รายงานวิจัยฉบับสมบูรณ์

โครงการ: การสร้างวิถีชีวสังเคราะห์ของสารผลิตภัณฑ์ธรรมชาติจาก *Menisporopsis theobromae* BCC 4162

Reconstruction of natural product biosynthetic pathways from *Menisporopsis theobromae* BCC 4162

โดย ผศ. ดร. ปกรณ์ วรรธนะอมร และคณะ

มีนาคม 2562

สัญญาเลขที่ DBG5980009

รายงานวิจัยฉบับสมบูรณ์

โครงการ: การสร้างวิถีชีวสังเคราะห์ของสารผลิตภัณฑ์ธรรมชาติจาก *Menisporopsis theobromae* BCC 4162

Reconstruction of natural product biosynthetic pathways from *Menisporopsis theobromae* BCC 4162

คณะผู้วิจัย	สังกัด
ผศ. ดร. ปกรณ์ วรรธนะอมร	ภาควิชาเคมี มหาวิทยาลัยเกษตรศาสตร์
Dr. Colin Lazarus	School of Biological Sciences, University of Bristol

สนับสนุนโดยสำนักงานคณะกรรมการอุดมศึกษาและสำนักงานกองทุนสนับสนุนการวิจัย

บทคัดย่อ

จากปรากฏการณ์ดื้อยาของเชื้อก่อโรคในมนุษย์ที่พบมากขึ้นในปัจจุบัน ทำให้มีความต้องการอย่างเร่งด่วนในการค้นหายาชนิดใหม่ โดยในช่วงทศวรรษที่ผ่านมาจะเห็นได้ว่าอัตราในการค้นพบยาชนิดใหม่นั้นน้อยลง จึงทำให้ต้องมีการพัฒนากระบวนการในการค้นหายาแบบใหม่ จากเทคโนโลยีทางด้าน omics ที่มีการพัฒนาอย่างรวดเร็วทำให้ปัจจุบันการหาลำดับยีนในสิ่งมีชีวิตเป็นไปได้ง่ายขึ้น โดยงานวิจัยชิ้นนี้ได้นำเทคนิคในการผลิตสารออกฤทธิ์ทางธรรมชาติจากเชื้อราโดยใช้กระบวนการ heterologous expression ของยีนที่เกี่ยวข้องกับการผลิตสารผลิตภัณฑ์ธรรมชาติจากเชื้อรา *Menisporopsis theobromae* BCC 4162 ที่ไม่มีการแสดงออกในสภาวะการเลี้ยงทั่วไป โดยในงานวิจัยชิ้นนี้ได้ทำจากศึกษากลุ่มยีนที่ไม่แสดงออกที่คล้ายกับกลุ่มยีนที่สังเคราะห์สารประกอบ tenellin และกลุ่มยีนที่ประกอบด้วย highly reducing polyketide synthase (HR-PKS) โดยการโคลนยีนดังกล่าวเข้าสู่ multigene expression vector แล้วทำการถ่ายเข้าสู่ *Aspergillus oryzae* NSAR1 ซึ่งจะใช้เป็น host ในการผลิตสาร จากผลการทดลองพบว่า ไม่มีสารออกฤทธิ์ชนิดใหม่ผลิตขึ้นจาก *A. oryzae* NSAR1 ที่ได้ถูกถ่ายทอดกลุ่มยีนที่มีความคล้ายกับกลุ่มยีนที่สังเคราะห์สารประกอบ tenellin ซึ่งน่าจะเป็นเหตุผลมาจาก intron ซึ่งเป็น non-encoding nucleotide ที่ยังคงอยู่ในยีนและไม่ได้ถูกกำจัดออกโดย host ดังนั้น intron ที่อยู่ในยีนของกลุ่มยีนดังกล่าว จึงได้ถูกกำจัดออกโดยใช้เทคนิค PCR ก่อนที่จะทำการโคลนเข้าสู่ expression vector โดยในขณะนี้อยู่ในระหว่างการวิเคราะห์สารเมแทโบไลต์ที่ผลิตขึ้นจากกลุ่มยีนดังกล่าวที่ปราศจาก intron สำหรับยีน HR-PKS นั้นภายหลังจากที่ทำการโคลนและถ่ายเข้าสู่ *A. oryzae* NSAR1 แล้ว พบว่ามีสารใหม่ที่ผลิตขึ้นมาเป็นสารหลักจำนวน 2 ชนิด แต่มีปริมาณที่น้อยมาก จึงจะต้องทำการเพิ่มปริมาณเชื้อที่เลี้ยงเพื่อให้ได้สารทั้ง 2 ชนิดในปริมาณที่มากพอที่จะใช้ในการวิเคราะห์โครงสร้าง และการทดสอบการออกฤทธิ์ทางชีวภาพต่อไป

สำหรับชีวสังเคราะห์ของสาร menisporopsin A นั้น ภายหลังจากการวิเคราะห์ transcriptome ของ *M. theobromae* 4162 ในช่วงที่มีการผลิตสาร menisporopsin A พบว่ามี PKS ยีนจำนวน 2 ยีนคือ reducing PKS gene (*men1*) และ non-reducing PKS gene (*men2*) ที่เกี่ยวข้องกับชีวสังเคราะห์ของสารประกอบดังกล่าวโดยพบว่า PKS ทั้งสองสามารถสังเคราะห์สาร ascotrichalactone A ซึ่งมีความคล้ายคลึงกับ menisporopsin A โดยวิถีชีวสังเคราะห์ในการสร้างสาร menisporopsin A นั้น อาจจะต้องมีการใช้เอนไซม์ ketoreductase (KR) เพิ่มขึ้นอีกหนึ่งเอนไซม์ เพื่อที่จะทำให้วิถีชีวสังเคราะห์ของสาร menisporopsin A สมบูรณ์

ผลงานที่ได้จากงานวิจัยชิ้นนี้สามารถที่จะเป็นรากฐานในการผลิตสารเมแทโบไลต์ชนิดใหม่จากกลุ่มยีนที่ไม่แสดงออก รวมทั้งยังสามารถนำมาใช้ในการศึกษาชีวสังเคราะห์ของสารผลิตภัณฑ์ธรรมชาติจากเชื้อรา ที่จะนำไปสู่การพัฒนาโครงสร้างของสารออกฤทธิ์ให้มีประสิทธิภาพมากขึ้นโดยอาศัยกลไกทางชีวสังเคราะห์

คำสำคัญ: menisporopsin A, macrocyclic polylactone, polyketide synthase

Abstract

With the increased prevalence of drug-resistant human pathogens, new drugs are urgently needed. Over the last decade the rate of drug discovery has declined requiring new approaches for drug discovery. Advances in omics technology have sped up the search for previously undiscovered genes responsible for the biosynthesis of novel natural products. In this work, heterologous expression was used to produce new metabolites from the cryptic natural product biosynthetic gene clusters and also to investigate the menisporopsin A biosynthesis from *Menisporopsis theobromae* BCC 4162. The cryptic tenellin-like and unknown highly reducing polyketide synthase (HR-PKS) gene clusters were constructed into the multiple expression vectors and introduced into the heterologous host, *Aspergillus oryzae* NSAR1, for metabolite production. Unfortunately, no new metabolites could be observed from the transformants containing tenellin-like gene cluster. This may be a result of the existence of introns in the genes that cannot be removed by the host. Therefore, the introns found in the genes of this cluster were removed by PCR technique prior to cloning into the expression vector. Currently, the transformants containing tenellin-like gene cluster are being analyzed for metabolite production. For the HR-PKS, two new major metabolites could be produced by the transformants, however, the amounts of these metabolites are very low. The medium used for culturing the transformants will be scaled up to afford enough metabolites for structure determination and biological activity tests.

The biosynthesis of menisporopsin A was also investigated. Transcriptome analysis of *M. theobromae* during production phase of menisporopsin A identified two PKS genes involved in its biosynthesis i.e. reducing PKS gene (*men1*) and non-reducing PKS gene (*men2*). Our results show that only these two PKSs are sufficient to catalyze the formation of ascotrichalactone A, a structural derivative of menisporopsin A. An additional ketoreductase (KR) is possibly required to complete the menisporopsin A biosynthesis. Results from this project provide us a platform for production of novel metabolites from the cryptic biosynthetic gene clusters and also an approach for biosynthetic studies of fungal natural products. This could help to improve their biological activities by structural modification based on biosynthetic machinery.

Keywords: menisporopsin A, macrocyclic polylactone, polyketide synthase

Executive summary

Natural products have served as the source of drugs for human possibly since early man. They can be found in a variety of natural sources such as animals, plants and micro-organisms. The latter are a major source of antibiotics and other biologically relevant compounds such as anticancer, antimalarial and immunosuppressive agents. During 1981-2010, 34% of small molecule approved drugs are based on the natural products and their derivatives obtained by semi-synthesis [1]. In 2014, World Health Organisation (WHO) reported that the prevalence of antimicrobial resistant pathogenic microorganisms is growing while the rate of novel antibiotics is plummeting [2]. The need for novel drugs is undoubtedly unavoidable.

In this work, we used heterologous expression approach as a tool for discovery of new metabolites from two cryptic biosynthetic gene clusters found in the seed fungus *Menisporopsis theobromae* BCC 4162 i.e. tenellin-like biosynthetic gene cluster and a gene cluster containing highly reducing type I polyketide synthase (HR-PKS) gene. Furthermore, heterologous expression of two PKS gene candidates based on the transcriptome result for menisporopsin A biosynthesis was also conducted. Menisporopsin A is a macrocyclic polylactone produced by *M. theobromae* BCC 4162 and exhibits a broad spectrum of biological activities including antimalarial activity. Identification of the PKSs involved in the biosynthesis of menisporopsin A will pave the way to study the mechanism on how the fungus programmes the multiple esterfication and cyclolactonization. This could lead to structural modification of the macrocyclic polylactone and may help to enhance their biological activities.

For reconstruction of biosynthetic gene clusters, homologous recombination in yeast and Gateway® cloning system were used to construct the expression vectors containing the target biosynthetic gene clusters. This was followed by the introduction of these expression vectors into heterologous host, *Aspergillus oryzae* NSAR1, for metabolite production. The transformants obtained were further extracted with ethyl acetate and the extracts were analyzed by HPLC-MS

for novel metabolite production. The new metabolites were further isolated using HPLC equipped with C-18 semipreparative column (10.0 mm x 250 mm). Structures of isolated compounds were determined by NMR and HRMS.

For the tenellin-like gene cluster, it consists of 5 genes i.e. *menS* which is a megasynthase gene encoding for a hybrid polyketide synthase - non-ribosomal peptide synthetase (PKS-NRPS), *menC* encoding for trans-acting enoyl reductase (ER), *menA* and *menB* encoding for cytochrome P450 oxidase and *menD* encoding for flavin oxidoreductase. Firstly, each gene was amplified directly from genomic DNA and cloned into expression vector. However, no novel metabolites could be obtained and this could be a result of the introns (non-encoding nucleotide region) remaining in the auxiliary gene. The intron removal in *A. oryzae* may not be sufficient. As each gene in the tenellin-like gene cluster contains some introns except *menC* and to overcome this problem, the introns were removed from each gene by PCR using primer with overlapping region without intron sequences. The expression vector containing genes without introns was successfully constructed and then introduced into *A. oryzae* NSAR1 for metabolite production. The screening for new metabolites produced by transfromants containing this gene cluster without introns is currently being investigated.

The HR-PKS gene was also successfully constructed and introduced into *A. oryzae* NSAR1. The ethyl acetate extract from transformants were analyzed by HPLC and our results show that some new metabolites could be observed comparing to transformants without HR-PKS gene. However the amounts of these new metabolites are very low (less than 1 mg from 1 L of culture). To obtain enough metabolites for structure determination and biological test, the cultures used for metabolite production have to be scaled up.

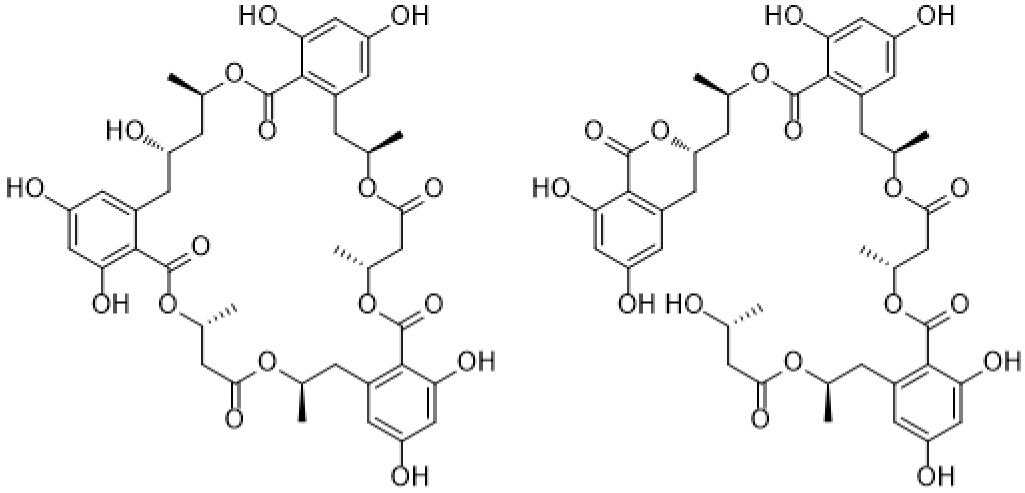
For the heterologous expression of the PKS gene candidates for menisporopsin A biosynthesis, two PKS genes were chosen based on transcriptomic result during the production phase of menisporopsin A. They are highly reducing and non-reducing PKS genes which are named *men1* and *men2*, respectively. The extracts from the transformants containing these two genes were analyzed by HPLC and there are 3 major metabolites produced by the transformants. These new metabolites were isolated and identified as (-)-orthosporin, (-)-6-hydroxymellein and ascotrichalactone A by NMR, HRMS and their optical activities. Both (-)-orthosporin and (-)-6-hydroxymellein are expected to be shunt products during the biosynthesis of menisporopsin A. Surprisingly, menisporopsin A was not found but its structural derivative, ascotrichalactone A, could be observed. The remaining of carbonyl group on ascotrichalactone A indicates that an additional ketoreductase (KR) is required to complete the biosynthesis of menisporopsin A. Currently, the candidate KRs for the biosynthesis of menisporopsin A are being investigated.

In summary, tenellin-like gene cluster and HR-PKS were successfully constructed and introduced into *A. oryzae* NSAR1 for metabolite production. However, the intron problem found in tenellin-like gene cluster may inhibit the production of new metabolites. With the new expression vector containing genes without introns, this may resolve the problem of metabolite production by this gene cluster. For the gene cluster containing HR-PKS, the production medium used for culturing the transformants will be scaled up to increase the yield of metabolites. In addition, our results also show that only two PKS genes, *men1* and *men2*, identified by transcriptomic data require for the biosynthesis of macrocyclic polylactone, ascotrichalactone A. The presence of carbonyl group in ascotrichalactone A indicates the need of additional *trans*-acting KR to complete the biosynthesis of menisporopsin A. These results will provide us a platform for mechanistic studies on the formation of macrocyclic polylactone and could be used to modify the structure of this class of metabolites for better biological activities.

Introduction

In order to increase the number of novel drugs, many approaches have been developed in order to achieve this aim including high-throughput screening (HTS). Nevertheless the successful rate of the discovery of new drugs by this technique has declined in the past decade. This may be results of several factors including the limitation of chemical libraries [3]. Therefore efforts toward the discovery of novel metabolites isolated from natural sources have re-emerged. This can be achieved by several approaches including searching for new species of bioactive compound producing micro-organisms. After the micro-organisms were isolated and cultured, extraction and purification using chromatographic methods are generally applied to obtain the metabolites. This will be followed by screening the compounds for biological activities. However only some metabolites can be detected and isolated from culture medium while a large number of metabolites from cryptic or silent biosynthetic pathways remain unexplored. This may be due to the growth medium and conditions not being optimal for metabolite production. With the advent of rapid genome sequencing technique, it has now become much easier to identify natural product biosynthetic genes from micro-organisms. The genomic data obtained can be utilised as a tool to predict metabolites based on biosynthetic gene cluster. This will be followed by optimizing the growth conditions to obtain the compounds or by expressing the gene cluster in heterologous host which is proven to be efficient in both bacterial and fungal systems. The latter technique can also be used to study the biosynthetic pathway of complex metabolites, including structural modifications, to obtain unnatural bioactive compounds.

Several biological active metabolites were discovered by traditional screening as mentioned above including a macrocyclic polylactone, menisporopsin A, and a linear polyester, menisporopsin B, from the seed fungus *Menisporopsis theobromae* BCC 4162 (Figure 1) [4, 5]. The latter indicates the lack of the final esterification to form menisporopsin A. Both compounds show a broad spectrum of biological activities such as cytotoxicity against cancer cell lines, antimalarial and antimycobacterial activities. The biosynthesis of menisporopsin A was also proved to be derived from 20 units of acetate [6].



Menisporopsin A

Menisporopsin B

Figure 1 Structures of menisporopsins A and B

To date, *M. theobromae* BCC 4162 has been shown to produce mainly menisporopsins A and B. Typically, micro-organisms are a host of biologically active compounds which makes us believe that a large number of cryptic biosynthetic pathways and therefore bioactive compounds are still undiscovered in this organism. Therefore genome sequencing is very crucial to reveal novel secondary metabolites produced by *M. theobromae* BCC 4162. From our genomic data of *M. theobromae* BCC 4162, there are 102 gene clusters possibly involved in the biosynthesis of secondary metabolites. These gene clusters were identified using antibiotics & Secondary Metabolites Analysis Shell (antiSMASH) programme [7] to search a genome sequence for secondary metabolite biosynthetic gene clusters. The majority of genes encode type I polyketide synthases (PKSs). Some of the gene clusters show the similarity to several natural product biosynthetic pathways such as tenellin [8], TAN-1612 [9], asperfuranone [10] and dehydrocurvularin [11] (Figure 2). Heterologous expression of these gene clusters will allow us to identify novel secondary matabolites produced by *M. theobromae* BCC 4162

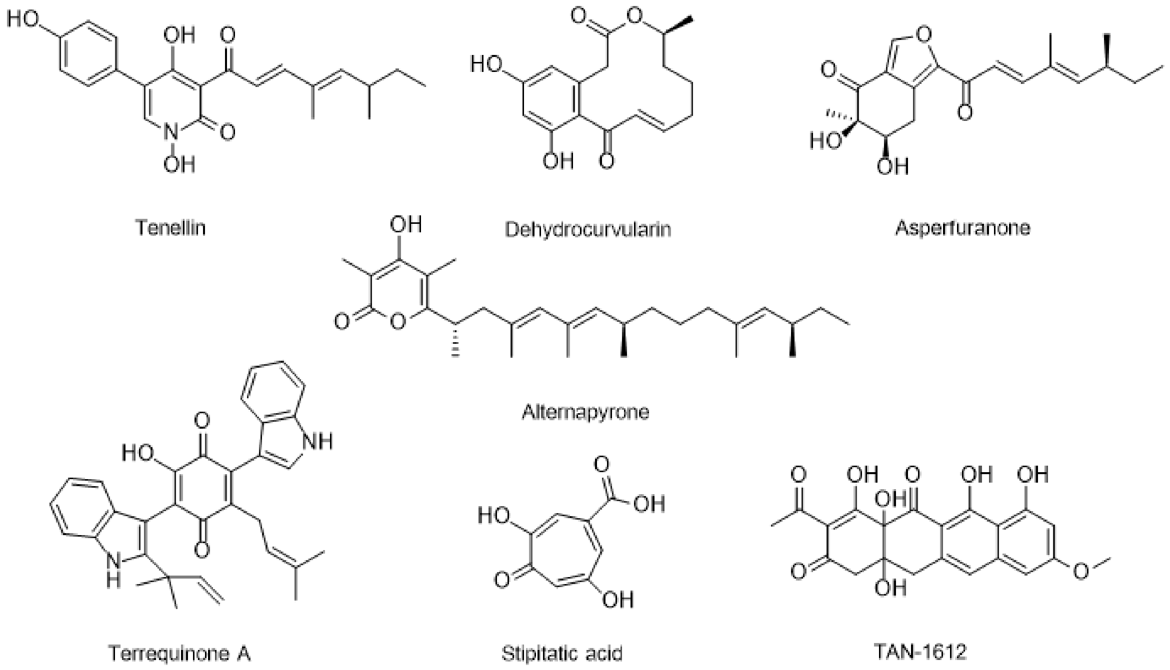


Figure 2 Structures of metabolites with similar biosynthetic gene clusters found in *M. theobromae* BCC 4162

In this work, two cryptic biosynthetic gene clusters were constructed for the production of new metabolites. The first gene cluster is similar to that of tenellin biosynthesis. Tenellin is a fungal metabolite containing 2-pyridone moiety. This compound was first isolated from *Beauveria bassiana* exhibiting cytotoxicity to mammalian erythrocytes and also acts as an iron chelator [12]. Several metabolites containing 2-pyridone also show antimalarial activity such as torrubiellone A [13], cordypyridones A and B [14]. Therefore the metabolites produced by this cryptic biosynthetic gene found in *M. theobromae* BCC 4162 may be invaluable 2-pyridone compound. The biosynthesis of tenellin is the most studied among 2-pyridone natural products [8]. Its biosynthetic gene cluster consists of four genes, namely *tenS* encoding for highly reducing type I PKS fused with nonribosomal peptide synthase (PKS-NRPS), *tenC* encoding for trans-acting enoyl reductase (ER), *tenA* and *tenB* both encoding for cytochrome P450 monooxygenases. An enoyl reductase (ER) domain in *tenS* is inactive, so the missing function is supplied by a *trans*-acting enoyl reductase encoded by *tenC.* The genes *tenA* and *tenB* are involved in the last two oxidation steps in tenellin biosynthesis (Figure 3).

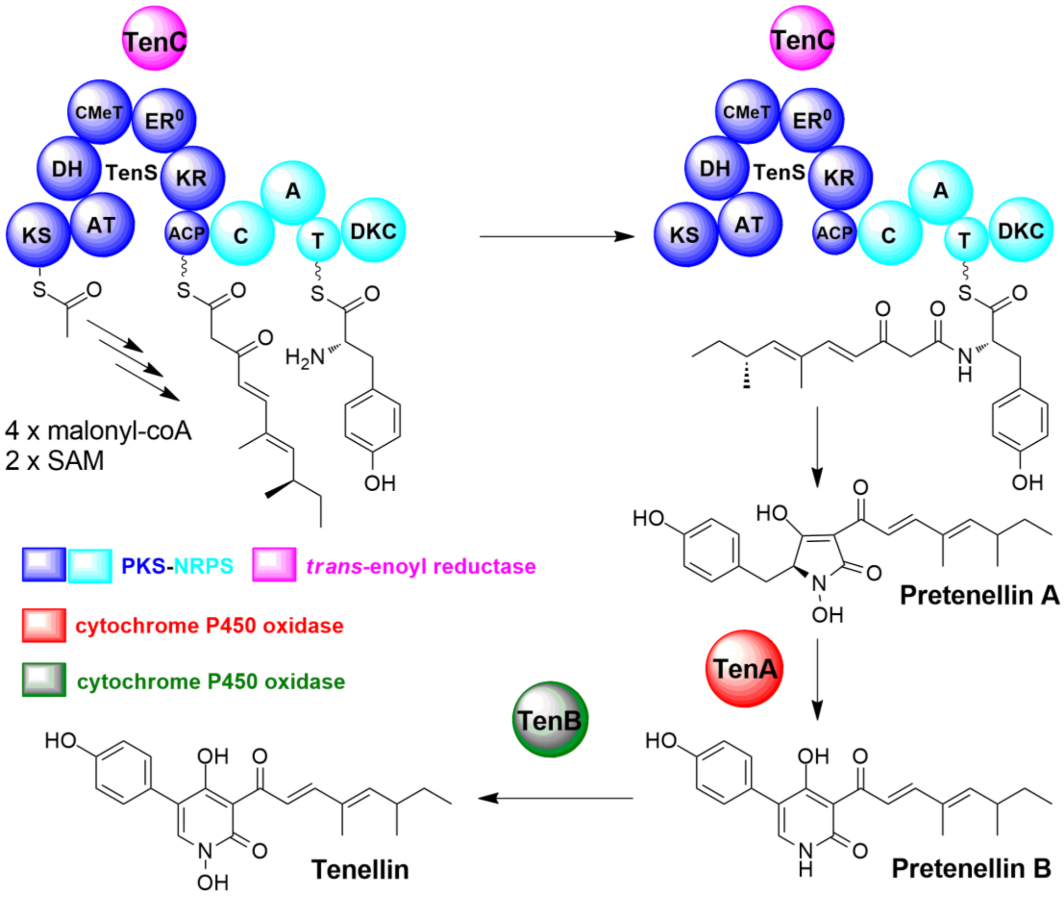


Figure 3 Biosynthetic pathway of tenellin from *B. bassiana*

Based on tenellin biosynthesis, the PKS-NRPS catalyses the formation of pentaketide intermediate followed by condensation with an aromatic amino acid [8]. One of the gene cluster found in *M. theobromae* BCC 4162 shows 80% similarity in gene contents to that of tenellin and each gene is named according to similarity to each gene involved in tenellin biosynthesis (Figure 4). However, an additional auxiliary gene, *menD*, was found in *M. theobromae* BCC 4162 and this gene possibly encodes for flavin oxidoreductase. Reconstruction of this gene cluster may result in the production of novel bioactive compounds which have similar structure to tenellin.

