
Peer reviewed version

Link to published version (if available): 10.1039/c7cc03303f

Link to publication record in Explore Bristol Research

PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via RSC at http://pubs.rsc.org/en/Content/ArticleLanding/2017/CC/C7CC03303F#!divAbstract. Please refer to any applicable terms of use of the publisher.

**University of Bristol - Explore Bristol Research**

**General rights**

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms
Genetic and Chemical Characterisation of the Cornexistin Pathway Provides Further Insight into Maleidride Biosynthesis

Katherine Williams, Agnieszka J. Szwalbe, Claire Dickson, Tim R. Dessonn, Nicholas P. Mulholland, Jason L. Vincent, John M. Clough, Andrew M. Bailey, Craig P. Butts, Christine L. Willis, Thomas J. Simpson and Russell J. Cox.

The biosynthesis of the herbicide cornexistin in the fungus Paecilomyces variotii was investigated by full sequencing of its genome, knockout of key genes within its biosynthetic gene cluster and isolation and identification of intermediate compounds. The general biosynthetic pathway resembles that of byssochlamic acid and other nonadrides in the early stages, but differs in requiring fewer enzymes in the key nonadride dimerisation step, and in the removal of one maleic anhydride moiety.

Weeds are significant pests of wheat, maize and soybean due to competition for inorganic nutrients. Herbicides which target weeds without harming these crops could improve food security. Cornexistin 1, produced by the fungus Paecilomyces variotii, is a potent wide-spectrum herbicide against weeds but with low activity against Zea mays (maize), and thus shows considerable promise as a commercial herbicide. Cornexistin 1 is a maleidride: these are 7-, 8- and 9-membered carbocyclic rings fused to one or (usually) two maleic anhydrides. Related compounds include agnestadride A 2 (heptadride), viburspiran 3 (octadride) and byssochlamic acid 4 (nonadride). Maleic anhydrides have been associated with a wide range of useful biological properties. We recently elucidated the complete biosynthetic pathway of byssochlamic acid 4. The first step of the pathway is catalysed by a fungal highly reducing polyketide synthase (hrPKS), and subsequent steps involve a citrate synthase (CS) homologue and a 2-methyl citrate dehydratase (2MCD) to create the key intermediate 5. Oikawa also reported a similar early pathway towards the phomoidrides 6 and Tang and coworkers have shown that the biosynthesis of rubratoxins follows the same pathway. Construction of the carbocyclic ring appears to be catalysed by pairs of enzymes with homology to phosphatidyl-ethanolamine binding proteins (PEBP) and ketosteroid isomerase-like enzymes (KI).

The titre of 1 is low (typically < 10 mg L\(^{-1}\) in our hands) and no total synthesis of 1 has been reported. We rationalised that the biosynthesis of 1 should be encoded by a biosynthetic gene cluster (BGC) highly similar to that involved in the production of 4 in Byssochlamys fulva. We thus set out to investigate the biosynthesis of 1 via genome sequencing and gene disruption experiments with the aim of increasing its titre and producing structurally related compounds.

We sequenced the genome of P. variotii, using Illumina paired-end and mate-pair technology to reveal a genome of approximately 35 Mb contained on 41 scaffolds with an N50 of 3.3 Mb. BLAST analyses identified a single putative maleidride-type BGC (Figure 1). Predicted coding sequences were annotated using BLAST and InterPro (Table 1). The putative cornexistin 1-BGC was compared to the previously identified byssochlamic acid BGC, as well as to two putative maleidride BGCs from Talaromyces stipitatus and one identified from several Cochliobolus species, using Artemis (see ESI).

This analysis showed that each of the key maleidride biosynthetic proteins encoded by the B. fulva cluster has a similar (> 38% identity) counterpart in the P. variotii cluster (see ESI). Comparisons among all five clusters show that the hrPKS, hydrolase, CS, 2MCD, PEBP-like and KI-like genes occur in all clusters. The biosynthesis of 1, however, is expected to be more complex than that of 4, requiring removal of one maleic anhydride ring and several oxidative modifications. We were not surprised, therefore, to find several additional
putative oxidase encoding genes and genes of unknown function in the putative cornexistin BGC.

