One pathway, many compounds: heterologous expression of a fungal biosynthetic pathway reveals its intrinsic potential for diversity†

Zahida Wasił,‡a Khomaizon A. K. Pahirulzaman,†b Craig Butts,a Thomas J. Simpson,a Colin M. Lazarus*bc and Russell J. Cox*ac

Heterologous expression of genes from the proposed aspyridone biosynthetic cluster from Aspergillus nidulans in the host Aspergillus oryzae led to the production of eight different compounds in addition to aspyridone A 1, one of the previously observed products. The pathway genes were incapable of producing aspyridone B 2, the previously accepted final product of the pathway. Expression of restricted sets of genes in addition to the core polyketide synthase – nonribosomal peptide synthetase (PKS-NRPS) genes apdA and apdC revealed: that apdE encodes a cytochrome P450 enzyme with ring-expanding and unprecedented dephenylation activity; that apdB encodes an N-hydroxylase, an activity not previously suspected; that the productivity of ApdA and ApdC proteins appears to be significantly enhanced in the presence of the downstream ApdE oxidase; and no obvious chemical roles for ApdD and ApdG. Furthermore, the ApdC enoyl reductase appears to operate with different stereoselectivity in different PKS cycles. All of these features illustrate the inherent diversity of compounds potentially produced by the apd pathway and the high utility of a whole pathway expression strategy for investigating and revealing new biosynthetic chemistry in fungi.

Introduction

The advent of rapid and economical genome sequencing of bacteria and fungi has revealed thousands of biosynthetic gene clusters encoding the synthesis of secondary metabolites. In the majority of cases the gene clusters can be rapidly annotated (e.g. using bioinformatic tools such as Anti Smash) and the class of compound they specify can be identified, but it is not yet possible to predict the precise structure of the encoded metabolite for any given natural product biosynthetic gene cluster. Fungi, in particular, present a problem because they can often contain more than 50 biosynthetic gene clusters and because the core synthases and synthetases are often iterative, displaying complex and cryptic programming. Gene clusters which have not been linked to the biosynthesis of a particular compound are referred to as orphan clusters. The vast majority of fungal biosynthetic gene clusters currently fall into this category.

Attempts have been made to link fungal biosynthetic gene clusters with particular compounds. For example knockout or silencing strategies can quickly achieve this if the compound of interest is observable. However most fungal biosynthetic gene clusters appear to be silent – that is, not transcribed at a sufficient level to produce an observable chemical product. In the case of such silent clusters, knockout and silencing strategies cannot be used. A further significant problem lies in the complex way in which fungal gene clusters are regulated. Fungi are eukaryotes and thus each structural gene (for example encoding an individual biosynthetic enzyme) is controlled by its own promoter. These promoters are activated under unique (often unknown) circumstances.

Some silent fungal biosynthetic gene clusters encode a transcriptional regulator. In a limited number of cases, for example that of the aspyridone A 1 biosynthetic gene cluster in Aspergillus nidulans, it has proven possible to ‘switch on’ biosynthesis by overexpressing this transcription factor. This has the effect of activating the promoters of the biosynthetic genes and thus causing transcription of the structural genes, in turn resulting in the production of the biosynthetic proteins themselves and the synthesis of a new compound. However, because not all silent biosynthetic gene clusters contain a transcription factor this method is limited.
Furthermore, switching on a range of genes simultaneously does not allow the effect (or the chemical programme) of each individual biosynthetic protein to be observed, and it is also possible that the induction of transcription factors may activate unknown genes outside the biosynthetic gene cluster, producing unpredictable chemical effects. Numerous other methods of activating silent fungal gene clusters have been reported, including chemical epigenetic modification, manipulation of signalling cascades and alteration of growth conditions (e.g. the OSMAC approach), but none are systematic, general, widely applicable in a range of fungi or, most crucially, predictably applicable to a gene cluster of interest.

We have been seeking to develop a systematic and general method for investigating the biosynthetic chemistry of fungal gene clusters based on heterologous expression – i.e. rapidly moving genes into a tractable host organism in a cumulative manner and observing chemical changes. We recently reported the development of a generally applicable plasmid system in which structural biosynthetic genes can be cloned downstream of strong promoters, thus overcoming the cryptic native transcription systems. The cloning vectors (Fig. 1A) feature a fungal selection marker (e.g. argB) for use in the arginine auxotroph of Aspergillus oryzae M-2-3) as well as a Gateway cassette downstream of the strong inducible amyl promoter (Pamyb) into which large fungal synth(et)ase genes can be rapidly cloned using in vitro recombination. The vector also contains colE1 and ampR for replication and selection in E. coli and a 2μ origin of replication and URA3 for selection in yeast. Three further strong fungal promoters (Padin, Ppada and Peno) are also included so that up to three tailoring genes can be expressed in parallel from the same system (Fig. 1B). Here we demonstrate how this system can be used to obtain detailed chemical information from a silent fungal gene cluster and provide insights into new and unprecedented reactivity.

Results

Hertweck and coworkers reported the discovery of a potential biosynthetic gene cluster in the filamentous fungus A. nidulans centred on a large open reading frame (apdA) encoding an iterative highly-reducing polyketide synthase (hrPKS) fused to a single module nonribosomal peptide synthetase (NRPS). Such hrPKS-NRPS genes, first described for fusaric acid and desmyethylbassinian, are now known to be a common feature in the genomes of many fungi. Also clustered with apdA were seven other ORFs encoding cytochrome P450 enzymes (apdB and apdE), an enoyl reductase (apdC), a FAD-dependent monoxygenase (apdD), an exporter (apdF), a dehydrogenase (apdG) and a transcription factor (apdR). Ectopic expression of apdR downstream of the inducible alcohol dehydrogenase promoter (Pade) in the A. nidulans strain SB4.1 led to the production of 1 and 2 and the up-regulation of genes apdA, apdB, apdC, apdD, apdE and apdF implicating these in biosynthesis. While these results showed that the apd cluster was most likely involved in aspyridone biosynthesis the precise chemical steps and intermediates could not be established. More recently Tang and coworkers proved that apdA and apdC encode proteins which work together to form the precursor aspyridone A by in vitro reconstitution of the proteins, but these experiments did not reveal information about the later steps of biosynthesis. We reasoned that the apd cluster would make a useful testbed for our heterologous expression technology, and the results could also allow us to delineate the precise chemical intermediates and steps of biosynthesis (Fig. 2).

