DOI: 10.1002/cbic.200700390

Authentic Heterologous Expression of the Tenellin Iterative Polyketide Synthase Nonribosomal Peptide Synthetase Requires Coexpression with an Enoyl Reductase

Laura M. Halo,[a] James W. Marshall,[a] Ahmed A. Yakasai,[a] Zhongshu Song,[a] Craig P. Butts,[a] Matthew P. Crump,[a] Mary Heneghan,[b] Andrew M. Bailey,[b] Thomas J. Simpson,[a] Colin. M. Lazarus,[b] and Russell J. Cox*[a]

The tenS *gene encoding tenellin synthetase (TENS), a 4239-residue polyketide synthase nonribosomal-peptide synthetase (PKS-NRPS) from* Beauveria bassiana*, was expressed in* Aspergillus oryzae *M-2-3. This led to the production of three new compounds, identified as acyl tetramic acids, and numerous minor metabolites. Consideration of the structures of these compounds indicates that the putative C-terminal thiolester reductase (R) domain does not act as a reductase, but appears to act as a Dieckmann cyclase (DKC). Expression of* tenS *in the absence of a trans-acting ER component encoded by* orf3 *led to errors in assembly of the polyketide component, giving clues to the mode of programming of highly reducing fungal PKS. Coexpression of* tenS *with* orf3 *from the linked gene cluster led to the production of a correctly elaborated polyketide. The NRPS adenylation domain possibly shows the first identified fungal signature sequences for tyrosine selectivity.*

Introduction

The filamentous fungi are prodigious producers of secondary metabolites, and compounds derived from the polyketide and nonribosomal peptide pathways are particularly common. Fungi also produce numerous compounds derived by a combination of these two pathways. For example, 2-pyrrolidinones such as fusarin C (**1**, *Fusarium venenatum*),[1] equisetin (**2**, *Fusarium heterosporum*),[2] chaetoglobosin A (**3**, *Penicillium expansum*)[3,4] and pseurotin A (**4**, *Pseudeurotium ovalis*[5] and *Aspergillus fumigatus*[6]) have been shown by classical isotope feeding experiments to be derived from polyketide and amino acid components. Other compounds such as the fungal 2-pyridones including tenellin (**5**, *Beauveria bassiana*),[7] leporin A (**6**, *Aspergillus leporis*)[8] and aspyridone A (**7**, *Aspergillus nidulans*)[9] are also derived from amino acids and polyketides, as are a very wide range of similar compounds.[10]

Recently we showed that precursors of fusarin C (**1**) and tenellin (**5**) are synthesised by giant enzymes consisting of an iterative type I polyketide synthase (PKS) fused to a single nonribosomal peptide synthetase (NRPS) module (Scheme 1). Thus, the first steps of fusarin biosynthesis are carried out by the fusarin synthetase (FUSS),[11] while tenellin synthetase (TENS) must carry out the first steps of tenellin biosynthesis.[12] Parallel results have been obtained in the case of the 2-pyrrolidinones equisetin (**2**),[2] pseurotin A[6] (**4**) and the 2-pyridone aspyridone A (**7**).[9] Very recently Schümann and Hertweck have linked the biosynthesis of chaetoglobosin A (**3**) with a similar PKS-NRPS.[13] These synthetases are encoded by ca 12 kb genes which occur frequently in the known genomes of filamentous fungi. In some cases such genes have been linked to surprising and new biological activities—a good example is the PKS-NRPS gene *ace1* which is involved in avirulence signalling in the plant pathogen *Magnaporthe grisea*.[14]

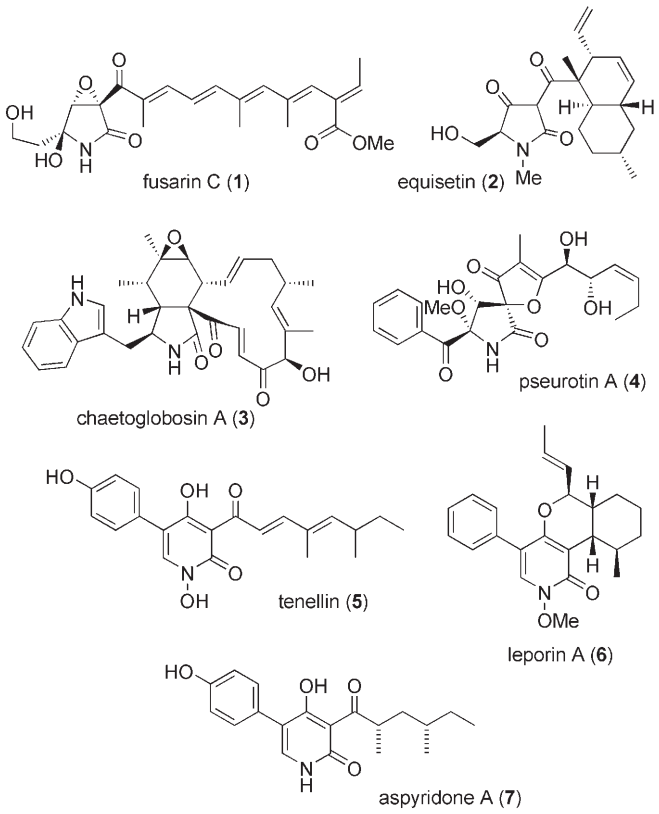
The PKS-NRPS genes generally occur within extended gene clusters. In the case of tenellin (Figure 1), at least three other putative biosynthetic genes are clustered along with the PKS-NRPS (*orf4*, *tenS*). In *B. bassiana* these genes encode two proteins homologous to cytochrome P450 monooxygenases (ORF1 and ORF2) and a putative Zn-dependent oxidoreductase (ORF3) which is a member of the family of trans-acting enoyl reductases.[15]

Sequence analysis of the PKS-NRPS genes suggests that the biosynthetic proteins consist of separate catalytic domains (Figure 1). At the N terminus is the iterative type I PKS usually consisting of β-ketoacyl ACP synthase (KAS), acyl transferase (AT), dehydratase (DH), C-methyltransferase (C-MeT), enoyl-reductase (ER), ketoreductase (KR) and acyl carrier protein (ACP). In this respect the PKS is similar to other highly reducing PKSs (HR-PKSs) known from the biosynthesis of lovastatin[16] and squalestatin,[17] for example. Immediately downstream of the

[a] *L. M. Halo, J. W. Marshall, A. A. Yakasai, Dr. Z. Song, Dr. C. P. Butts, Dr. M. P. Crump, Prof. T. J. Simpson, Dr. R. J. Cox*
School of Chemistry, University of Bristol, Cantock's Close
Bristol, BS8 1TS (UK)
Fax: (+44) 117-929-8611
E-mail: r.j.cox@bris.ac.uk

[b] *Dr. M. Heneghan, Dr. A. M. Bailey, Dr. C. M. Lazarus*
School of Biological Sciences, University of Bristol
Woodland Road, BS8 1UG (UK)

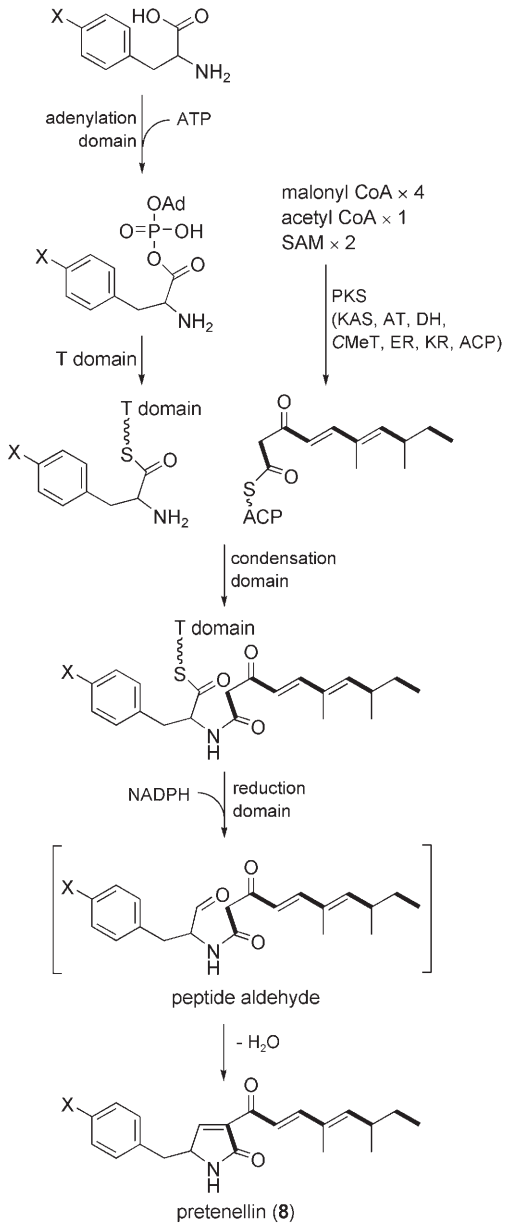
Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.



