Structure revision of cryptosporioptides and determination of the genetic basis for dimeric xanthone biosynthesis in fungi†

Claudio Greco, a Kate de Mattos-Shipley, a Andrew M. Bailey, b Nicholas P. Mulholland, c Jason L. Vincent, c Christine L. Willis, d e Russell J. Cox, f g * and Thomas J. Simpson f, g * a

Three novel dimeric xanthones, cryptosporioptides A–C were isolated from Cryptosporiopsis sp. 8999 and their structures elucidated. Methylation of cryptosporioptide A gave a methyl ester with identical NMR data to cryptosporioptide, a compound previously reported to have been isolated from the same fungus. However, HRMS analysis revealed that cryptosporioptide is a symmetrical dimer, not a monomer as previously proposed, and the revised structure was elucidated by extensive NMR analysis. The genome of Cryptosporiopsis sp. 8999 was sequenced and the dimeric xanthone (dmx) biosynthetic gene cluster responsible for the production of the cryptosporioptides was identified. Gene disruption experiments identified a gene (dmxR5) encoding a cytochrome P450 oxygenase as being responsible for the dimerisation step late in the biosynthetic pathway. Disruption of dmxR5 led to the isolation of novel monomeric xanthones. Cryptosporioptide B and C feature an unusual ethylmalonate subunit: a hrPKS and acyl CoA carboxylase are responsible for its formation. Bioinformatic analysis of the genomes of several fungi producing related xanthones, e.g. the widely occurring ergochromes, and related metabolites allows detailed annotation of the biosynthetic genes, and a rational overall biosynthetic scheme for the production of fungal dimeric xanthones to be proposed.

Introduction

Anthraquinones and xanthones are among the most common and earliest discovered fungal secondary metabolites, and they often have interesting and useful bioactivities. These compounds are polyketides, the skeletons of which are produced from acetyl and malonyl CoA by non-reducing polyketide synthases (nr-PKS). Examples include: the anthraquinones emodin 1 and its reduced congener chrysophanol 2 which is a precursor of shamilxanthone 3,3 and monodictytxanthone 4 from Aspergillus nidulans. Emodin 1 is a precursor of the antifungal agent geodin 5,7 which while not strictly a xanthone, belongs to the same biosynthetic family. We have also recently isolated other xanthones such as agnestin A 6 from Paecilomyces variotii and linked it to its biosynthetic genes. Cladofulvin 7 from the tomato pathogen Cladosporium fulvum is a dimeric anthraquinone also derived from 2. Dimeric xanthones are also known, and they often have varied biological activities. They include the antimalarial ascherxanthone A 8, the antimicrobial and anticancer dicerandrol C 9,12 and the mycotoxins secalonic acids B 10 and D 11 which are also known as ergochromes (Fig. 1). Although many of the dimeric xanthones, particularly the widespread secalonic acid family, have been studied for many years, key aspects of their biosynthesis such as the mechanisms and timing of the dimerisation process, and isolation and analysis of their biosynthetic gene clusters remain largely unknown.

In the course of our ongoing studies on fungal maleidride biosynthesis we analysed fermentations of the endophytic fungus Cryptosporiopsis sp. 8999, which produces the unusual octadride, viburspiran. This strain was also reported to produce a metabolite named cryptosporioptide, which was assigned the monomeric xanthone methyl ester structure 12 (Fig. 2). This structure contains an unusual ring-contracted xanthone which is difficult to rationalise biosynthetically, and an unprecedented N-malonyl aminal bridge. Further studies of this strain also led to the isolation of the corresponding free acid, named cryptosporioptide A 13, and cryptosporioptide B 14, which lacks the malonic acid ester amide bridge.
Subsequently, isolation of an analogue 15, in which the malonate bridge has been replaced by a succinate, was reported from the insect parasite *Cordyceps gracilloides*. Confusingly, this was also named cryptosporioptide A. The structure of the parent cryptosporioptide 12 was assigned on the basis of FAB-HRMS which indicated a molecular formula $C_{19}H_{19}NO_{10}$. Detailed analysis of 1D and 2D $^1$H and $^{13}$C NMR spectra gave the connectivities and relative stereochemistry on which the structures were assigned. The absolute configuration was assigned by circular dichroism allied to time-dependent density functional theory (TDDFT) computational procedures. The related structures 13 and 15 were assigned on the basis of the similarity of their NMR spectra and optical rotations to those originally reported for cryptosporioptide.