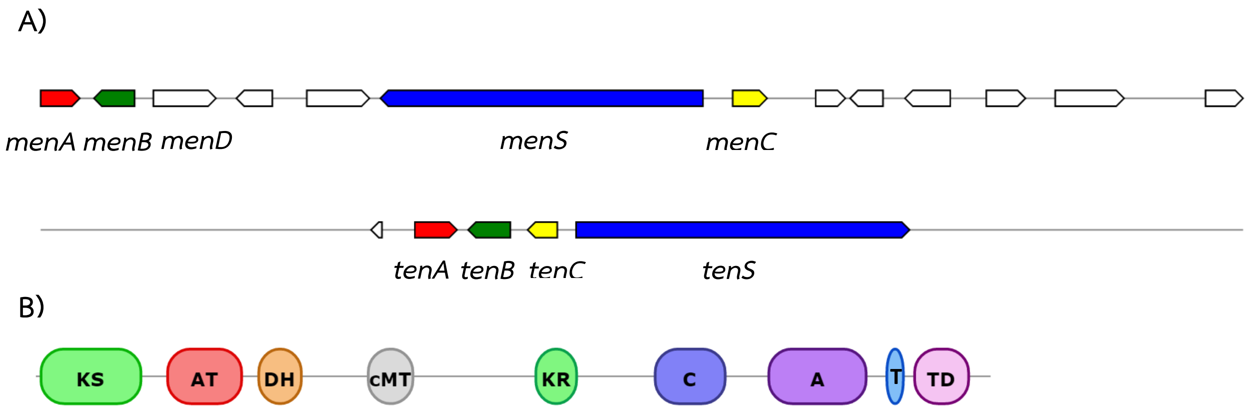


Figure 4 A) A biosynthetic gene cluster found in *M. theobromae* BCC 4162 (top) compared with that of tenellin (below). **B)** The domain organization of *menS* consists of ketosynthase (KS), acyltransferase (AT), dehydratase (DH), C-methyltransferase (cMT), *β*-keto reductase (KR) linked to a nonribosomal peptide synthetase module comprising condensation (C), adenylation (A), thiolation (T, also called peptidyl carrier protein) and terminal reductase (TD) domains..

The second gene cluster as shown in Figure 5 contains highly reducing PKS (HR-PKS) gene with no related natural product biosynthetic gene cluster. In order to investigate its function, this gene was also constructed and introduced into *A. oryzae* for metabolite production. This

HR-PKS may involve in the production of novel bioactive compound that cannot be produced from *M. theobromae* BCC 4162 under general medium and growth condition.

Figure 5 A biosynthetic gene cluster contains highly reducing polyketide synthase gene (coloured in brown) found in *M. theobromae* BCC 4162. The regulatory gene is highlighted in green.

In addition, the genomic data obtained from *M. theobromae* BCC 4126 was firstly used as a tool to reveal the biosynthetic pathway of menisporopsin A. This compound is uniquely constructed from five polyketide subunits via ester bonds. This class of metabolite is also found in several fungi such as *Periconia byssoides* [15], *Penicillium verruculosum* [16] and *Hypoxylon oceanicum* [17]. Understanding how these pathways are programmed to assemble each subunit together will give us a better insight on its biosynthetic pathway and may provide us information to modify their structures. From our previous work, we were able to amplify reducing and non-reducing PKS genes during the production phase of menisporopsin A. This led us to identify two potential PKS gene candidates which are R-PKS and NR-PKS genes. Unfortunately, only full length of NR-PKS gene could be obtained from genome sequencing and no metabolites could be detected from the transformants containing the NR-PKS gene. However, transcriptomic data was used in this work and successfully identified two PKS genes that are truly responsible for the biosynthesis of menisporopsin A.

Research objectives

1.1 To discover the novel metabolites from the cryptic biosynthetic gene clusters of *M. theobromae* BCC 4162 using heterologous expression.

1.2 To identify the PKS genes responsible for the menisporopsin A biosynthesis using heterologous expression.

Research methodology

1. Transcriptome sequencing

M. theobromae BCC 4162 was first grown in potato dextrose broth (PDB) for 7 days and then seed culture was transferred to menisporopsin A production medium, FMSM, consisting of 1% w/v fructose, 2.5 % w/v meat extract, 0.0025% w/v KH_2PO_4 and 0.0025% w/v $MgSO_4.7H_2O$ [18]. Mycelia for RNA extraction were collected both before and after growing in FMSM for 4 days in order to compare the expression level of PKS genes during production of menisporopsin A. Total RNA was extracted using RNeasy plant mini kit (QIAGEN) and treated with DNase I, then mRNA was enriched using oligo dT and fragmented before conversion to cDNA; adapters were subsequently connected to the fragments. After quality control by Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System, the transcriptome library was sequenced on the Illumina HiSeq 4000 platform by BGI Tech Solutions (Hong Kong). Paired-end read data were filtered to remove low quality and adapter-polluted sequences and *de novo* assembly of the clean data was performed using Trinity [19]. The assembled transcripts were then annotated with 7 functional databases (NR, NT, GO, COG, KEGG, Swissprot and Interpro). The transcripts were compared using Blast v2.2.23 against NT, NR, COG, KEGG and SwissProt databases and were annotated using Blast2GO v2.5.0 and InterProScan5 v5.11-51.0 by comparing them against GO and InterPro databases, respectively. Only the transcripts encoded for polyketide synthases were chosen for analysis.

2. Reconstruction of biosynthetic gene clusters in multigene expression vector

Due to the large size of hybrid PKS-NRPS and other type I polyketide synthase genes, the genes encoding these enzymes were amplified into 2-3 fragments by polymerase chain reaction (PCR). Each fragment was joined by homologous recombination in yeast to reporter genes contained within the *E. coli*-yeast shuttle vectors such as pEYA2dsRED and pEYA2eGFP containing reporter genes encoding DsRED and eGFP fluorescent proteins respectively (Figure 6).

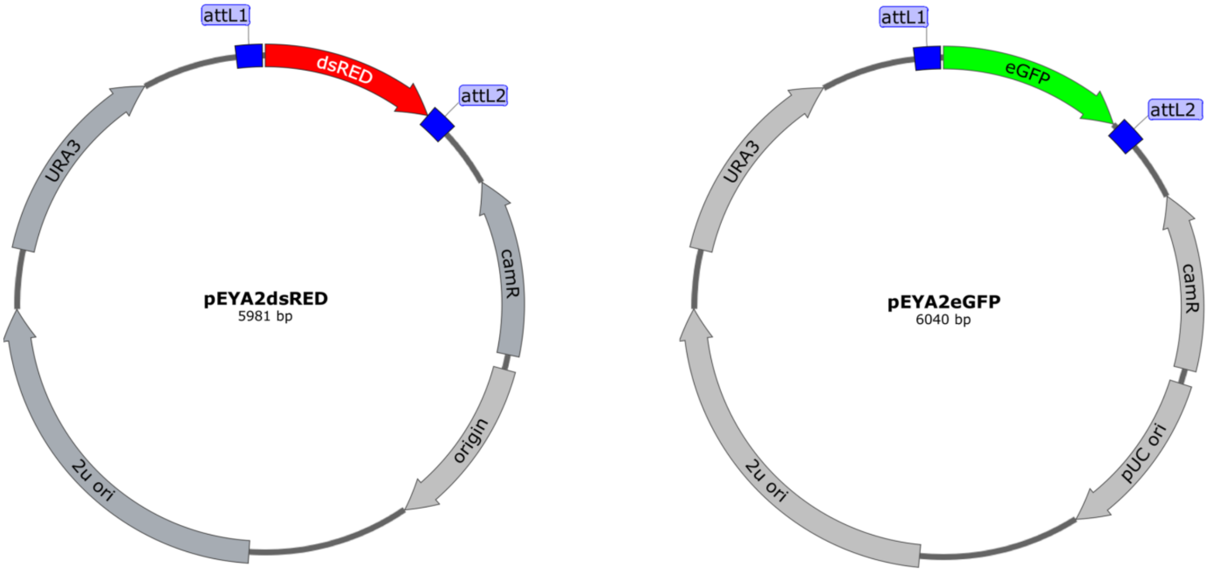


Figure 6 Plasmid maps of *E. Coli*-yeast shuttle vectors used for assembly of megasynthase genes such as hybrid PKS-NRPS and PKS genes.

For The primers used in this work are shown in Table S1, and all PCRs were performed using KAPA HiFi HotStart DNA polymerase (KAPABIOSYSTEMS, USA); the programme used for amplification of each gene was: initial denaturation at 95°C for 5 min followed by 35 cycles of 98°C for 20 sec, 65-68 °C for 15 sec and 72°C for 1.5 min, with a final extension at 72°C for 5 min. *Saccharomyces cerevisiae* YPH 499 was transformed with vectors linearized upstream of their reporter genes plus overlapping PCR products using the TRAFO method **[20]** and recombinant plasmids were shuttled to *E. coli* DH5α.

gene	primer	Sequence 5′-3′
men1	*Fragment 1*	
	Men1-1-fwd	TAATGCCAACTTTGTACAAAAAAGCAGGCTATGGGCAGCGTCTACGATAG
	Men1-1-rev	AGCTCCGCGACGAGGGATGG
	Fragment 2	
	Men1-2-fwd	TCAATCGCGCCCATCTTACG
	Men1-2-rev	ATATCAACACGCGGGACGTG
	Fragment 3	
	Men1-3-fwd	TCCAAGCCCAGAAGAAGTCG
	Men1-3-rev	CTTGATGACGTCCTCGGAGGAGGCCATCTGCTTGGTGGTGGCGGCGGTAG
men2	*Fragment 1*	
	men2-1-fwd	TAATGCCAACTTTGTACAAAAAAGCAGGCTATGGCGGGCCCCGTTGCAAC
	men2-1-rev	AGGCGCCGAGCACCTGCTCG
	Fragment 2	
	men2-2-fwd	ATGCAGGAACGGTGCGAGTC
	men2-2-rev	TGAACAGCTCCTCGCCCTTGCTCACCATGTTGTCCATGGCGTCCTTAAGG
HR-PKS	*Fragment 1*	
	pEYA2-C89F	TAATGCCAACTTTGTACAAAAAAGCAGGCTATGGGCAGCGTCTACGATAG
	C89-mid-R2	AGCTCCGCGACGAGGGATGG
	Fragment 2	
	C89-mid-F2	TCAATCGCGCCCATCTTACG
	C89-mid-R3	ATATCAACACGCGGGACGTG
	Fragment 3	
	C89-mid-F3	TCCAAGCCCAGAAGAAGTCG
	C89-dsRED-R	CTTGATGACGTCCTCGGAGGAGGCCATCTGCTTGGTGGTGGCGGCGGTAG
menS	*Fragment 1*	
	menS-1-fwd	TTTGTACAAAAAAGCAGGCTCCGCGGCCGCATGACGCAGAATCAGGAGAC
	menS-1-rev	CGGTAGATCTCCACGTTCGC
	Fragment 2	

	menS-2-fwd	GTGCTGAGCGAAGTTCTGCT
	menS-2-rev	CCACTACAATTGCCAGCGGTT
	Fragment 3	
	menS-3-fwd	CGCCGACTATCTGCTTGCAC
	menS-3-rev	CCTTGACCGCGTCCACGAT
	Fragment 4	
	menS-4-fwd	GCAGATCTCGAGCCAGCAGATCACCATTG
	menS-4-rev	CCGGTGAACAGCTCCTCGCCCTTGCTCACCATTGGCTCTGCACCGAGATAAG
menA	Padh-orf0-F2	TTCTTTCAACACAAGATCCCAAAGTCAAAATGTTTCTCATACACAATTTC
	orf0-Tadh-R	TCATTCTATGCGTTATGAACATGTTCCCTCTACTTGTACAGCGCTACC
menB	Peno-B-F	GTCGACTGACCAATTCCGCAGCTCGTCAAAATGGTGCCGCGGGTAGGCG
	Peno-B-R	GGTTGGCTGGTAGACGTCATATAATCATACCTAGAGCTTGCG
menC	Pgpd-C-F	AACAGCTACCCCGCTTGAGCAGACATCACCATGGCAGTCTTTGCC
	Tgpd-C-R	ACGACAATGTCCATATCATCAATCATGACCCTACGCAGCACATAGCACCTCG
menD	Padh-MenD-F	TCTTTCAACACAAGATCCCAAAGTCAAAATGGAACTTAACATCGCGAAGC
	MenD-Tadh-R	TTCATTCTATGCGTTATGAACATGTTCCCTTCATGTTAGCATTGGGTACC

Table 1 Primers used for constructing *men1* and *men2*. Underlined nucleotides are regions of overlapping sequence used for homologous recombination.

The Gateway® LR recombination reaction (Invitrogen, USA) was performed to insert PKS gene fusions into the amyB expression cassette of vectors such as pTYGSarg, pTYGSade and pTYGSmet which have different selection marker. For the tenellin-like gene cluster, auxilary genes were cloned into expression vector using homologous recombination in yeast prior to introduction of the PKS-NRPS gene into the expressionvector by Gateway® cloning. The expression vector used in this method has been developed by Dr. Lazarus at School of Biological Science, University of Bristol. This expression vector can accommodate up to 4 genes (Figure 7) under inducible promoter, P*amy*B and three constitutive promoters i.e. P*adh*, P*gpd*A and P*eno* [21]. Using different selection markers, a large number of genes can be simultaneously expressed from more than one expression vector.

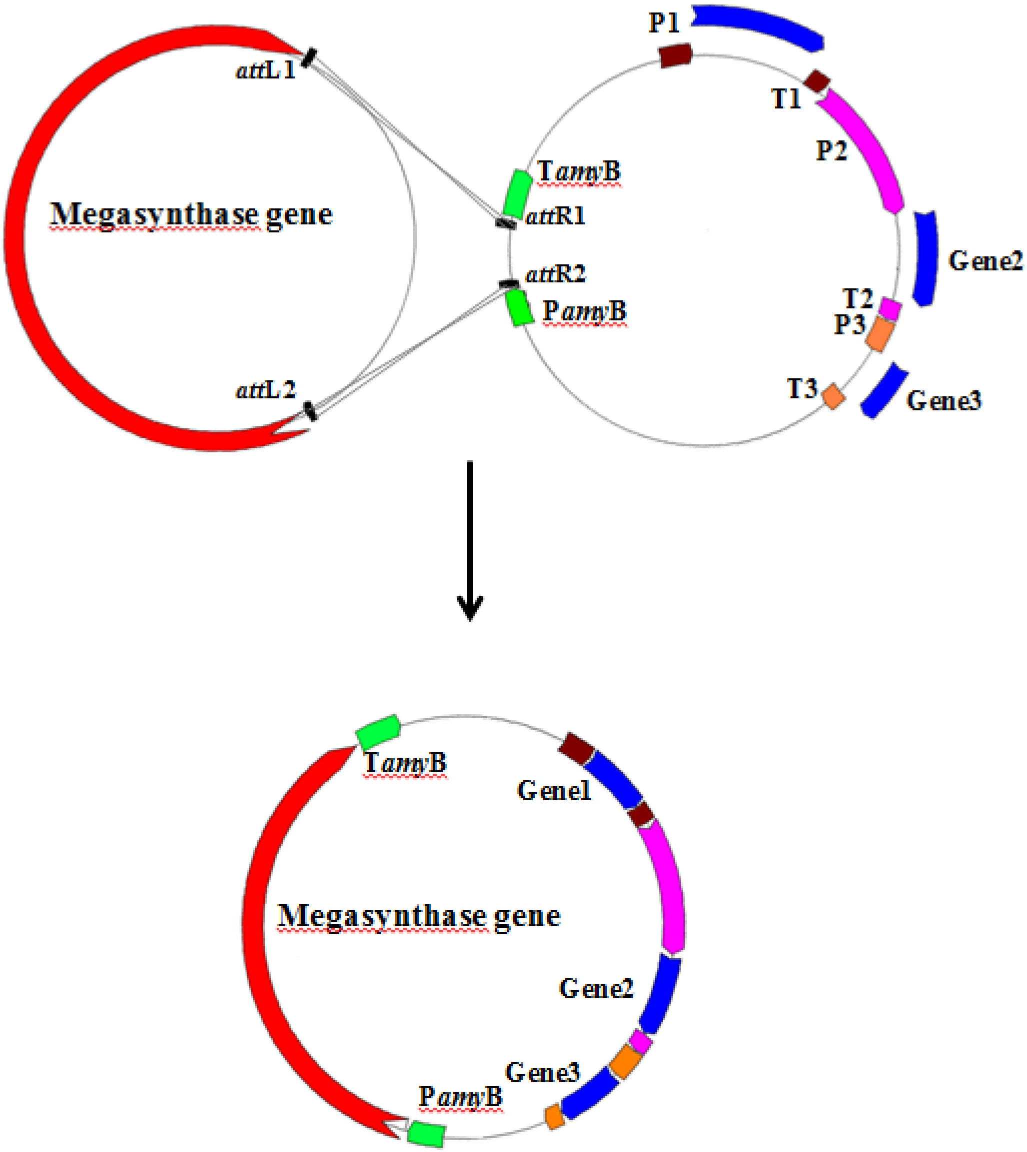


Figure 7 Multigene expression vector containing 4 expression cassettes with 4 promoters i.e. PamyB, P*adh* (P1), P*gpd*A (P2) and P*eno* (P3). The megasynthase gene can be cloned into the vector using Gateway® LR reaction and auxillary genes can be cloned into each expression cassette by homologous recombination technique.

3. Transformation of *Aspergillus oryzae* and analysis of transformants

The expression vector containing biosynthetic genes were introduced into *A. oryzae* NSAR1 as an expression host by protoplast-mediated transformation [22]. The selected transformants were subcultured serially on appropriately supplemented Czapek-Dox plates until genetically pure. Expression of megasynthase gene fused with reporter gene was induced by culturing in CMP medium containing 3.5% Czapek-Dox, 2% maltose and 1% peptone for metabolite production and monitored for expression by fluorescence or confocal microscopy. Seven-day-old CMP cultures were homogenized using a hand blender and the pH adjusted to 3. Ethyl acetate was added to the blended culture and stirred for 10 min. The cell debris was removed by vacuum filtration through Whatman No.1 filter paper. The filtrate was shaken vigorously in a separating funnel and the ethyl acetate layer was collected. Extraction was repeated twice, and combined extracts were dried with anhydrous Na_2SO_4. The ethyl acetate was then removed under reduced pressure rotary evaporation and the solid residue dissolved in HPLC grade methanol. HPLC analysis of metabolites was performed on an Agilent Technologies 1200 system using a C-18 Kaseisorb LC ODS 2000 column (4.6 mm x 150 mm); elution was with a $CH_3CN:H_2O$ gradient and peaks were detected with UV at 254, 280, 315 and 450 nm for metabolites produced by transformants. For big scale production, the transformants were cultured in 1 L of medium and extracts were fractionated on a semi-preparative C-18 Kaseisorb LC ODS 2000 column (10.0 mm x 250 mm) and the structures of isolated metabolites were determined by NMR and HRMS.

Results and discussion

1. Reconstruction of the cryptic biosynthetic gene clusters found in *M. theobromae* BCC 4162

Both tenellin-like gene cluster and HR-PKS were successfully amplified from genomic DNA by PCR and constructed into the expression vectors as shown in Figure 8.

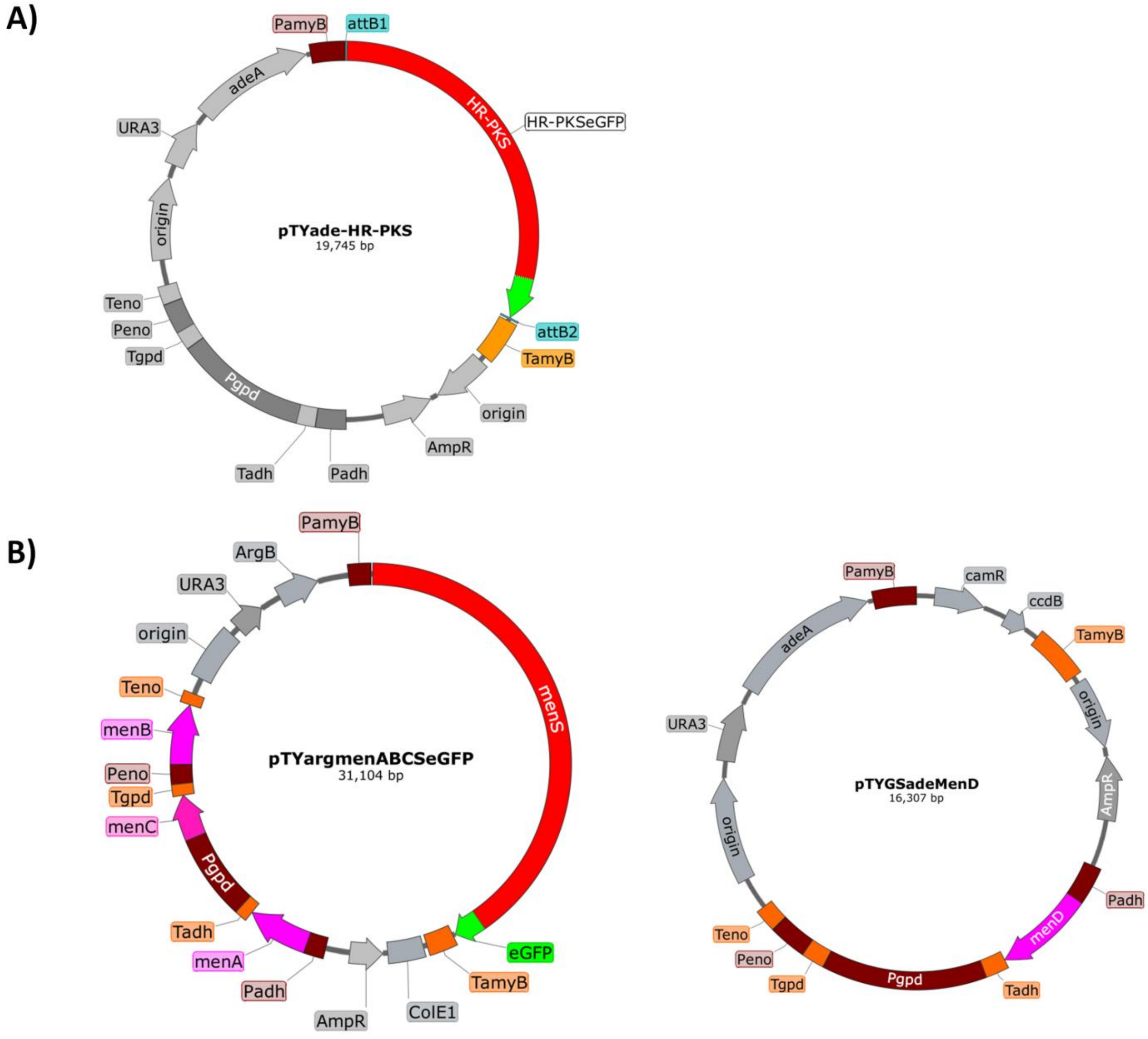


Figure 8 Plasmid maps of expression vectors containing A) HR-PKS and B) tenellin-like gene cluster consisting of *menS*, *menA*, *menB*, *menC* and *menD*.

For tenellin-like gene cluster, expression vectors containing PKS-NRPS with auxiliary genes for each step during the biosynthesis were also constructed. Both PKS-NRPS gene from tenellin-like gene cluster and HR-PKS gene with unknown function are fused with reporter gene encoding for eGFP fluorescent protein. All expression vectors were introduced into *A. oryzae* NSAR1 for metabolite production. The transfromants containing the tenellin-like gene cluster were screened for new metabolite production however no new metabolite could be observed

(Figure 9) from the extract of transformants although green fluorescent of eGFP fused with PKS-NRPS could be detected under fluorescence microscope. This may indicate that some of the auxiliary genes could not be expressed due to the existence of introns (non-encoding nucleotide region) in the genes.