The apdA gene encodes a hrPKS-NRPS. In the case of many fungal hrPKS-NRPS it is known that the enoyl reductase (ER) domain is catalytically inactive (for example the cases of tenellin and desmyethylbassinian), but Vederas and coworkers have also demonstrated similar behaviour in the hrPKS system responsible for lovastatin biosynthesis. In all these cases the missing ER functionality is supplied by a trans-acting protein usually encoded within the same gene cluster as the PKS itself. The apdA-encoded hrPKS-NRPS falls into this category and the apdC gene is known to encode the trans-acting ER. Because hrPKS lacking functional ER domains are known to function poorly and erratically on their own we expressed apdA and apdC together in the fungal host Aspergillus oryzae. Transformants and untransformed controls were grown in nutrient broth medium, acidified, extracted with ETOAc, and concentrated residues were examined by LCMS. Three new peaks were observed in transformants as compared to controls. These compounds were purified by automated mass-directed reverse-phase chromatography (Table 1).

The major component (61 mg L⁻¹) was shown to possess molecular formula C₁₉H₂₅NO₄ by HRMS (observed 332.1852; calculated 332.1856 for M[+H]). ¹H NMR indicated the presence of an aliphatic group consistent with the 2,4-dimethylhexanoyl sidechain of aspyridone and HSQC and HMBC spectra confirmed the product to be the expected 5 based on the in vitro work of Tang and coworkers. The minor second component (6, 0.6 mg L⁻¹) possessed an identical molecular formula and a near identical set of NMR data to 5. ¹H NMR in CDCl₃ revealed that 5 possesses a 10-consuming 1,3-diol and a 1,2-diol and a 1,4-diol and an amide functionality. ¹H NMR and ²H NMR indicated that 6 possesses a 10-consum-
fractional differences (<0.05 ppm) in chemical shifts for the side-chain protons, but ¹H NMR in DMSO-d₆ revealed a significant difference in the coupling between methine H-5 and methylene H-14 protons consistent with 5 and 6 being epimeric at C-5. Minor component 6 forms crystals on standing, while 5 does not, and X-ray analysis of 6 showed it to be the 7,9-anti isomer. Previous papers have shown a syn-dimethyl side chain for aspyridone. Finally, a minor component with formula C₁₉H₂₅NO₃ (observed HRMS 330.1692, calculated 330.1700 for [M]Na⁺) was observed in low yield. The same compound was produced in higher titre in a subsequent experiment and NMR analysis of this showed it to be the non-hydroxylated analogue 7.

In a second experiment apdA, apdC and apdE were coexpressed. The apdE gene encodes a cytochrome P450 enzyme, but its function is unknown. Transformants expressing these three genes produced three new compounds. The major component (184 mg L⁻¹) possessed a molecular formula of C₁₉H₂₅NO₃ – a single oxygen more than 5 (observed HRMS 370.1625, calculated 370.1625 for [M]Na⁺). 1D and 2D NMR analysis showed this compound to be 14-hydroxypreaspyridone A 8. The second major compound (120 mg L⁻¹) had a molecular formula of C₁₉H₂₃NO₄ (observed HRMS 330.1692, calculated 330.1700 for [M]H⁺) and a distinctive 2-pyridone proton in the ¹H NMR. Inspection of the 1D and 2D NMR, and comparison with the literature values, showed this to be aspyridone A 1 itself. Crystals of 1 were prepared and the X-ray structure showed that the side-chain pendant methyls matched those of 6 in being arranged anti. The third compound (72 mg L⁻¹) recovered from this experiment had a molecular formula of C₁₃H₁₉NO₃ (observed HRMS 238.1429, calculated 238.1438 for [M]H⁺), corresponding to the loss of phenoxide from 1. This was confirmed by 1D and 2D NMR and the structure was shown to be the dephenylated 2-pyridone 9. Crystals of 9 could be prepared, and the X-ray crystal structure confirmed the structure determined by NMR and again showed an anti arrangement of the side-chain pendant methyl groups (Fig. 3).

Experiments to determine the origin of the amino acid component of aspyridone A 1 were conducted by supplementing growing cultures of A. oryzae + apdACE with [1-¹³C]-tyrosine during the fermentation. Isolated aspyridone A 1 was specifically ¹³C-labelled at C-4 (5% incorporation), while isolated 9 was also labelled at C-4 (13% incorporation).

Next, apdA, apdC and apdB were coexpressed. Once again apdB encodes a cytochrome P450 monooxygenase with no known function. In this experiment the same three compounds were produced as by expression of apdA and apdC alone, although the overall titre was around 2.5-fold higher. The gene tenA also encodes a cytochrome P450 enzyme known to be involved in the oxidative ring expansion of acyl tetramic acids to

Table 1  Titres of compounds isolated from expression experiments; all values mg L⁻¹; tr. trace

<table>
<thead>
<tr>
<th>Transformant (total crude)</th>
<th>1</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>Total yield</th>
<th>Approx. fold increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>apdAC (430)</td>
<td></td>
<td>60.7</td>
<td>0.6</td>
<td>tr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61.3</td>
<td>—</td>
</tr>
<tr>
<td>apdACE (1126)</td>
<td>119.8</td>
<td></td>
<td>15.0</td>
<td>183.6</td>
<td>72.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>390.6</td>
<td>6.5</td>
</tr>
<tr>
<td>apdABC (805)</td>
<td></td>
<td>141.0</td>
<td>2.2</td>
<td>tr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>143.2</td>
<td>2</td>
</tr>
<tr>
<td>apdAC + tenA (653)</td>
<td></td>
<td>121.0</td>
<td>2.2</td>
<td>tr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>123.2</td>
<td>2</td>
</tr>
<tr>
<td>apdACEB (300)</td>
<td>1.8</td>
<td>34.6</td>
<td>5.7</td>
<td>tr.</td>
<td>2.6</td>
<td>6.2</td>
<td>2.7</td>
<td>4.0</td>
<td>57.6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>apdACED (448)</td>
<td>5.1</td>
<td></td>
<td>0.3</td>
<td>8.2</td>
<td>49.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.0</td>
<td>1</td>
</tr>
<tr>
<td>apdACEBG (622)</td>
<td>12.6</td>
<td>98.0</td>
<td></td>
<td>tr.</td>
<td>42.6</td>
<td>15.8</td>
<td>8.2</td>
<td>177.2</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2  Aspyridone A 1 and B 2, related 2-pyridones tenellin 3 and demethylbassianin 4 and the proposed apd biosynthetic gene cluster.
2-pyridones. However, expression of tenA with apdA and apdC led to the production of the same compounds as the expression of apdA and apdC alone, although again, the overall titre was around 2-fold higher than expression of the two genes alone.