We now report the isolation and structure elucidation of a novel dimeric tetrahydroxanthone metabolite 16, methylation of which gives a compound with identical NMR properties to those reported for cryptosporioptide, along with related dimeric structures, 21 and 22 were also isolated. Comparison with genome sequences of fungi known to produce xanthones and related structures, allows identification and annotation of the biosynthetic gene clusters (BCGs) for several monomeric and dimeric xanthones, *inter alia* the secalonic acids.

### Results

*Cryptosporiopsis* sp. 8999 was grown on a range of different media. Three related metabolites were produced in good yields when the fungus was grown on brown rice (e.g. Fig. 4B). The major metabolite (900 mg kg $^{-1}$) had almost identical IR, UV and $^1$H and $^{13}$C NMR data apart from the lack of a methoxy group when compared to that reported for cryptosporioptide A 13. When esterified using TMS-diazomethane, the resulting methyl ester had identical NMR data to those reported for cryptosporioptide 12 by Saleem et al., for which the reported FAB-HRMS data was 444.0781 $[M + Na]^+$ ($C_{19}H_{19}NO_{10}Na$). However, when the molecular weight for the derivatised compound was re-determined using FAB-HRMS and ESI-HRMS, we obtained HRMS of 801.1506 $[M + Na]^+$, and 778.1828 $[M]^+$ respectively, both consistent with a molecular formula $C_{38}H_{34}O_{18}$ and notably lacking...
nitrogen. The $^1$H NMR spectrum showed 9 hydrogen environments, but 19 different carbon environments were observed in the $^{13}$C NMR spectrum consistent with a symmetrical dimer.

The presence of a tetrasubstituted aromatic ring was indicated by ortho-coupled hydrogens ($\delta_h$ 7.70 and 6.43 ppm, $J = 8.5$ Hz) in the $^1$H NMR spectrum, which also showed two methyl singlets ($\delta_h$ 1.65 and 1.59 ppm), two oxygen bearing methine singlets ($\delta_h$ 3.42 and 5.64 ppm), two low field exchangeable hydrogens ($\delta_h$ 11.69 and 14.02 ppm) and geminally coupled methylenes ($\delta_h$ 3.39 and 3.44 ppm, $J = 15.8$ Hz). The $^{13}$C NMR spectrum also showed other signals attributable to a phenolic ring ($\delta_c$ 159.2, 157.2, 140.6 and 106.1 ppm), four ester/enolic carbons ($\delta_c$ 165.5, 166.7, 170.0 and 104.4 ppm) and a benzo-phenone carbonyl ($\delta_c$ 187.9 ppm). Signals at $\delta_c$ 56.0, 58.8, 73.9 and 78.9 ppm are consistent with the presence of an epoxide, secondary alcohol and ether. The connectivities were determined from extensive HMBC correlations which confirmed the presence of the aromatic ring, and a highly substituted cyclohexene (Fig. 3), the key signal being H-5 (5.64 ppm) which correlates with a total of eight carbons ($6, 7, 8a, 9, 10a, 11, 12$ and 13) at $\delta_c$ 58.8, 56.0, 170.0, 187.9, 78.9, 28.6, 18.2 and 165.5 ppm respectively). The H-5/C-13 correlation establishes the position of attachment of the malonate moiety, the other malonate carbonyl (C-15) correlating with the methoxyl ($\delta_h$ 3.65, $\delta_c$ 52.7 ppm). The connectivities between the two carbocyclic rings were confirmed by observation of HMBC correlations (see ESI Fig. S15†) from both H-5 (5.64 ppm) and H-4 (6.43 ppm) to the C-9 ketone ($\delta_c$ 187.9 ppm), and NOEs between the 12-methyl and H-4 in the aromatic ring. The relative stereochemistry C-5 to C-7 and C-11 was established by NOEs between H-5 and H-7 and the 11- and 12-methyls.