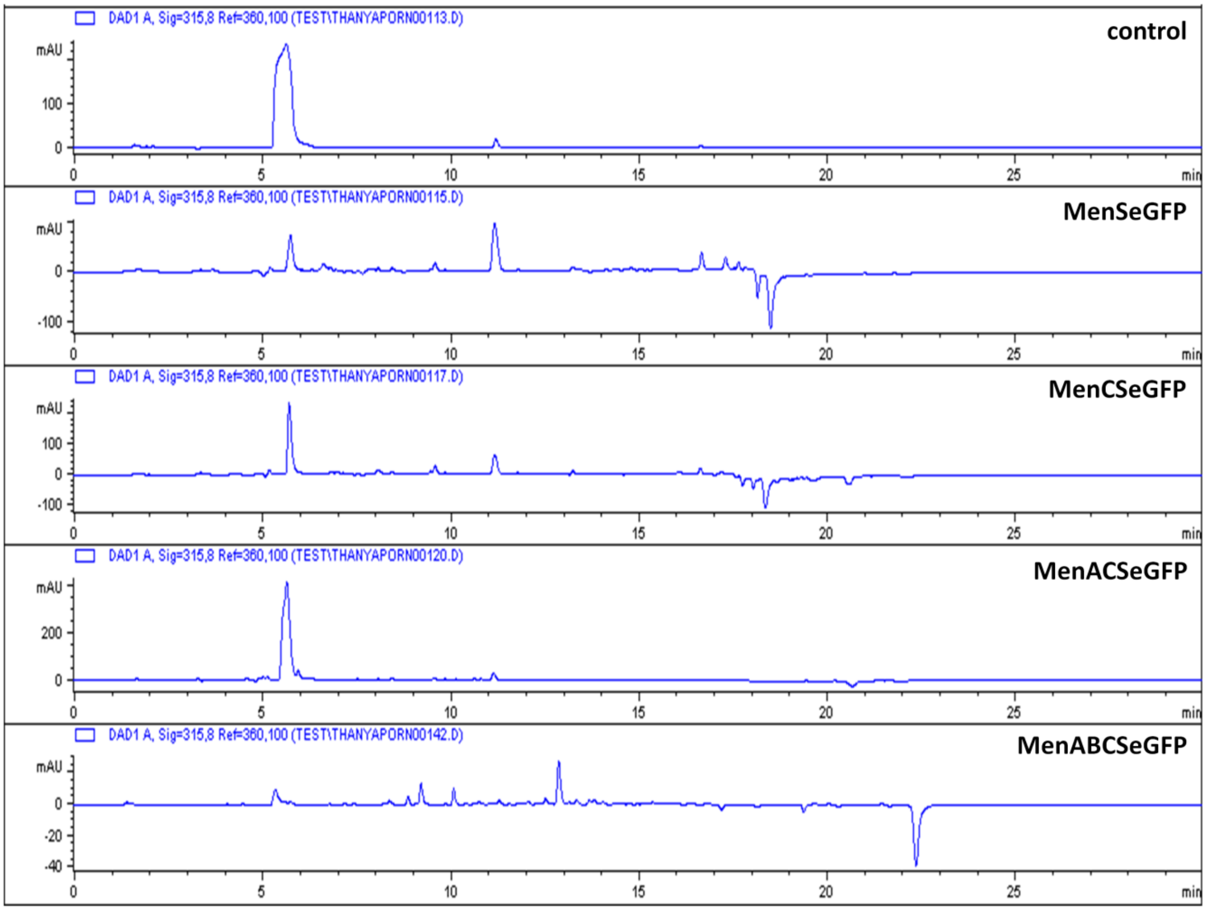


Figure 9 Comparison of HPLC chromatograms of the ethyl acetate crude extracts from *A. oryzae* transformants containing genes from tenellin-like gene cluster. Peaks were detected with UV at 315 nm.

For the transformants containing unknown HR-PKS gene, high expression level of this gene could be observed from the emission of eGFP proteins that are fused with HR-PKS. Also new metabolites could be observed comparing with negative control (transformants without

HR-PKS) (Figure 10). Unfortunately, the amounts of metabolites produced by these transformants were very low (less than 1 mg from 1 L of culture) and could not be elucidated for their chemical structures. Currently transformants are being cultured in higher volume of medium to obtain enough metabolites for structure determination and biological tests.

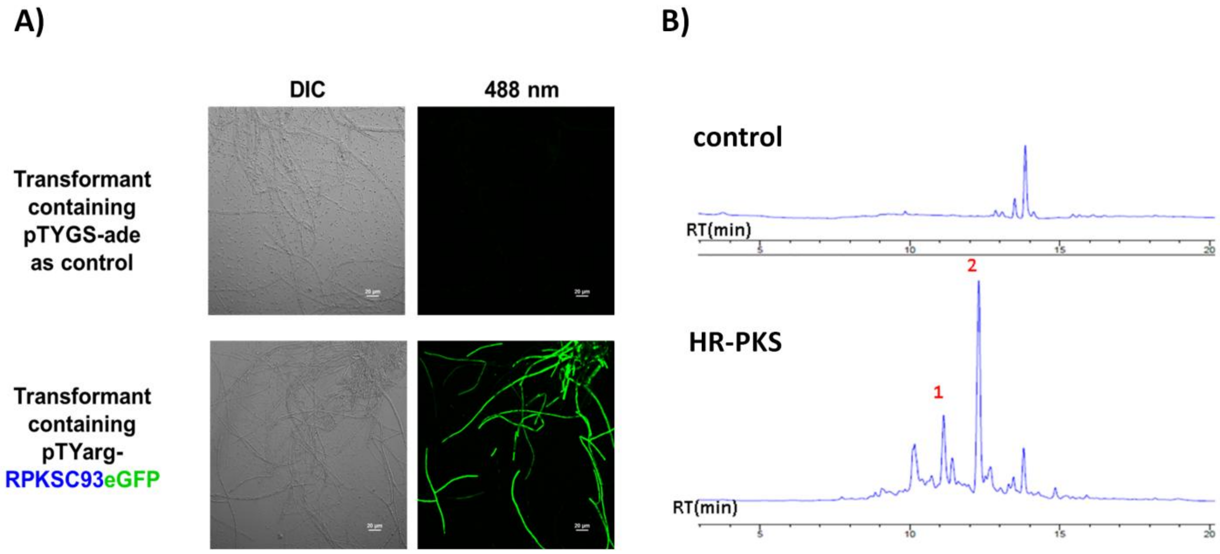


Figure 10 A) Differential interference contrast (DIC) and fluorescence imaging of *A. oryzae* transformants containing HR-PKS gene (bottom) and control (top) using wavelength at 488 nm to excite eGFP fused with HR-PKS. B) HPLC chromatograms of the ethyl acetate crude extract from *A. oryzae* transformants containing HR-PKS gene. Major metabolites are labelled as Peaks 1 and 2 were produced from the transformants containing HR-PKS gene. UV absorption of these compounds was recorded at 450 nm.

2. Intron removal of genes in tenellin-like biosynthetic gene cluster

Due to the lack of new metabolites produced by the transfromants containing tenellin-like biosynthetic gene cluster, this may be a result of the existence of introns in the genes. To overcome this problem, introns in each gene were removed by PCR using primers containing overlapping sequences without intron regions. These primers were designed based on prediction by comparing with nucleotide sequences of similar gene found in tenellin biosynthesis. The expression vectors containing genes without introns (Figure 11) were constructed using homologous recombination in yeast and Gateway$^{\circledR}$ cloning as described previously.

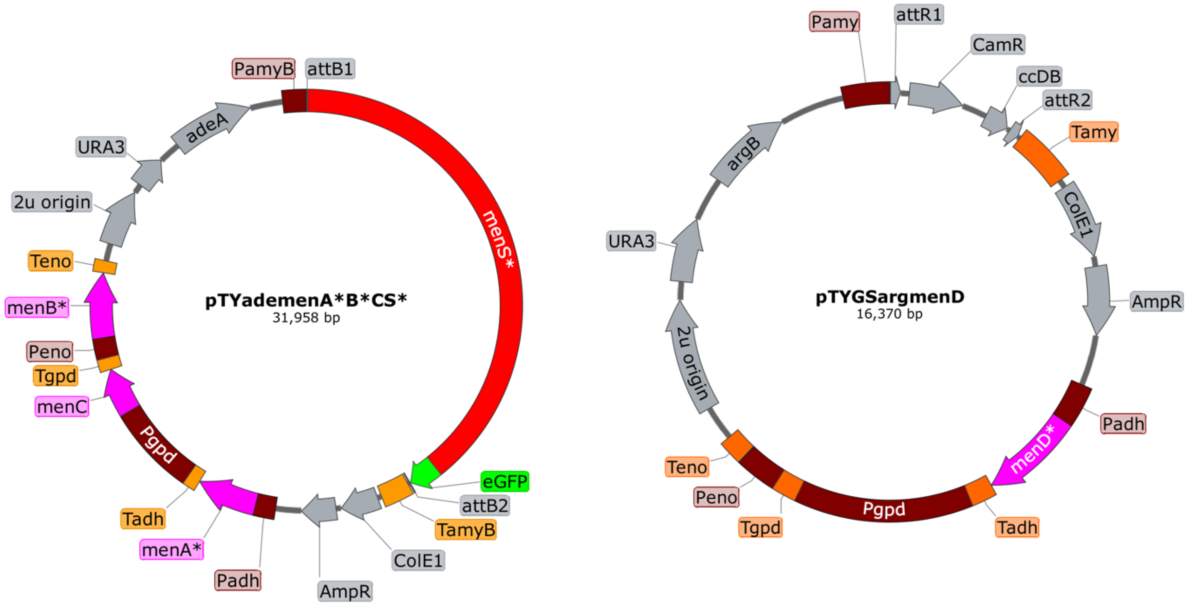


Figure 11 Maps of expression vectors containing genes without introns from tenellin-like gene cluster.

These expression vectors were introduced into *A. oryzae* NSAR1 and the transformants containing these genes are currently being investigated for metabolite production.

3. Reconstruction of biosynthetic gene cluster involved in menisporopsin A biosynthesis

From previous work, two PKS genes were identified during the production phase of menisporopsin A which are reducing (R-PKS) and non-reducing (NR-PKS) polyketide synthase genes. The full length of NR-PKS gene was obtained by genome sequencing. However the nucleotide sequence of R-PKS still remains uncompleted. Nevertheless, the NR-PKS was cloned into expression vector and introduced into *A. oryzae* NSAR1 for metabolite production but no new metabolites could be observed. Subsequently, transcriptome sequencing was performed during menisporopsin A production phase and two highly-expressed candidate PKS transcripts, highly reducing and non-reducing PKSs, were identified and designated as *men1* (highly reducing PKS gene) (contig: CL2915) and *men2* (non-reducing PKS gene) (contig: unigene363) as shown in Table 2.

Contigs	FPKM value during production phase	FPKM value before production phase
Unigene363_All	51.63	21.38
CL2915.Contig4_All	44.55	27.25
CL2915.Contig7_All	42.15	28.3
Unigene6408_All	18.48	14.2
CL4699.Contig3_All	18.32	8.06
CL4817.Contig1_All	14.63	4.4
Unigene80_All	5.1	2.27
Unigene5941_All	5.1	0.59
CL4817.Contig2_All	4.13	2.61

Table 2 Expression level of some PKS genes before and during the production phase of menisporopsin A.

Both *men1* and *men2* were cloned into the expression vectors, pTYGSmet and pTYGSarg, respectively. The resulting expression vectors are named pTYmetMen1dsRED and pTYargMen2eGFP (Figure 12).

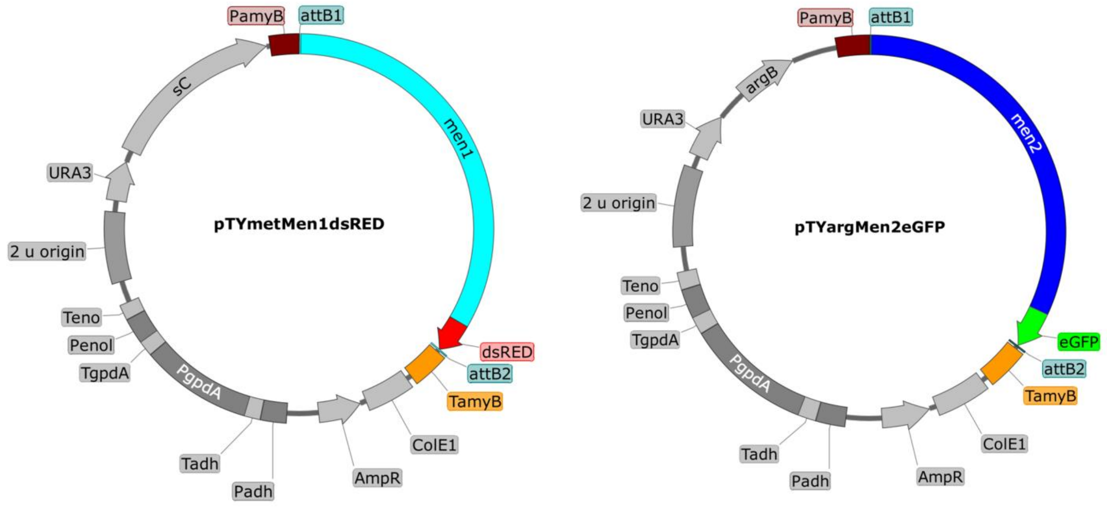


Figure 12 Plasmid maps of expression vectors containing *men1* and *men2.*

Each functional domain of Men1 and Men2 were identified by SMART [23] and the core domain of Men1 was predicted by sequence alignment with other reducing PKSs from resorcylic acid lactone biosynthesis (Figure 13).

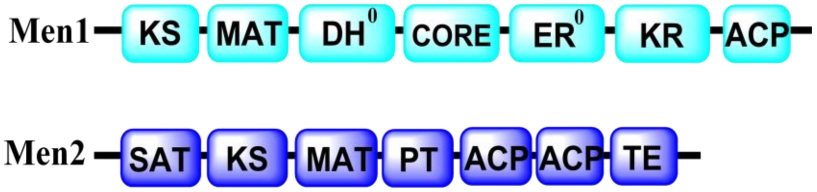


Figure 13 Domain boundaries of Men1 and Men2

The predicted domains of Men2 are in agreement with other NR-PKSs, however Men1 containing a complete set of reductive domains i.e. ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER). Therefore Men1 is classified as a highly reducing (HR) PKS. Based on structure of menisporopsin A, it indicates that only reduction of β-carbonyl intermediates can occur during the biosynthesis and therefore both DH and ER domains of Men1 are inactive.

From the sequence analysis, the active site of DH domains is typically the His/Asp catalytic dyad whilst ER domains contain GGVG motif for NADPH binding. In Men1, the catalytic residues of DH are equivalent to His/Glu instead of His/Asp (Figure 14). The acidic residue Glu of the DH catalytic dyad is usually found in type II FAS, FabZ. For ER domain, the NADPH binding motif is replaced with AVVS (Figure 15). It is quite clear that ER of Men1 lacks NADPH binding site therefore it is inactive.

Although, the catalytic dyad is found in DH domain but the motif and residue possible involved in interaction with ACP are not observed as can be seen in DH from modular type I PKS. The LPFXW motif (Figure 14) and Arg residue locating near the active site are possibly involved in the interaction with ACP [24]. In non-functional DH of modular PKS, this motif and Arg are not conserved. This may indicate that the dehydration cannot occur due to the lack of interaction with ACP domain. For Men2, it consists of typical domains found in most NR-PKSs i.e starter-unit acyl carrier protein transacylase (SAT) for starter unit selection, ketosynthase (KS) for decarboxylative condensation, acyltransferse (AT) for loading extender units, acyl carrier protein (ACP) for shuffling intermediates during the biosynthesis, product template (PT) for aromatization and thioesterase for catalyzing hydrolysis or cyclization of final products. However, ACP found in Men2 is a tandem doublet, this is not uncommon as can be seen in other fungal NR-PKSs [25-27].

Both pTYmetMen1dsRED and pTYargMen2eGFP were introduced into *A. oryzae* NSAR1 as heterologous host and the transformants were analyzed by HPLC. The results show that three major metabolites were produced by the transfromants as shown in Figure 16.

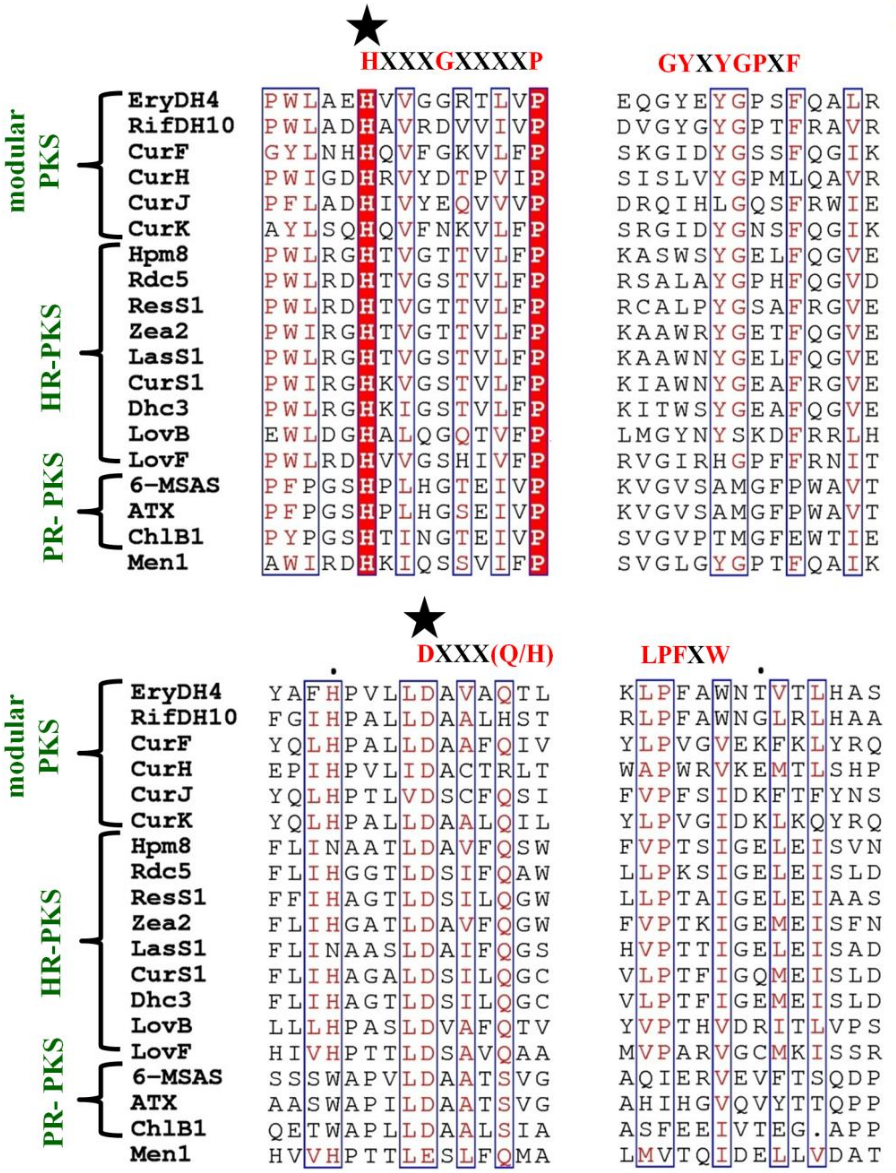


Figure 14 Sequence alignment of several conserved residues found in DH domains of partially reducing (PR), highly reducing (HR) and modular PKSs. Residues marked with star are the catalytic dyad of DH domains.

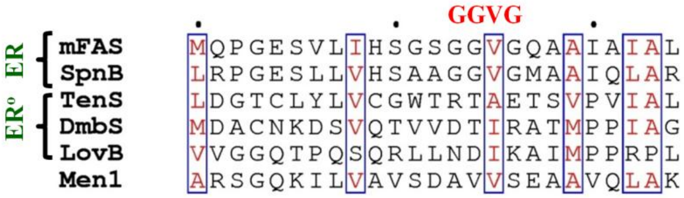


Figure 15 Sequence alignment of/around the conserved NADPH-binding site (GGVG) of functional and inactive ER domains. This motif in Men1 is replaced with AVVS.

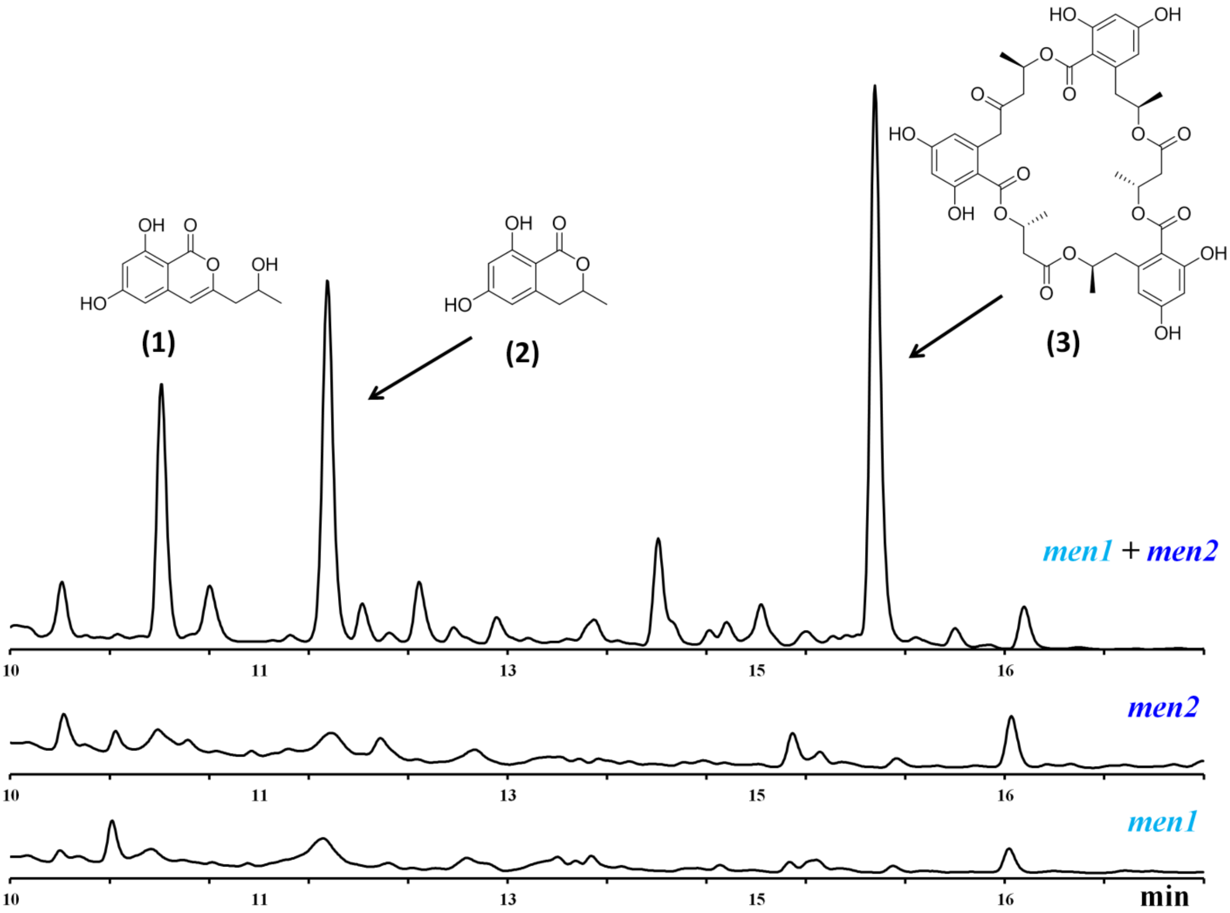


Figure 16 HPLC chromatograms of *A. oryzae* transformants containing *men1* and *men2* (top), *men2* (middle) and *men1* (bottom). Peaks were detected with UV at 254 nm.

The metabolites were then isolated using HPLC equipped with semi-preparative C-18 column to yield three compounds which were identified by NMR, HRMS and their optical activities as the isocoumarin (-)-orthosporin **(1)**, the dihydroisocoumarin, (-)-6-hydroxymellein **(2)** and ascotrichalactone A **(3)** as shown in Figure 16.

To confirm the expression of both *men1* and *men2*, fluorescence imaging of transformants containing both genes were also performed (Figure 17) and the results show that both green and red fluorescence could be observed as a result of fusion protein between Men1 and Men2 with dsRED and eGFP proteins.