Next, a series of 4-way coexpression experiments were performed. We reasoned that apdB might function as apdE and so in the first of these experiments apdA, apdC and apdE were coexpressed with apdB. This resulted in many of the previously-observed compounds being produced, together with three new compounds in low titre. The major of these had a molecular formula of C_{13}H_{19}NO_{4} (observed HRMS of 276.1219, calculated 276.1206 for [M]Na\(^{+}\)), one oxygen more than the dephenylated 2-pyridone 9. \(^{1}\)H NMR indicated that none of the C–H protons of 9 were missing, and significant chemical shift perturbations of H-5 and H-6 showed that the oxidation must be on the pyridone nitrogen. This conclusion was confirmed by 2D NMR. The second compound (3 mg L\(^{-1}\)) had a similar \(^{1}\)H NMR to 10 featuring an additional 3-proton singlet at 3.93 ppm. HRMS indicated a molecular formula of C_{14}H_{22}NO_{4} consistent with a methylation of 10 (observed HRMS of 290.1368, observed 290.1363 for [M]Na\(^{+}\)). Further 1D and 2D NMR experiments indicated that the new methyl was attached to the N-hydroxyl.

Finally another minor compound was purified (4 mg L\(^{-1}\)). This proved to be the dehydrated compound 12 (observed HRMS of 352.1528, calculated 352.1520 for [M]Na\(^{+}\)). In order to determine the geometry of the C-5/C-14 olefin of 12 several methods were attempted. Although crystals of 12 could be prepared, they did not diffract X-rays sufficiently well to be able to determine the structure. NOESY and ROESY NMR experiments were also performed. These suggested that the NH was proximal to the phenyl protons but the exchangeability of the NH meant that unambiguous evidence for the geometry could not be gained this way. Finally, the \(^{3}\)J\(_{HC}\) coupling constant between H-14 and C-4 was measured using an EXSIDE experiment.\(^{21}\) The observed value of 3.5 Hz is consistent with a cis relationship of these atoms, and DFT calculations of the \(^{3}\)J\(_{HC}\) values for the two possible E/Z isomers of 12 reinforced this conclusion with calculated values of 4.5 Hz and 7.5 Hz for the cis and trans geometries respectively.

The apdD gene encodes an FAD-dependent monooxygenase and Hertweck and coworkers speculated that this might be involved in hydroxylation of aspyridone A \(^{1}\) to form aspyridone B \(^{2}\). We thus expressed apdA, apdC and apdE along with apdD. Once again this experiment resulted in no new compounds, and the overall titre was again lowered, this time to ca 70 mg L\(^{-1}\).

The role of apdG was investigated in a five-way experiment featuring expression of apdA, apdC and apdE (which make aspyridone A \(^{1}\)) along with apdB and apdG. This experiment was achieved by expressing apdABCE as previously described, and introducing a second vector carrying apdG driven by \(P_{\text{gpdA}}\). This vector carries a selection marker consisting of the bleomycin resistance gene \(ble\) downstream of \(P_{\text{gpdA}}\). However, this
The experiment did not result in the production of any new compounds, giving approximately the same distribution of compounds as expression of apdABCE alone.

Finally we compared the productivity of the A. nidulans SB4.1 strain with that of the A. oryzae apdACEB strain. A. nidulans SB4.1 was grown in fermentation broth and extracted with EtOAc (affording 1069 mg L\(^{-1}\) of crude extract). LCMS analysis was performed and the results compared to the A. oryzae extracts. Comparison of the LCMS traces showed that in our hands the A. nidulans SB4.1 strain produced a similar range of compounds to the A. oryzae apdACEB strain. The major components were the two preaspyridone A diastereomers 5 and 6, with lesser amounts of the 14-hydroxy compound 8 and the anhydro compound 12. Hertweck and coworkers reported aspyridone A 1 as the major component, but in our hands the A. nidulans SB4.1 strain produced this compound in low titre, and did not appear to produce aspyridone B 2 at all. The dephenylated compounds were also not observed by us or the previous investigators in the SB4.1 strain (Fig. 4).

**Discussion**

Combinations of the putative aspyridone biosynthetic genes and a heterologous combination of aspyridone genes with tenA from the tenellin biosynthetic cluster were easily constructed in our expression vector. Transcription of each gene was confirmed by qRT-PCR,† confirming our previous observation that all promoters are active in this system.\(^\text{13}\) Expression of apdA and apdC alone led to the formation of preaspyridone A diastereomers 5 and 6 in agreement with the in vitro results of Tang and coworkers.\(^\text{17}\) Similar biosynthetic chemistry has been shown to operate during the biosynthesis of pretenellin A 13 and many other related compounds including the cytochalasans.\(^\text{22,23}\) In brief, the hrPKS-NRPS builds a polyketide, in this case a dimethyltetramic acid which is linked to L-tyrosine by the NRPS, which then catalyses a Dieckmann cyclisation to form the acyl tetramic acid ring and release it from the synthetase.\(^\text{20,24}\) In

---

**Scheme 1** Biosynthetic pathways of aspyridone A 1 and tenellin 3.
the case of the aspyridone synthetase Tang and coworkers have shown in in vitro experiments that the NRPS module displays extended substrate selectivity, utilising 3-phenylalanine, 3-fluorophenylalanine and 3-tryptophan in addition to 3-tyrosine. Our results tend to confirm this observation, insofar as the formation of 7 is most easily explained as the result of utilisation of phenylalanine in vivo. The incorporation of 3-tyrosine most simply explains the formation of 6.