Figure 1. Cornexistin biosynthetic gene cluster.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Putative function</th>
<th>Genes</th>
<th>Putative function</th>
</tr>
</thead>
<tbody>
<tr>
<td>pv-pks1</td>
<td>Polyketide synthase (PKS)</td>
<td>pvL1</td>
<td>Diene lactone hydrolase</td>
</tr>
<tr>
<td>pvL1</td>
<td>Hydrolase</td>
<td>pvL2</td>
<td>MFS transporter</td>
</tr>
<tr>
<td>pvL2</td>
<td>2-methylcitrate dehydratase (2MCD)</td>
<td>pvL3</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>pvL3</td>
<td>Ketosteroid isomerase-like (Ki-like)</td>
<td>pvL4</td>
<td>Transketolase</td>
</tr>
<tr>
<td>pvL4</td>
<td>MFS transporter</td>
<td>pvL5</td>
<td>Aldo/keto reductase</td>
</tr>
<tr>
<td>pvL5</td>
<td>Non-heme iron doxynogenase</td>
<td>pvL6</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>pvL6</td>
<td>Citrate synthase-like (CS-like)</td>
<td>pvL7</td>
<td>Fumarylacetoacetate hydrolase</td>
</tr>
<tr>
<td>pvL7</td>
<td>Gluconolactonase</td>
<td>pvL8</td>
<td>NRPS-like enzyme</td>
</tr>
<tr>
<td>pvL8</td>
<td>C6 transcription factor</td>
<td>pvL9</td>
<td>Phosphatidylethanolamine-binding protein-like (PEBP-like)</td>
</tr>
<tr>
<td>pvL9</td>
<td>Serine carboxypeptidase</td>
<td>pvL10</td>
<td>MFS transporter</td>
</tr>
</tbody>
</table>

Table 1. Analysis of the cornexistin BGC from *P. variotii*.

The most highly produced compounds in WT *P. variotii* under cornexistin production conditions are a set of xanthone related compounds, including monodictyphenone, monodictyxanthone, chrysophanol, and three unidentified congeners A, B and C (Figure 2A). These polyketides require malonyl CoA for their construction by a fungal non-reducing PKS (nrPKS) which may compete with cornexistin biosynthesis, and be at least partially responsible for the low titre of 1.

![Figure 2](Image)

Figure 2. ELSD chromatograms showing the comparison of crude extracts from 11 day old cultures of *P. variotii*; A, WT strain; B, AgnΔpks strain. See ESI for detailed detection of 4 using HR-LCMS.

The xanthone nrPKS was identified by comparison with other known xanthone pks genes (such as *mdpg* from *Aspergillus nidulans* and *clarG* from *Cladosporium fulvum*) and then disrupted using a geneticin resistance cassette to produce strain PVAgνΔpks. As expected this strain produced no xanthone related compounds 6-8 and A-C (Figure 2B), but produces, on average, a 2.7 fold higher cornexistin titre compared to WT *P. variotii*. Cornexistin was crystallised and its structure confirmed by X-ray analysis. The analysis also established the absolute configuration of cornexistin (Flack parameter 0.05) as 3S,6S,7R for the first time (Figure 3).

Figure 3. X-ray structure and absolute configuration of cornexistin 1.

The removal of overlapping xanthone-related peaks also allowed the isolation and characterisation of several new nonadride compounds: 2,2'-dihydro-2-hydroxycornexistin 15, isolated as its hemiacetal 9; 2,2'-dihydrocornexistin 10; and 6-dehydroxy-2,2'-dihydro-cornexistin 11 (Figure 2B). The relative stereochemistry of 9 was determined by comparison of chemical shift data obtained from *ab initio* calculations with the experimental values using the DP4 method described by Goodman and coworkers. Bysscholamic acid 4 itself was also detected by HR-LCMS analysis (see ESI).