While the absolute configuration of the original cryptosporioptide structure was established by comparison of the measured circular dichroism (CD) with that predicted by computational methods, the complete revision of the structure renders this assignment invalid, and indeed suggests caution should be exercised when applying these methods which are being increasingly used for configurational assignment. The absolute configuration at the C-10a stereocentres of secalonic acid B 10 and related structures, e.g. dicerandrol C 9 have been assigned on the basis of a positive n-π* CD band at 330–340 nm as 10aR and 10aS respectively (note change in CIP designation due to priority inversion for CO$_2$Me and CH$_2$OAc).20–22 The absolute configuration of dicerandrol C 9 has also been confirmed by total synthesis.24 Cryptosporioptide A dimethyl ester 17 has a reported CD (341 nm, $\Delta\varepsilon = +5.0$)17 which is consistent with the 10aR configuration. Thus cryptosporioptide can be designated as 5R,6S,7S,10aR.

Finally, the connectivity between the individual xanthone monomers was established by an intra-dimer NOE between 1-OH and H-3. The point of dimerisation was confirmed by isolation of monomers 21 and 22 following KO of DmxRS, the cytochrome P450 responsible for oxidative coupling (see below). Thus the previously reported structures 12, and 13 have been revised to 17 and 16 respectively, and we propose renaming them cryptosporioptide A dimethyl ester, and cryptosporioptide A respectively. We have not observed any trace of the methyl ester 17 in any of our extracts, and the reported isolation of 17 as a natural product is possibly an artefact of the purification involved (Sephadex LH20 eluted with methanol).

The remaining two dimeric metabolites, cryptosporioptides B 18 and C 19, differ from cryptosporioptide A 16 in the malonyl subunit, their UV and NMR spectra being otherwise identical (see ESI Tables S1 and S2†). HRMS showed their molecular formulae to be C$_{38}$H$_{34}$O$_{18}$ and C$_{40}$H$_{38}$O$_{18}$ respectively. The NMR spectra showed the latter 19 to be also essentially symmetrical.
with signals attributable to the ethylmalonate moieties ($\delta_{H} 3.29, 2H, m; 1.81, 4H, m; 0.85, 6H, m$). Doubling of some signals was observed, probably due to facile epimerisation of the ethylmalonyl substituent, giving diastereomers at C-13. The spectra for 18 are more complex with many signals doubled due to the presence of only one ethylmalonate moiety removing the symmetry. Again on the basis of the close similarities of the NMR spectra, the C. graciloides metabolite\textsuperscript{19} 15 has been renamed as cryptosporioptide D 20.

While the biosynthetic pathways towards monomeric xanthones and related compounds such as geodin 5 have been investigated previously, there is much less information on the dimeric systems despite these having been the subject of intensive study over many years.\textsuperscript{16} The main question is the timing of ring cleavage to give the xanthone moiety relative to dimerisation. The recently reported\textsuperscript{20} co-occurrence of the monomeric benzenolides (e.g. ergochromene B) with their symmetrical dimer, secalonic acid B 10 supports, but does not unequivocally prove that during secalonic acid biosynthesis, xanthone formation precedes dimerisation. In those systems where the anthraquinonoid carbonyl (i.e. C-10) is retained, it is often found as a methyl ester as in the secalonic acids (note C-12 in SA numbering), or partially reduced to a hydroxymethyl equivalent, e.g. in dicerandrol C 9, or fully reduced to a methyl group as in the ascherxanthones, e.g. 8 and cryptosporioptides 16–20. Other questions include: which of the original anthraquinone rings retains the cleaved carbonyl C-10 (C-12 in SA nomenclature) – that containing the polyketide starter unit as in secalonic acids, e.g. 10 and cryptosporioptides or the non-starter ring as in, e.g. ascherxanthones 8; and the location of the site of dimerisation (Fig. 1).