PKS is the NRPS consisting of condensation (C), adenylation (A), thiolation (T) and putative thiolester reduction (R) domains.[12]

With this knowledge in hand it is possible to speculate about the catalytic activities of the PKS-NRPS proteins. For example, we have hypothesised that pretenellin (**8**) should be the product of TENS (Scheme 1).[12] However, such predictions cannot yet be made with high confidence because of ambiguities in the results of feeding experiments and the present lack of knowledge regarding the *selectivities* of the catalytic domains. For tenellin these uncertainties encompass the adenylation domain because isotope feeding experiments suggest that both phenylalanine and tyrosine are efficient precursors.[18] Furthermore, there are no firmly established rules for predicting A domain selectivity in fungal NRPS systems as there are for bacterial A domains. There is also uncertainty regarding the cyclisation step—it is not known how, or when, this occurs. Questions also surround the PKS as the structure of tenellin requires an ER-mediated step in the first condensation cycle during tenellin biosynthesis but sequence analysis (vide infra) suggests that the ER domain may be inoperative.

In order to clarify the exact role of the TENS PKS-NRPS protein we expressed *B. bassiana tenS* in the heterologous host *Aspergillus oryzae* M-2-3 using the pTAex3 expression system.[19] The use of *A. oryzae* M-2-3 together with pTAex3 to determine the role of various fungal PKS genes has been pioneered by Ebizuka, Fujii and others,[20] and we have used it to express the tetraketide synthases involved in squalestatin and methylorcinaldehyde biosynthesis.[17,21] However, this strategy has not yet been reported in the case of the PKS-NRPS genes.



Scheme 1. The previously proposed catalytic activities and product of TENS. See text for domain abbreviations. CoA, coenzyme A ; SAM, *S*-adenosylmethionine.

Results

We used a dual recombination strategy for first assembling the required PKS-NRPS gene and then transferring it to a derivative of pTAex3 (See the Supporting Information). *A. oryzae* M-2-3 (*argB*$^-$) was transformed with the expression plasmid pTAex3-*tenS* for expression trials; transformation with an empty pTAex3 vector provided a control. Selection was carried out on minimal medium lacking arginine, allowing discrimination of the desired clones which carry the *argB* gene, and 11 independent transformants were isolated. Three of these were confirmed then transferred to liquid medium and grown in shake culture. After 10 days the mycelia and medium were homogenised, acidified to pH 3 and the mixture extracted with ethyl

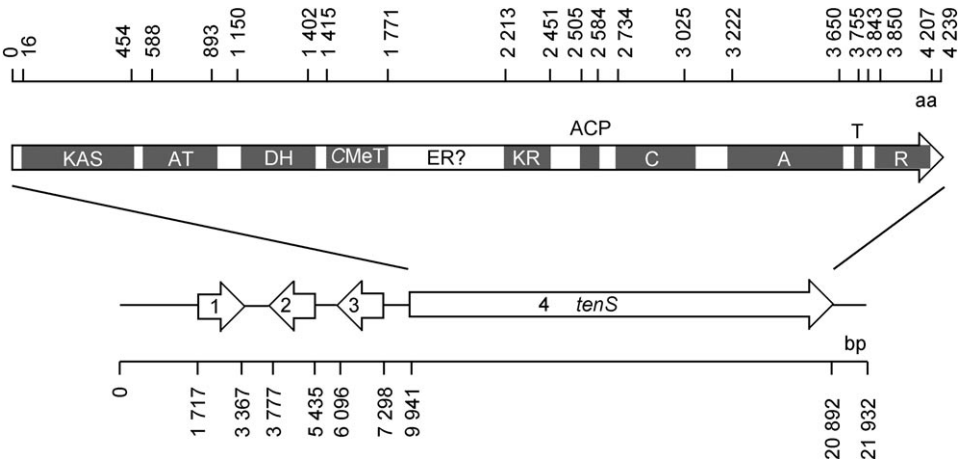


Figure 1. The organisation of the tenellin biosynthetic gene cluster in *B. bassiana. tenS* encodes a 4239-amino acid synthetase. The approximate domain boundaries are shown.[12]

acetate (EtOAc). The organic extracts were concentrated, defatted, reextracted from water into EtOAc, evaporated again and dissolved in methanol (MeOH) prior to LCMS analysis. Chromatograms of extracts from *A. oryzae* pTAex3-*tenS* expression clones were compared with chromatograms of extracts made from the *A. oryzae* pTAex3 control.

LCMS analysis identified several new peaks that were not present in untransformed (UT) *A. oryzae* M-2-3 or in *A. oryzae* M-2-3 pTAex3 grown under identical conditions (Figure 2). These compounds proved difficult to isolate by normal phase (flash) chromatography, but a combination of preparative thin-layer and reverse-phase chromatography eventually yielded three new compounds (Figure 3). Prototenellin A (**9A**; EI, $[M]^+$, $C_{21}H_{23}NO_4$, calcd 353.1627, found 353.1631) was isolated as the major metabolite in 3.8 mg L^{-1} yield. Prototenellin B (**9B**; EI, $[M+Na]^+$, $C_{18}H_{19}NO_4Na$, calcd 336.1206, found 336.1220) was the next most abundant compound (3.6 mg L^{-1}). Prototenellin C (**9C**; EI, $[M+Na]^+$, $C_{21}H_{25}NO_6Na$, calcd 410.1580, found 410.1595) was more difficult to isolate, but was eventually obtained in 3.0 mg L^{-1}. Further compounds were present in lower titres, but these could not be isolated in sufficient quantity to be characterised.

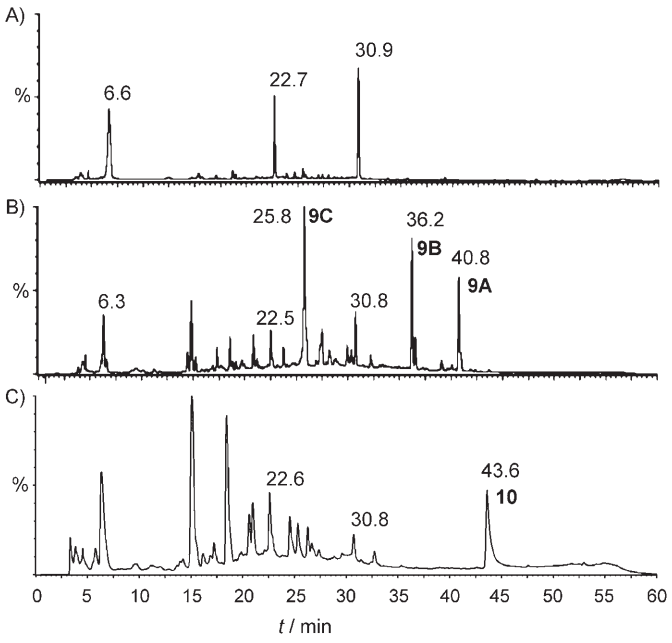


Figure 2. Comparison of HPLC traces of organic extracts from: untransformed *A. oryzae* M-2-3 (A); *A. oryzae* M-2-3 pTAex3-*tenS* (B); and A. oryzae M-2-3 pTAex3-*tenS*/pTAex3-*orf3* (C) strains (diode array, 200–400 nm). See the Experimental Section for extraction and HPLC details.