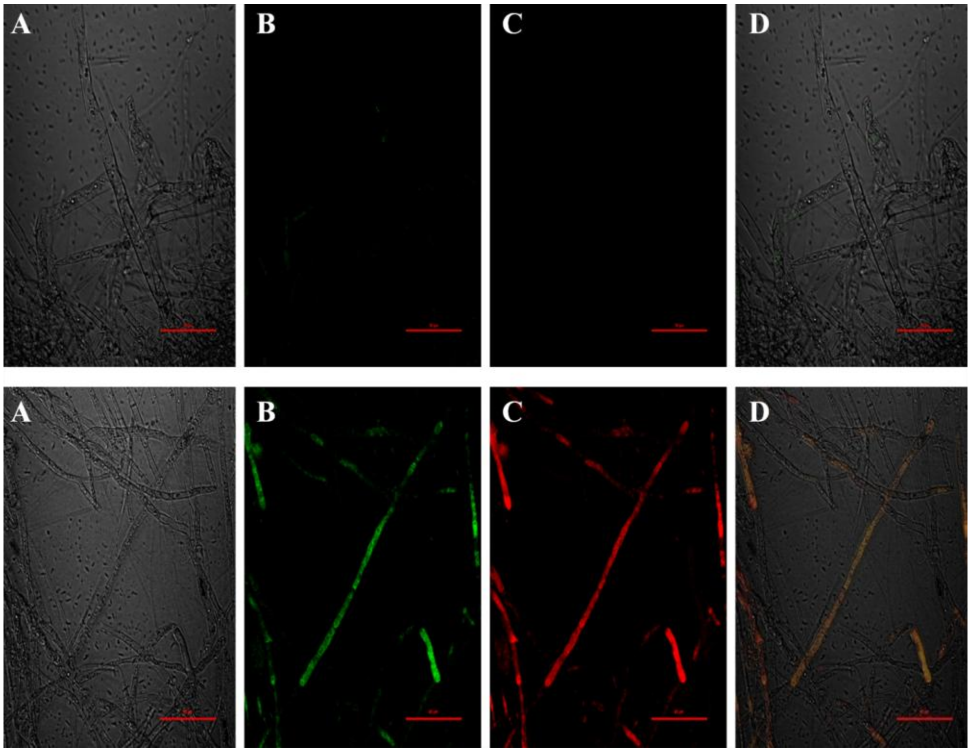
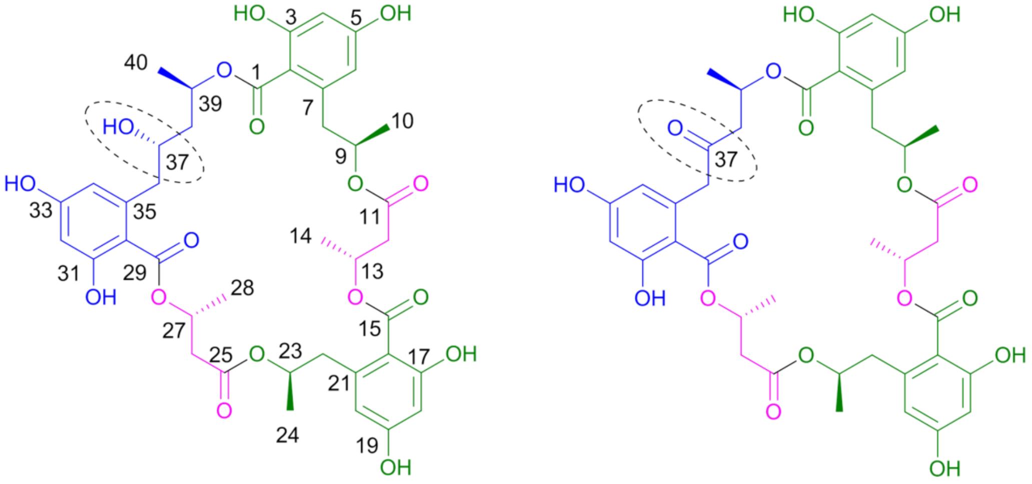


Figure 17 A) Differential interference contrast (DIC) and fluorescence imaging of *A. oryzae* transformants using wavelengths at B) 488, C) 561 and D) 488 and 561 nm to excite eGFP, DsRed and both eGFP and DsRed, respectively. Top panels (scale bars = 50 μm): double transformant with pTYGSmet and pTYGSarg; Bottom panels (scale bars = 60 μm): double transformant with pTYmetMen1dsRED and pTYargMen2eGFP.

Both (-)-orthosporin **(1)** and (-)-6-hydroxymellein **(2)** are shunt products formed by cyclization of the aromatic subunits of menisporopsin A and were previously found in the fungus *Ascotricha* sp. Zj-M-5 [28]. This fungus also produces a variety of linear and macrocyclic polyesters including ascotrichalactone A **(3)**. This compound differs from menisporopsin A only in the presence of a keto group at C-37 (Figure 18).



menisporopsin A

ascotrichalactone A

Figure 18 Structures of menisporopsin A comparing to ascotrichalactone A

This indicates that the keto group at C-37 of ascotrichalactone A cannot be recognized by the internal KR domain in Men2. This suggests that a *trans*-acting KR is required to reduce the keto group at C-37 of ascotrichalactone A to complete the menisporopsin A biosynthesis. There is a large number of *trans*-acting KRs required for natural product biosynthesis such as in the biosyntheses of SIA7248 from *Streptomyces* sp. A7248 [29], jawsamycin from *Streptoverticillium fervens* HP-891 [30], antimycin from *Streptomyces albus* J1074 [31] and neoantimycin from *Streptomyces conglobatus* [32]. Interestingly, all of them are found in bacteria and are biosynthesized by modular type I PKSs or NRPSs.

In this work, our results show that only reducing and non-reducing PKSs are required for the biosynthesis of a macrocyclic polylactone. This is not in agreement with our previous suggestion that an additional NRPS-like enzyme could involve in the formation of multiple esterification and cyclolactonization. These activities could be catalysed by the TE domain of Men2, which is similar to *E. coli* EntF involved in the formation of the cyclotrimerizing lactone, enterobactin (Figure 19). This type of reaction is also found in macrodiolide biosynthesis, as in the case of elaiophylin from *Streptomyces violaceusniger* DSM 4137 [33] and conglobatin from *Streptomyces conglobatus* (Figure 20) [34] but not in fungi.

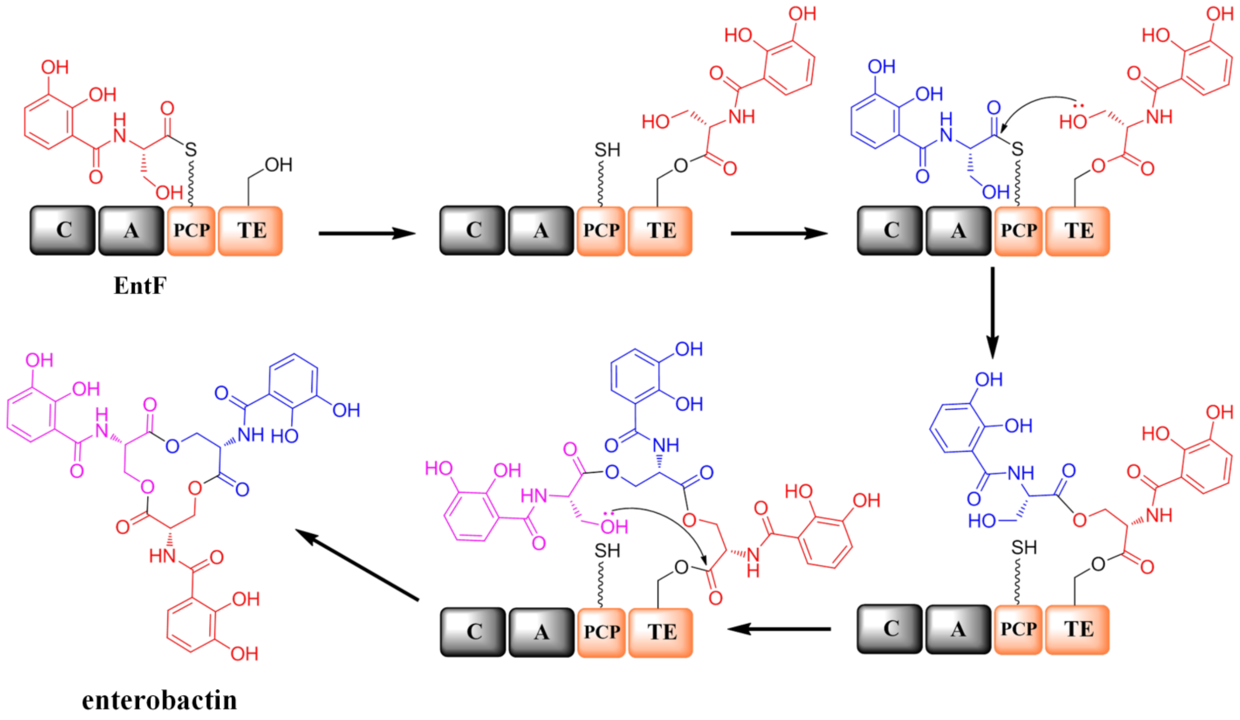


Figure 19 Proposed mechanism for the cyclolactonization of enterobactin from *E. coli* catalysed by NRPS, EntF.

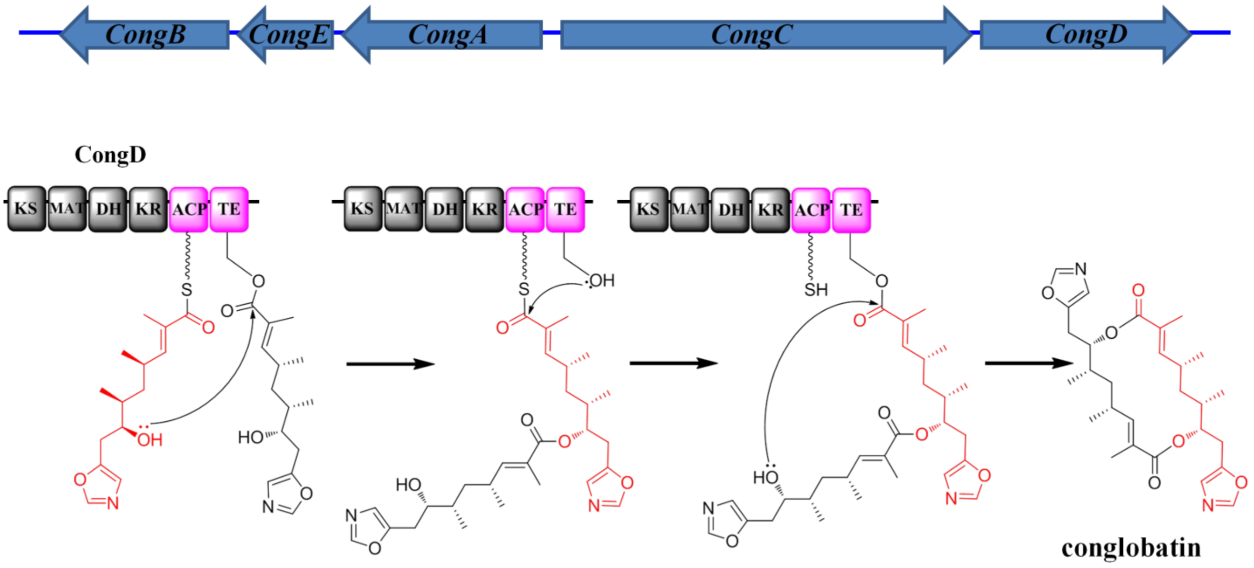


Figure 20 Proposed mechanism for the macrodiolide formation of conglobatin catalysed by CongD from *S. conglobatus* with its biosynthetic gene cluster showing on the top panel.

From our results, we are tempting to propose a scheme for the biosynthesis of menisporopsins A and B based on macrodiolide formation in congoblatin biosynthesis as shown in Figure 21.

Summary

In this work, several natural product biosynthetic gene clusters were constructed such as tenellin-like gene cluster with and without introns and HR-PKS with unknown function. Currently metabolites produced by these gene clusters are being investigated. For the biosynthesis of menisporopsin A, our results show that only two polyketide synthases require for the production of ascotrichalactone A, structural derivative of menisporopsin A. The *trans*-acting ketoreductase is proposed to be an additional enzyme for the menisporopsin A biosynthesis. Our results provide a platform for future study on the mechanism of multiple esterification and cyclolactonization of menisporopsin A. This will give us more insights into the biosynthesis of macrocyclic polylactone including structural modification to enhance their biological activities.

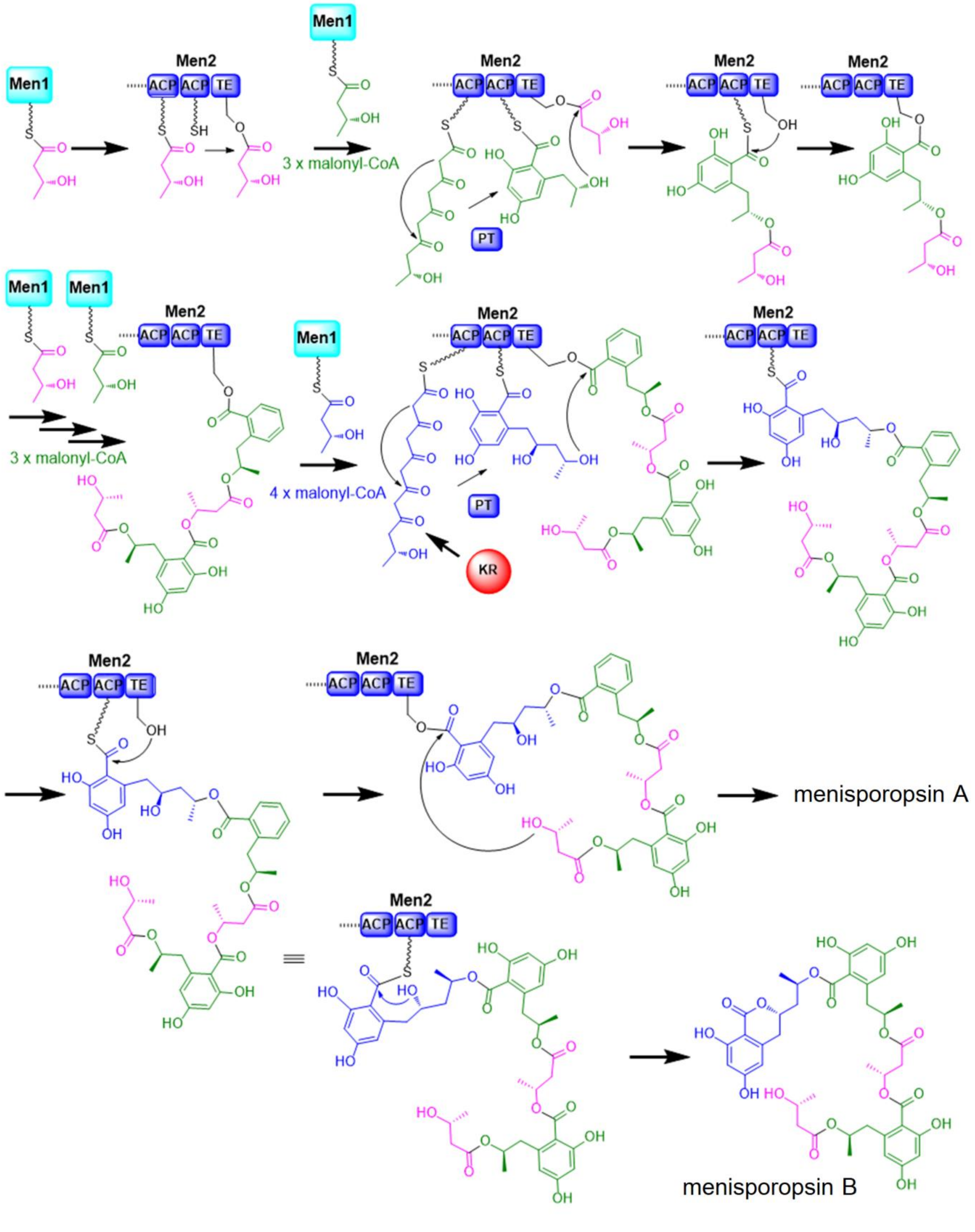


Figure 21 Proposed biosynthesis of menisporopsins and B

Outputs

1. A part of results from the heterologous expression of genes involved in menisporopsin A biosynthesis has been published in Organic and Biomolecular Chemistry as described below. This work was also presented as oral presentation in PERCH-CIC Congress X: 2018 International Congress for Innovation in Chemistry at Jomtien Palm Beach Resort Pattaya during 4-7 July 2018.

Bunnak, W., Wonnapinij, P., Sriboonlert, A., Lazarus, C.M., Wattana-Amorn, P. **(2019)** Heterologous biosynthesis of a fungal macrocyclic polylactone requires only two iterative polyketide synthases. *Org. Biomol. Chem.* 17, 374-379.

2. A part of results from tenellin-like gene cluster was presented as a poster at the International Conference of Agriculture and Natural Resources during 26-28 April 2018 at Hotel Winsor Suites and Convention, Bangkok, Thailand.

3. Laboratory workshop on fungal natural product molecular genetics

The laboratory workshop on fungal natural product molecular genetics given by Dr. Lazarus was held at Faculty of Science, Kasetsart University on 9-13 Jan 2017. There are 14 participants attending this workshop from different departments and research institutes such as Department of Genetics and Department of Biochemistry, Faculty of Science, Kasetsart University and Chulabhorn Research Institute. Students from my research group were also participated in this workshop. During the workshop, Dr. Lazarus gave the background on the system he created for biosynthetic gene cluster reconstruction. This allowed the participants to understand the principle of approaches used in the workshop. As we are currently using the same methods to reconstruct our biosynthetic gene clusters so this workshop gave me and my students an opportunity to discuss with Dr. Lazarus about the problems we experienced in our work. Suggestions and advice from Dr. Lazarus during the workshop are invaluable and allow us to solve technical problems we previously had in our work.

As the workshop we organised is the first practical workshop in Thailand using molecular biology tools to reconstruct a fungal biosynthetic gene cluster in a fungal heterologous host. This led to several discussions between Thai researchers and Dr. Lazarus. We also invited a research team (Professor Chris Willis, Dr. Paul Race and Dr. Andy Bailey) from University of Bristol who visited Thailand during our workshop to give research talks on natural products. This gave a chance for Thai natural product researchers not only from Kasetsart University but also from other research institutes (Chulabhorn Research Institute and National Center for Genetic Engineering and Biotechnology) to discuss natural product research with a team from University of Bristol. This could possibly lead to the establishment of new links between researchers in Thailand and UK.

Figure 22 Activities during laboratory workshop on fungal natural product molecular genetics on 9-13 Jan 2017 at Faculty of Science, Kasetsart University.

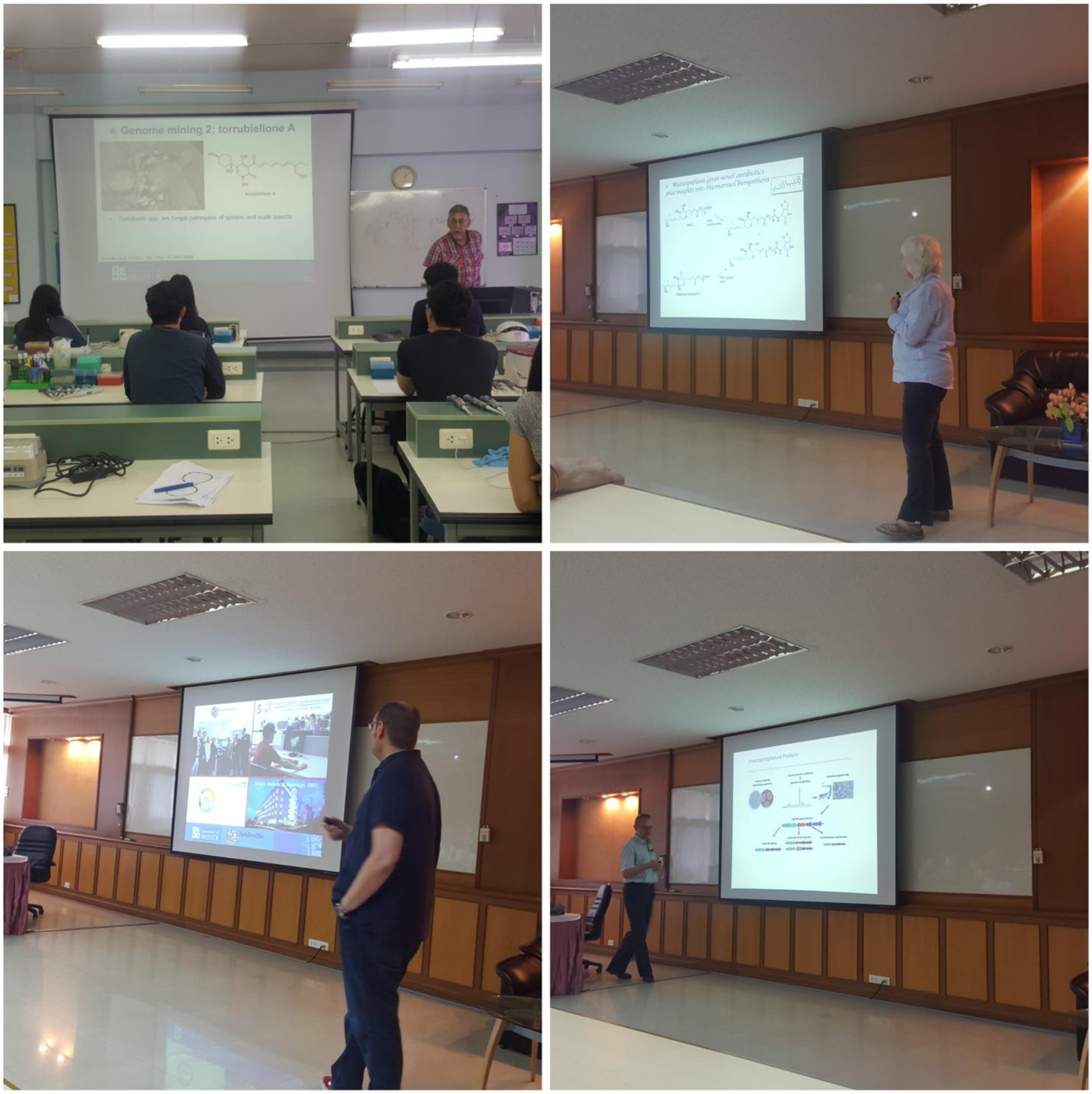


Figure 23 Research talks on natural products during the workshop given by Dr. Colin Lazarus, Professor Chris Willis, Dr. Paul Race and Dr. Andy Bailey from University of Bristol.

4. Symposium on Microbial Natural Products: Discovery of Novel Bioactive Compounds and Biosynthetic Pathways

This symposium was held at the Faculty of Science, Kasetsart University on 27 November 2018. Total participants attending this symposium were 31 people from different universities and research institutes. This was a joint symposium with JSPS-NRCT programme between Kitasato University and Kasetsart University. Four invited speakers who are the experts in the bacterial natural product area are from Kitasato Univeristy and Kasetsart University. Dr. Colin Lazarus, co-researcher for this project, and I also gave presentations relating to this work.

Figure 24 Activities during symposium on 28 Nov 2018 at Faculty of Science, Kasetsart University.

5. Research visit

Two PhD students from my research group supported by this funding visited Dr. Colin Lazarus's lab at the School of Biological Sciences, University of Bristol. These research visits were very useful for students as they learnt several techniques for fungal transformation such as CRISPR-cas9, biolistic and agrobacterium-mediated fungal transformation techniques. I also visited Dr. Lazarus's lab for one month to discuss about current research and also reconstructed an expression vector to use for activation of metabolite production.

6. A MSc student from my research group who worked in this project has recently graduated. Her thesis title is "Reconstruction of a predicted gene cluster from *Menisporopsis theobromae* BCC 4162 involved in the biosynthesis of 2-pyridone derivatives".

References

1. Newman, D. J. & Cragg, G. M. (2012) Natural products as sources of new drugs over the 30 years from 1981 to 2010, *J Nat Prod.* **75**, 311-35.
2. (2014) Antimicrobial resistance: Global Report on Surveillance, *World Health Organization*.
3. Koehn, F. E. & Carter, G. T. (2005) The evolving role of natural products in drug discovery, *Nat Rev Drug Discov.* **4**, 206-20.
4. Chinworrungsee, M., Kittakoop, P., Isaka, M., Maithip, P., Supothina, S. & Thebtaranonth, Y. (2004) Isolation and structure elucidation of a novel antimalarial macrocyclic polylactone, menisporopsin A, from the fungus *Menisporopsis theobromae*, *J Nat Prod.* **67**, 689-92.
5. Ruanglek, V., Chokpaiboon, S., Rattanaphan, N., Madla, S., Auncharoen, P., Bunyapaiboonsri, T. & Isaka, M. (2007) Menisporopsin B, a new polyester from the seed fungus *Menisporopsis theobromae* BCC 4162, *J Antibiot (Tokyo).* **60**, 748-51.
6. Wattana-amorn, P., Juthaphan, P., Sirikamonsil, M., Sriboonlert, A., Simpson, T. J. & Kongkathip, N. (2013) Biosynthetic origins of menisporopsin A, *J Nat Prod.* **76**, 1235-7.
7. Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., Lee, S. Y., Fischbach, M. A., Muller, R., Wohlleben, W., Breitling, R., Takano, E. & Medema, M. H. (2015) antiSMASH 3.0-a

comprehensive resource for the genome mining of biosynthetic gene clusters, *Nucleic Acids Res.* **43**, W237-43.

8. 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**, 289-97.

9. Li, Y., Chooi, Y. H., Sheng, Y., Valentine, J. S. & Tang, Y. (2011) Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an alpha-hydroxylation-dependent Claisen-like cyclization catalyzed by a dimanganese thioesterase, *J Am Chem Soc.* **133**, 15773-85.

10. Chiang, Y. M., Szewczyk, E., Davidson, A. D., Keller, N., Oakley, B. R. & Wang, C. C. (2009) A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*, *J Am Chem Soc.* **131**, 2965-70.

11. Xu, Y., Espinosa-Artiles, P., Schubert, V., Xu, Y. M., Zhang, W., Lin, M., Gunatilaka, A. A., Sussmuth, R. & Molnar, I. (2013) Characterization of the biosynthetic genes for 10,11-dehydrocurvularin, a heat shock response-modulating anticancer fungal polyketide from *Aspergillus terreus*, *Appl Environ Microbiol.* **79**, 2038-47.