![Scheme 1](Image)

Scheme 1. Proposed biosynthesis of cornexistin 1. Note that absolute configuration of 1 and 4 are known, but all others are inferred.

Two diastereomers of dihydrocornexistin were generated by catalytic hydrogenation of cornexistin 1, one of which was identical to the natural product 10 isolated from *P. variotii*. It was not possible to assign the stereochemistry at C-2 by
analysis of the $^1$H NMR data, but detailed DFT calculations (See ESI) for the two hydrogenation products showed that the naturally occurring compound 10 is most likely the 25 epimer.

The pvpsks1 gene encoding the hrPKS was disrupted\(^{16}\) by insertion of a hygromycin resistance cassette to produce PVΔpks strains. Correctly integrated transformants no longer produced 1, confirming that the cluster is associated with its biosynthesis (Figure 4B). Disruptions of the oxygenases encoded by pvL13 and pvL5 were also achieved in the WT strain which produces xanthones and various unrelated compounds (Figure 4). However, despite the presence of the unrelated compounds, disruption of pvL13 gave a strain which clearly no longer produced cornexistin 1, or hemiacetal 9, but accumulated 2,2'-dihydrocornexistin 10 and a trace of 6-dehydroxy-2,2'-dihydrocornexistin 11 (Figure 4C). Strains with a disrupted pvL5 no longer produced cornexistin 1, or diol 10, but accumulated 6-dehydroxy-2,2'-dihydrocornexistin 11 instead (Figure 4D).

Figure 4. ELSD comparison of extracts from WT P. variotii with extracts from the putative late gene disruptions and P. variotii PVΔpks1. *, xanthone related; +, unrelated.

Disruption of the Kl-like gene gave strain (AgnΔpks-PVΔL3) in which 5, 5' and only a trace of 13' were produced (Fig. 5E), again consistent with previous heterologous expression experiments with the Kl genes in the byssochlamic acid BGC.\(^5\) Knockout of pvL11, encoding a putative dienolate hydrolase, was performed to produce AgnΔpks-PVΔL11 strains which still produced cornexistin 1, albeit at significantly reduced titre (2.6 fold less – see ESI). Disruptions of pvL7, pvL9 and pvL14 (putatively identified as a gluconolactonase, a serine carboxypeptidase and a transketolase respectively) all led to the cessation of cornexistin 1 biosynthesis, but no intermediates could be identified (see ESI).

These results show that maleidride BGCs appear to be well-represented in the genomes of filamentous fungi. The core genes encode: a PKS and a hydrolase responsible for polyketide biosynthesis and probable release; and a citrate-synthase-like protein. These genes are homologous in all the systems examined to date and are also common with the squalestatin S1 BGC.\(^{17}\) A dehydratase then makes the key carboxymethyl maleic anhydride precursor 5. In the next step the carboxymethyl maleic anhydride precursor 5 must be dimerised to form either a heptadride or a nonadride. In the case of byssochlamic acid 4, four proteins seem to be involved in this step: a pair of Kl-like proteins; and a pair of PEBP-like proteins. In previous work\(^5\) we showed that both PEBP and both Kl encoding genes must be expressed for high production of 4. In the case of cornexistin 1, however, only one copy each of the Kl and PEBP genes is present and required for biosynthesis. The Kl-like protein(s) appears to be absolutely required for biosynthesis, but the presence of the PEBP-like protein(s) appears to improve titre during the biosynthesis of both 1 (KO experiments described here) and 4.\(^5\)

Figure 5. ELSD comparison of P. variotii AgnΔpks with extracts from the putative early gene disruptions. *Unrelated compounds. Note the ratios of 5, 5' and 13' depend on fermentation and extraction conditions. See references 3 and 5 for more details.
Dimerisation of the carboxymethyl maleic anhydride precursor 5 would be expected to lead to the formation of byssochlamic acid 4 as an intermediate during the biosynthesis of 1.\(^5\) A very low titre of 4 was observed in \textit{P. variotii} (see ESI), but it could not be isolated and further characterised. In lieu of this we fed 4 isolated from WT \textit{B. fulva} to the maleidride PKS KO strain of \textit{P. variotii} but cornexistin biosynthesis was not restored, presumably because 4 cannot penetrate either the cells, or an internal compartment where biosynthesis occurs.