Prototenellin A (**9A**) displayed characteristic ^{1}H NMR signals corresponding to the tyrosine-derived moiety of the molecule—in particular, diagnostic aromatic doublets indicated the presence of the *para*-substituted phenyl ring, while diastereotopic benzylic protons showed distinctive coupling to the adjacent methine proton in the COSY spectrum (Table 1). In the polyketide moiety a distinctive coupling pattern showed that H-9 was adjacent to both H-8 and H-10. Long-range coupling between H-14 and H-10 and H-12 placed methyl-14 as attached to C-11, and similar long-range coupling between H-15 and H-8 placed methyl-15 attached to C-7. These assignments were further confirmed by COSY H,H and HMBC, HSQC and HSQC-DEPT C,H correlations (Figure 3).

Prototenellin B (**9B**) was identified by a similar series of NMR experiments in which the tyrosine-derived moiety was rapidly identified and the structure of the simplified polyketide was determined by short-range and long-range ^{1}H,^{1}H coupling. Once again HMBC and HSQC correlation spectra confirmed the connectivity.

Prototenellin C (**9C**) was more difficult to fully identify because of the lack of material and instability during purification. HRMS analysis suggested that prototenellin C (**9C**) was related to prototenellin A (**9A**) by the formal addition of two water molecules and by one oxidation. Prototenellin C (**9C**) was insoluble in $CDCl_3$, and spectra obtained in CD_3OD were difficult to interpret because of extensive exchange of deuterium and the presence of tautomers. Spectra obtained in CD_3CN also suggested that the differences between **9A** and **9C** were located in the polyketide moiety of the compound, but significant signal overlap in the aliphatic region of the spectrum prevented further use of CD_3CN as a useful NMR solvent.

When **9C** was dissolved in CD_3OD, ^{1}H NMR showed the presence of the tyrosine-derived moiety, and chemical shifts for H-16 and H-5 suggested that the tetramic acid moiety was also intact. ^{1}H and ^{13}C NMR signals and HMBC correlations for positions 11–14 supported the presence of the hydroxyl butyryl moiety shown for **9C** in Figure 3, although possible deuterium exchange occurred at C-9 and C-11 (See the Supporting Information). However, doubling of signals possibly due to diastereoisomerism and/or tautomerism prevented further elucida-

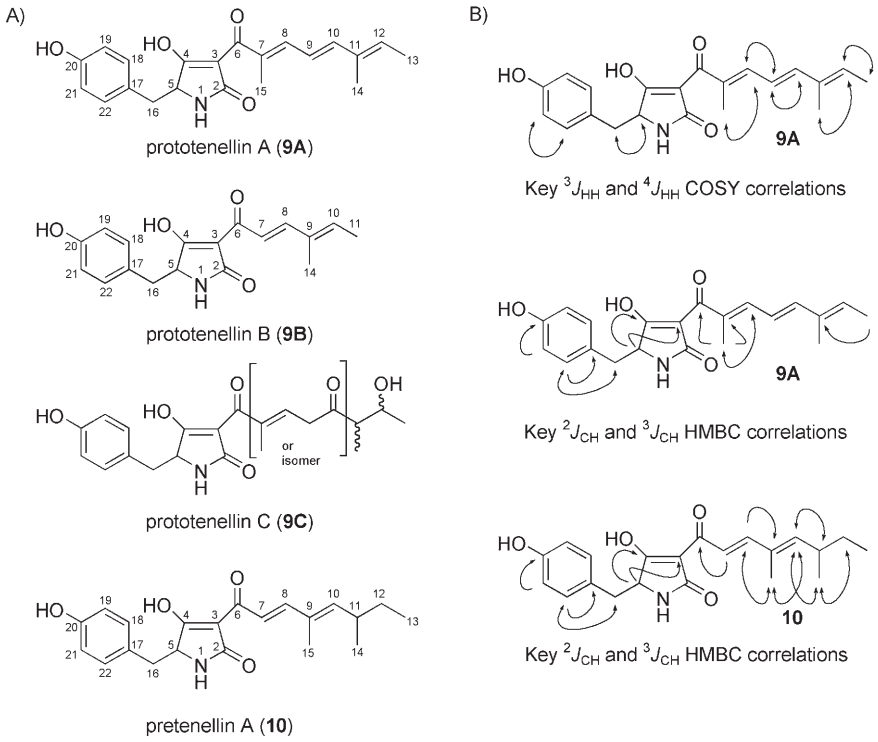


Figure 3. A) Structures of major compounds isolated from *A. oryzae* M-2-3 transformed to express *tenS* (**9A**–**C**) or to co-express *tenS* and *orf3* (**10**). B) Key correlations from 2D NMR experiments for **9A** and **10**.

tion of ^1H and ^{13}C chemical shifts for the remaining polyketide atoms. It is likely that **9C** has the structure shown in Figure 3.

Other compounds were produced by TENS in *A. oryzae* but low concentration prevented adequate NMR characterisation. LCMS analysis (Figure 4), however, indicated that two of the minor compounds eluting close to **9A** (*m/z* 354) differed from it by 14 mass units corresponding to a single methyl group. The compound eluting at 40.4 min appeared to be monomethylated (*m/z* 340), while that eluting at 42.7 min appeared to be trimethylated (*m/z* 368). Three compounds elute with *m/z* 354 corresponding to dimethylated compounds (39.4, 41.1 and 42.3 min). There also appears to be an isomer of **9B**, eluting at 36.9 min, possibly representing a methylation regioisomer. Isomers of **9C** were also observed, although there was no *m/z* evidence for related compounds with differing numbers of methyl groups in this case.

In order to probe the activity of the putative trans-acting ER encoded by *orf3*, we constructed a pTAex3 derivative in which selection was provided by a basta-resistance (bar) cassette, and the *argB* gene was inactivated by truncation. This allowed co-transformation of the *tenS*/argB vector with the *orf3* bar vector into *A. oryzae*. Dual selection was achieved by plating the

Table 1. ^1H (400 MHz, $CDCl_3$) and ^{13}C (100 MHz, $CDCl_3$) NMR data for new compounds **9A**, **9B** and **10**.