Preaspyridone A 5 is the substrate for oxidative modification catalysed by the cytochrome P450 monooxygenase ApdE. Again, similar enzymes are known – for example TenA converts the tetramic acid pretenellin A 13 to the 2-pyridone pretenellin B 14 (Scheme 1). In the case of apdE, however, the chemistry is more diverse with three products being formed. The expected 2-pyridone 1 is observed, but so are the 14-hydroxylated compound 8 and the 5-dephenylated pyridone 9. The formation of both compounds can be explained by oxidative mechanisms. In the case of TenA it is thought that hydrogen atom abstraction from the benzylic carbon leads to a benzylic radical 16 which can rearrange and eliminate to form the observed 2-pyridone nucleus (Scheme 2). Benzyl alcohols such as 8 are not thought to be intermediates in the ring expansion in the case of TenA. In the case of ApdE it appears that a benzylic radical could also be formed. One reaction pathway would lead to the observed 2-pyridone by the same rearrangement that occurs for tenellin, but the benzylic radical could also be hydroxylated directly (rebound mechanism) to form the observed alcohol 8. It appears unlikely that this oxidation could be catalysed by an adventitious enzyme in A. oryzae as the 14-hydroxylated compound 8 is not formed when apdA and apdC are expressed alone. Furthermore, the equivalent compound in the tenellin series, prototenellin-D 15 is not observed when tenS, tenC and the P450-encoding ten genes are expressed in A. oryzae, but is only observed in extracts of the native host Beauveria bassiana, again suggesting that B. bassiana contains a similar adventitious enzyme not present in A. oryzae.

It also appears that the enzyme ApdE is responsible for the oxidative dephenylation of preaspyridone A 5. Hydroxylation at the phenyl-bridgehead carbon, leads to a putative hydroxyquinone 17 (Scheme 2) which could undergo a coupled ring-expansion and dephenylation. Such hydroxyquinones are known in natural products such as jacaglabroside B 18. Similar compounds have also been synthesised from para-hydroxyaromatic species by mild oxidants including bisacetoxyiodobenzene and oxone. The product of the dephenylation was shown to be derived from tyrosine, like its precursor preaspyridone A 5, because of a high incorporation of label from [1-13C]-tyrosine. Thus, unlike TenA which is a highly selective ring-expandase,
ApdE appears to be much less selective in its oxidation chemistry. The experiments also showed that TenA appears to be highly selective for its substrate as it could not act on the preaspyridone A diastereomers 5/6 which are very similar to pretenellin A 13.

The production of 9 and its derivatives also sheds light on the biosynthesis of a class of fungal metabolites known as the atpenins 19–22 which are potent and specific inhibitors of mitochondrial succinate-ubiquinone oxidoreductase (Fig. 5) and important tools for studying respiration.27 Although clearly related to other fungal 2-pyridones, the biosynthesis of the atpenins has not been reported. Our results suggest that they are likely to be the products of PKS-NRPS systems very similar or identical to the apd cluster, but containing additional genes for oxidation, methylation and chlorination of intermediate 9.

The polyketide moieties of 1, 6, 9 and all the atpenins 19–22 have been determined to feature pendant methyl groups which are arranged anti. Pendant methyl groups are derived from S-adenosyl methionine (SAM) during polyketide biosynthesis in fungi, catalysed by a C-methyltransferase (CMet) domain of the PKS. However, the stereochemistry of the carbon bearing the methyl is set by the enoyl reductase (ER, Scheme 3). In the case of ApdA the ER of the PKS is known to be inactive, but ER activity is provided in trans by ApdC. Fungal PKS are iterative, in each case to produce the observed metabolite, and containing additional genes for oxidation, methylation and chlorination of intermediate 9. The polyketide moieties of 1, 6, 9 and all the atpenins 19–22 have been determined to feature pendant methyl groups which are arranged anti. Pendant methyl groups are derived from S-adenosyl methionine (SAM) during polyketide biosynthesis in fungi, catalysed by a C-methyltransferase (CMet) domain of the PKS. However, the stereochemistry of the carbon bearing the methyl is set by the enoyl reductase (ER, Scheme 3). In the case of ApdA the ER of the PKS is known to be inactive, but ER activity is provided in trans by ApdC. Fungal PKS are iterative, in each case to produce the observed metabolite, and containing additional genes for oxidation, methylation and chlorination of intermediate 9.

The role of apdB was also demonstrated. This gene encodes another cytochrome P450 monoxygenase. Initial experiments showed that ApdB appears not to act on the tetramic acids 5 or 6. However, it does appear to N-hydroxylate the dephenylated 2-pyridone 9, forming 10 and 11. These compounds were only observed in experiments where apdB was expressed and so this is very unlikely to be a native reaction of the A. oryzae host. Interestingly, although northern blotting experiments suggested that apdB was expressed during the biosynthesis of aspyridones in the native host A. nidulans, N-oxygenated compounds were not observed and our experiments also appear to rule out a role for apdB during aromatic ring hydroxylation to form aspyridone B 2. Furthermore, the formation of the N-O-methyl compound 11 must be catalysed by an endogenous A. oryzae O-methyl transferase as none of the expressed genes was predicted to encode proteins with methyltransferase activity.

We were then left with the possibility of either ApdD or ApdG being involved in ring-hydroxylation of 1 to form aspyridone B 2. However expression of apdD or apdG with combinations of the other genes which produce 1 did not lead to the formation of 2 or related compounds. It thus appears possible that aspyridone B 2 is formed in A. nidulans by an enzyme encoded outwith the predicted apd gene cluster.