With these questions in mind, the 54 Mbp draft genome of Cryptosporiopsis sp. 8999 was obtained using Illumina MiSeq and assembled and annotated using Newbler v2.9. The total number of contigs was 2166 with an N50 of 204 Kb. Bioinformatic studies using antiSMASH\textsuperscript{21} identified ten BGCs containing non-reducing polyketide synthase (nr-PKS) genes. A putative gene cluster for xanthone biosynthesis\textsuperscript{22} was readily identified by BLAST analysis using protein sequences from: the shamixanthone 3-monocetetone 4 BGC from Aspergillus nidulans;\textsuperscript{23,24} the agnestin 6 BGC from Paecilomyces variotii;\textsuperscript{6} the recently (but partially, vide infra) described secalonic acid 10 BGC from Claviceps purpurea;\textsuperscript{27} the geodin 5 BGC from Aspergillus terreus;\textsuperscript{2} and the cladofulvin 7 BGC from Cladosporium fulvum.\textsuperscript{2} Development of a transformation system, targeted knockout of the putative cryptosporioptide (dmx) PKS using the bipartite method of Neilson and coworkers,\textsuperscript{28} and subsequent observation of abolition of all cryptosporioptide production (Fig. 4) proved this assignment to be correct. We then used the cryptosporioptide BGC to screen the genomes of other fungi for similar clusters. Penicillium oxalicum\textsuperscript{29} and Aspergillus aculeatus\textsuperscript{30} are both known to produce secalonic acid D 11,\textsuperscript{31} and their genomes have been sequenced.\textsuperscript{32} In both cases we found a BGC (P. oxal, A. aculea Table 1) featuring homologs of many of the genes in the dmx BGC.

The dmx BGC spans approximately 84 Kb surrounding dmxPKS which encodes a fungal nr-PKS homologous to the monocetetone PKS (MdpG, 67% identity) from A. nidulans\textsuperscript{6} and the analogous polyketide synthases encoded by the geodin 5 (ged), agnestin 6 (agn) and secalonic acid 10, 11 (sec) clusters (Table 1, Fig. 5). Homologies to all genes which encode proteins required for the synthesis of emodin 1 (dmxPks, dmxR1, dmxR15 and dmxR16) are present, as are the genes required for the synthesis of chrysophanol 2 (dmxR7, dmxR17 and dmxR18). The geodin pathway, which does not proceed via chrysophanol 2, lacks homologs of dmxR17 and dmxR18.

Oxidative ring-opening of chrysophanol 2 is specified by dmxR6 which encodes a Baeyer–Villiger monoxygenase (BVO) homologous to those encoded by the mdp, ged, agn and sec clusters where this transformation is required. Notably cladofulvin, which is a dimeric anthraquinone, does not require this chemistry and the cla cluster lacks a dmxR6 homolog. Beyond this point the content of the clusters varies significantly.

The dmx BGC has a series of unique genes not present in the other clusters such as dmxR2, dmxR8, dmxR11–13 and dmxL1–3. In particular dmxL2 encodes a highly reducing PKS (hr-PKS) homologous to the lovastatin diketide synthase\textsuperscript{33} and squalestatin tetraketide synthase.\textsuperscript{34} The gene dmxL1 encodes an acyl-CoA carboxylase. The gene dmxR5 encodes a cytochrome P450 enzyme homologous to ClaM which is known to dimerise natale-emodin to form cladofulvin 7.\textsuperscript{35,36} This gene is missing from the clusters of the monomeric compounds, but is present in the sec cluster (CPUR_05419). Finally, dmxR13 encodes an O-acyl transferase, and again this is absent from the other clusters.