	Prototenellin A (**9A**)			Prototenellin B (**9B**)			Pretenellin A (**10**)		
Pos.	δ_H	*J* [Hz]	δ_C	δ_H	*J* [Hz]	δ_C	δ_H	*J* [Hz]	δ_C
1	5.79	s	–	5.95	s	–	5.84	br s	–
2	–		177.7	–		175.5	–		176.5
3	–		96.1	–		98.4	–		98.6
4[a]	–		192.5	–		194.2	–		194.3
5	3.96	dd, 10.1, 3.4	62.6	3.98	dd 10.2, 3.2	64.5	4.00	dd, 10.5, 2.9	63.6
6	–		184.8	–		176.0	–		176.5
7	–		127.4	7.19	d, 15.5	115.1	7.16	d, 15.5	115.8
8	8.06	d, 11.2	144.6	7.55	d, 15.5	151.7	7.58	d, 15.5	151.5
9	6.54	dd, 15.0, 11.2	121.3	–		135.0	–		133.5
10	6.78	d, 15.0	147.9	6.20	q, 7.1	142.1	5.91	d, 9.3	153.2
11	–		135.5	1.88	d, 7.1	16.0	2.53	m	35.0
12	5.91	q, 7.3	134.2	–		–	1.27	m	29.4
							1.45	m	
13	1.84	d, 7.3	14.5	–		–	0.89	t, 7.3	11.7
14	1.86	s	12.4	1.91	s	11.3	1.04	d, 6.6	19.8
15	2.05	s	11.9	–		–	1.94	s	11.9
16	3.26	dd, 13.9, 3.4;	37.8	3.20	dd, 13.9, 3.2	38.0	3.25	dd, 14.9, 2.9	37.5
	2.62	dd, 13.9, 10.1		2.60	dd, 13.9, 10.2		2.64	dd, 10.5, 14.9	
17	–		130.7	–		128.4	–		129.9
18/22	7.09	d, 8.0	130.3	7.05	d, 7.2	130.4	7.09	br m	130.7
19/21	6.80	d, 8.0	115.8	6.78	d, 7.2	115.5	6.80	br m	115.8
20	–		154.7	–		155.4	–		156.1

[a] C-4 was weak in the 1D ^{13}C spectrum for both **9A** and **9B** but was located in 2D spectra by HMBC correlation to H-5. ^{13}C chemical shifts were also verified by comparison to literature values for related compounds.[18,43] Sample concentrations were between 1 and 3 mg mL^{-1}.

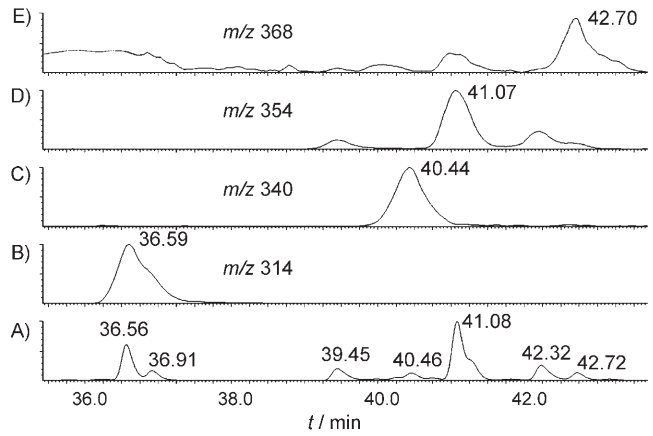


Figure 4. LCMS chromatograms of secondary metabolites derived from *A. oryzae* M-2-3 pTAex3-*tenS*. A) Diode array trace (200–400 nm); B)–E), single ion monitoring (ES^+) for protonated ions ($[M+H]^+$) corresponding to singly methylated tetraketides (*m/z* 314), monomethylated pentaketides (*m/z* 340), dimethylated pentaketides (*m/z* = 354) and trimethylated pentaketides (*m/z* 368).

transformation mixture on a minimal medium lacking arginine, and overlaying after 24 h with basta, followed by subculturing single colonies onto basta-containing minimal medium lacking arginine. Eleven clones were isolated which showed both basta resistance and arginine prototrophy. These clones were examined by PCR of genomic DNA which confirmed the presence of both *tenS* and *orf3*. Five clones were grown in production medium and organic extracts were examined by LCMS. Four clones produced a single new compound (*m/z* 356.1842 ($[M+H]^+$), calcd for $C_{21}H_{26}NO_4$ 356.1862) which was purified by HPLC. Pretenellin A (**10**), was identified by ^1H and ^{13}C NMR. In comparison with tenellin itself, **10** showed very similar ^1H chemical shifts and an identical coupling pattern for the polyketide moiety (C-7 to C-15). The tetramic acid moiety showed near identical chemical shifts to **9A**–**C** (C-3 to C-5 and C-16 to C-20). Both polyketide and tetramic acid moieties showed very similar ^1H and ^{13}C chemical shifts to a synthetic analogue of **10** lacking the aromatic *p*-hydroxy group.[18] Final structural confirmation was provided by HSQC-DEPT and HMBC experiments.

Discussion

Expression of *tenS* in *A. oryzae* has allowed us to determine the function of a fungal PKS-NRPS in isolation for the first time. Other recent work from the groups of Schmidt,[2] Hertweck[13] and Turner[6] have investigated the activity of fungal PKS-NRPS proteins in the context of other proteins from the biosynthetic pathway. Our experiments, however, show that TENS behaves differently in the presence and absence of the ORF3 trans-acting ER, and examination of the chemical structure of products can also shed light on the proposed activities of other biosynthetic proteins encoded in the tenellin biosynthetic gene cluster. The structures of compounds **9A**–**C** and **10** reveal key aspects of the programming of fungal PKS-NRPS proteins, especially in relation to polyketide biosynthesis, the ring closing reaction and amino acid selection.

The programming of polyketide biosynthesis

Fungal polyketide synthases can be classified into three types: nonreducing synthases (NR-PKS) responsible for the biosynthesis of compounds such as tetrahydroxynaphthalene and norsolorinic acid; partially reducing synthases (PR-PKS) which make 6-methylsalicylic acid; and highly reducing synthases (HR-PKS) which make complex polyketides such as squalestatin and lovastatin.[22]

The PKS components of TENS closely match other HR-PKS. In addition to KAS, AT and ACP domains, TENS has domains putatively responsible for C-methyl transfer, keto-reduction, dehydration and enoyl-reduction (ER). Other HR-PKS, such as the squalestatin tetraketide synthase (SQTKS)[17] and the lovastatin diketide synthase (LDKS),[16] are known to possess functional ER domains. The ER domain of the lovastatin nonaketide synthase (LNKS), however, is known to be inoperative; the necessary enoyl-reductase activity is supplied by a protein encoded by *lovC* which has been shown to act *in trans* with LNKS to produce the reduced nonaketide intermediate.[16] Close inspection of TENS suggests that its ER domain might also be inoperative. In support of this, phylogenetic analysis shows that the putative ER domain from TENS more closely matches that of LNKS and aspyridone synthetase (ASPS), both of which possess inoperative ER domains, than those of LDKS, SQTKS and other synthases for which the ER domain has been shown to be functional (Figure 5).

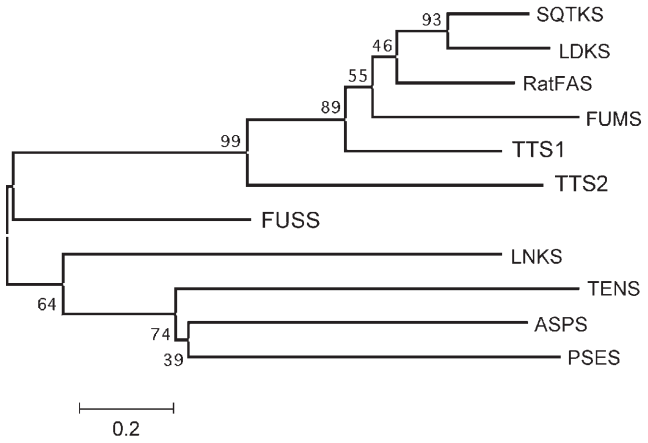


Figure 5. Phylogenetic analysis of putative ER domains of fungal PKS and PKS-NRPS: Protein names (peptide sequence used for comparison): SQTKS, squalestatin tetraketide synthase; LDKS, lovastatin diketide synthase; RatFAS, rat fatty acid synthase; FUMS, fumonisin synthase; TTS1, T-toxin synthase 1; TTS2, T-toxin synthase 2; FUSS, fusarin synthetase; LNKS, lovastatin nonaketide synthase; TENS, tenellin synthetase; ASPS, aspyridone synthetase; PSES, pseurotin synthetase. Selected ER domain sequences were extracted and aligned using ClustalX. The alignment was then subjected to boot-strapped phylogenetic analysis using the Neighbour-joining analysis in MEGA3.[41,42]

A single enoyl-reduction step is required after the first chain-extension condensation reaction during the biosynthesis of the tenellin polyketide. It is clear from the fully unsaturated side-chains of prototenellin A (**9A**) and prototenellin B (**9B**) that TENS itself is not capable of carrying out this step. The

tenellin gene cluster contains a gene encoding a *lovC* homologue (ORF3) and we have assumed that this trans-ER is responsible for correct reduction during tenellin biosynthesis.[12] Vederas and co-workers observed that expression of *lovB* encoding LNKS in *A. terreus* led to the production of polyunsaturated pyrones from truncated polyketide intermediates.[16] Thus, it appears that in the absence of the lovC protein, LNKS is incapable of extending the chain to the required nonaketide, producing only a hexaketide and a heptaketide. However, Vederas observed that LNKS was able to carry out the correct C-methylation event during the third chain extension (the first cycle requiring the action of ER), but thereafter, in the absence of lovC, correct programming broke down.