12. Jirakkakul, J., Cheevadhanarak, S., Punya, J., Chutrakul, C., Senachak, J., Buajarern, T., Tanticharoen, M. & Amnuaykanjanasin, A. (2015) Tenellin acts as an iron chelator to prevent iron-generated reactive oxygen species toxicity in the entomopathogenic fungus *Beauveria bassiana*, *FEMS Microbiol Lett.* **362**, 1-8.

13. Isaka, M., Chinthanom, P., Supothina, S., Tobwor, P. & Hywel-Jones, N. L. (2010) Pyridone and tetramic acid alkaloids from the spider pathogenic fungus *Torrubiella* sp. BCC 2165, *J Nat Prod.* **73**, 2057-60.

14. Isaka, M., Tanticharoen, M., Kongsaeree, P. & Thebtaranonth, Y. (2001) Structures of cordypyridones A-D, antimalarial N-hydroxy- and N-methoxy-2-pyridones from the insect pathogenic fungus *Cordyceps nipponica*, *J Org Chem.* **66**, 4803-8.

15. Yamada, T., Iritani, M., Doi, M., Minoura, K., Ito, T. & Numata, A. (2001) Absolute stereostructures of cell-adhesion inhibitors, macrosphelides C, E-G and I, produced by a Periconia species separated from an Aplysia sea hare, *J Chem Soc., Perkin Trans. 1*, 3046-3053.

16. Ito, M., Tsuchida, Y., Mizoue, K. & Hanada, K. (1992) NG-011 and NG-012, novel potentiators of nerve growth factor. II. The structure determination of NG-011 and NG-012, *J Antibiot (Tokyo).* **45**, 1566-72.

17. Schlingmann, G., Milne, L. & Carter, G. T. (2002) Isolation and identification of antifungal polyesters from the marine fungus *Hypoxylon oceanicum* LL-15G256, *Tetrahedron.* **58**, 6825-6835.

18. Madla, S., Kittakoop, P. & Wongsa, P. (2006) Optimization of culture conditions for production of antimalarial menisporopsin A by the seed fungus *Menisporopsis theobromae* BCC 4162, *Lett Appl Microbiol.* **43**, 548-53.

19. Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. & Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome, *Nat Biotechnol.* **29**, 644-52.

20. Gietz, R. D. & Woods, R. A. (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method, *Methods Enzymol.* **350**, 87-96.

21. Pahirulzaman, K. A., Williams, K. & Lazarus, C. M. (2012) A toolkit for heterologous expression of metabolic pathways in *Aspergillus oryzae*, *Methods Enzymol.* **517:241-60**.

22. Heneghan, M. N., Yakasai, A. A., Halo, L. M., Song, Z., Bailey, A. M., Simpson, T. J., Cox, R. J. & Lazarus, C. M. (2010) First heterologous reconstruction of a complete functional fungal biosynthetic multigene cluster, *Chembiochem.* **11**, 1508-12.

23. Letunic, I. & Bork, P. (2018) 20 years of the SMART protein domain annotation resource, *Nucleic Acids Res.* **46**, D493-D496.

24. Keatinge-Clay, A. (2008) Crystal structure of the erythromycin polyketide synthase dehydratase, *J Mol Biol.* **384**, 941-53.

25. Mayorga, M. E. & Timberlake, W. E. (1992) The developmentally regulated *Aspergillus nidulans* wA gene encodes a polypeptide homologous to polyketide and fatty acid synthases, *Mol Gen Genet.* **235**, 205-12.

26. Takano, Y., Kubo, Y., Shimizu, K., Mise, K., Okuno, T. & Furusawa, I. (1995) Structural analysis of PKS1, a polyketide synthase gene involved in melanin biosynthesis in *Colletotrichum lagenarium*, *Mol Gen Genet.* **249**, 162-7.

27. Yu, J. H. & Leonard, T. J. (1995) Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type I polyketide synthase, *J Bacteriol.* **177**, 4792-800.

28. Zhang, J.-L., Wang, W.-J., Xu, X.-M., Li, D.-Y., Hua, H.-M., Ma, E.-L. & Li, Z.-L. (2016) From macrocyclic to linear and further: naturally degradable polyesters from the fungus *Ascotricha* sp. ZJ-M-5, *Tetrahedron.* **72**, 4895-4901.

29. Wang, H., Zhang, H., Zou, Y., Mi, Y., Lin, S., Xie, Z., Yan, Y. & Zhang, H. (2014) Structural insight into the tetramerization of an iterative ketoreductase siam through aromatic residues in the interfaces, *PLoS One.* **9**, e97996.

30. Hiratsuka, T., Suzuki, H., Kariya, R., Seo, T., Minami, A. & Oikawa, H. (2014) Biosynthesis of the structurally unique polycyclopropanated polyketide-nucleoside hybrid jawsamycin (FR-900848), *Angew Chem Int Ed Engl.* **53**, 5423-6.

31. Liu, J., Zhu, X., Seipke, R. F. & Zhang, W. (2015) Biosynthesis of antimycins with a reconstituted 3-formamidosalicylate pharmacophore in *Escherichia coli, ACS Synth Biol.* **4**, 559-65.

32. Zhou, Y., Lin, X., Williams, S. R., Liu, L., Shen, Y., Wang, S. P., Sun, F., Xu, S., Deng, H., Leadlay, P. F. & Lin, H. W. (2018) Directed Accumulation of Anticancer Depsipeptides by Characterization of Neoantimycins Biosynthetic Pathway and an NADPH-Dependent Reductase, *ACS Chem Biol.* **13**, 2153-2160.

33. Zhou, Y., Prediger, P., Dias, L. C., Murphy, A. C. & Leadlay, P. F. (2015) Macrodiolide Formation by the Thioesterase of a Modular Polyketide Synthase, *Angew Chem Weinheim Bergstr Ger.* **127**, 5321-5324.

34. Zhou, Y., Murphy, A. C., Samborskyy, M., Prediger, P., Dias, L. C. & Leadlay, P. F. (2015) Iterative Mechanism of Macrodiolide Formation in the Anticancer Compound Conglobatin, *Chem Biol.* **22**, 745-54.

Appendix

A copy of publication resulting from this project

Organic & Biomolecular Chemistry

PAPER

Cite this: *Org. Biomol. Chem.*, 2019, **17**, 374

Heterologous biosynthesis of a fungal macrocyclic polylactone requires only two iterative polyketide synthases†

Waraporn Bunnak,[a] Passorn Wonnapinij,[b] Ajaraporn Sriboonlert,[b] Colin M. Lazarus[c] and Pakorn Wattana-Amorn *[a]

Received 7th November 2018,
Accepted 11th December 2018

DOI: 10.1039/c8ob02773k

rsc.li/obc

Menisporopsin A is a bioactive macrocyclic polylactone produced by the fungus *Menisporopsis theobromae* BCC 4162. A scheme for the biosynthesis of this compound has been proposed, in which reducing (R) and non-reducing (NR) polyketide synthases (PKSs) would catalyze the formation of each menisporopsin A subunit, while an additional non-ribosomal peptide synthetase (NRPS)-like enzyme would be required to perform multiple esterification and cyclolactonization reactions. Transcriptome analysis of *M. theobromae* identified an R-PKS gene, *men1*, and an NR-PKS gene, *men2*, which both exhibited highest expression levels during the menisporopsin A production phase. These were cloned into separate vectors for heterologous expression in *Aspergillus oryzae* NSAR1. Unexpectedly, coexpression of the two PKSs alone was sufficient to catalyze the formation of the macrocyclic polylactone, ascotrichalactone A, a structural derivative of menisporopsin A. The unanticipated esterification and cyclolactonization activities could reside in the unusual thioesterase domain of the NR-PKS, which is similar to that of the NRPS catalyzing elongation and cyclization of trilactone in enterobactin biosynthesis and that of modular PKSs catalyzing macrodiolide formation in elaiophylin and conglobatin biosyntheses.

Introduction

Different types of polyketide synthases (PKSs) produce the core structures of an important and diverse class of secondary metabolites. The macrocyclic polylactone menisporopsin A is a polyketide natural product isolated from *Menisporopsis theobromae* BCC 4162,[1] while the linear polyester menisporopsin B is expected to be an intermediate in menisporopsin A biosynthesis.[2] Both compounds exhibit a broad spectrum of biological properties such as cytotoxic and antimalarial activities.[1,2] Our previous work showed that menisporopsin A is assembled from 20 units of acetate and that its biosynthesis is likely to be catalyzed by reducing (R) and non-reducing (NR) PKSs.[3] The mechanism for multiple esterifications remained unknown. Macrocyclic polylactones are also found in several other fungi, some being composed of subunits similar to those of menisporopsin A.[4–7] Each subunit can be assembled into a variety of patterns (Fig. 1). Identification of the enzymes involved in the biosynthesis of the distinctive structure of menisporopsin A should lead to a better understanding of how fungi programme multiple esterifications and cyclolactonization in the final step of macrocyclic polylactone biosynthesis.

Results and discussion

Transcriptomic analysis of *M. theobromae* in the menisporopsin A production phase identified two highly-expressed candidate PKS transcripts, designated *men1* (CL2915) and *men2* (unigene363) (ESI Files S2† for more detailed information). The transcriptomic data have been deposited in the NCBI Sequence Read Archive (SRA) database with accession number PRJNA494652. The functional domain boundaries of each of the encoded enzymes were identified by SMART (Fig. S1 and 2†),[8] with the exception of the core domain of *men1* which was identified by sequence alignment with PKSs involved in resorcylic acid lactone biosynthesis. The predicted domains of Men2 are consistent with an expected NR-PKS nature, whereas Men1 contains a complete set of reductive domains *i.e.* keto-

---

[a] Department of Chemistry, Special Research Unit for Advanced Magnetic Resonance and Center of Excellence for Innovation in Chemistry, Faculty of Science, Kasetsart University, Bangkok, 10900, Thailand. E-mail: fscipwa@ku.ac.th

[b] Department of Genetics, Faculty of Science, Kasetsart University, Bangkok, 10900, Thailand

[c] School of Biological Sciences, University of Bristol, Bristol, BS8 1TQ, UK

† Electronic supplementary information (ESI) available: Primers used for this work, amino acid sequences of Men1 and Men2, NMR data for isolated metabolites and expression vector maps. See DOI: 10.1039/c8ob02773k

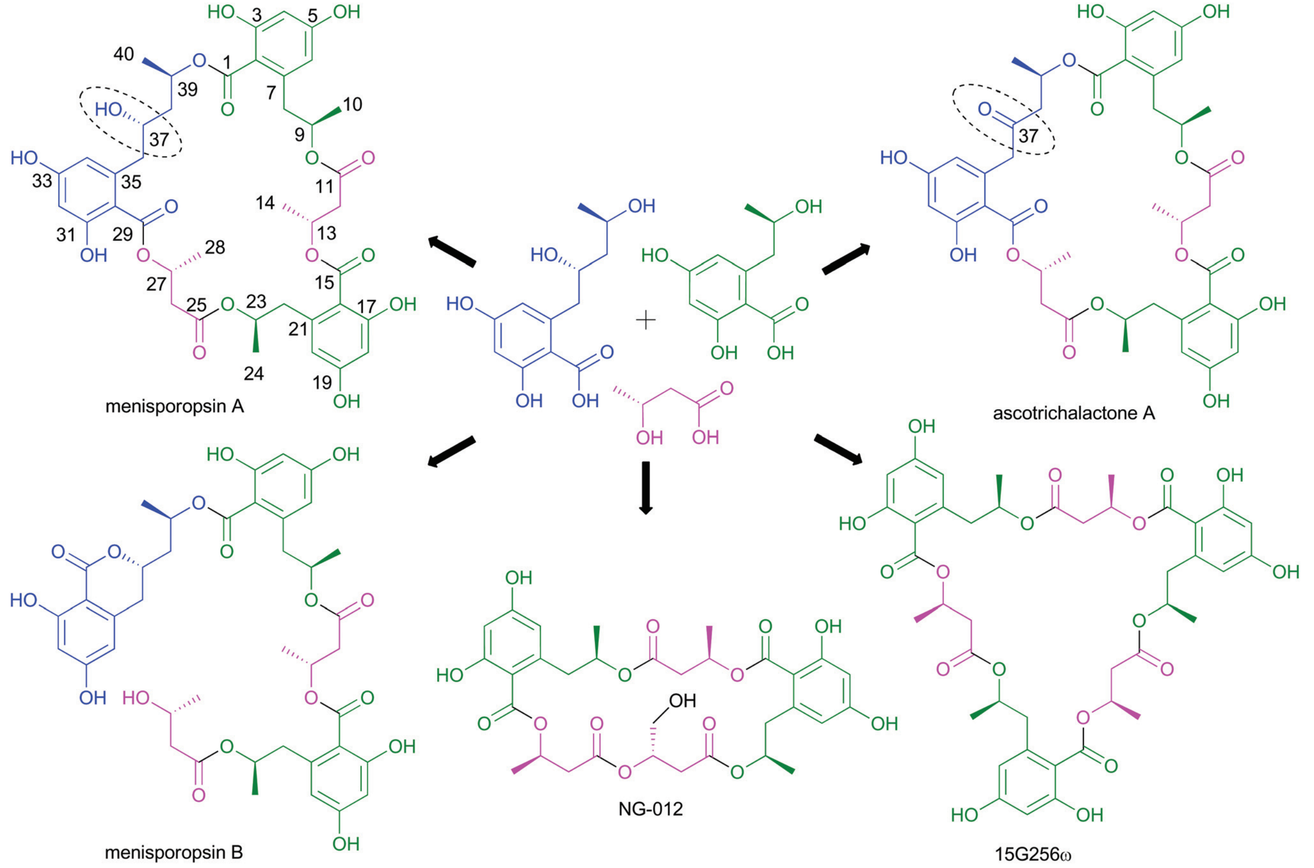


Fig. 1 Example of macrocyclic polylactones found in fungi. Some contain the same subunits but differ in the number and assembly pattern of each subunit.

reductase (KR), dehydratase (DH) and enoyl reductase (ER), and so could be classified as a highly reducing (HR) PKS. However, the structure of menisporopsin A indicates that reduction of β-carbonyl intermediates goes no further than production of hydroxyl groups, indicating that the DH and ER domains of Men1 are inactive. The active site of DH domains typically contains a His/Asp catalytic dyad whilst ER domains contain a GGVG motif for NADPH binding. In the Men1 DH domain, replacement of His/Asp with His/Glu (Fig. S3†) may not alone abolish activity since this latter dyad is present in the type II FAS, FabZ.[9] However the conserved LPFXW motif and Arg residue located near the active site are not observed. LPFXW and Arg are believed to interact with the ACP domain and are not conserved in non-functional DH domains of modular PKSs (Fig. 2).[10] Although this motif is also not highly conserved in functional DH domains of other R-PKSs, some similarities can be found among them (Fig. S3†). The ER domain of Men1 is clearly inactive, since the NADPH binding motif is replaced with AVVS (Fig. 3). Thus, Men1 has the domain structure of an HR-PKS but should function as a (par-

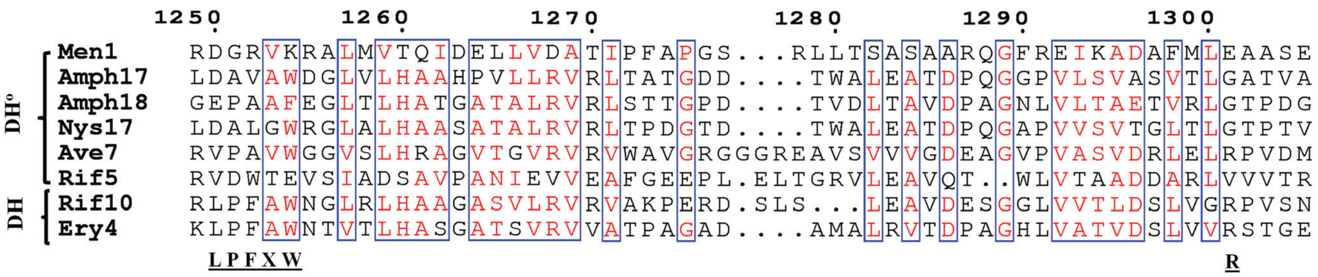

```
                 1250      1260      1270         1280      1290      1300
     Men1     RDGRVKRALMVTQIDELLVDATIPFAPGS...RLLTSASAARQGFREIKADAFMLEAASE
     Amph17   LDAVAWDGLVLHAAHPVLLRVRLTATGDD....TWALEATDPQGGPVLSVASVTLGATVA
DH°  Amph18   GEPAAFEGLTLHATGATALRVRLSTTGPD....TVDLTAVDPAGNLVLTAETVRLGTPDG
     Nys17    LDALGWRGLALHAASATALRVRLTPDGTD....TWALEATDPQGAPVVSVTGLTLGTPTV
     Ave7     RVPAVWGGVSLHRAGVTGVRVRVWAVGRGGGREAVSVVVGDEAGVPVASVDRLELRPVDM
     Rif5     RVDWTEVSIADSAVPANIEVVEAFGEEPL.ELTGRVLEAVQT..WLVTAADDARLVVVTR
DH   Rif10    RLPFAWNGLRLHAAGASVLRVRVAKPERD.SLS...LEAVDESGGLVVTLDSLVGRPVSN
     Ery4     KLPFAWNTVTLHASGATSVRVVATPAGAD....AMALRVTDPAGHLVATVDSLVVRSTGE
                  LPFXW                                                  R
```

Fig. 2 Sequence alignment of the predicted ACP interaction site (LPFXW and R) of functional and inactive DH domains of modular PKSs compared to that of Men1. Rif10 (AEK39124) and Rif5 (AAC01711) from *Amycolatopsis mediterranei* S699; Ery4 (CAM00064) from *Saccharopolyspora erythraea* NRRL 2338; Ave7 (BAA84478) from *Streptomyces avermitilis*; Nys17 (AAF71767) from *Streptomyces noursei* ATCC 11455; Amph17 (AJE44526) and Amph18 (AJE39059) from *Streptomyces nodosus*. GenPept accession numbers are given in brackets.

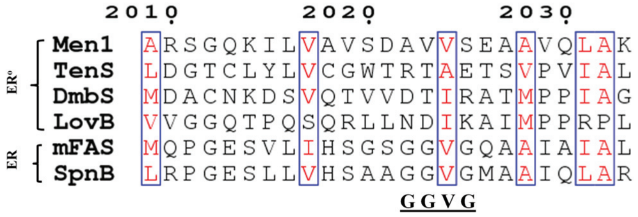


Fig. 3 Sequence alignment of/around the NADPH-binding motif (GGVG) found in functional (mFAS and SpnB) and inactive (Men1, TenS, DmbS and LovB) ER domains. TenS (A0JJU1) and DmbS (ADN43685): hybrid PKS-NRPSs from *Beauveria bassiana*; LovB (Q9Y8A5): PKS from *Aspergillus terreus*; mFAS (NP_001093400): mammalian FAS from *Sus scrofa*; SpnB (AAG23265): PKS extender module 2 from *Saccharopolyspora spinosa*. GenPept accession numbers are given in brackets.

tially-reducing) R-PKS. Men2, in contrast, has the typical complement of domains found in most NR-PKSs. This includes the presence of the ACP as a tandem doublet, which has been observed in some other fungal NR-PKSs.[11–13]

To determine whether Men1 and Men2 are involved in menisporopsin A biosynthesis, the transcript coding regions were assembled by homologous recombination in yeast, moved to expression vectors by Gateway transfer and used to transform *Aspergillus oryzae* NSAR1 (Fig. S4†).[14] This has proved to be a successful platform in several natural product biosynthetic studies including tenellin.[15] Expression of the two PKSs in transformants was confirmed by fluorescence microscopy, since Men1 was fused to DsRED and Men2 to eGFP (Fig. 4). Organic extracts of transformants expressing *men1*, *men2* or both genes were analyzed by HPLC (Fig. 5). Metabolites (compounds **1–3**) produced by transformants were isolated by HPLC and identified by NMR spectroscopy, HR-MS

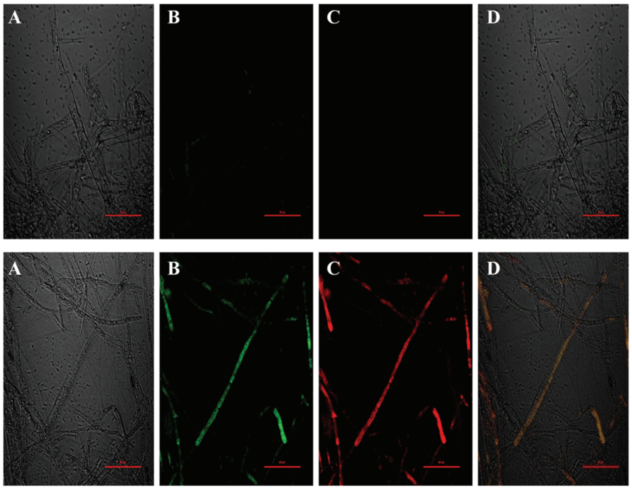


Fig. 4 (A) Differential interference contrast (DIC) and fluorescence imaging of *A. oryzae* transformants using wavelengths at (B) 488, (C) 561 and (D) 488 and 561 nm to excite eGFP, DsRed and both eGFP and DsRed, respectively. Top panels (scale bars = 50 μm): double transformant with pTYGSmet and pTYGSarg; bottom panels (scale bars = 60 μm): double transformant with pTYmetMen1dsRED and pTYargMen2eGFP.

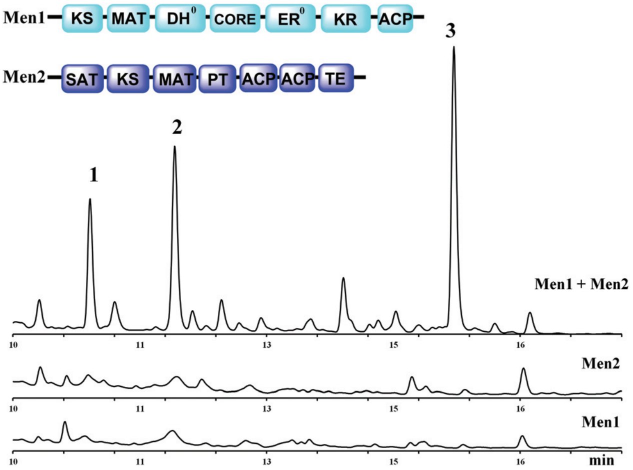
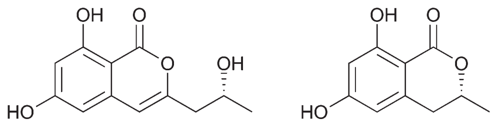


Fig. 5 HPLC chromatograms of *A. oryzae* transformants expressing Men1 (bottom), Men2 (middle) and Men1 plus Men2 (top). The domain organizations of the encoded PKSs are shown above the chromatograms.

and their optical activities. Compounds **1** and **2** were identified as the isocoumarin (−)-orthosporin and the dihydroisocoumarin (−)-6-hydroxymellein, respectively (Fig. 6).

Both compounds are cyclized products of aromatic subunits of menisporopsin A and were previously found in the fungus *Ascotricha* sp. Zj-M-5.[7] This fungus also produces a variety of linear and macrocyclic polyesters including ascotrichalactone A, which has the same structure as compound **3**. This compound differs from menisporopsin A only in the presence of a keto group at C-37 (Fig. 1). This indicates that the KR domain of Men1 cannot recognize the keto group at C-37, and suggests that a *trans*-acting KR must be required to complete menisporopsin biosynthesis.

Discrete KRs required for natural product biosynthesis are found in some bacteria, as in the biosyntheses of SIA7248 from *Streptomyces* sp. A7248,[16] jawsamycin from *Streptoverticillium fervens* HP-891,[17] antimycin from *Streptomyces albus* J1074[18] and neoantimycin from *Streptomyces conglobatus*.[19] These compounds are biosynthesized by modular type I PKSs or NRPSs that either lack a KR domain or have an inactive one. Our results show that only two PKSs are required for the biosynthesis of a macrocyclic polylactone, in contrast to our previous suggestion that an additional NRPS-like enzyme activity would be needed to catalyze cyclolactonization.[3] The cyclolactonization activity could reside in the TE domain of the NR-PKS, which is similar to *E. coli* EntF involved in the formation of the cyclotrimerizing lactone, entero-



(-)-orthosporin (**1**)

(-)-6-hydroxymellein (**2**)

Fig. 6 Structures of (−)-orthosporin and (−)-6-hydroxymellein.

