGTATC CTCCAATCTG GCAAAGGAAT GGGTAAGATG GTCTTTGTGC
2255   Y  S  Q  V   E  E  G   L  R  I   L  Q  S  G   K  G  M   G  K  M   V  F  V  P·

7141  CAACACCAGA CGATATGCTC CCCATCGTCC CGCCACGAGC TCCTATATAT CGACTCAAGG AAGAGGCTAC
2279  ·  T  P  D   D  M  L   P  I  V  P   P  R  A   P  I  Y   R  L  K  E   E  A  T·

7211  GTATGTGATG GCTGGTGGTC TCGGTGGTAT AGGACGAAGC TTTGCAAGAT GGATGGCCAA GCATGGAGCT
2302  ·  Y  V  M   A  G  G  L   G  G  I   G  R  S   F  A  R  W   M  A  K   H  G  A

7281  CGCAATCTTG TGTTCCTTTC CAGTTCTGGC AATATCACGT CGGCAGTCTC CGAGCTGGTG AGAGATCTCG
2325   R  N  L  V   F  L  S   S  S  G   N  I  T  S   A  V  S   E  L  V   R  D  L  E·

7351  AATTCGACAA GTGCAAAGTG GAAATCATCC GCTGCGATGT CAGCGATAAA GAAAAGCTCA AAGCGGTTCT
2349  ·  F  D  K   C  K  V   E  I  I  R   C  D  V   S  D  K   E  K  L  K   A  V  L·

7421  CGAAGAGTGC CAAAGGAAAA TGCCGCCAAT CAAGGGTGTT GTACAAGGAG CCATGAAATT GAA<b>Ggtgagt</b>     Intron V
2372  ·  E  E  C   Q  R  K  M   P  P  I   K  G  V   V  Q  G  A   M  K  L   K

7491  gaatgaaaac gccgtacagg gaaatgatct gacatgtt<b>ta gG</b>ATACAATG TTCGAAAATA TGCACTACGA
+2 2393                                            D  T  M   F  E  N  M   H  Y  E·

7561  AGAATTCGAG CTCGCCGTAA AGCCCAAAGT ACAGGGGTCC TGGAACCTCC ATGAACTCCT TCCACGAGAT
2403  ·  E  F  E   L  A  V  K   P  K  V   Q  G  S   W  N  L  H   E  L  L   P  R  D

7631  ATGGATTTTT TCGTTATGCT TTCATCCGCC ACTGGTGTTT TAGGCAATCG AGCACAGGCA AATTATGCAG
2426   M  D  F  F   V  M  L   S  S  A   T  G  V  L   G  N  R   A  Q  A   N  Y  A  A·

7701  CTGGAAACAC CTACCAGGAT ACGCTCGCCC ACCACCGCCG ATTGCTGGGC TTGCCAGCTT CGACCATTGA
2450  ·  G  N  T   Y  Q  D   T  L  A  H   H  R  R   L  L  G   L  P  A  S   T  I  D·

7771  TTTGGGTCAG GTCCTCTCTG TCGGTTATGT TGCTGAAAAC CGAGACCGTG CTATGGTGGC CAAGCATCTT
2473  ·  L  G  Q   V  L  S  V   G  Y  V   A  E  N   R  D  R  A   M  V  A   K  H  L

7841  GCAACTGTGC TTGAGGTCAT CAGGGAAGAT GAGATTCATG TTCTTCTCGA GTATCTGATG GACCCTCGCT
2496   A  T  V  L   E  V  I   R  E  D   E  I  H  V   L  L  E   Y  L  M   D  P  R  C·

7911  GCTCACCTCC GCCCCAGCTT GTTTCTGGTC TGACAACAGC TCAAATCTAC CGCCAGCGAG GAATGCCAAC
2520  ·  S  P  P   P  Q  L   V  S  G  L   T  T  A   Q  I  Y   R  Q  R  G   M  P  T·

7981  GCCAACTTAC TTGAGCTTTC CTCTCTTCAC ACACCTTCGG AACACAACTT CATCACGAAA TGGCGGCAGC
2543  ·  P  T  Y   L  S  F  P   L  F  T   H  L  R   N  T  T  S   S  R  N   G  G  S

8051  GATGGAGACG GTGGGATGTT AATCGAAGCA CTACTCAACT CAGCAAATAC ACTCGACGAG GCTACGGAGA
2566   D  G  D  G   G  M  L   I  E  A   L  L  N  S   A  N  T   L  D  E   A  T  E  I·

8121  TCGTTCTCGA AGCGGTTATC GCGAAATTAT CATCTTTACT TGCTGTGCCG CCAGAGAACA TAGATCGAGG
2590  ·  V  L  E   A  V  I   A  K  L  S   S  L  L   A  V  P   P  E  N  I   D  R  G·

8191  AAAGTCCGTC AGCTCAAATG GCGTTGATTC TTTGGTCGCC ATGGAGTTCA GAACATTCCT AGCGAAGGAA
2613  ·  K  S  V   S  S  N  G   V  D  S   L  V  A   M  E  F  R   T  F  L   A  K  E

8261  GTCAAGGCTG ACGTTCCTGT TCTCGAGATT CAAGGCACCA TGAGTGTCAC AGATCTCTCC AGGAAGATTG
2636   V  K  A  D   V  P  V   L  E  I   Q  G  T  M   S  V  T   D  L  S   R  K  I  A·

8331  CAAGCGTGAG TAGCGCGGTT CAGATCACTG AAGAGAAGTA G
2660  ·  S  V  S   S  A  V   Q  I  T  E   E  K  *
```