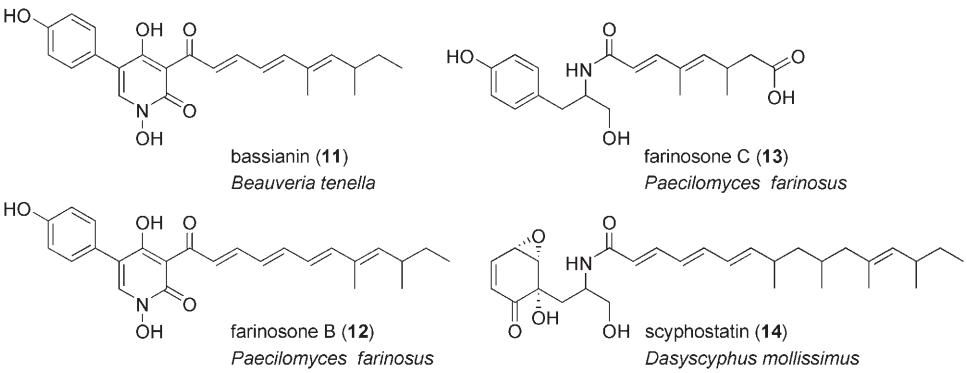
TENS appears to act similarly in the absence of the trans-ER that is the likely product of *orf3* expression—both pentaketide (e.g., **9A**) and tetraketide (e.g., **9B**) products are observed. The timing of the C-methylation programming also breaks down, with TENS performing its second methylation after the *third* extension, rather than after the second extension as required for tenellin biosynthesis. LCMS analysis of crude extracts also indicates that the fidelity of methylation is disrupted with monomethyl and trimethyl isomers of **9A** detected at low concentration, along with possible dimethyl regioisomers.

The structure of prototenellin C (**9C**) has not been fully elucidated, but HRMS and partial ^{1}H NMR analysis indicates it is most likely to be one of the three isomers shown in Figure 3 bearing a more highly oxygenated side chain than tenellin. Regardless which of these is the actual structure of **9C**, it is clear that the production of **9C** indicates further loss of control in programming by TENS in the absence of ORF3. Here, in common with the LNKS observations,[16] the polyketide moiety is more highly oxidised indicating the failure of later KR and DH steps after the first extension when the ER step is not carried out.

Coexpression of *tenS* with *orf3* leads to the production of a single compound, pretenellin (**10**). This compound can be considered as a genuine precursor of tenellin. LCMS analysis did not reveal the presence of either isomeric or chain-shortened analogues of **10** or any of the unsaturated prototenellin compounds **9A**–**C**. This shows that as well as performing the required ER reaction after the first condensation, programming and fidelity of TENS are returned to normal in the presence of the ORF3 protein.

Thus in the case of tenellin biosynthesis, at least, lack of the first enoyl reduction can affect the programming of all subsequent PKS steps. Presence of the trans-ER appears to be crucial for correct PKS programming in TENS. Two possible ways of affecting the PKS programming can be envisaged for this protein. In one scenario the ER protein could be involved in the first modification cycle only, and the subsequent changes in programming would be guided by the structure of the growing polyketide itself. The second possibility would involve the ER remaining bound to TENS during all rounds of extension and processing, allowing its influence to be felt throughout the entire biosynthetic cycle. Further experiments will have to be performed to differentiate between these possibilities.

Differences in the trans-ER may also explain the wide range of compounds known which are related to tenellin. These include bassianin (**11**, *Beauveria tenella*)[7] and farinosone B (**12**, *Paecilomyces farinosus*).[23] Consideration of these compounds, and the diverse range of compounds isolated from *A. oryzae* expressing *tenS*, indicates that the NRPS domains of TENS and related PKS-NRPS proteins operate independently of the PKS as amino-acid addition and ring closing reactions appear to proceed correctly with all observed polyketides.



Ring closing

The NRPS modules of TENS, FUSS and other fungal PKS-NRPS synthetases terminate with a putative reductive (R) domain. In their entirety, the A, T and R domains show end-to-end homology to the fungal α-aminoadipate reductase (AAR) complex involved in lysine biosynthesis encoded by *lys2*.[24,25] The A domain of AAR selects α-aminoadipic acid, activates the ε-carboxylic acid to an acyl adenylate and transfers it to the phosphopantetheine (PP) thiol of the T domain to form an enzyme-linked thiolester. The R domain then utilises NADPH to perform a reduction of the thiolester to release aminoadipate-ε-semialdehyde, regenerating the PP thiol (See Scheme S1 in the Supporting Information). By analogy to this chemistry we speculated that the TENS PKS-NRPS complexes would activate and attach phenylalanine or tyrosine to its own PP thiol. This must then be condensed with the ACP-linked polyketide, by the C domain, to form a T domain-linked amide. Action of the R domain could then reductively release this as a peptide-aldehyde which would be reactive for a Knoevenagel-type ring-forming reaction—yielding a compound such as **8** (Scheme 1).

Reductive release mechanisms have been observed for a number of NRPS-type systems: for example Müller and co-workers described the MxcG component of the myxochelin

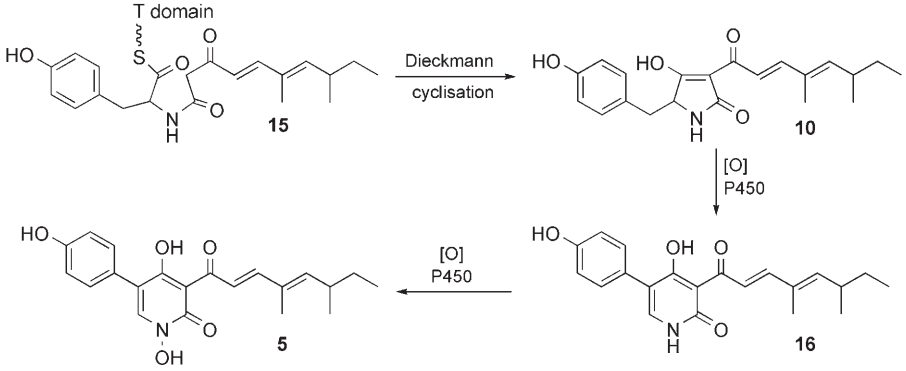
synthetase from the myxobacterium *Stigmatella aurantica* Sg a15 which reduces a lysine-derived carboxylate twice to form a primary alcohol.[26] In the NRPS responsible for gramicidin biosynthesis in *Bacillus brevis* ATCC 8185, the terminal R domain of LgrD[27] releases the product as an aldehyde, which is further reduced by a second reductase encoded by *LgrE* located upstream of the NRPS genes.[28] We have also recently reported a fungal NR-PKS (methylorcinaldehyde synthase, MOS) from *Acremonium strictum* which uses a reductive release mechanism catalysed by a C-terminal R domain.[21]