Precise titres of all purified compounds were measured by making standard solutions of the purified metabolites and using these to calibrate a standard HPLC method (Table 1). Individual fungal expression clones were grown in triplicate and the average titre of each metabolite determined by HPLC of exhaustive organic extracts of each fermentation. Consideration of the overall productivity of each clone (Table 1) shows that as well as the chemical make-up of the extract varying by expression clone, the total amount of produced material also varies. Previous work in Bristol, as well as by Vederas and coworkers has already shown that hrPKS enzymes are more productive when expressed together with their cognate trans-acting ER, suggesting a direct protein–protein interaction between the hrPKS and the ER.19,20 The results of the experiments reported here show that the presence of a downstream cytochrome P450 can also boost the productivity of the system, from 60 mg L−1 to nearly 400 mg L−1, again suggesting a higher-order protein–protein organisation. Interestingly this phenomenon also appears to occur when proteins are present which are not acting as chemical catalysts. For example, the probable presence of ApdB appears to increase the productivity of ApdA + ApdC approximately 2-fold, despite the fact that ApdB has no substrate to act on. Conversely, the probable presence of ApdB...
appears to lower the influence of ApdE, perhaps indicating an effect of competition. However such arguments must remain inconclusive in the absence of information about actual protein concentrations in vivo.

Conclusion

Heterologous expression of biosynthetic genes from the aspyridone pathway has allowed the individual catalytic activities of each encoded enzyme to be investigated and the intermediates to aspyridone A 1 to be determined. It has also allowed the unprecedented oxidative chemistry catalysed by ApdE to be observed. ApdE shares 48% amino acid identity and 64% similarity to TenA, but nevertheless these enzymes show different substrate selectivities and different fidelities. The activity of ApdE gives clues about the biosynthesis of the aptenins and based on our results it should now be possible to rapidly discover a putative atpenin biosynthetic gene cluster in Penicillium sp. FO125, the producing organism.

Our results also raise questions about the identity and definition of fungal biosynthetic gene clusters, their control and their chemical products. Fungal biosynthetic gene clusters can be defined by a number of means – including cotemporal transcription, as demonstrated by Hertweck and coworkers in their chemical products. Fungal biosynthetic gene clusters can be identified by many bioinformatic means (some even claiming accuracy)8). It is clear that in the case of the apd genes, apdA to apdG are cotemporally transcribed, seemingly in response to the expression of apdR. However, not all of these genes encode proteins which are catalytically active – only apdA, apdC, and apdE are required for the synthesis of aspyridone A 1. There appear to be insufficient genes in the cluster to encode the synthesis of aspyridone B 2, and transcription and bioinformatic analyses are inadequate for determining the role of apdB, or for predicting the new chemistry enacted by ApdE. However, all these features were revealed by our heterologous expression experiments.

Furthermore, our experiments show that heterologous expression of genes using strong constitutive promoters can significantly enhance titres of the produced compounds – for example expression of apdA and apdC alone affords one major compound at 61 mg L
-1, but coexpression with apdE increases this significantly to 390 mg L
-1 (but for three compounds). The origin of this effect is not directly discernible from our experiments, but suggests that the productivity of the aspyridone biosynthetic enzymes are enhanced when they work together. The ability of genes from the apd cluster to encode the synthesis of diverse compounds is exemplified by our coexpression of apdABC with which produced at least seven different compounds attributable to the apd genes.

This result emphasises our previous observation that differential gene expression levels during temellin biosynthesis in the native host Beauveria bassiana can affect both titre and biosynthetic programming,54 in turn indicating that fine control over gene expression can affect the chemical structures of the compounds produced from a single biosynthetic gene cluster. This can occur in eukaryotes where each structural gene is controlled by its own promoter and where complex networks of control elements are involved in the activation of genes within each pathway. In our experiments where this fine control has been bypassed we assume that the relative ratios of the produced proteins is different from that produced under native conditions. It has long been known that the production of individual secondary metabolites by filamentous fungi is very sensitive to growth and media conditions, but our results also suggest that changes in gene expression could also change the compounds produced from a single gene cluster.

Clear lessons emerge from this analysis: first, that the definition of fungal biosynthetic gene clusters, however achieved, is at best speculative; second, that heterologous expression can reveal a significant number of the details of secondary metabolic pathways in fungi; third, that the ability of fungal pathways to produce (many) more than one compound is inherent – and that a single compound should not be regarded as the sole product of a fungal biosynthetic gene cluster; fourth, that genome mining in fungi will remain problematic while so many gene expression pathways, enzymatic activities and protein–protein interactions remain highly unpredictable; and finally that heterologous expression offers a systematic and broadly applicable method for investigating the chemistry of silent and orphan fungal biosynthetic gene clusters, in addition to providing access to new natural products.

Experimental

1 Cloning procedures are given in the ESL†

1.1 Construction of pTAYAargAC (apdA, apdC, apdE). P
apdA
 and P
eno were amplified from pTAYAargASP (ABCE) as a single fragment using primers newgpdA-F, eno/asc/Page-R, while P
adh
 was amplified using primers adh/asc-F and adh/asc-R. These fragments were reassembled by homologous recombination in S. cerevisiae with Asc1-cut pTAYAArgPage and shuttled back into E. coli to create pTAYAArgApcG. apdAEGFP was then transferred into pTAYAargApcG by Gateway LR recombination to create pTAYAArgAC. Transformation of A. oryzae with this plasmid yielded 42 transformants and metabolites were extracted from 9.

1.2 Construction of pTAYAargACE (apdA, apdC, apdE). The apdB gene was removed from pTAYAargASP by homologous recombination in yeast. pTAYAargASP was linearized with MluI which cuts at a unique site within apdB. Amplification of P
eno
 with primers eno/asc/page-F and eno/asc/page-R produced a ‘patch’ fragment that could recombine both up- and downstream of the apdB coding region. This created pTAYAargACE, which carries only apdA, apdC and apdE. Analytical digestion with HindIII yielded 3 fragments of 23 699, 1071, 3955 bp confirming the removal of apdB, which contains an additional HindIII site. This plasmid was used to transform A. oryzae M-2-3. Metabolites were extracted from 11 of the 12 transformants recovered.

1.3 Construction of pTAYAargABC (apdA, apdB, apdC). Three fragments, P
apdA
 + apdB (3631 bp), apdB (2264 bp) and P
adh
 (500 bp) were amplified from pTAYAargASP using the primers: PgpdA plus apdC, newgpdA-F and gpdA-apdC-R; apdC,
enoapdB-F and enoapdB-R; apdE, adh/asc-F and adh/asc-R. These fragments were reassembled in yeast with AscI-cut pTAYAGSargPage to create pTAYAGSargBC. Finally, apdAeGFP (see Construction of pTAYAargASP) was inserted to create pTAYAargABC, which was used to transform A. oryzae. Thirty three transformants were recovered, metabolites were extracted from 8 of these, with 7 of them showing high titres of preapyridone.