In order to gain evidence for the function of these genes we devised knockout experiments. Disruption of dmxR6 (BVO) gave high titres of chrysophanol 2 as expected, showing that dimerisation occurs after anthraquinone ring cleavage and xanthone formation (Fig. 6). Deletion of dmxR5 (putative dimerase) abolished dimer production and two novel metabolites with similar UV spectra to the cryptosporioptides but with molecular weights consistent with monomers were isolated (Fig. 7). Full NMR analysis (see ESI)\textsuperscript{1} confirmed that these were hemi-cryptosporioptide 21 (C$_{19}$H$_{18}$O$_{9}$), and an analogue 22 (C$_{19}$H$_{18}$O$_{8}$) containing a tertiary alcohol at C-6 which was isolated as a mixture of epimers at C-5.

When monomer 21, but not 22, was re-fed to the ΔdmxPKS mutant, cryptosporioptide A 16 production was restored (Fig. 8), confirming this as a pathway intermediate.
Disruptions of dmxL2 and dmxL1 both abolished production of cryptosporioptide B 18 and C 19, to give exclusively cryptosporioptide A 16 (Fig. 9B and C). Knockout of dmxR13 gave a similar result, abolishing cryptosporioptide C 18 and D 19 production but retaining cryptosporioptide A 16 in smaller quantities than previously (Fig. 9D). A number of new metabolites were produced in small quantities that precluded structure elucidation although mass spectrometry indicated molecular

Table 1  The cryptosporioptide dimeric xanthone (dmx) cluster and similarities of encoded proteins with those of the monodictyphenone (mdp), geodin (ged), agnestin (agn), cladofulvin (cla) and secalinic acid (sec; P. oxa and A. acu) clusters

<table>
<thead>
<tr>
<th>Chemistry</th>
<th>Predicted function</th>
<th>Dmx</th>
<th>mdp</th>
<th>% ID</th>
<th>Ged</th>
<th>% ID</th>
<th>Agn</th>
<th>% ID</th>
<th>Cla</th>
<th>% ID</th>
<th>Sec</th>
<th>% ID</th>
<th>P. oxa</th>
<th>% ID</th>
<th>A. acu</th>
<th>% ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steps to emodin hydroquinone</td>
<td>Non-reducing PKS Hydrolyase/thioesterase Decarboxylase Anthrone oxidase Oxidoreductase</td>
<td>DmxPKS MdpG 66.6 GedC 65.1 Agnl1 66.4 ClaG 53.6</td>
<td>05437 58.5</td>
<td>09227 62.4</td>
<td>AausL 62.8</td>
<td>DmxR1 MdpF 68.5 GedB 68.7 Agnl7 65.1 ClaF 55.3</td>
<td>05436 65.5</td>
<td>09238 64.7</td>
<td>AausM 64.1</td>
<td>DmxR15 MdpH1 60.1 GedC 65.1 Agnl1 62.8 ClaH1 75.7</td>
<td>05434 76.0</td>
<td>09224 61.4</td>
<td>AausD 62.2</td>
<td>DmxR16 MdpH2 41.2 GedH 42.4 Agnl2 46.7 ClaH2 52.7</td>
<td>05435 37.6</td>
<td>09222 61.2</td>
</tr>
<tr>
<td>Steps to chrysophanol 2</td>
<td>Dehydratase Reductase</td>
<td>DmxR17 MdpB 64.1</td>
<td>05428 69.6</td>
<td>09226 58.5</td>
<td>AausK 57.2</td>
<td>DmxR18 MdpC 79.7</td>
<td>05429 85.6</td>
<td>09239 71.6</td>