Multiple alignment of the thiolester reductive domains from *Candida glabrata* AAR, LgrD, MxcG and the *Acremonium strictum* MOS R domains with the TENS putative R domain support the idea that this domain could be involved in a reductive process (See the Supporting Information). In particular an NAD(P)H binding motif (ILLTGATGFLGAXLLEXLL) is well conserved in TENS. A similar sequence is present in the known thiolester reductases, as well as in an aldehyde reductase from *Sporobolomyces salmonicolor* (1Y1P),[29] which has had its structure solved by X-ray crystallography.[29]

However, a problem with the reductive-release hypothesis is that this mechanism would provide a product (e.g., **8**) at a lower oxidation level than observed in the final metabolite. In the case of tenellin, reoxidation at C-4 would be required to provide a substrate suitable for further oxidative ring expansion.[12] In the case of equisetin (**2**), the final product is also at the ester oxidation level at C-4, again indicating that if a reductive release and cyclisation were used by the equisetin synthetase (EQS) a further oxidation step would be needed. The recently reported aspyridone synthetase (ASPS) also has an apparent C-terminal reductase but produces a product at the tetramic acid/hydroxypyridone oxidation level. Hertweck and co-workers have proposed the involvement of one of two P450-encoding genes in the ASPS biosynthetic cluster in the oxidation of the proposed pyrrolinone ring intermediate.[9]

The new compounds **9A**–**C** isolated after the expression of *tenS* alone, and pretenellin A (**10**) from *tenS*/*orf3* coexpression, all retain the enol functionality at C-4, which would be consistent with direct formation of the tetramic acid ring by a Dieckmann condensation of the thiolester, without intervention of the reductase. It is therefore unclear what role the putative R domain plays during tenellin biosynthesis. It is conceivable that there could be a reductive release, followed by a reoxidation process (either by the R domain itself or by adventitious enzymes in vivo), although this would not be an optimal release strategy. Intriguingly the NADPH binding motif of the TENS putative R domain is similar to that of *Pseudomonas aeruginosa* UDP-*N*-acetylglucosamine 4-epimerase, which achieves its reaction by an oxidation/reduction sequence.[30] Adventitious reoxidation by a cytoplasmic enzyme cannot be ruled out at this stage, but we regard this as unlikely because of the similarities in proposed cyclisation in a number of different fungi.

However, it is more likely that the putative R domain catalyses a Dieckmann cyclisation (DKC) of the bound *N*-β-ketoacyl α-aminothiolester intermediate (Scheme 2) which would achieve both cyclisation and release. Such DKC reactions have analogues in synthetic chemistry where *N*-acetoacetyl α-amino-esters are good substrates for base-catalysed cyclisation.[10,18] Such a process, however, would have no requirement for NADPH.



Scheme 2. Proposed biosynthesis of tenellin in *B. bassiana*.

An alternative and additional potential role for the N-terminal domain could be as an editing device. For the Dieckmann cyclisation to occur the polyketide bound to the PKS ACP must possess a β-carbonyl. In cases in which the PKS does not provide a polyketide correctly functionalised at the β-carbon atom, the PKS-NRPS would be blocked by a covalently linked acyl group attached to the T domain. Reduction of this would release an aldehyde or alcohol and free-up the synthetase. Tentative support for this hypothesis comes from the known fungal metabolites farinosone C (**13**, *Paecilomyces farinosus*)[23] and scyphostatin (**14**, *Dasyscyphus mollissimus*).[31]

Amino acid selection

Both phenylalanine and tyrosine have been shown to be incorporated into tenellin.[18,32,33] Phenylalanine could be a direct substrate for TENS giving a benzyl-substituted 2-pyrollidinone which could be hydroxylated later—possibly by one of the two P450 monooxygenases encoded in the tenellin cluster. Alternatively, phenylalanine could be hydroxylated 'early' to give tyrosine which is the substrate for TENS. It is also conceivable that TENS could accept both phenylalanine and tyrosine as substrates, and that products derived from phenylalanine could be oxidised later by one of the P450 monooxygenases.

Expression of *tenS*, in the absence of other genes from the tenellin gene cluster, shows that TENS selects tyrosine as its substrate. Amino acid selection is performed by the adenyla-

Table 2. Signature residues for amino acid selection by adenylation domains. Sequence numbering from Gramicidin-S synthetase (GrsA) A domain 1.[36]

	Signature position									
Sequence (selectivity)	235	236	239	278	299	301	322	330	331	517
gramicidin-S GrsA A1 (Phe)[36]	D	A	W	T	I	A	A	I	C	K
tyrocidine A TycA A1 (Phe)[37]	D	T	F	T	I	A	A	I	C	K
tyrocidine A TycC A3 (Tyr/Trp)[37]	D	A	L	T	T	G	E	V	V	K
fengycin FenA A3 (Tyr)[38]	D	G	T	I	T	A	E	V	A	K
TENS (Tyr)[12]	D	M	V	I	P	P	C	A	A	K
ASPS (Tyr)[9]	D	M	V	I	Y	W	C	A	A	K
A. oryzae (?)	D	M	A	I	R	W	C	A	A	K
PSES (Phe)[6]	D	A	Y	T	S	W	A	I	C	K

tion (A) domain. Comprehensive structure and sequence analysis of *bacterial* A domains by Marahiel[34] and Challis and Townsend[35] has shown that the amino-acid binding pocket is composed of ten key amino acids (Table 2). Because these amino acids are involved in substrate selection it has been possible to draw up signature sequences which can predict which amino acids a given A domain will select. For example a bacterial A domain with selectivity for phenylalanine is the gramicidin-S GrsA A1 domain.[36] The signature sequence of this domain shows clear similarities to other bacterial A domains which select phenylalanine, such as tyrocidine TycA A1 (Table 2).[37] Similarly, A domains with selectivity for tyrosine are known, such as tyrocidine TycC A3[37] and fengycin FenA A3.[38] In these cases, Challis and Townsend have suggested that E322 may be involved in hydrogen-bonding to the aromatic hydroxyl of tyrosine.[35]

However, in the case of fungal A domains this analysis has been harder to perform because fungi appear to use signature sequences which have diverged significantly from bacterial sequences. Additionally, there is a relative lack of data relating fungal A domains with their known substrates—although limited progress has been made recently. For this reason comparison of the TENS A domain signature sequence to those of phe- and tyr- activating bacterial A domains is unproductive (Table 2).

Hertweck and co-workers very recently showed that a cryptic PKS-NRPS biosynthetic gene cluster in *A. nidulans* encodes proteins which synthesise aspyridone A (**7**).[9] This compound appears to utilise tyrosine as its amino acid component. We therefore compared the A domain signature sequences of TENS and the aspyridone synthetase (ASPS, Table 2). These show remarkable similarity, and for the first time may indicate a fungal A-domain signature sequence for tyrosine. Genome sequencing has revealed the sequences of a number of other fungal PKS-NRPS biosynthetic genes. Among these, *A. oryzae* itself encodes a cryptic PKS-NRPS whose A domain also contains a remarkably similar signature sequence (Table 2) to the TENS and ASPS A domains. Transcript analysis by us (data not shown) and others[39] indicates that this PKS-NRPS is not expressed during *tenS* expression in *A. oryzae* M-2-3. It appears likely that the *A. oryzae* PKS-NRPS would also utilise tyrosine. Turner and co-workers have recently identified the pseurotin A PKS-NRPS (PSES) from *Aspergillus fumigatus*, which appears to select phenylalanine.[6] Interestingly, analysis of the PSES A-domain signature sequence shows that this is more similar to the bacterial Phe-selecting A domains than to the fungal Tyr-selecting A domains.