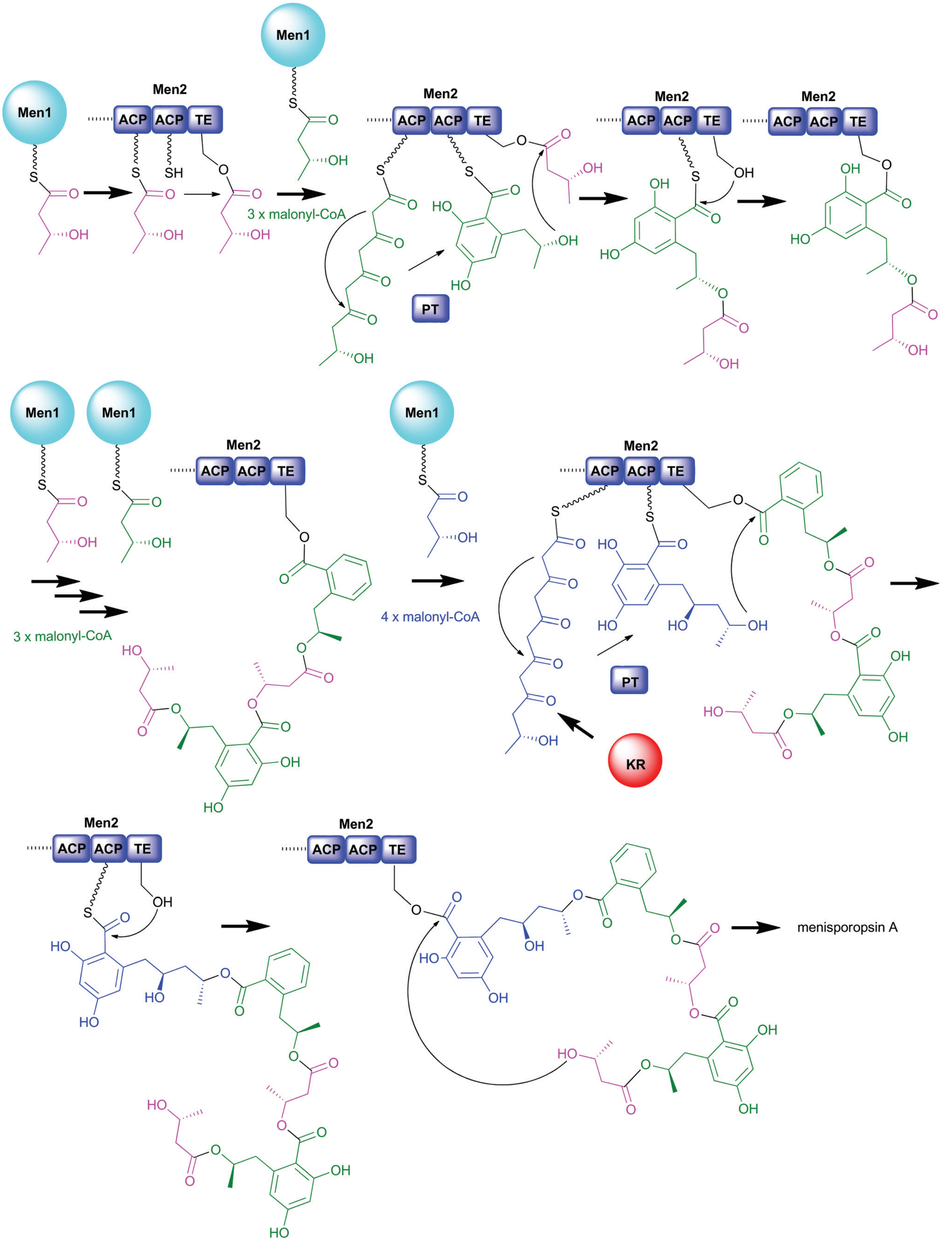


Fig. 7 Proposed biosynthesis of menisporopsin A.

bactin.[20] This type of reaction is also found in macrodiolide biosynthesis, as in the case of elaiophylin from *Streptomyces violaceusniger* DSM 4137[21] and conglobatin from *Streptomyces conglobatus*,[22] but not in fungi. Our observations do not exclude the unlikely possibility of endogenous cyclolactonization activity in the *A. oryzae* expression host, but we would rather propose a scheme for the biosynthesis of menisporopsin A based on macrodiolide formation in congoblatin biosynthesis (Fig. 7).

First, a reduced diketide intermediate, 3-hydroxybutyryl-*S*-ACP is produced by Men1 and transferred to Men2; this is followed by a second reduced diketide which is further elongated using 3 units of malonyl-coA to form a reduced pentaketide. The cyclization of this intermediate by the PT domain forms the second subunit, 2,4-dihydroxy-6-(2-hydroxy-*n*-propyl)benzoyl-*S*-ACP. The TE domain then esterifies the secondary hydroxyl group on the side chain of the second subunit with the acyl-TE of the first subunit to form the first ester intermediate. This process occurs iteratively to form a linear tetraester intermediate. The final subunit is formed by a similar process, except that an extra malonyl-CoA is required in an additional elongation step to form a reduced hexaketide intermediate, and the carbonyl group next to the secondary hydroxyl group is reduced by a *trans*-acting ketoreductase. Again, the PT domain catalyzes cyclization to form the largest subunit, 2,4-dihydroxy-6-(2,4-dihydroxy-*n*-pentyl) benzoyl-*S*-ACP. Then the linear pentaester intermediate is formed. In this step, if the intermediate transfer rate is slow, intramolecular cyclization involving the secondary hydroxyl group of the pentaester intermediate may occur to form menisporopsin B. Such a reaction has been observed in radicicol biosynthesis, when inactivation of the TE resulted in isocoumarin production.[23] Alternatively, transfer of the pentaester intermediate to the TE domain would allow cyclolactonization to be catalyzed by the TE to form menisporopsin A.

Conclusions

We identified an HR- and an NR-PKS that are necessary and sufficient for the biosynthesis of a fungal macrocyclic polylactone. This information provides a platform to study the mechanism by which an NR-PKS performs multiple esterification and cyclolactonization reactions, and may lead to directed structural modification of this class of metabolites to enhance their biological activities.

Experimental

Materials and methods

Transcriptome sequencing. *M. theobromae* BCC 4162 was first grown in 25 ml potato dextrose broth (PDB) for 7 days and then 1.25 ml of culture was transferred to 25 mL of menisporopsin A production medium[24] (FMSM, containing 1% w/v fructose, 2.5% w/v meat extract, 0.0025% w/v KH_2PO_4 and 0.0025% w/v $MgSO_4·7H_2O$). Mycelia for RNA extraction were collected both before the culture was transferred and after growing in FMSM for 4 days in order to compare the expression level of PKS genes involved in menisporopsin A biosynthesis. Total RNA was extracted using RNeasy plant mini kit (QIAGEN) and treated with DNase I, then mRNA was enriched using oligo dT and fragmented before conversion to cDNA; adapters were subsequently connected to the fragments. After quality control by Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System, the transcriptome library was sequenced on the Illumina HiSeq 4000 platform by BGI Tech Solutions (Hong Kong). Paired-end read data were filtered to remove low quality and adapter-polluted sequences and *de novo* assembly of the clean data was performed using Trinity.[25] The assembled transcripts were then annotated with 7 functional databases (NR, NT, GO, COG, KEGG, Swissprot and Interpro). Blast v2.2.23 (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to compare the transcripts against NT, NR, COG, KEGG and SwissProt databases using default parameters. Blast2GO v2.5.0 (https://www.blast2go.com) and InterProScan5 v5.11-51.0 (https://www.ebi.ac.uk/interpro/interproscan.html) were used for annotating the transcripts by comparing them against GO and InterPro databases, respectively. Only the transcripts annotated as polyketide synthases were chosen for analysis (ESI File S2† for transcriptomic result analysis).

Plasmid construction. The 7.8 kb HR-PKS gene, *men1*, and the 6.6 kb NR-PKS gene, *men2*, were amplified as 3 and 2 overlapping fragments, respectively, from *M. theobromae* BCC 4162 genomic DNA. Fragments were joined by homologous recombination in yeast to reporter genes contained within the *E. coli*-yeast shuttle vectors pEYA2dsRED (*men1*) and pEYA2eGFP (*men2*). The primers used in this work are shown in Table S1,† and all PCRs were performed using KAPA HiFi HotStart DNA polymerase (KAPABIOSYSTEMS, USA); the programme used for amplification of *men1* and *men2* fragments was: initial denaturation at 95 °C for 5 min followed by 35 cycles of 98 °C for 20 s, 65 °C for 15 s and 72 °C for 1.5 min, with a final extension at 72 °C for 5 min. *Saccharomyces cerevisiae* YPH 499 was transformed with vectors linearized upstream of their reporter genes plus overlapping PCR products using the TRAFO method,[26] and recombinant plasmids were shuttled to *E. coli* DH5α. The Gateway® LR recombination reaction (Invitrogen, USA) was performed to insert the *men1dsRED* and *men2eGFP* gene fusions into the amyB expression cassette of vectors pTYGSmet and pTYGSarg, respectively.

***Aspergillus oryzae* NSAR1 transformation.** The resulting expression plasmids, pTYmetMen1dsRED and pTYarg Men2eGFP (Fig. S4†), were introduced into *A. oryzae* NSAR1 by protoplast-mediated transformation.[15] Transformants were selected and subcultured serially on appropriately supplemented Czapek-Dox plates until genetically pure. Expression of *men1dsRED* and *men2eGFP* from the amyB promoter was induced by culturing in CMP medium (3.5% Czapek-Dox, 2% maltose and 1% peptone) and monitored by confocal microscopy.

Transformant analysis. Seven-day-old CMP cultures were homogenized using a hand blender and the pH adjusted to 3.

Ethyl acetate (0.5 vol) was added to the blended culture and stirred for 10 min, after which debris was removed by vacuum filtration through Whatman No.1 filter paper. The filtrate was shaken vigorously in a separating funnel and the ethyl acetate layer was collected. Extraction was repeated twice, and combined extracts were dried with anhydrous Na_2SO_4. The organic solvent was then removed by rotary evaporation and the solid residue dissolved in HPLC grade methanol. HPLC analysis of metabolites was performed on an Agilent Technologies 1200 system using a C-18 Kaseisorb LC ODS 2000 column (4.6 mm × 150 mm); elution was with a $CH_3CN:H_2O$ gradient from 5–95% over 12 min and peaks were detected with UV at 254 nm. For purification, extracts were fractionated on a semi-preparative C-18 Kaseisorb LC ODS 2000 column (10.0 mm × 250 mm) to yield 3 isolated compounds (**1–3**), the structures of which were determined by NMR and HRMS.

Sequence alignments. Multiple sequence alignments were performed on BioEdit software[27] and rendered using ESPript 3.0.[28]

Structure determination of isolated metabolites. All NMR spectra were recorded and analyzed on an AVANCE III HD 400 spectrometer using acetone-d_6 as solvent. 1D (^{1}H and ^{13}C) and 2D NMR (COSY, HSQC and HMBC) data were used for structure determination of each compound. Their configurations were then assigned based on previously reported optical rotation value. The ^{1}H and ^{13}C chemical shifts are listed in the ESI File S1.†

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Royal Society, the Newton Fund (NA150457), the Thailand Research Fund (TRF) (DBG5980009) and the Center of Excellence for Innovation in Chemistry (PERCH-CIC), Office of the Higher Education Commission, Ministry of Education (OHEC). We also thank the Royal Thai Government for a DPST scholarship for W.B. and BIOTEC Culture Collection, Thailand for the seed fungus *M. theobromae* BCC 4162.

Notes and references

1 M. Chinworrungsee, P. Kittakoop, M. Isaka, P. Maithip, S. Supothina and Y. Thebtaranonth, *J. Nat. Prod.*, 2004, **67**, 689–692.

2 V. Ruanglek, S. Chokpaiboon, N. Rattanaphan, S. Madla, P. Auncharoen, T. Bunyapaiboonsri and M. Isaka, *J. Antibiot.*, 2007, **60**, 748–751.

3 P. Wattana-amorn, P. Juthaphan, M. Sirikamonsil, A. Sriboonlert, T. J. Simpson and N. Kongkathip, *J. Nat. Prod.*, 2013, **76**, 1235–1237.

4 M. Ito, Y. Tsuchida, K. Mizoue and K. Hanada, *J. Antibiot.*, 1992, **45**, 1566–1572.

5 G. Schlingmann, L. Milne and G. T. Carter, *Tetrahedron*, 2002, **58**, 6825–6835.

6 J.-W. He, Z.-Q. Mu, H. Gao, G.-D. Chen, Q. Zhao, D. Hu, J.-Z. Sun, X.-X. Li, Y. Li, X.-Z. Liu and X.-S. Yao, *Tetrahedron*, 2014, **70**, 4425–4430.

7 J.-L. Zhang, W.-J. Wang, X.-M. Xu, D.-Y. Li, H.-M. Hua, E.-L. Ma and Z.-L. Li, *Tetrahedron*, 2016, **72**, 4895–4901.

8 I. Letunic and P. Bork, *Nucleic Acids Res.*, 2018, **46**, D493–D496.

9 M. S. Kimber, F. Martin, Y. Lu, S. Houston, M. Vedadi, A. Dharamsi, K. M. Fiebig, M. Schmid and C. O. Rock, *J. Biol. Chem.*, 2004, **279**, 52593–52602.

10 A. Keatinge-Clay, *J. Mol. Biol.*, 2008, **384**, 941–953.

11 M. E. Mayorga and W. E. Timberlake, *Mol. Gen. Genet.*, 1992, **235**, 205–212.

12 Y. Takano, Y. Kubo, K. Shimizu, K. Mise, T. Okuno and I. Furusawa, *Mol. Gen. Genet.*, 1995, **249**, 162–167.

13 J. H. Yu and T. J. Leonard, *J. Bacteriol.*, 1995, **177**, 4792–4800.

14 K. A. Pahirulzaman, K. Williams and C. M. Lazarus, *Methods Enzymol.*, 2012, **517**, 241–260.

15 M. N. Heneghan, A. A. Yakasai, L. M. Halo, Z. Song, A. M. Bailey, T. J. Simpson, R. J. Cox and C. M. Lazarus, *ChemBioChem*, 2010, **11**, 1508–1512.

16 H. Wang, H. Zhang, Y. Zou, Y. Mi, S. Lin, Z. Xie, Y. Yan and H. Zhang, *PLoS One*, 2014, **9**, e97996.

17 T. Hiratsuka, H. Suzuki, R. Kariya, T. Seo, A. Minami and H. Oikawa, *Angew. Chem., Int. Ed.*, 2014, **53**, 5423–5426.

18 J. Liu, X. Zhu, R. F. Seipke and W. Zhang, *ACS Synth. Biol.*, 2015, **4**, 559–565.

19 Y. Zhou, X. Lin, S. R. Williams, L. Liu, Y. Shen, S. P. Wang, F. Sun, S. Xu, H. Deng, P. F. Leadlay and H. W. Lin, *ACS Chem. Biol.*, 2018, **13**, 2153–2160.

20 C. A. Shaw-Reid, N. L. Kelleher, H. C. Losey, A. M. Gehring, C. Berg and C. T. Walsh, *Chem. Biol.*, 1999, **6**, 385–400.

21 Y. Zhou, P. Prediger, L. C. Dias, A. C. Murphy and P. F. Leadlay, *Angew. Chem. Weinheim Bergstr. Ger.*, 2015, **127**, 5321–5324.

22 Y. Zhou, A. C. Murphy, M. Samborskyy, P. Prediger, L. C. Dias and P. F. Leadlay, *Chem. Biol.*, 2015, **22**, 745–754.

23 H. Zhou, K. Qiao, Z. Gao, J. C. Vederas and Y. Tang, *J. Biol. Chem.*, 2010, **285**, 41412–41421.

24 S. Madla, P. Kittakoop and P. Wongsa, *Lett. Appl. Microbiol.*, 2006, **43**, 548–553.

25 M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev, *Nat. Biotechnol.*, 2011, **29**, 644–652.

26 R. D. Gietz and R. A. Woods, *Methods Enzymol.*, 2002, **350**, 87–96.

27 T. A. Hall, *Nucleic Acids Symp. Ser.*, 1999, **41**, 95–98.

28 X. Robert and P. Gouet, *Nucleic Acids Res.*, 2014, **42**, W320–W324.

Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Electronic Supplementary Information

Heterologous biosynthesis of a fungal macrocyclic polylactone requires only two iterative polyketide synthases

Waraporn Bunnak, Passorn Wonnapinij, Ajaraporn Sriboonlert, Colin M. Lazarus and Pakorn Wattana-Amorn

Experimental

Table S1: Primers used for constructing *men1* and *men2*. Underlined nucleotides are regions of overlap with *E. coli*-yeast shuttle vectors (pEYA2eGFP and pEYA2dsRED)

gene	primer	Sequence 5′-3′
men1	*Fragment 1*	
	Men1-1-fwd	TAATGCCAACTTTGTACAAAAAAGCAGGCTATGGGCAGCGTCTACGATAG
	Men1-1-rev	AGCTCCGCGACGAGGGATGG
	Fragment 2	
	Men1-2-fwd	TCAATCGCGCCCATCTTACG
	Men1-2-rev	ATATCAACACGCGGGACGTG
	Fragment 3	
	Men1-3-fwd	TCCAAGCCCAGAAGAAGTCG
	Men1-3-rev	CTTGATGACGTCCTCGGAGGAGGCCATCTGCTTGGTGGTGGCGGCGGTAG
men2	*Fragment 1*	
	men2-1-fwd	TAATGCCAACTTTGTACAAAAAAGCAGGCTATGGCGGGCCCCGTTGCAAC
	men2-1-rev	AGGCGCCGAGCACCTGCTCG
	Fragment 2	
	men2-2-fwd	ATGCAGGAACGGTGCGAGTC
	men2-2-rev	TGAACAGCTCCTCGCCCTTGCTCACCATGTTGTCCATGGCGTCCTTAAGG

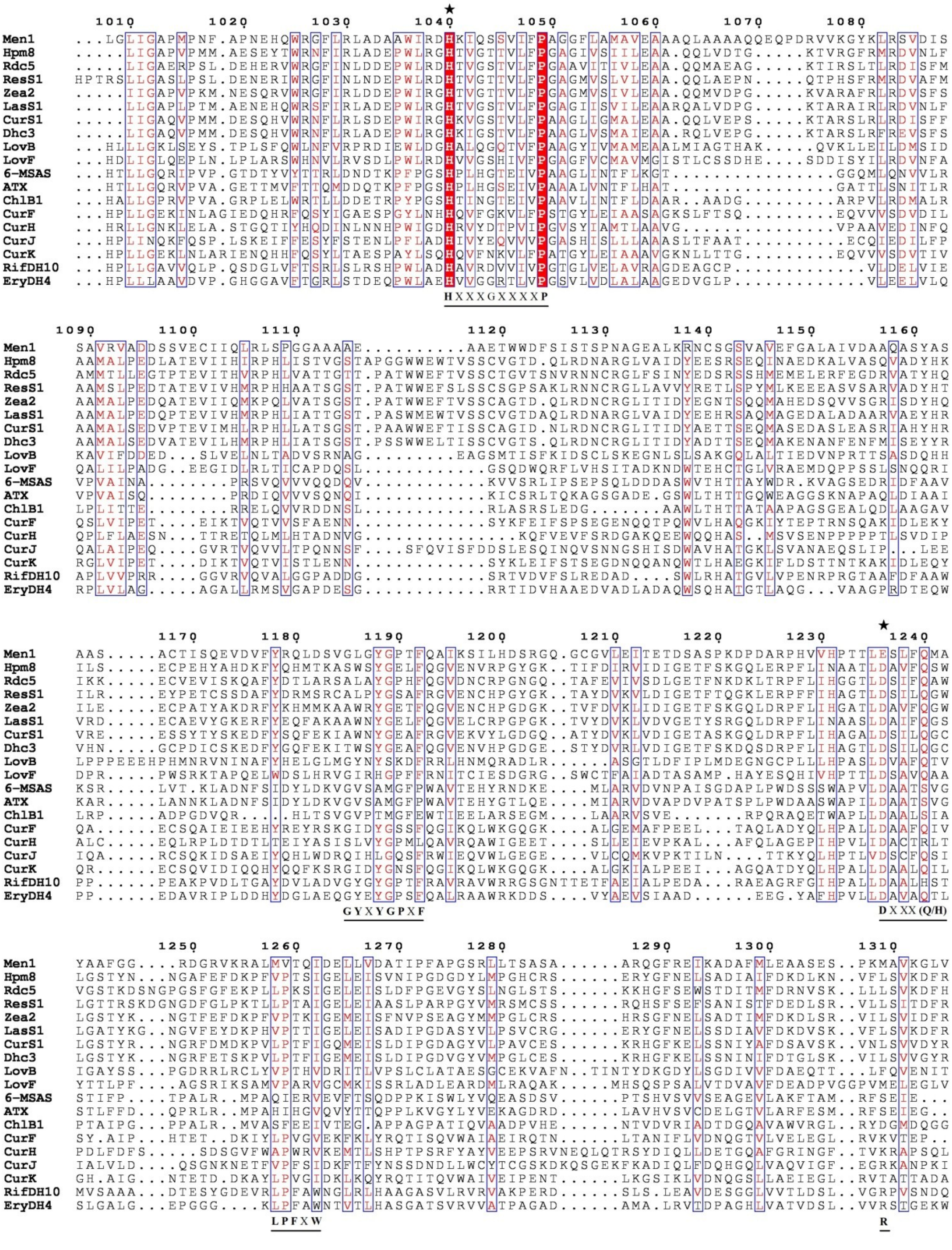
Results

MGSVYDSQSIAIVGLSCRLPGDADNAERFWNLMSEGRSAISSVPADRWNSKGFRDPTGKKRQNTSLTDRAHFVKGDISEFDANFFTISKAEADSM
DPQQRIMLEVAYEAFENAGLSMDSLAKSQTGCWVSSFSQDWKEMHFSDPDAAPKYAMSGMQPELLSNRVSYFFDLQGPSMTIETACSGSLVGLH
VACQSLRAGDCETALVGGANLFLNQNMFLALSNQSFLAPDGLCKAFDASANGYGRGEGFAAVILKPIEKAIRDGDHIRAVIRGTGTNQDGRTKGL
TMPNGHAQESLIRSTYAAAGLDLKDTAYFEAHGTGTQAGDFEELGAISRTVADARQKAGLEDLWVGSAKTNIGHLEAVAGLVGVLKAVLVLEN
GVIPPNLHFKNPNPRIPFGKWRIKVPTERIQWPSDGIRRVSVNSFGFGGSNAHAILDDADQYLSGRGIIRANGKSHHHHHQHQQQQLNGGNGGSNG
VNGTSEVNGTSGVNGTTTAITNGSTHVNGTAATAATAAAQQIIILNSYDQEGLGRQREALLRYAEKQQQQQQQQGGQGGADPEKLLGDLAFTLN
QKRSRLPWRTFFTASTLPELSRALEAASTFPAIRSGAATPRIAYVFTGQGAQWAQMGMDLLRFHVFRESVEAADRHLTQIGCPWSAVEELQRGDA
ESNIHISWYSQTLCTVIQVALVQLLESWNVRPRSVVGHSSGEMGAAFAIGALSREDAWTIAYWRGKLSSELTTIAPTQKGAMMAVGASHAQAQA
WVDGLTRGRCVVACVNSPSSVTVSGDESGLDELAAMLKEQGVFARKLKVSTAYHSHHMKAVAEAYLDALKGVRTRTVPAEGGAPQMFTSVSE
SLVDPAELGPAHWVANLISPVLFSNTVRELARPKGPDDEASGSAVDLMVEIGPHAALRGPVTQILQSHGLPALDYYSVLSRGANSVDTALAVVGE
LVCRGVPVDLGAVNRAHLTAEQQLQADRRPSLVAELPPYAWNHAKTYWSESRISRELKYRPAPQLGLIGAPMPNFAPNEHQWRGFLRLADAAWI
RDHKIQSSVIFPAGGFLAMAVEAAAQLAAAAQQEQPDRVVKGYKLRSVDISSAVRVADDSSVECIIQLRLSPGGAAAAEAAETWWDFSISTSPNA
GEALKRNCSGSVAVEFGALAIVDAAQASYASAASACTISQEVDVFYRQLDSVGLGYGPTFQAIKSILHDSRGQGCGVLEITETDSASPKDPDARPH
VVHPTTLESLFQMAYAAFGGRDGRVKRALMVTQIDELLVDATIPFAPGSRLLTSASAARQGFREIKADAFMLEAASESPKMAVKGLVCVEMPSAS
GMGGGGGGGLDADQASYSAMLSKFVWKPALELLSAPEQAKLLEDATRLPEDEAQRLASEATAELHAVKAVLESAQSKKIANLKLRNAAKWISQ
QLQASGIPGKPAENGAREGGSSSGFTAEVEKVLSGLAEADVLLGSKGSADHLVAQLPGMKMSLEKMYKLVNYMAHANPNLTVLEIVPGGAGVD
FSLPLSAKDIPSTIQYTYASPSADNVQQMQERLGGGSGDSALALALAPRFRVLEIEQDLADQGLDPGSFDIVIGCNLLSNAVNVEKTLSQAKSLLTE
GGKMALVELNKPSPAALPVLGILCDWWKRRDDGLRRPFTTDMVNESLAGQGFAIELATPDFTDPALQQSSLVLASCQPASAGKESAAQEVVSILV
RKDSSEAVNALASQLSQACNGAKTVTWEAGVDFKGQHLISLLEFDTPLLDRLTEEDFGLVKQLITQAASLQWVTAIPEPHASTVMGLARVARFEV
PSLRFQTVTLDPSSVLALDRAATLIIQAQKKSTSQDKEFKEVDSVLHVPRVDIDAPLNEQVTRLLLEEDVEPMPLGSGDAARKLCIRNPGMLNTLCF
EIDSLPSTVLAEDEVEMQVKASGLSPKDVAICLGQVSDTALGFEASGIVTRVGAGVAQFQAGDKICMMARGAHRTVLRSKSALCQRIPEGMSFEQ
AAAVPLAHGAVYHALVNIARARSGQKILVAVSDAVVSEAAVQLAKHLGLEAFVTTESQDRTPLIGTKEDYGISDDHIFYSRDPTYVKEITRLTNGA
GVDCVLSSVSGEALKHATSCLAPFGTFVDLGAKDVRSSAILDKHPEAMFAAINLERISELRPDMAGRIMDGTFALLREGAIKPVKLLAAYPASDLE
TAMQALHARSRQDKIVIAYSADQVVPVLHNPRESLRLPGDKTYLIAGGLGGIGRNIANLLVECGARHLAFVSRSGVTSEAQQKLVDNLTQRGAKI
AVYRCNIGDAQSLEQTLARCSAEMPPVKGVIHSAVVFRDAVIHNMTYAQWHELMESKLGGSWNLHALTTSYDLDFFLCIGSFMAIIGGLSQSNYA
AGGAFQDGLAHMRQSMGLPAATIDLGIVKGFGAVEEQGAVGHTLEWREPFGVDEDAVFALIKKALLGQMDKDGPGVPPQMINTVPTGGMVRES
GVGQPYYFEDPRFAIMAAIGTRNADGADGQASVALKEQLAQAESPEEAARLVSAAVAAKVAKLMQVGAEEIDAGKPLHAYGVDSLVAIEYVHW
AKKEVAAEITVFDVMASVPISAFASDLAKKGEWGTTAATTKQ

Fig. S1: Amino acid sequence of Men1. Domains were identified using SMART and highlighted as follows: KS (red), AT (dark green), DH (cyan), ER (bright green), KR (magenta) and ACP (dark grey).