<td>AausN 61.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BVMO</td>
<td>Baeyer–Villiger oxidase</td>
<td>DmxR6 MdpL 46.1 GedK 44.9 Agnl3 65.3</td>
<td>05427 53.7</td>
<td>09233 62.1</td>
<td>AausH 61.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-S reductase</td>
<td>SDR</td>
<td>DmxR3</td>
<td>---</td>
<td>05418 58.0</td>
<td>---</td>
<td>AausD 64.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimerase</td>
<td>Cytochrome P450</td>
<td>DmxR5</td>
<td>---</td>
<td>05419 39.7</td>
<td>---</td>
<td>AausE 45.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-S hydrox.</td>
<td>Monoxygenase</td>
<td>DmxR9 MdpD 68.5</td>
<td>05422 55.8</td>
<td>09229 55.8</td>
<td>AausC 56.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dmx, agn and sec only</td>
<td>Hypothetical</td>
<td>DmxR10</td>
<td>---</td>
<td>05417 20.0</td>
<td>---</td>
<td>AausO, AausP 47.3, 47.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ged and sec only</td>
<td>Methyltransferase</td>
<td>---</td>
<td>---</td>
<td>05424</td>
<td>09224</td>
<td>AausQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ged only</td>
<td>Glutathione-S-transferase</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mdp only</td>
<td>Acetyl-CoA synthase Glutathione-S-transferase</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mdp only</td>
<td>Methyltransferase</td>
<td>DmxR11 MdpL 64.1 MdpJ 44.9</td>
<td>05420</td>
<td>09235</td>
<td>AausJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mdp only</td>
<td>Acetyl-CoA synthase Glutathione-S-transferase</td>
<td>DmxR12 MdpC 68.5 GedE 46.2</td>
<td>05431</td>
<td>---</td>
<td>AausL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sec only</td>
<td>Hypothetical Monoxygenase SDR (C-S reductase)</td>
<td>DmxR14</td>
<td>---</td>
<td>05433 32.8</td>
<td>09243 27.0</td>
<td>AausR 28.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-catalytic</td>
<td>Transcription factor Coactivator</td>
<td>DmxR15 MdpA 69.1 GedD 31.2</td>
<td>05431</td>
<td>---</td>
<td>AausE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-catalytic</td>
<td>Transcription factor MFS transporter</td>
<td>DmxR16 MdpE 48.7</td>
<td>05421</td>
<td>---</td>
<td>---</td>
<td>AausS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-catalytic</td>
<td>Transcription factor</td>
<td>DmxL4</td>
<td>---</td>
<td>05422 47.9</td>
<td>09228 52.1</td>
<td>---</td>
<td>AausB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-catalytic</td>
<td>ABC transporter</td>
<td>DmxL5 29.4</td>
<td>05432</td>
<td>09224</td>
<td>---</td>
<td>AausA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5  Map of the cryptosporioptide dimeric xanthone (dmx) BGC coloured by proposed function (see Table 1). Not to scale.