Conclusions

Expression of *tenS* in *A. oryzae* M-2-3 reveals key information about fungal A domain selectivity, HR-PKS programming and the role of putative R domains during product release from fungal PKS-NRPS proteins. We have shown that in vivo the PKS of TENS, acting in concert with the trans-ER encoded by *orf3*, most likely produces a doubly methylated pentaketide attached to its ACP. In parallel, the A domain selects and activates tyrosine and attaches it to the T domain. The C domain would then link this to the polyketide, forming enzyme-bound amide **15**, and finally Dieckmann cyclisation, catalysed by the C-terminal DKC domain, would afford pretenellin A (**10**, Scheme 2). This compound requires only two oxidations to be converted to tenellin (**5**) via the 2-pyridone (**16**, Scheme 2).[12] These oxidations would be performed by the two cyctochrome P450 enzymes encoded in the tenellin cluster. Further questions remain to be answered, however, and we are currently examining the actual role of the C-terminal DKC domain during tenellin biosynthesis, knockout of the tailoring genes, as well as the expression of a number of other PKS-NRPS genes from fungi.

Experimental Section

Transformation.[40] *A. oryzae* M-2-3 was grown on MEA (Malt Extract Agar, Fluka) plates for 10 days. Spores were collected into water (1 mL) and inoculated into DPY (dextrin-peptone-yeast extract: 2% dextrin, 1% polypeptone, 0.5% yeast extract, 0.5% KH_2PO_4, 0.05% $MgSO_4$, 50 mL) and grown at 25 °C and 200 rpm for 12 h. Mycelia were collected by filtration and washed with water and saline (0.8 M). Protoplasting was performed using NovoZym (Novo Biolabs; 10 mg mL^{-1}) in saline (0.8 M) with gentle shaking at 25 °C for 2 h. Protoplasts were centrifuged at 707 *g* for 10 min and washed twice with saline (0.8 M) and once with solution 1 (0.8 M NaCl, 10 mM $CaCl_2$, 50 mM Tris-HCl, pH 7.5). Protoplasts were diluted to 3×10^7 mL^{-1} in solution 1 and divided into aliquots (200 μL) and pTAex3-*tenS* (20 μL, 5 μg) was added. The protoplasts were placed on ice and Solution 2 (1 mL, 60% PEG 6000, 0.8 M NaCl, 50 mM $CaCl_2$, 50 mM Tris-HCl, pH 7.5) was then added dropwise over 15 min. Then mixtures were incubated at RT for 20 min.

Czapek-Dox agar (5 mL) supplemented with sorbitol (1 M) was added to the transformation mixtures. The mixtures were poured onto Czapek-Dox plates supplemented with sorbitol (1 M) and grown at 32 °C for 7 days. Single transformant colonies were transferred onto fresh Czapek-Dox plates and grown at 32 °C.

PCR to confirm presence of *tenS*. Genomic DNA was prepared from *A. oryzae* M-2-3 pTAex3-*tenS* using the Extract-*N*-Amp Plant PCR kit (Sigma). A PCR to amplify a 734 bp fragment of *tenS* was performed using 5′-GAAGCTCCCGGCATCCTTGGTGTA-3′ and 5′-ATACATTTCCTCCCGGGCGTCGTC-3′ primers.

Construction of pTAex3-BAR-*orf3*. *orf3* was amplified from *B. bassiana* genomic DNA with 5′-CACCAAGGACAGCCATTCAACCCATC-3′ and 5′-TAGCATCAACTCAGGGCAGC-3′ primers using KOD DNA polymerase (Novagen). The PCR product was cloned into pENTR/D-TOPO Gateway entry vector (Invitrogen) and sequenced. *orf3* was transferred into the fungal expression vector pTAex3-Targ-BAR (C.M.L., unpublished results) using Gateway in vitro recombination (Invitrogen). pTAex3-Targ-BAR is a derivative of pTAex3, which has a Gateway destination cassette (Invitrogen) and BAR (basta resistance) gene, and the *arg*B gene is truncated at both ends (and thus inactive).

Cotransformation with pTAex3-*tenS* and pTAex3-BAR-*orf3*. Cotransformation with pTAex3-*tenS* and pTAex3-BAR-*orf3* was done as transformation with pTAex3-*tenS* except that after 24 h growth transformation plates were over-laid with Czapek-Dox agar supplemented with BASTA (final concentration 50 $\mu g\,mL^{-1}$). Cotransformants were cultivated on Czapek-Dox-BASTA plates at 30 °C

Fermentation. *A. oryzae*/pTAex3-*tenS* was grown on MEA plates for 7–10 days at 32 °C. Spores and aerial mycelia were collected into sterile water (1 mL). Czapek-dox supplemented with 3% maltose and 1% polypeptone (100 mL) was inoculated with the suspension (300 μL) and cultivated for 10 days, 200 rpm, at 25 °C. The same procedure was used for *tenS/orf3* cotransformants.

Extraction and purification. Ten 10 day-old cultures (each 100 mL in 500 mL Erlenmeyer flasks) were pooled and acidified to pH 3 with 37% HCl. The mixture was homogenised and extracted with ethyl acetate (2×500 mL). The organic extract was combined and concentrated in vacuo. The residue was defatted by partitioning between *n*-hexane (300 mL) and 10% aqueous methanol (300 mL). The methanol fraction was concentrated and the aqueous residue was further diluted with distilled water (300 mL) and re-extracted into ethyl acetate (2×300 mL). Upon evaporation the organic fraction yielded a dark yellow residue (206 mg).

The defatted extract (50 mg) was loaded on a glass-backed silica gel thin layer chromatography plate (20×20 cm). The plate was developed in chloroform/methanol/water/acetic acid (80:18:1.5:0.5) to yield seven bands. The band at $R_f = 0.21$ corresponded to prototenellin C. A combination of prototenellin A and B appeared as a yellow band at $R_f = 0.4$. These bands were carefully scraped off and extracted with acetone/ethanol solvent system (9:1). The resulting extracts were concentrated to yield semi-pure solids of prototenellin C and prototenellin A and B combined which were refined by preparative HPLC.

The semipure compounds, dissolved in HPLC grade methanol, were then subjected to HPLC preparative purification using a Phenomenex Luna 5μ C_{18} (2) (250×4.6 mm) reverse phase column. A gradient of water (+0.05% TFA) and acetonitrile (+0.05% TFA) solvent system was used. 75 μL of the extract solution was injected for each round of one hour HPLC program (0 min, 5% B; 5 min, 5% B; 45 min, 75% B; 46 min, 95% B; 50 min, 95% B; 55 min, 5% B; 60 min, 5% B). Eluted fractions were collected with an automatic fraction collector (Pharmacia LKB-FRAC-100): 0.5 mL per 30 s.

Prototenellin A ($R_t = 40.7$ min) was obtained from fractions 81 and 82; prototenellin B ($R_t = 36.2$ min) eluted in fractions 72–74 while fractions 52 and 53 contained prototenellin C ($R_t = 25.8$ min). The respective eluents were evaporated in order to remove acetonitrile and then extracted from water into ethyl acetate which was concentrated to yield the pure compounds.

Compound characterisation. NMR experiments were conducted with a Jeol ECP-400 instrument. ^1H chemical shifts were determined by ^1H and COSY NMR and ^{13}C chemical shifts were obtained by HSQC, HSQC-DEPT and HMBC experiments. NMR data are shown in Table 1 and in the Supporting Information. HRMS was measured at the University of Bristol, School of Chemistry HRMS facility and is described in the text. IR data were obtained using a Perkin–Elmer FTIR instrument for compounds mounted directly on the diamond cell: prototenellin A (**9A**), IR ν_{max} (neat) 3296, 2923, 2852, 1656, 1613, 1555,1516, 1445 cm^{-1}; prototenellin B (**9B**), IR ν_{max} (neat) 3273, 2924, 2853, 1646, 1610, 1568, 1516, 1442 cm^{-1}; pretenellin A (**10**), IR ν_{max} (neat) 3254, 2960, 2924, 1604, 1557, 1514, 1428 cm^{-1}.