MAGPVATSDRQQRLIFFGDQTVDALPCIKILTAQAHRLPALRRFLRDAADVIQVLLSSLEFDDHDHYRRFETICELAEIYSKQDGTHETIACALWTT
AQFGDLVMRAELNPSILTGGQQSADPTYVVGICGGLLPAAATATARDINELLDIGRKLVAVAFRLGVAQWRRAMDIEGKPGRWAVTIVNVPAKQ
IRTILDAFNEDMEIPKHRQFYISFLAKGWVTVSGPPSLFPELWAYSSTLNAASKMQLPLGTPAHAAHLPPVNVSEIVGTGDVLDLTVRENFFTVSTS
TCQPFQCQDLGSLLQESLRDITGKTLNIAGVNDYVISALDRNTPVRVSSFGPASQIASFKKTLEEAGYQVELDLGDPSLGNPEYPIDADSRDGSNLIA
VVGQSVRFPGSEDVETFWENIKAGRSFETEIPASRFDLAHHYDATGSKTSSVTTKYGSFLENPGLFDNRLFNVSPREAKQMDPIQRILMMCSYEAL
QAAGYSPDGSLSTNSMRIATYFGQSGDDWRQVRASQEVDIYYIPGTVRSFAPGKLNYHYKWGGGNYAVDSACAASTTTMMLASQALLARDCD
MALAGGGQLNAAPEPYAGLSRAGFLSKTTGGCKTFREDADGYCRGEGVGVVVLKRLEDALAENDNVLAVIRGADRNFSWDATSITHPSVSAQV
KLVKSVLRNTGVEPEEIGFVEMHGTGTQAGDGVEMETVTTVFGSRPKDNPLYVGAVKANIGHGEAAAGVASVIKAIEVLRHRTIPTQAGFPGPRD
PKFNHLDGMNIRIPESVVPFQPAPAPFSSDGKRKVLVNNFDASGGNNCVLLEEAPARDRETTADPRGVYTVAVSARTTNSLKNNISRLLGYLQSHP
DASVADVAYTTTARRMHEDLKKAYTVQTASELVSLLQADLKKDLTAVQYRSPHSVVFAFTGQGAQYAGMGKQLFDTSAAFRESVQAFHELAV
WQGFPEFLHLIADDQADVKAADPVQLQLAVVVLEMALANLWKSWGVEPGLVVGYSLGEYPALYVAGVLSVHDVIFLVGNRARLMQERCESGS
YAMLATQSSPQDLEQVLGAYPSCAVACKGAPRSTVVSGPTEDITQLHSELKEKNINGTLLNVPYGFHCAQVDPILDDFRDMADGIQFNKPRIPVAS
SLEGTVVTEEGVFSSAYLVRQTREPVNFIGAVKAAESSGRADNTTVWIEIGPKRVLSSLVKSTLSADQGRLLHTIDDKDSNWKTIAAAMTAGYTEG
MSIKWPKFHKLFSKHLTLLELPTYAFDLKDYWIPPAVPVTAAAPVAAPAADPSLPVIPVVPGFPTASLQQVRSEQINGDEAKVTFETVISHPALLAVI
RGHRVGGVDLFPASGFMDMAFSAAKYIHHRTKSGQPVPEISMKHLAITHPLTPSSGQSRQIVIVTASKRSGSSVVDVSFRSRDGSAEQDHGDCKLH
FDKRGSWDAEWAQTAHFINAAKKNVIANGTSPAGTGHRLPKSVVYKLFSSLVEYSGAFRAMEEVYVTDDFQKEAVASVVLPGGSSEFYVNPYW
SDALIHVCGFLLNSSPNLPSQDCFLFNGLEEMRLLSDDLQPGIPYTSYVYLTEPGSSPDSQAPRPKHARGDVYVFQGEKIVGVAKGVVFQRLTRRV
LATVLGGKLPGAAAPVREIAAPAPIRAVAPAPAPVAPRPVEMYRVPGVVGDEKADAAIGKILARAGANPAHITDATTFAEIGFDSLEWIELVREIRT
SLDLEVPASFFFEYPKVNGLRRAISELSLDYQGPASGSVSVSSSATTHGMTTPSSTSSAQSSQSSQTPDGPGIYANAVIDIVLSQTGFDKADLLPTTR
FDDMGLDSLCTMEVVGVVREQTGLDLPASFFHQNPTVAHVRRALGSDSDGDSKPKSAPAPPAPEPVVEVAAPAPAVQAPPAILDGDLASYHCDFF
LMQGSSDSTKTPLFFLPDGTGYPAVLLKLPPIFEGDNALFTCKSPLLNVAEGREVRCTIEGLAAAYAAAIQRARPHGPYLLAGYSLGGAYAFEVAKI
LADAGEVVQGLLFVDFNMAASVGKLHRDRNPVPVDLTVGAMEKTGWMQGIQNDDKDFNIPPAPPKIKFHALSVFKSLISYYPTPMTPSQRPRNT
YALWAGIGMQDLLGTKNAGFLPAYGIIDWQMGDRHENNGPAGWEEYIGGPVKCATMPCDHLSLLMSHHWIPKSAEVIKGLLKDAMDN

Fig. S2: Amino acid sequence of Men2. Domains were identified using SMART and are highlighted as follows: SAT (yellow), KS (red), AT (dark green), PT (cyan), ACP (dark grey) and TE (magenta).



```
           1010       1020      1030      1040      1050      1060      1070      1080
                                             ★
Men1     ...LGLIGAPMPNF.APNEHQWRGFLRLADAAWIRDHKIQSSVIFPAGGFLAMAVEAAAQLAAAAQQEQPDRVVKGYKLRSVDIS
Hpm8     .....LIGAPVPMM.AESEYTWRNFIRLADEPWLRGHTVGTTVLFPGAGIVSIILEAA.QQLVDTG.....KTVRGFRMRDVNLF
Rdc5     .....LIGAERPSL.DEHERVWRGFINLDDEPWLRDHTVGSTVLFPGAAVITIVLEAA.QQMAEAG.....KTIRSLTLRDISFM
ResS1    HPTRSLLGASLPSL.DENERIWRGFINLNDEPWLRDHTVGTTVLFPGAGMVSLVLEAA.QQLAEPN.....QTPHSFRMRDVAFM
Zea2     .....IIGAPVPKM.NESQRVWRGFIRLDDEPWIRGHTVGTTVLFPGAGMVSIVLEAA.QQMVDPG.....KVARAFRLRDVSFS
LasS1    .....LLGAPLPTM.AENEHQWRSFIRLADEPWLRGHTVGSTVLFPGAGIISVILEAARQALVDPG.....KTARAIRLRDVNLF
CurS1    .....IIGAQVPMM.DESQHVWRNFLRLSDEPWIRGHKVGSTVLFPAAGLIGMALEAA.QQLVEPS.....KTARSLRLRDISFF
Dhc3     .....LIGAQVPMM.DESQHVWRNFLRLADEPWLRGHKIGSTVLFPAAGLVSMAIEAA.RQLVEPG.....KTARSLRFREVSFS
LovB     ...HLLLGKLSEYS.TPLSFQWLNFVRPRDIEWLDGHALQGQTVFPAAGYIVMAMEAALMIAGTHAK.....QVKLLEILDMSID
LovF     ...HDLIGLQEPLN.LPLARSWHNVLRVSDLPWLRDHVVGSHIVFPGAGFVCMAVMGISTLCSSDHE....SDDISYILRDVNFA
6-MSAS   ...HTLLGQRIPVP.GTDTYVYTTRLDNDTKPFPGSHPLHGTEIVPAAGLINTFLKGT................GGQMLQNVVLR
ATX      ...HTLLGQRVPVA.GETTMVFTTQMDDQTKPFPGSHPLHGSEIVPAAALVNTFLHAT................GATTLSNITLR
ChlB1    ...HALLGPRVPVA.GRPLELWRTLLDDETRPYPGSHTINGTEIVPAAVLINTFLDAAR..AADG.........ARPVLRDMALR
CurF     ...HPLLGEKINLAGIEDQHRFQSYIGAESPGYLNHHQVFGKVLFPSTGYLEIAASAGKSLFTSQ.........EQVVVSDVDIL
CurH     ...HRLLGNKLELA.STGQTIYHQDINLNNHPWIGDHRVYDTPVIPGVSYIAMTLAAVG...............VPAAVEDINFQ
CurJ     ...HPLINQKFQSP.LSKEIFFESYFSTENLPFLADHIVYEQVVVPGASHISLLLAAASLTFAAT.........ECQIEDILFP
CurK     ...HPLLGEKLNLARIENQHHFQSYLTAESPAYLSQHQVFNKVLFPATGYLEIAAAVGKNLLTTG.........EQVVVSDVTIV
RifDH10  ...HPLLGAVVQLP.QSDGLVFTSRLSLRSHPWLADHAVRDVVIVPGTGLVELAVRAGDEAGCP.............VLDELVIE
EryDH4   ...HPLLLAAVDVP.GHGGAVFTGRLSTDEQPWLAEHVVGGRTLVPGSVLVDLALAAGEDVGLP.............VLEELVLQ
                                             HXXXGXXXXP

       1090      1100      1110                1120      1130      1140      1150      1160
Men1     SAVRVADDSSVECIIQLRLSPGGAAAAE..........AAETWWDFSISTSPNAGEALKRNCSGSVAVEFGALAIVDAAQASYAS
Hpm8     AAMALPEDLATEVIIHIRPHLISTVGSTAPGGWWEWTVSSCVGTD.QLRDNARGLVAIDYEESRSEQINAEDKALVASQVADYHK
Rdc5     AMMTLLEGTPTEVITHVRPHLVATTGTT.PATWWEFTVSSCTGVTSNVRNNCRGLFSINYEDSRSSHMEMELERFEGDRVATYHQ
ResS1    AAMSLPEDTATEVIVHMRPHHAATSGST.PATWWEFSLSSCSGPSAKLRNNCRGLLAVVYRETLSPYMLKEEEASVSARVADYHT
Zea2     AAMALPEDQATEVIIQMKPQLVATSGST.PATWWEFTVSSCAGTD.QLRDNCRGLITIDYEGNTSQQMAHEDSQVVSGRISDYHQ
LasS1    AAMALPEDQPTEVIVHMRPHLIATTGST.PASWMEWTVSSCVGTDAQLRDNARGLVAIDYEEHRSAQMAGEDALADAARVAEYHR
CurS1    AAMALSEDVPTEVIMHLRPHLLATSGST.PAAWWEFTISSCAGID.NLRDNCRGLITIDYAETTSEQMASEDASLEASRIAHYHR
Dhc3     AAMALSEDVATEVILHMRPHLIATSGST.PSSWWELTISSCVGTS.QLRDNCRGLITIDYADTTSEQMAKENANFENFMISEYYR
LovB     KAVIFDDED...SLVELNLTADVSRNAG.........EAGSMTISFKIDSCLSKEGNLSLSAKGQLALTIEDVNPRTTSASDQHH
LovF     QALILPADG..EEGIDLRLTICAPDQSL............GSQDWQRFLVHSITADKNDWTEHCTGLVRAEMDQPPSSLSNQQRI
6-MSAS   VPVAINA........PRSVQVVVQQDQV............KVVSRLIPSEPSQLDDDASWVTHTTAYWDR.KVAGSEDRIDFAAV
ATX      VPVAISQ........PRDIQVVVSQNQI............KICSRLTQKAGSGADE.GSWLTHTTGQWEAGGSKNAPAQLDIAAI
ChlB1    LPLITTE........RRELQVVRDDNSL............RLASRSLEDG.......AAWLTHTTATAAPAGSGEALQDLAAGAV
CurF     QSLVIPET....EIKTVQTVVSFAENN.............SYKFEIFSPSEGENQQTPQWVLHAQGKIYTEPTRNSQAKIDLEKY
CurH     QPLFLAESN...TTRETQLMLHTADNVG...............KQFVEVFSRDGAKQEEWQQHAS..MSVSENPPPPPTLSVDIP
CurJ     QALAIPEQ....GVRTVQVVLTPQNNSF....SFQVISFDDSLESQINQVSNNGSHISDWAVHATGKLSVANAEQSLIP...LEE
CurK     RGLVIPET....DIKTVQTVISTLENN.............SYKLEIFSTSEGDNQQANQWTLHAEGKIFLDSTTNTKAKIDLEQY
RifDH10  APLVVPRR....GGVRVQVALGGPADDG............SRTVDVFSLREDAD....SWLRHATGVLVPENRPRGTAAFDFAAW
EryDH4   RPLVLAG.....AGALLRMSVGAPDESG............RRTIDVHAAEDVADLADAQWSQHATGTLAQG...VAAGPRDTEQW

                 1170      1180      1190      1200       1210      1220      1230      1240
                                                                                      ★
Men1     AAS.....ACTISQEVDVFYRQLDSVGLGYGPTFQAIKSILHDSRGQ.GCGVLEITETDSASPKDPDARPHVVHPTTLESLFQMA
Hpm8     ILS.....ECPEHYAHDKFYQHMTKASWSYGELFQGVENVRPGYGK..TIFDIRVIDIGETFSKGQLERPFLINAATLDAVFQSW
Rdc5     IKK.....ECVEVISKQAFYDTLARSALAYGPHFQGVDNCRPGNGQ..TAFEVIVSDLGETFNKDKLTRPFLIHGGTLDSIFQAW
ResS1    ILR.....EYPETCSSDAFYDRMSRCALPYGSAFRGVENCHPGYGK..TAYDVKVLDIGETFTQGKLERPFFIHAGTLDSILQGW
Zea2     ILE.....ECPATYAKDRFYKHMMKAAWRYGETFQGVENCHPGDGK..TVFDVKLIDIGETFSKGQLDRPFLIHGATLDAVFQGW
LasS1    VRD.....ECAEVYGKERFYEQFAKAAWNYGELFQGVELCRPGPGK..TVYDVKLVDVGETYSRGQLDRPFLINAASLDAIFQGS
CurS1    VRE.....ESSYTYSKEDFYSQFEKIAWNYGEAFRGVEKVYLGDGQ..ATYDVKLVDIGETASKGQLDRPFLIHAGALDSILQGC
Dhc3     VHN.....GCPDICSKEDFYGQFEKITWSYGEAFQGVENVHPGDGE..STYDVRLVDIGETFSKDQSDRPFLIHAGTLDSILQGC
LovB     LPPPEEEHPHMNRVNINAFYHELGLMGYNYSKDFRRLHNMQRADLR......ASGTLDFIPLMDEGNGCPLLLHPASLDVAFQTV
LovF     DPR.....PWSRKTAPQELWDSLHRVGIRHGPFFRNITCIESDGRG..SWCTFAIADTASAMP.HAYESQHIVHPTTLDSAVQAA
6-MSAS   KSR.....LVT.KLADNFSIDYLDKVGVSAMGFPWAVTEHYRNDKE....MLARVDVNPAISGDAPLPWDSSSWAPVLDAATSVG
ATX      KAR.....LANNKLADNFSIDYLDKVGVSAMGFPWAVTEHYGTLQE....MIARVDVAPDVPATSPLPWDAASWAPILDAATSVG
ChlB1    LRP.....ADPGDVQR....HLTSVGVPTMGFEWTIEELARSEGM....LAARVSVE.......RPQRAQETWAPLLDAALSIA
CurF     QA......ECSQAIEIEEHYREYRSKGIDYGSSFQGIKQLWKGQGK....ALGEMAFPEEL...TAQLADYQLHPALLDAAFQIV
CurH     ALC.....EQLRPLDTDTLTEIYASISLVYGPMLQAVRQAWIGEET....SLLEIEVPKAL...AFQLAGEPIHPVLIDACTRLT
CurJ     IQA.....RCSQKIDSAEIYQHLWDRQIHLGQSFRWIEQVWLGEGE....VLCQMKVPKTILN....TTKYQLHPTLVDSCFQSI
CurK     QR......ECSQVIDIQQHYQQFKSRGIDYGNSFQGIKQLWKGQGK....ALGKIALPEEI...AGQATDYQLHPALLDAALQIL
RifDH10  PP......PEAKPVDLTGAYDVLADVGYGYGPTFRAVRAVWRRGSGNTTETFAEIALPEDA...RAEAGRFGIHPALLDAALHST
EryDH4   PP......EDAVRIPLDDHYDGLAEQGYEYGPSFQALRAAWRKDDS....VYAEVSIAAD......EEG.YAFHPVLLDAVAQTL
                                   GYXYGPXF                                             DXXX(Q/H)

                 1250      1260      1270      1280            1290      1300        1310
Men1     YAAFGG....RDGRVKRALMVTQIDELLVDATIPFAPGSRLLTSASA......ARQGFREIKADAFMLEAASES..PKMAVKGLV
Hpm8     LGSTYN...NGAFEFDKPFVPTSIGELEISVNIPGDGDYLMPGHCRS......ERYGFNELSADIAIFDKDLKN..VFLSVKDFR
Rdc5     VGSTKDSNGPGSFGFEKPLLPKSIGELEISLDFPGEVGYSLNGLSTS......KKHGFSEWSTDITMFDRNVSK..LLLSVKDFH
ResS1    LGTTRSKDGNGDFGLPKTLLPTAIGELEIAASLPARPGYVMRSMCSS......RQHSFSEFSANISTFDEDLSR..VLLSITDFR
Zea2     LGSTYK...NGTFEFDKPFVPTKIGEMEISFNVPSEAGYMMPGLCRS......HRSGFNELSADTIMFDKDLSR..VILSVIDFR
LasS1    LGATYKG..NGVFEYDKPHVPTTIGELEISADIPGDASYVLPSVCRG......ERYGFNELSSDIAVFDKDVSK..VFLSVKDFR
CurS1    LGSTYR...NGRFDMDKPVLPTFIGQMEISLDIPGDAGYVLPAVCES......KRHGFKELSSNIYAFDSAVSK..VNLSVVDYR
Dhc3     LGSTYK...NGRFETSKPVLPTFIGEMEISLDIPGDVGYVMPGLCES......KRHGFKELSSNINIFDTGLSK..VILSVVGYR
LovB     IGAYSS...PGDRRLRCLYVPTHVDRITLVPSLCLATAESGCEKVAFN..TINTYDKGDYLSGDIVVFDAEQTT...LFQVENIT
LovF     YTTLPF....AGSRIKSAMVPARVGCMKISSRLADLEARDMLRAQAK.....MHSQSPSALVTDVAVFDEADPVGGPVMELEGLV
6-MSAS   STIFP....TPALR..MPAQIERVEVFTSQDPPKISWLYVQEASDSV......PTSHVSVVSEAGEVLAKFTAM..RFSEIE...
ATX      STLFFD...QPRLR..MPAHIHGVQVYTTQPPLKVGYLYVEKAGDRD......LAVHVSVCDELGTVLARFESM..RFSEIEG..
ChlB1    PTAIPG...PPALR..MVASFEEIVTEG.APPAGPATIQVAADPVHE......NTVDVRIADTDGQAVAWVRGL..RYDGMDQGG
CurF     SY.AIP...HTET..DKIYLPVGVEKFKLYRQTISQVWAIAEIRQTN......LTANIFLVDNQGTVLVELEGL..RVKVTEP..
CurH     PDLFDFS.....SDSGVFWAPWRVKEMTLSHPTPSRFYAYVEEPSRVNEQLQTRSYDIQLLDETGQAFGRINGF..TVKRAPSQL
CurJ     IALVLD....QSGNKNETFVPFSIDKFTFYNSSDNDLLWCYTCGSKDKQSGEKFKADIQLFDQHGQLVAQVIGF..EGRKANPKI
CurK     GH.AIG...NTETD.DKAYLPVGIDKLKQYRQTITQVWAIVEIPENT......LKGSIKLVDNQGSLLAEIEGL..RVTATTADA
RifDH10  MVSAAA...DTESYGDEVRLPFAWNGLRLHAAGASVLRVRVAKPERD......SLS.LEAVDESGGLVVTLDSL..VGRPVSNDQ
EryDH4   SLGALG...EPGGG....KLPFAWNTVTLHASGATSVRVVATPAGAD......AMA.LRVTDPAGHLVATVDSL..VVRSTGEKW
                            LPFXW                                                       R
```

Fig. S3. Multiple sequence alignment of DH domains from several reducing PKSs Similar conserved residues are colored in red. The catalytic dyads are marked by stars and the underlined residues are common motifs found in DH domains. Highly reducing PKSs: Hpm8 from *Hypomyces subiculosus* (B3FWT3), Rdc5 from *Pochonia chlamydosporia* (B3FWU0), ResS1 from *Sarocladium zeae* (AHV78252), Zea2 from *Gibberella zeae* PH-1 (A0A098D8A0), LasS1 from *Lasiodiplodia theobromae* (AHV78245), CurS1 from *Aspergillus terreus* (L7X8J4), Dhc3 from *Alternaria cinerariae* (A0A0N7D4P6), LovB from *Aspergillus terreus* (Q9Y8A5), LovF from *Aspergillus terreus* (Q9Y7D5); partially reducing PKSs: 6-MSAS from *Penicillium griseofulvum* (P22367), ATX from *Aspergillus terreus* NIH2624 (Q0CJ59), ChlB1 from *Streptomyces antibioticus* (AAZ77673); modular PKSs: CurF (AAT70101), CurH (AAT70103), CurJ (AAT70105), CurK (AAT70106) from *Lyngbya majuscula* , RifDH10 from *Amycolatopsis mediterranei* (AEK39124), EryDH4 from *Saccharopolyspora erythraea* NRRL 2338 (CAM00064). GenPept accession numbers are given in brackets.

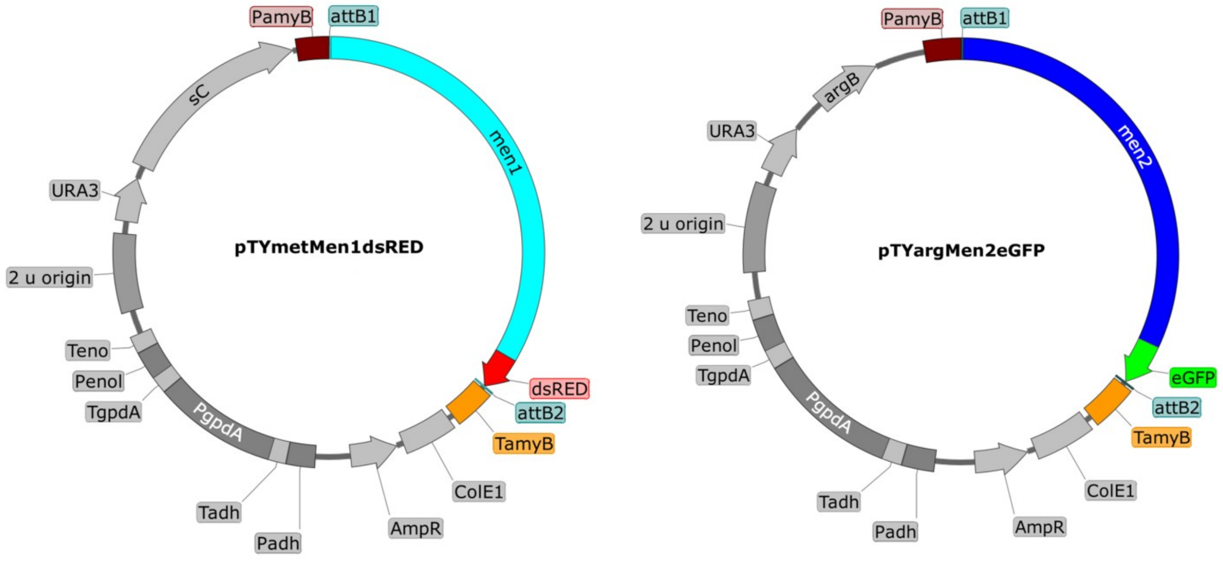
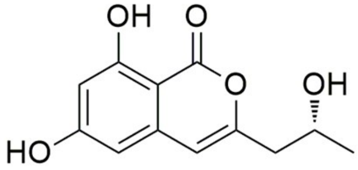


Fig. S4: Plasmid maps of pTYmetMen1dsRED and pTYargMen2eGFP generated using SnapGene software (from GSL Biotech; available at snapgene.com)

The ^{1}H and ^{13}C chemical shifts of compounds **(1)** – **(3)** are listed below.