Fig. 6  Knockout of dmxR6 (BVMO). (A) LCMS analysis (DAD) of organic extract of Cryptosporiopsis sp. 8999 dmxR6 fermented on brown rice; (B) LCMS analysis (DAD) of organic extract of wild type Cryptosporiopsis sp. 8999 fermented on brown rice. LCMS method 50–90% CH3CN:H2O gradient, 20 min. * = unrelated compound.
weights consistent with monomeric xanthones, in particular non-malonylated monomeric xanthone 35 (Scheme 1).

Discussion

From these results, a biosynthetic sequence for cryptosporiopptides can be proposed (Scheme 1) and more general conclusions for the formation of dimeric xanthones made. The BGCs encoding the biosynthesis of shamixanthone 3 (mdp),6 geodin 5 (ged),7 agestin 6 (agn)8 and cladofulvin 7 (cla) are all known. Tbudynski and coworkers recently revealed a putative secalonic acid (sec) BGC in Claviceps purpurea.27 In their analysis fifteen genes were described, but here, by comparison to the other known BGCs we extend this analysis to include six more genes in the sec cluster and two additional gene clusters from Penicillium oxalicum and Aspergillus aculeatus (Table 1 and see ESI).

All analysed clusters encode proteins with high homologies (41–76% orf for orf) to proteins known to be involved in the biosynthesis of emodin hydroquinone 26, and all BGCs except that for geodin contain genes which advance the pathway to chrysophanol 2 (Table 1). Thus, the nrPKS (DmxPKS) produces an enzyme-bound octaketide 23 which is released (DmxR1, giving 24), decarboxylated (DmxR15, giving 25) and oxidised (DmxR16) to give emodin 1, followed by reduction (DmxR7) to emodin hydroquinone 26. These steps are fully consistent with previous experimental results. A-ring reduction (DmxR18 giving 27), dehydration (DmxR17) and probable spontaneous re-oxidation, results in overall deoxygenation to chrysophanol 2, again consistent with previous results. Baeyer–Villiger oxidation (DmxR6) would then be expected to give monodicty lactone 28 in equilibrium with monodictyphenone 29 which could give the electron-rich reduction of C-12 to an alcohol 30 and hydroxylation at C-5 showing that a specific methyltransferase NsrG directs monoester 29b towards the biosynthesis of the heterodimeric xanthone neosartorin.27 This is consistent with the presence of methyltransferase-encoding genes (CPUR_05424, P_oxa-09242 and aacuQ) in the secalonic acid BGCs but not the others (Table 1). Alternatively, reduction at C-12 to form an aldehyde 29b would direct the pathway towards the shamixanthone group. In the case of the cryptosporiopptides, however, reduction of C-12 to an alcohol 30 and hydroxylation at C-5 (likely DmxR9, see below) could give the electron-rich aromatic 31 which could eliminate H2O to form the ortho-quinonemethide 32, followed by tautomeration to para-quinonemethide 33 and complete the formal reduction to produce the 10-methyl group.

C-5 hydroxylation is required in the cryptosporiopside 16, shamixanthone 3 and secalonic acid 10,11 pathways. In shamixanthone 3, this is proposed to be carried out by the monooxygenase MdpD, which shows 68% homology to DmxR9, while CPUR_05423 encodes a homolog in the sec cluster. Homologs of DmxR9 are also encoded in the P. oxalicum and A. aculeatus
clusters. Notably homologs of \textit{dmxR9} are missing from the \textit{ged}, \textit{agn} and \textit{cla} clusters where this chemistry is not required.

In an early application of doubly $^{13}$C-labelled acetate feeding experiments, Vining and coworkers showed\textsuperscript{38} that during secalonic acid biosynthesis in \textit{Pyrenochaeta terrestris}, the initial product of anthraquinone ring cleavage must have a symmetrical 1,3-dihydroxyphenyl ring as in \textit{29–33}. Our suggested pathway is in agreement with this observation.

We propose that conjugate addition of C-4a-OH to the resulting para-quinone \textit{33} then gives cyclohexadienone \textit{34}, which is then reduced at C-5 to give the dihydroxanthone \textit{35}. The ring closing reaction may be performed by DmxR10. This protein has no close BLAST hits with proteins of known function, but structural analysis using Phyre-2 (ref. 39) (see ESI\textsuperscript{†}) shows that it contains a SnoAL domain\textsuperscript{40} and is related to proteins such as PhzB\textsuperscript{41} and Trt14 (ref. 42) known to be involved in secondary metabolite ring-forming reactions.

In this step the stereochemistry at C-10a is set, and it is known that both epimers at this position can be formed in the cases of the secalonic acids (e.g. see Fig. 10). It is interesting to note that the three secalonic acid clusters examined here (Table 1) appear to encode multiple copies of this protein, conceivably explaining the presence of both 10a-epimers in these systems. However further experimental work will be required to confirm this hypothesis.

The 6,7-epoxide in the cryptosporioptides could be introduced by the cytochrome P450 monooxygenase encoded by
$dmxL3$, which is unique to the $dmx$ cluster, to give $36$. Our results suggest that the hrPKS ($DmxL2$) manufactures butyrate. This is consistent with domain analysis of $DmxL2$ (NCBI CDD)$^{33}$ which shows it to consist of N-terminal KS and AT domains, followed by a DH domain. Highly-reducing PKS often feature a C-methyltransferase domain but this appears to be absent or inactive in $DmxL2$, consistent with the lack of methylation observed. Canonical C-terminal ER, KR and ACP domains make up the rest of the hrPKS. Butyrate is then carboxylated ($DmxL1$) to form ethylmalonate. It is not yet clear whether the carboxylation occurs while the butyrate is attached to the ACP of $DmxL2$, but this unusual fungal metabolite could then be esterified to O-5 by $DmxR13$.