1.4 Construction of pTAYAACtenA (apdA, apdC, tenA). $P_{gpdA}$ apdC and $P_{eno}$ were amplified from pTAYAargASP as a single fragment using primers newgpdA-F and eno/asc/Page-R, while tenA was amplified together with $P_{adh}$ from pTAYAGSargPageTenEllin13 using primers adh/asc-F and trpof1-R. These fragments were reassembled in AscI-cut pTAYAGSargPage by homologous recombination in S. cerevisiae, creating pTAYAGSargCtenA. Plasmids were recovered in E. coli and the correct clone was used as a vector for insertion of apdAeGFP via LR recombination to create pTAYAACtenA. This plasmid was used to transform A. oryzae. Forty two transformants were recovered and metabolites were extracted from 14. Of these, 11 produced new compounds.

1.5 Construction of pTAYAargASP (apdA, apdB, apdC, apdE). The genes for the three tailoring enzymes of aspyridone biosynthesis were amplified using the following primers: apdE, adh-adpD-E and adh-adpD-E; apdC, gpdA-apdC-F and gpdA-apdC-R; apdB, eno-apdB-F, eno-apdB-R. The primers produced overlaps with the promoters and flanking vector sequences and the coding regions were each amplified together with 300 bp of 3′-UTR from genomic DNA of A. nidulans; the product sizes were 1980 bp, 1461 bp and 2264 bp. Recombination in yeast was used to combine the fragments and the resultant plasmids were shuttled into E. coli. Following analytical PCR, plasmids from two colonies were extracted and restriction analysis indicated that all 3 genes were present in both.

The apdA gene was amplified from A. nidulans genomic DNA as three overlapping fragments of ~4 kb each using primer pairs: apdA1-F and apdA1-R; apdA2-F and apdA2-R; apdA3-F and apdA3-R. The forward primer for fragment 1 and reverse primer for fragment 3 contained 30 bases homologous to the vector region for recombination. Homologous recombination in S. cerevisiae was used to reassemble apdA between GATEWAY attL sites to allow easy transfer into any expression vector that contains a GATEWAY destination cassette. pEYATenSeGFP13 was digested with NotI and XhoI and used as the vector. The tenS gene was replaced by apdA, while the eGFP remained to facilitate identification of A. oryzae transformants that successfully expressed the new hybrid gene. Plasmids extracted from S. cerevisiae colonies were shuttled into E. coli TOP10 cells. Four colonies were extracted and the plasmids digested with NotI plus Kpn1 to prove that apdA had been correctly assembled to create pEYAapdAeGFP. GATEWAY LR recombination between pEYAapdAeGFP and pTAYAGSargPageASP resulted in production of pTAYAargASP ready to transform A. oryzae.

Fifty nine colonies were recovered following transformation of A. oryzae with pTAYAargASP. All transformants were subcultured onto selective medium. In order to confirm all the genes in the pTAYAargASP were expressed, qRT-PCR was performed (see ESI†). Forty four transformants were cultured for metabolite extraction. All 14 had a yellow pigmented morphology, but one was discarded due to contamination. In all 13 extracts analysed, new compounds were detected.

1.6 Construction of pTAYAargACED (apdA, apdC, apdE, apdD). Plasmid pTAYAargACED was constructed from pTAYAargASP using the same strategy as that used to create pTAYAargACE. apdD was amplified using primers apdD-F and apdD-R which flanked the PCR product with sequences overlapping the enolase promoter and vector sequences respectively. Following recombination in S. cerevisiae and plasmid recovery in E. coli, three colonies were miniprepped and the plasmids were confirmed by restriction digestion. This plasmid was used to transform A. oryzae, and metabolites were extracted from 5 transformants. Of these, 4 produced new compounds. The presence of apdD in the transformant was confirmed by PCR.

1.7 Co-transformation of A. oryzae with pTAYAargASP (carrying apdABCE) and pTAYAGSbleG (carrying apdG), pTAYAGSblebarPage was created by homologous recombination in S. cerevisiae. AgeI-cut pTAYAGSargPage was recombined with PCR fragments of $P_{esp}$ + $bar$ (YA-TrpC-F and Bar-YA-R) and $P_{gpdA}$ (newgpdA-F and newgpdA-R) and plasmids were shuttled back into E. coli. Age1 cuts twice in both the arg gene and $P_{gpdA}$. Thus, while the $P_{esp}$ + $bar$ fragment replaced the arg gene, the $P_{gpdA}$ fragment provided a ‘patch’ to allow the plasmid to recircularize. Clones containing the desired plasmid pTAYAGSblebarPage were identified by PCR. This plasmid was further modified by replacing the bar gene with ble.

The ble gene was amplified and inserted into NgoMIV-cut pTAYAGSblebarPage by homologous recombination in yeast to create pTAYAGSbleblePage. The plasmid was confirmed to carry the ble-resistance gene by restriction analysis with NgoMIV + AscI.

pTAYAGSbleblePage was then used as a vector to co-express apdG together with pTAYAargASP in A. oryzae. The apdG gene was amplified using the primers apdG-F and apdG-R and placed under control of $P_{gpdA}$ in AscI-cut pTAYAGSbleblePage. Two promoter fragments ($P_{gpdA} + P_{adh}$ and $P_{eno}$) were also amplified ($P_{gpdA} + P_{adh}$ and $P_{gpdA} + P_{adh}$), adh/asc-F and newgpdA-R, $P_{eno}$ eno/asc/Page-F and eno/asc/Page-R) and inserted to ‘patch’ the Asc1 sites downstream of P_adh and P_eno in AscI-cut pTAYAGSbleblePage. The resulting plasmid, pTAYAGSblebleG was recovered in E. coli and confirmed by restriction digestion with NgoMIV.

Twelve transformants were obtained by co-transformation of A. oryzae with pTAYAargASP (carrying apdABCE) and pTAYAGSblebleG (carrying apdG). Transformants were selected on a minimal medium containing 100 μg mL$^{-1}$ zeocin (a cheaper alternative to bleomycin or phleomycin). Metabolites were extracted from all 12 transformants, but novel peaks were detected in only 5.