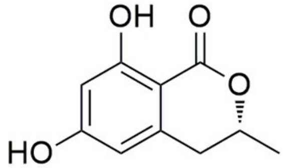
(-)-orthosporin **(1)**

^{1}H NMR (400 MHz, CD_3COCD_3): δ 11.16 (s, 1H), 6.43 (s, 1H), 6.41 (d, *J* = 2.2 Hz, 1H), 6.37 (d, *J* = 2.2 Hz, 1H), 4.16 (m, 1H), 2.59 (m, 2H), 1.23 (d, *J* = 6.2 Hz, 3H).

^{13}C NMR (100 MHz, CD_3COCD_3): δ 167.2, 166.5, 164.6, 156.6, 141.0, 106.5, 103.4, 102.3, 100.0, 65.6, 44.0, 23.7.

ESITOFMS *m/z* calcd for $C_{12}H_{12}O_5Na$ $[M+Na]^+$ = 259.0577, found = 259.0595

Both ^{1}H and ^{13}C NMR data and optical activity are in agreement with those reported in the literatures.[1, 2]



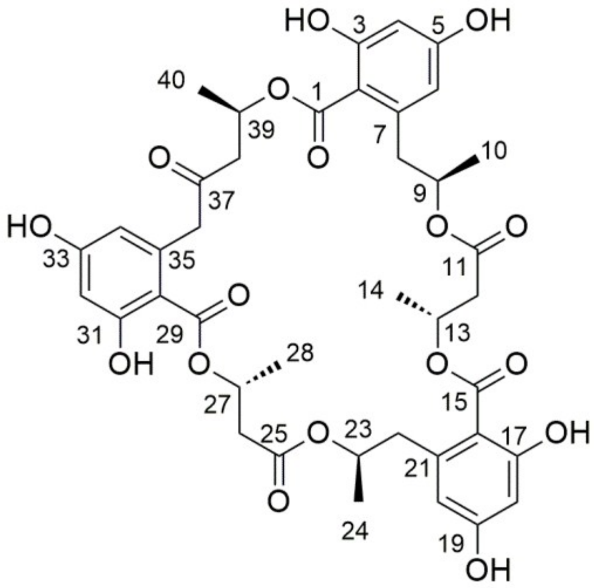
(-)-6-hydroxymellein **(2)**

^{1}H NMR (400 MHz, CD_3COCD_3): δ 11.28 (s, 1H), 6.29 (br, 1H), 6.26 (d, *J* = 1.9 Hz, 1H), 4.69 (m, 1H), 2.81-2.97 (m, 2H), 1.45 (d, *J* = 6.3 Hz, 3H).

^{13}C NMR (100 MHz, CD_3COCD_3): δ 170.8, 165.5, 165.4, 143.3, 107.6, 102.0, 101.7, 76.4, 35.1, 20.9.

ESITOFMS *m/z* calcd for $C_{10}H_{10}O_4Na$ $[M+Na]^+$ = 217.0471, found = 217.0471

Both ^{1}H and ^{13}C NMR Data and optical activity are in agreement with those reported in the literature.[3]



Compound **(3)** was identified as ascotrichalactone A and NMR data in acetone-d_6 are shown in the Table S2 and Figure S1 and S2. This was also confirmed when NMR data of compound **(3)** were collected in DMSO-d_6 and data are in accordance with those reported in the literature.[2]

ESITOFMS *m/z* calcd for $C_{40}H_{44}O_{17}Na$ $[M+Na]^+$ = 819.2476, found = 819.2472

Table S2: NMR data for compound **(3)**, ascotrichalactone A, in acetone-d_6:

Position	δ_C (ppm)	δ_H (ppm)
2,4-dihydroxy-6-(2-hydroxy-*n*-propyl)benzoic acid		
1, 15	171.4, 171.3	
2, 16	106.3, 106.2	
3, 17	166.0, 165.9	
4, 18	102.8, 102.7	6.28 (d, *J* = 2.4 Hz, 1H), 6.24 (d, *J* = 2.7 Hz, 1H)
5, 19	163.2, 163.3	
6, 20	113.1, 112.4	6.36 (d, *J* = 2.4 Hz, 1H), 6.33 (d, *J* = 2.4 Hz, 1H)
7, 21	143.3, 143.6	
8, 22	42.0, 41.7	3.58 (dd, *J* = 6.9, 3.3 Hz, 1H-8a), 2.93 (dd, *J* = 6.8, 12.3 Hz, 1H-8b)
		3.55 (dd, *J* = 7.6, 4.7 Hz, 1H-22a), 2.96 (dd, *J* = 13.6, 7.2 Hz, 1H-22b)
9, 23	72.8, 72.7	5.15-5.24 (m, 2H)
10, 24	19.9, 20.1	1.21 (d, *J* = 6.2 Hz, 3H), 1.22 (d, *J* = 6.2 Hz, 3H)
3- hydroxybutyric acid		
11, 25	170.4, 170.4	
12, 26	41.4, 41.0	2.94 (dd, *J* = 15.8, 7.0 Hz, 1H-12a), 2.82 (dd, *J* = 15.8, 6.3 Hz, 1H-12b)
		2.76 (br d, *J* = 6.3 Hz, 2H)
13, 27	70.1, 70.1	5.56 (m, 1H), 5.45 (m, 1H)
14, 28	20.2, 19.8	1.43 (d, *J* = 6.3 Hz, 3H), 1.34 (d, *J* = 6.4 Hz, 3H)
2,4-dihydroxy-6-(4-hydroxy-2-oxo-*n*-pentyl)benzoic acid		
29	171.1	
30	105.9	
31	166.4	
32	102.9	6.29 (d, *J* = 2.4 Hz, 1H)
33	163.5	
34	113.9	6.24 (d, *J* = 2.8 Hz, 1H)
35	140.5	
36	51.5	4.25 (d, *J* = 17.8 Hz, 1H-36a), 3.94 (d, *J* = 17.8, 1H-36b)
37	204.7	
38	48.1	3.29 (dd, *J* = 17.3, 5.2 Hz, 1H-38a), 3.03 (dd, *J* = 17.3, 7.8 Hz, 1H-38b)
39	69.9	5.58-5.65 (m, 1H)
40	20.7	1.47 (d, *J* = 6.2 Hz, 3H)
OH		11.49 (br), 11.41 (br), 11.29 (br), 9.29 (br)

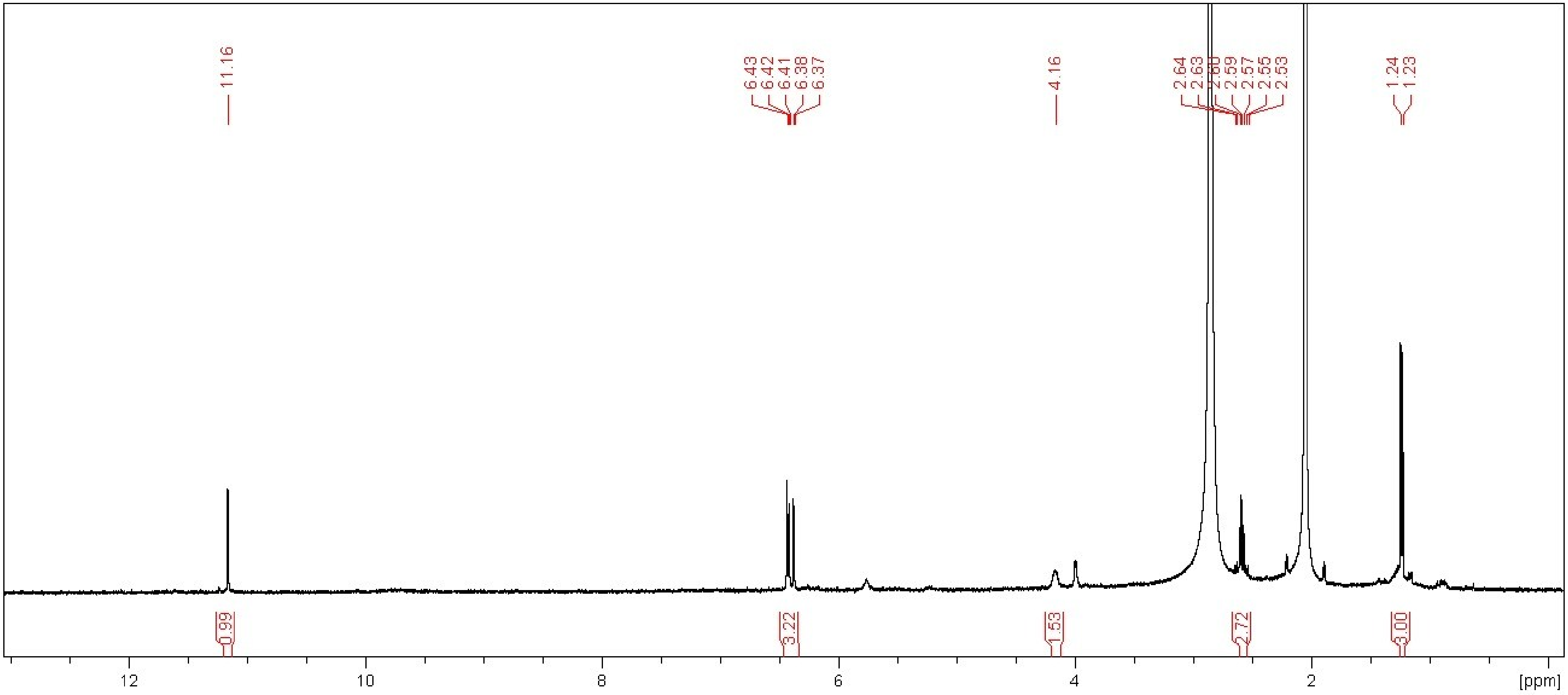


Fig. S5: ^{1}H-NMR spectrum of compound **(1)**, (-)-orthosporin, in acetone-d_6

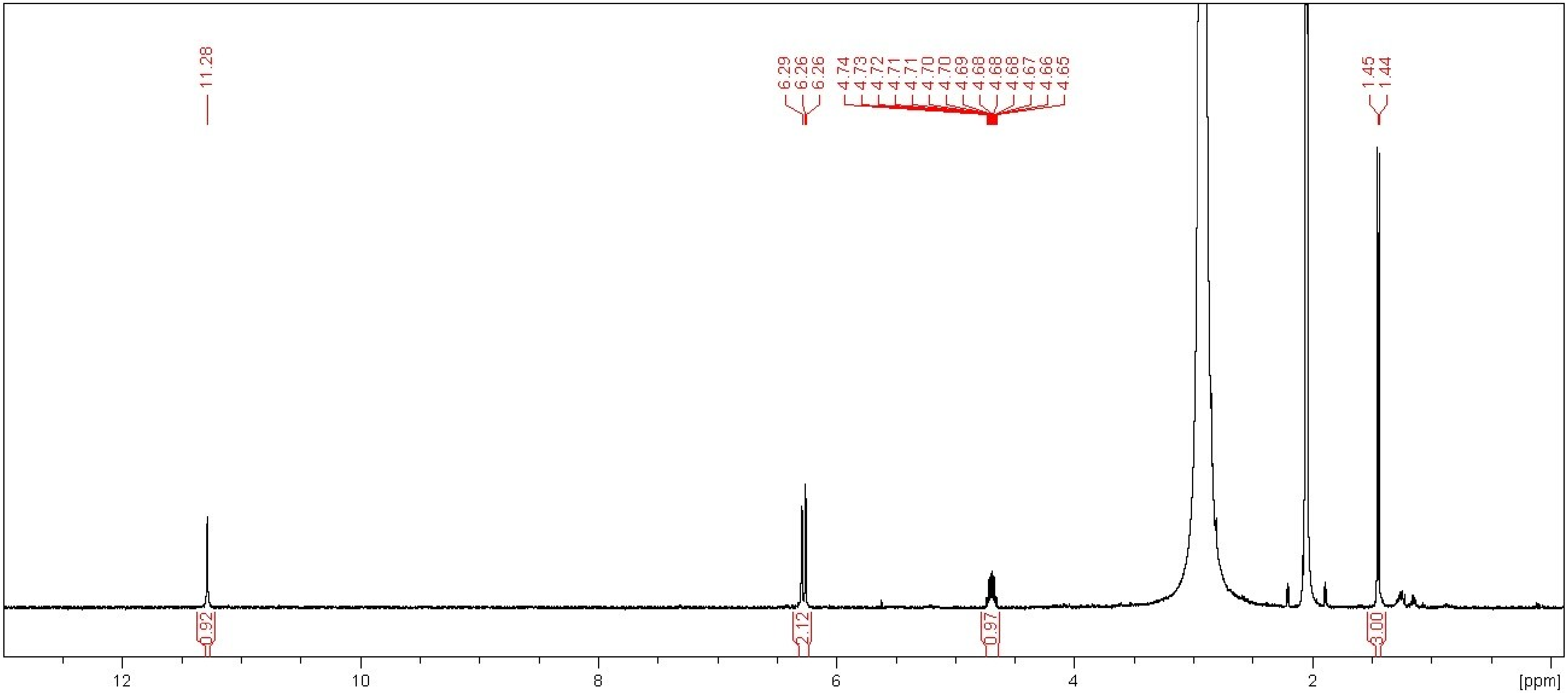


Fig. S6: ^{1}H-NMR spectrum of compound **(2)**, (-)-6-hydroxymellein, in acetone-d_6

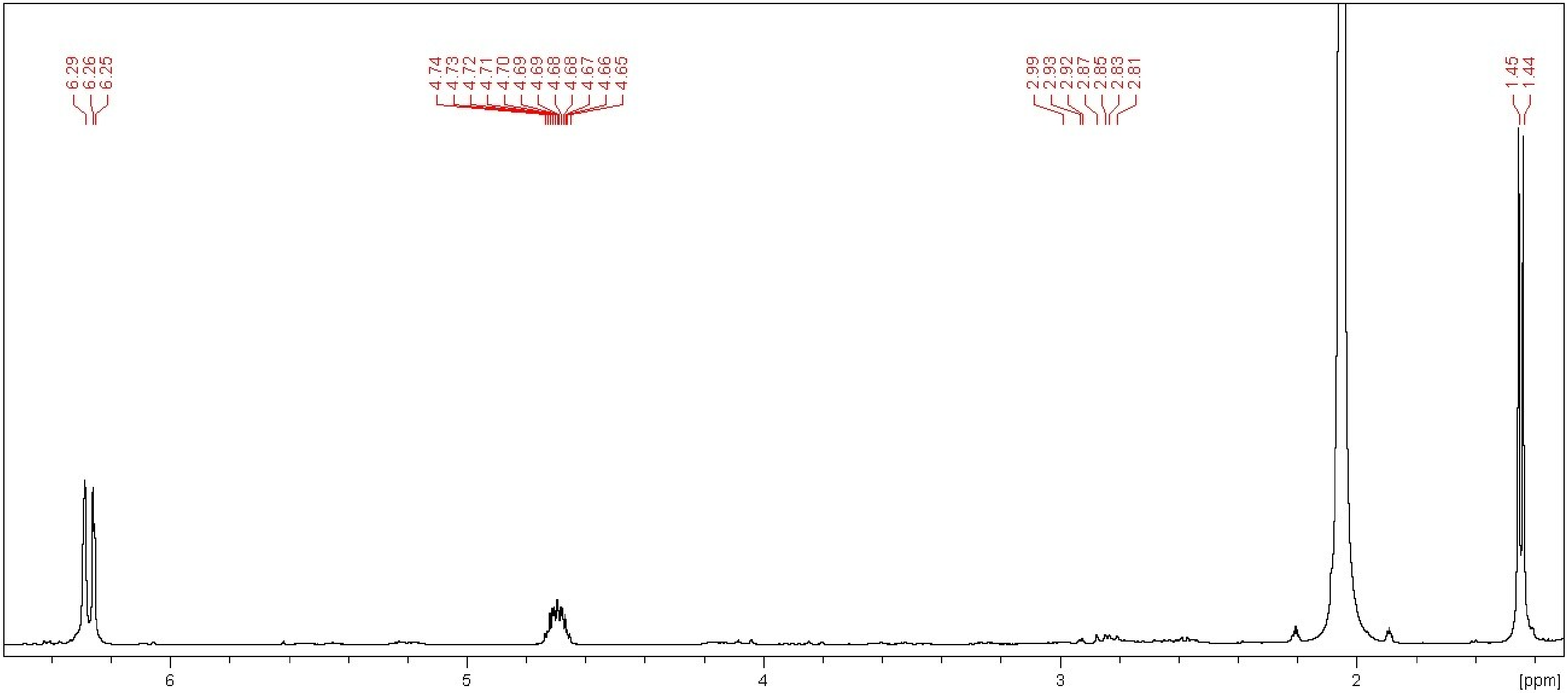


Fig. S7: ^{1}H-NMR spectrum with water suppression of compound **(2)**, (-)-6-hydroxymellein, in acetone-d_6

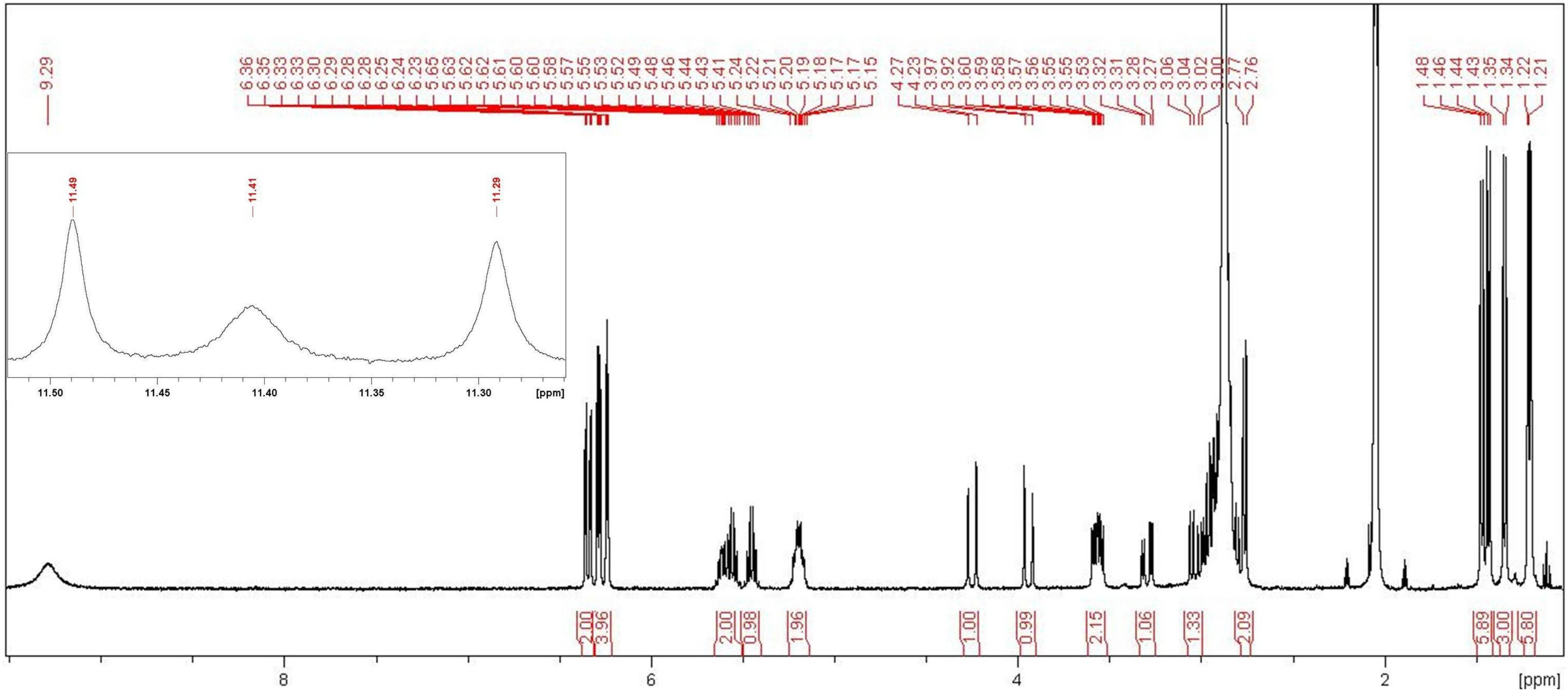


Fig. S8: ^{1}H-NMR spectrum of compound **(3)**, ascotrichalactone A, in acetone-d_6

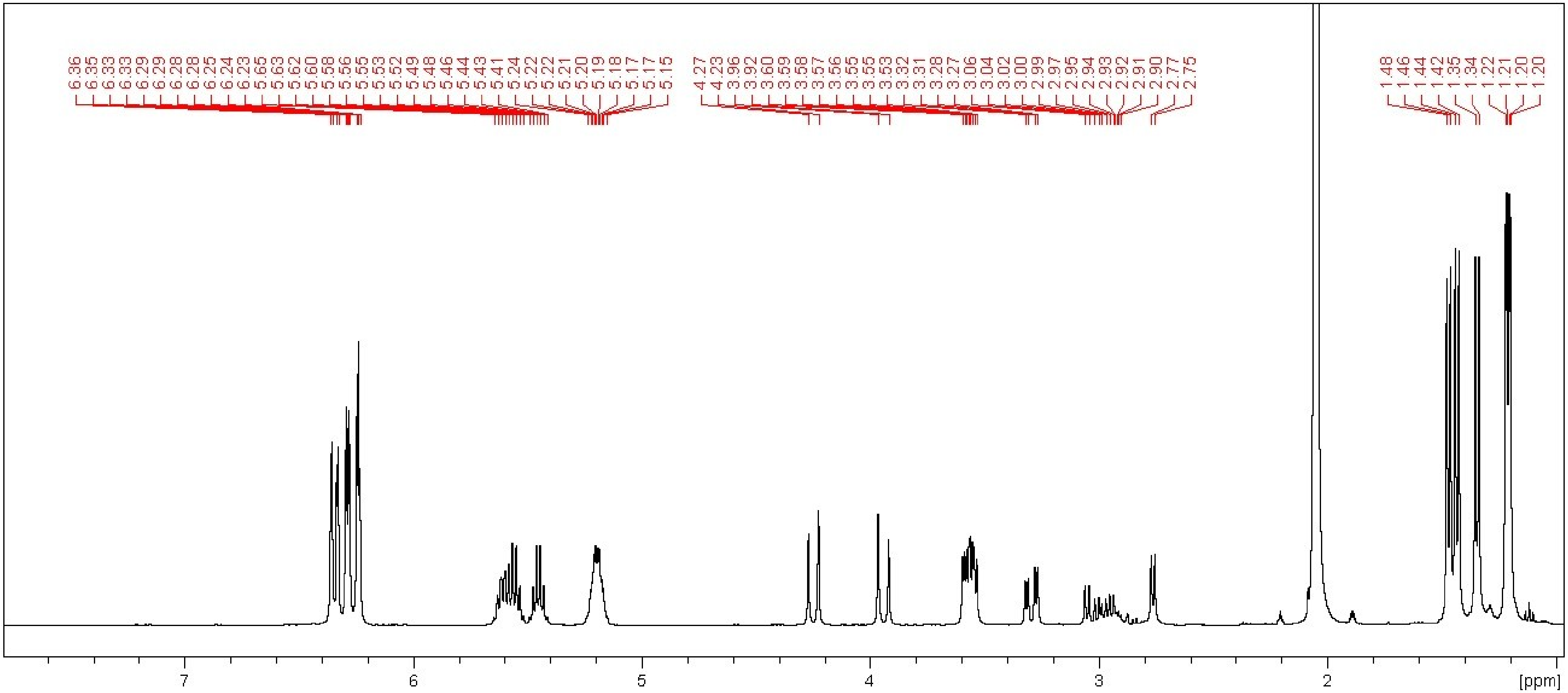


Fig. S9: ^{1}H-NMR spectrum with water suppression of compound **(3)**, ascotrichalactone A, in acetone-d_6

References

1. Y. F. Hallock, J. Clardy, D. S. Kenfield and G. Strobel, *Phytochemistry*, 1988, **27**, 3123-3125.
2. J.-L. Zhang, W.-J. Wang, X.-M. Xu, D.-Y. Li, H.-M. Hua, E.-L. Ma and Z.-L. Li, *Tetrahedron*, 2016, **72**, 4895-4901.
3. M. S. Islam, K. Ishigami and H. Watanabe, *Tetrahedron*, 2007, **63**, 1074-1079.