In the absence of ethylmalonate or the acyl transferase $DmxR13$, malonyl CoA can be used, possibly transferred by $DmxR2$. Finally, dimerisation ($DmxR5$) gives the observed dimers $16$, $18$ and $19$ as the final products of the pathway. A homolog of $dmxR5$ is found in the $C. purpurea$ sec cluster (CPUR\_05419, again previously unrecognised) and $P. oxalicum$ and $A. aculeatus$ clusters, and the cladofulvin cluster where $claM$ has been already shown to direct oxidative dimerisation.$^9$ However, $dmxR5$ is missing from the $agn$, $mdp$ and $ged$ clusters as expected. The lower relative homology of the cladofulvin dimerase may reflect the fact that cladofulvin is a dimeric anthraquinone rather than a dimeric xanthone, and that cladofulvin is an asymmetric dimer in contrast to the symmetric cryptosporiopiptide and secalonic acid dimers.
A key step in the pathways to the dimeric xanthone metabolites is reduction at C-5. This could result in either 5S or 5R products depending on the selectivity of the reductase. Both types of xanthones are known (Fig. 10) and are, as shown above, likely to be direct precursors of the secalonic acids. Indeed, the ergochrome nomenclature explicitly recognises this, such that secalonic acid A 38 is also known as ergochrome AA showing that it is a dimer of two xanthone A 39 units. Xanthones A 39, B 40, C 41 and D 42 have C-5S stereochemistry, while xanthones E 43 and F 44 and blennolide B 45 have 5R stereochemistry. C. purpurea produces secalonic acids which are derived from xanthones A–D and it thus must reduce to give exclusively 5S stereochemistry. In contrast P. oxalicum produces secalonic acid D 11 which derives from xanthone E 43 which has 5R stereochemistry. Aspergillus aculeatus, however, produces both secalonic acids B 10 and D 11 which are made from xanthones B 40 and E 43 and it must therefore be able to reduce at C-5 to give both possible stereoisomers. In accord with this observation the sec cluster encodes a single SDR, which is homologous to DmxR3 (58%). The cryptosporioptides 16, 18 and 19 and secalonic acids A 38 and B 10 possess 5S stereochemistry, so we propose that DmxR3 is a 5S reductase. The P. oxalicum cluster possesses a different SDR encoded by P_oxa-09231 and we propose that this is a 5R selective reductase. In agreement with these ideas the A. aculeatus cluster encodes two SDRs, one of which (AacuD) is homologous to the 5S reductase DmxR3, while the other (AacuF) is homologous to the 5R selective P_oxc-09231. Homologous SDRs are not encoded by the agn, geo or cla clusters where the pathways do not require C-5 reduction, again consistent with this hypothesis.

Conclusions

Here we have shown that the originally reported monomeric structures of the cryptosporioptides must be reassigned to a series of dimeric xanthones featuring unusual malonate substituents and full reduction at C-10 (=C-12 in SA nomenclature). Isolation of the cryptosporioptide BGC and its verification by directed knockout experiments allowed the unusual biosynthesis and attachment of the ethylmalonyl substituents to be determined. Furthermore, comparison with known, but homologues of the pathways revealed here should allow further targeted engineering in known dimeric xanthone clusters to be designed and performed, and it should also allow clusters encoding new dimeric xanthones to be more rapidly recognised and discovered.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank BBRC (BB/J006289/1, BB/L01386X/1) and Syngenta for funding. LCMS instruments were provided by EPSRC (EP/F006104/1) and DFG (INST 187/621). 500 MHz NMR (EP/L01999/1) was provided by EPSRC. Cryptosporiopsis sp. 8999 was a kind gift from Dr Barbara Schulz and its genome was sequenced at the University of Cambridge Sequencing Centre. Markiyan Samborskyj is thanked for assistance with bioinformatics. Katherine Williams and Zhongshu Song are thanked for technical assistance.

Notes and references