1.8 Construction of vector pTAYAGSargPage. The Padh, PgpdA and Peno promoters were amplified by PCR using sets of primers (adh/asc-F, adh/asc-R, newgpdA-F, newgpdA-R, eno/asc-Page-F and eno/asc-Page-R respectively) with 5′ extensions that provided overlap between adjacent promoters and the vector insertion site, as well as providing Ascl sites downstream of each promoter. They were inserted into NgoMIV-cut
pTAYAGSarg<sup>13</sup> via homologous recombination in yeast, and resultant plasmids were shuttled into <i>E. coli</i>. Four <i>E. coli</i> colonies were screened by restriction digestion with AscI. Of these, three showed the expected restriction pattern for pTAYAGSargPage, while the fourth lacked the Peno fragment. Analytical PCR was further done on the clones to verify successful insertion of the promoters in the correct orientation using combinations of primers: adh/asc-F and gpdA-R (for Padh and PgdA), gpdA-F and eno/asc-R (PgdA and Peno).

2 Fermentation conditions

The <i>Aspergillus oryzae</i> transformants were first grown on Czapek Dox Agar (minimal media, see ESI<sup>†</sup>) and for further production of spores, they were inoculated and grown on DPY solid media for 7–10 days at 25 °C.

2.1 Culturing and inoculation of spores of <i>A. oryzae</i> transformants in liquid media. 10 mL of deionized water was added on DPY plates having 10 days old growing spores of the transformants. The spores were made to pass into the deionized water by careful scratching the surface with sterilized loop. The spore suspension (1 mL) from the plate was added in each 100 mL liquid medium contained in 250 mL Erlenmeyer flask. The spores were allowed to grow in the liquid culture for 7 days on shakers at 200 rpm at 28 °C.

2.2 Extraction. Cells and media (1 L) were homogenized using a hand-held electric blender and then acidified to pH 4.0 using 37% HCl. An equal volume of ethyl acetate was added and stirred for 10 min. The resulting mixture was vacuum filtered through Whatman no. 1 filter paper. The filtrate was transferred into a separating funnel and shaken vigorously. The mixture was allowed to stand to separate the layers. The organic layer was washed once with concentrated brine solution and then with deionized water. The organic phase was dried (MgSO<sub>4</sub>), filtered and evaporated to dryness. The crude extract was dissolved in 10% aqueous methanol and defatted by extraction with hexane. The methanolic layer was evaporated to dryness and then made into a solution of 10 mg mL<sup>−1</sup> in HPLC grade methanol and analysed by LCMS.

For purification of compounds the crude extract was made into a solution of 50 mg mL<sup>−1</sup> in HPLC grade methanol and 200 µL aliquots were injected in each run of mass-directed HPLC preparative purification.

3 Analytical LCMS

Samples as prepared above (20 µL) were analysed using a Waters 2795HT HPLC system. Detection was achieved by uv between 200 and 400 nm using a Waters 998 diode array detector, and by simultaneous electrospray (ES) mass spectrometry using a Waters ZQ spectrometer detecting between 150 and 600 m/z units. Chromatography (flow rate 1 mL min<sup>−1</sup>) was achieved using a Phenomenex Kinetex column (2.6 µ, C<sub>18</sub>, 100 Å, 4.6 × 100 mm) equipped with a Phenomenex Security Guard pre-column (Luna C<sub>5</sub> 300 Å). Solvents were: A, HPLC grade H<sub>2</sub>O containing 0.05% formic acid; B, HPLC grade MeOH containing 0.045% formic acid; and C, HPLC grade CH<sub>3</sub>CN containing 0.045% formic acid. Gradients were as follows.

- Method 1. MeOH: 0 min, 10% B; 10 min, 90% B; 12 min, 90% B; 13 min, 10% B; 15 min, 10% B.
- Method 2. CH<sub>3</sub>CN: 0 min, 10% C; 10 min, 90% C; 12 min, 90% C; 13 min, 10% C; 15 min, 10% C.

4 Purification and characterization of compounds

Purification of compounds was generally achieved using a Waters mass-directed autopurification system comprising of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex LUNA column (5µ, C<sub>18</sub>, 100 Å, 10 × 250 mm) equipped with a Phenomenex Security Guard pre-column (Luna C<sub>3</sub> 300 Å) eluted at 4 mL min<sup>−1</sup> for simultaneous analysis by diode array detector (Waters 2998), evaporative light scattering (Waters 2424) and ESI mass spectrometry in positive and negative modes (Waters Quatro Micro).

Solvent used were: A, HPLC grade water containing 0.05% formic acid; and Solvent B, HPLC grade CH<sub>3</sub>CN containing 0.045% formic acid. Extracts from (apdACEB) were purified with Method 1, (apdACED) with Method 2, (apdACE) with Method 3 and (apdACEBG) with Method 4.

Method 1: 0 min, 50% B; 22 min, 75% B; 24 min, 95% B; 26 min, 95% B; 27 min, 50% B; 30 min, 50% B.
Method 2: 0 min, 55% B; 22 min, 60% B; 24 min, 95% B; 26 min, 95% B; 27 min, 55% B; 30 min, 55% B.
Method 3: 0 min, 55% B; 22 min, 75% B; 24 min, 95% B; 26 min, 95% B; 27 min, 55% B; 30 min, 55% B.
Method 4: 0 min, 50% B; 22 min, 65% B; 24 min, 95% B; 26 min, 95% B; 27 min, 50% B; 30 min, 50% B.

Detected peaks were collected into glass test tubes. Combined fractions were evaporated under a flow of dry N<sub>2</sub> gas and residues weighed and dissolved directly in NMR solvent for NMR analysis.