The central steps of the pathway involve the intriguing oxidative removal of one of the maleic anhydride moieties. The \textit{pvL11}-encoded dienelactone hydrolase may be involved in this phase of biosynthesis. Its deletion reduced, but did not abolish, the biosynthesis of 1, indicating a possible spontaneous step. A similar diene lactone hydrolase (Tropl) is implicated in maleic anhydride hydrolysis during stipitatic acid biosynthesis in \textit{Talaromyces stipitatus} and its KO shows a similar reduction in titre.\(^16\) Proteins encoded by \textit{pvL7}, \textit{pvL9} and \textit{pvL14} are also implicated in the central steps of biosynthesis because their deletion leads to loss of cornexistin, but they do not appear to be involved in either the early or late steps of biosynthesis. Unfortunately no new intermediates were observed in these experiments, but it is plausible that a diene lactone hydrolase and a thiamine pyrophosphate (TPP) dependent transketolase are involved in these steps (see ESI). Further \textit{in vitro} work will be required to confirm these proposed activities. We were also unable to discover the source of the 3-hydroxylation and no gene annotated in the present cluster appears suitable for this oxygenation step.

The final steps of the pathway involve decoration of the carbon skeleton by oxygenases encoded by \textit{pvL13} and \textit{pvL5}. These steps involve selective hydroxylation at the 6-position by a non-heme iron dependent monooxygenase (pvL5) and dehydrogenation of the 2′-2′ position. This may involve initial P450 (\textit{pvL13}) hydroxylation at C-2 to give putative intermediate 15 which was not observed (the more stable hemi-acetal 9 being isolated). It is not clear how 9 is converted to 1 as it appeared stable under various conditions in our hands. It is possible that the \textit{pvL13} encoded P450 can also catalyse the required elimination to form 1 - such multifunctional cytochrome P450 oxidases are known in other contexts, for example LovA apparently catalyses dual epoxidation and elimination during lovastatin biosynthesis\(^19\) and AusE which catalyses three separate oxidative steps in spiro-lactone formation during austinol biosynthesis.\(^20\)

We have thus elucidated many of the key steps involved in the biosynthesis of the herbicide cornexistin in \textit{P. variotii}. Knockout of the xanthone biosynthetic pathway led to a significant increase in cornexistin titre, presumably by removal of competition for PKS building blocks. Several cornexistin pathway KO experiments produced new compounds which may be useful in SAR investigations and which were not previously available via synthetic chemistry approaches. The pathway confirms and amplifies our previous discoveries regarding the biosynthesis of maleidrides, but our results have also produced other intriguing possibilities surrounding the mechanism of nonadride production and the maleic anhydride moiety. Experiments to probe these questions are ongoing in our laboratories.

We thank BBSRC (KW, BB/J006289/1) and Syngenta (AS) for funding. EPSRC (EP/FO06104/1) and DFG (INST 187/621) are thanked for LCMS equipment. Sian Deller (nucleic acid preparation), Mark Montgomery (X-ray), Phil Sidebottom (NMR) and John Delaney (modelling) are also thanked for technical assistance. \textit{P. variotii} K5103 was a gift from Syngenta and was sequenced and assembled at The Genome Analysis Centre (Norwich, UK) under contract to Syngenta through Genome Enterprises Ltd.

**Notes and References**

11. Full cornexistin BGC uploaded to genbank. Accession: MF197864