Acknowledgements

We thank the EU FP6 MC-EST programme BRISENZ for funding (LMH). A.Y. thanks Bayero University/MacArthur Foundation for funding. J.M. gratefully acknowledges the support of the BBSRC and the M. Evans Scholarship. The work was supported by BBSRC grant BB/E007791/1.

Keywords: *Beauveria* · biosynthesis · fungi · heterologous expression · natural products

[1] P. S. Steyn, R. Vleggaar, *J. Chem. Soc. Chem. Commun.* **1985**, 1189–1191.
[2] J. W. Sims, J. P. Fillmore, D. D. Warner, E. W. Schmidt, *Chem. Commun.* **2005**, 186–188.
[3] W. Graf, J.-L. Robert, J. C. Vederas, C. Tamm, P. H. Solomon, I. Miura, K. Nakanishi, *Helv. Chim. Acta* **1974**, *57*, 1801–1815.
[4] J.-L. Robert, C. Tamm, *Helv. Chim. Acta* **1975**, *58*, 2501–2504.
[5] P. Mohr, C. Tamm, *Tetrahedron* **1981**, *37*, 201–212.
[6] S. Maiya, A. Grundmann, X. Li, S.-M. Li, G. Turner, *ChemBioChem* **2007**, *8*, 1736–1743.
[7] C. K. Wat, A. G. McInnes, D. G. Smith, J. L. C. Wright, L. C. Vining, *Can. J. Chem.* **1977**, *55*, 4090–4098.
[8] M. R. Tepaske, J. B. Gloer, D. T. Wicklow, P. F. Dowd, *Tetrahedron Lett.* **1991**, *32*, 5687–5690.
[9] S. Bergmann, J. Schumann, K. Scherlach, C. Lange, A. A. Brakhage, C. Hertweck, *Nat. Chem. Biol.* **2007**, *3*, 213–217.
[10] B. Royles, *Chem. Rev.* **1995**, *95*, 1981–2001.
[11] Z. Song, R. J. Cox, C. M. Lazarus, T. J. Simpson, *ChemBioChem* **2004**, *5*, 1196–1203.
[12] K. L. Eley, L. M. Halo, Z. Song, H. Powles, R. J. Cox, A. M. Bailey, C. M. Lazarus, T. J. Simpson, *ChemBioChem* **2007**, *8*, 289–297.
[13] J. Schümann, C. Hertweck, *J. Am. Chem. Soc.* **2007**, *129*, 9564–9565.
[14] H. U. Bohnert, I. Fudal, W. Dioh, D. Tharreau, J. L. Notteghem, M. H. Lebrun, *Plant Cell* **2004**, *16*, 2499–2513.
[15] R. Simeone, P. Constant, C. Guilhot, M. Daffe, C. Chalut, *J. Bacteriol.* **2007**, *189*, 4597–4602.
[16] J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas, C. R. Hutchinson, *Science* **1999**, *284*, 1368–1372.
[17] R. J. Cox, F. Glod, D. Hurley, C. M. Lazarus, T. P. Nicholson, B. A. M. Rudd, T. J. Simpson, B. Wilkinson, Y. Zhang, *Chem. Commun.* **2004**, 2260–2261.

[18] M. C. Moore, R. J. Cox, D. R. Duffin, D. O'Hagan, *Tetrahedron* **1998**, *54*, 9195–9206.
[19] Y. Sugano, R. Nakano, K. Sasaki, M. Shoda, *Appl. Environ. Microbiol.* **2000**, *66*, 1754–1758.
[20] A. Watanabe, Y. Ono, I. Fujii, U. Sankawa, M. E. Mayorga, W. E. Timberlake, Y. Ebizuka, *Tetrahedron Lett.* **1998**, *39*, 7733–7736.
[21] A. M. Bailey, R. J. Cox, K. Harley, C. M. Lazarus, T. J. Simpson, E. Skellam, *Chem. Commun.* **2007**, 4053–4055.
[22] R. J. Cox, *Org. Biomol. Chem.* **2007**, *5*, 2010–2026.
[23] Y. Cheng, B. Schneider, U. Riese, B. Schubert, Z. Li, M. Hamburger, *J. Nat. Prod.* **2004**, *67*, 1854–1858.
[24] K. D. An, H. Nishida, Y. Miura, A. Yokota, *BMC Evol. Biol.* **2003**, *3*, 9.
[25] T. M. Zabriskie, M. D. Jackson, *Nat. Prod. Rep.* **2000**, *17*, 85–97.
[26] N. Gaitatzis, B. Kunze, R. Müller, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11136–11141.
[27] N. Kessler, H. Schuhmann, S. Morneweg, U. Linne, M. A. Marahiel, *J. Biol. Chem.* **2004**, *279*, 7413–7419.
[28] N. Schracke, U. Linne, C. Mahlert, M. A. Marahiel, *Biochemistry* **2005**, *44*, 8507–8513.
[29] S. Kamitori, A. Iguchi, A. Ohtaki, M. Yamada, K. Kita, *J. Mol. Biol.* **2005**, *352*, 551–558.
[30] N. Ishiyama, C. Creuzenet, J. S. Lam, A. M. Berghuis, *J. Biol. Chem.* **2004**, *279*, 22635–22642.
[31] M. Tanaka, F. Nara, K. Suzuki-Konagai, T. Hosoya, T. Ogita, *J. Am. Chem. Soc.* **1997**, *119*, 7871–7872.
[32] J. L. C. Wright, L. C. Vining, A. G. McInnes, D. G. Smith, J. A. Walter, *Can. J. Biochem.* **1977**, *55*, 678–685.
[33] E. Leete, N. Kowanko, R. A. Newmark, L. C. Vining, A. G. McInnes, J. L. C. Wright, *Tetrahedron Lett.* **1975**, *16*, 4103–4106.
[34] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493–505.
[35] G. L. Challis, J. Ravel, C. A. Townsend, *Chem. Biol.* **2000**, *7*, 211–224.
[36] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO J.* **1997**, *16*, 4174–4183.
[37] R. Weckermann, R. Furbass, M. A. Marahiel, *Nucleic Acids Res.* **1988**, *16*, 11841–11841.
[38] S. Steller, D. Vollenbroich, F. Leenders, T. Stein, B. Conrad, J. Hofemeister, P. Jacques, P. Thonart, J. Vater, *Chem. Biol.* **1999**, *6*, 31–41; S. Steller, D. Vollenbroich, F. Leenders, T. Stein, B. Conrad, J. Hofemeister, P. Jacques, P. Thonart, J. Vater, *Chem. Biol.* **1999**, *6*, R156.
[39] T. Akao, M. Sano, O. Yamada, T. Akeno, K. Fujii, K. Goto, S. Ohashi-Kunihiro, K. Takase, M. Yasukawa-Watanabe, K. Yamaguchi, Y. Kurihara, J. Maruyama, P. R. Juvvadi, A. Tanaka, Y. Hata, Y. Koyama, S. Yamaguchi, N. Kitamoto, K. Gomi, K. Abe, M. Takeuchi, T. Kobayashi, H. Horiuchi, K. Kitamoto, Y. Kashiwagi, M. Machida, O. Akita, *DNA Res.* **2007**, *14*, 47–57.
[40] I. Fujii, Y. Ono, H. Yada, K. Gomi, Y. Ebizuka, U. Sankawa, *Mol. Gen. Genet.* **1996**, *253*, 1–10; I. Shibuya, K. Tsuchiya, G. Tamura, T. Ishikawa, S. Hara, *Biosci. Biotechnol. Biochem.* **1992**, *56*, 1674–1675.
[41] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, D. G. Higgins, *Nucleic Acids Res.* **1997**, *25*, 4876–4882.
[42] S. Kumar, K. Tamura, M. Nei, *Briefings Bioinf.* **2004**, *5*, 150–163.
[43] P. S. Steyn, P. L. Wessels, *Tetrahedron Lett.* **1978**, *19*, 4707–4710.

Received: July 14, 2007
Published online on February 12, 2008