4.1 Aspyridone A 1 (ref. 8) Light brown solid, IR (neat): 1<sup>1</sup>H NMR (CD3OD, 500 MHz) δ = 0.89 (d, 3H, J = 7.5 Hz, 3H, H-12), 0.92 (d, 3H, J = 6.5 Hz, 3H, H-13), 1.15 (d, 3H, J = 6.5 Hz, 3H, H-14), 1.18 (m, 1H, H-11a), 1.35 (m, 1H, H-9a), 1.37 (m, 1H, H-11b), 1.43 (m, 1H, H-10), 1.65 (m, 1H, H-9b), 4.40 (q, 2H, J = 6.8 Hz, 1H, H-8), 6.83 (d, 1H, J = 8.5 Hz, 1H, H-17, H-19), 7.29 (d, 1H, J = 8.5 Hz, 1H, H-16-H-20) 7.49 (s, 1H, H-6);<sup>13</sup>C NMR (CD3OD, 125 MHz): δ = 11.7 (C-12), 17.6 (C-14), 19.4 (C-13), 30.9 (C-11), 33.6 (C-10), 41.2 (C-9), 41.8 (C-8), 107.0 (C-3), 116.0 (C-5), 116.5 (C-17, C-19), 125.1 (C-15), 131.5 (C-16, C-20) 140.6 (C-6), 158.3 (C-18), 163.9 (C-4), 177.6 (C-2), 214.3 (C-7).

4.2 Preaspyridone A (major diastereomer) 5 (ref. 17) Pale brown solid, IR (neat): 1<sup>1</sup>H NMR (DMSO d6) δ = 0.80 (m, 3H, H-12), 0.79 (m, 3H, H-11), 0.95 (d, 3H, J = 7 Hz, 3H, H-13), 1.07 (m, 1H, H-10a), 1.28 (m, 1H, H-9), 1.29 (m, 1H, H-10b), 1.29 (m, 1H, H-8a), 1.37 (m, 1H, H-8b), 2.82 (d, 1H, J = 4.5 Hz, 2H, H-14), 3.51 (q, 2H, J = 6.6 Hz, 1H, H-7), 4.08 (brs, 1H, H-5), 6.59 (d, 1H, J = 8.5 Hz, 2H, H-17, H-19), 6.89 (d, 1H, J = 8 Hz, 2H, H-16, H-20), 8.92 (s, 1H, H-1), 9.16 (brs, 1H, H-1), 11.6 (brs, 1H, H-2).
13C NMR (125 MHz, DMSO-d6): δ 11.7 (C-12), 17.3 (C-14), 19.3 (C-13), 31.0 (C-11), 33.5 (C-10), 40.6 (C-9), 42.1 (C-8), 98.9 (C-5), 107.3 (C-3), 142.3 (C-6), 175.8 (C-2), 213.8 (C-7).

4.8 1-Methoxy, 4-hydroxy-3-(2,4-dimethylhexanoyl) 2-pyridone 11. Light brown solid, IR (neat): vmax 3667, 3190, 2961, 2876, 2366, 1691, 1584 cm⁻¹. 1H NMR (500 MHz, DMSO-d6) δ = 0.80 (t, J = 7.2 Hz, 3H, H-12), 0.84 (d, J = 6.5 Hz, 3H, H-13), 1.03 (d, J = 7 Hz, 3H, H-14), 1.13 (m, 1H, H-11a), 1.25 (m, 1H, H-9a), 1.25 (m, 1H, H-11b), 1.37 (m, 1H, H-10), 1.49 (m, 1H, H-9b), 3.93 (s, 3H, H-15), 4.15 (q, J = 7 Hz, 1H, H-8), 5.99 (d, J = 7.5 Hz, 1H, H-5), 8.24 (d, J = 7.5 Hz, 1H, H-6); 13C NMR (125 MHz, DMSO-d6): δ 11.1 (C-12), 16.5 (C-14), 18.6 (C-13), 29.4 (C-11), 31.8 (C-10), 39.6 (C-9), 39.9 (C-8), 64.4 (C-15), 98.4 (C-5), 146.2 (C-6), 109.9 (C-3), 159.9 (C-2), 214.5 (C-7).

4.9 Z-5,14-Anhydropreaspyridone A 12. Bright yellow solid, IR (neat): vmax 3667, 3190, 2961, 2876, 2366, 1691, 1584 cm⁻¹. 1H NMR (500 MHz, DMSO-d6) δ = 0.79 (m, 3H, H-11), 0.82 (m, 3H, H-12), 1.05 (d, J = 6.5 Hz, 3H, H-13), 1.08 (m, 1H, H-10a), 1.32 (m, 1H, H-9), 1.32 (m, 1H, H-10b), 1.38 (m, 1H, H-8a), 1.47 (m, 1H, H-8b), 3.75 (q, J = 6.6 Hz, 1H, H-7), 6.37 (s, 1H, H-14), 6.78 (d, J = 8.5 Hz, 2H, H-17, H-19), 7.46 (d, J = 8.5 Hz, 2H, H-16, H-20); 13C NMR (125 MHz, DMSO-d6): δ = 10.9 (C-11), 16.8 (C-13), 18.8 (C-12), 28.7 (C-10), 31.5 (C-9), 40.4 (C-8), 110.4 (C-14), 115.6 (C-17, C-19), 124.8 (C-15), 131.5 (C-16, C-20), 158.8 (C-18), 184.0 in CD2OD (C-4).

4.10 Crystal structure data. Crystal structure data has been deposited with the CCDC under the following accession numbers: 6, 941137; 7, 941138; 9, 941139.

5 3JCCH coupling constant calculations

The values of 3JCCH for the cis and trans isomers of 12 were established by computation of a truncated analogue with the polyketide sidechain removed (see ESI†). The geometries were optimised with DFT/B3LYP functional using 6-31g basis set before calculation of coupling constants using the Gauge Independent Atomic Orbital (GIAO) approach with a 6-31g+(d,p) basis set. All calculations were performed in Gaussian09. Geometry optimisation and 3JCCH calculations were repeated for the corresponding enol forms and found to give coupling constant values in agreement with the corresponding enone forms, so no further account of enolisation was considered in this analysis.

Acknowledgements

ZW Thanks the Government of Pakistan (HEC) for funding. KAKP thanks Majlis Amanah Rakyat (MARA) Malaysia for funding. EPSRC is thanked for LCMS equipment (EP/F066104/1). We also thank Professors C. Hertweck and A. Brakhage for the provision of A. nidulans SB4.1.

Notes and references


