

Cite this: *Chem. Sci.*, 2019, 10, 233

All publication charges for this article have been paid for by the Royal Society of Chemistry

Characterisation of the biosynthetic pathway to agnestins A and B reveals the reductive route to chrysophanol in fungi†

Agnieszka J. Szwalbe,^a Katherine Williams,^a Zhongshu Song,^a Kate de Mattos-Shipley,^a Jason L. Vincent,^b Andrew M. Bailey,^c Christine L. Willis,^d Russell J. Cox^{ade} and Thomas J. Simpson^{id}*^a

Two new dihydroxy-xanthone metabolites, agnestins A and B, were isolated from *Paecilomyces variotii* along with a number of related benzophenones and xanthenes including monodictyphenone. The structures were elucidated by NMR analyses and X-ray crystallography. The agnestin (*agn*) biosynthetic gene cluster was identified and targeted gene disruptions of the PKS, Baeyer–Villiger monooxygenase, and other oxido-reductase genes revealed new details of fungal xanthone biosynthesis. In particular, identification of a reductase responsible for *in vivo* anthraquinone to anthrol conversion confirms a previously postulated essential step in aromatic deoxygenation of anthraquinones, e.g. emodin to chrysophanol.

Received 23rd August 2018
Accepted 23rd November 2018

DOI: 10.1039/c8sc03778g

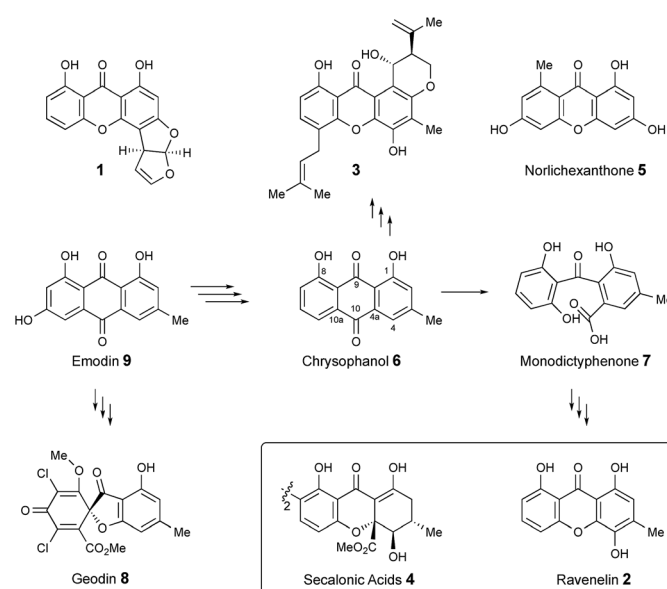
rsc.li/chemical-science

Xanthenes and related benzophenones are produced by a variety of filamentous fungi.¹ Examples (Scheme 1) include desmethyl-sterigmatocystin 1,² a key intermediate to the aflatoxin group of mycotoxins produced by *Aspergillus flavus*, ravenelin 2³ from *Drechslera ravenelii* and prenylated xanthenes, e.g. shamixanthone 3 from *Aspergillus variegator*.⁴ More complex derivatives include the dimeric xanthone ergochromes (secalonic acids) e.g. 4 from *inter alia* *Claviceps purpurea*⁵ and *Penicillium oxalicum*.⁶

While some fungal xanthenes are produced *via* simple folding of a polyketide chain, e.g. norlichexanthone 5 from *Lecanora straminea*,⁷ most are produced *via* Baeyer–Villiger-type oxidation of anthraquinones such as chrysophanol 6 to give benzophenones such as monodictyphenone 7 which are subsequently cyclized to xanthenes.⁸ Recent interest in these pathways has been stimulated by the results of genome sequencing and bioinformatic analysis⁹ which has allowed identification of the gene clusters in, e.g. *Aspergillus nidulans* for biosynthesis of shamixanthone 3.^{10,11} Other related compounds include geodin 8¹² which is derived from emodin 9, the

precursor of chrysophanol 6,¹³ while cladofulvin from the tomato pathogen *Cladosporium fulvum* is a dimeric naphthoquinone also derived from 9.¹⁴ Despite these advances key questions remain about the precise genes, proteins and chemical steps involved.

In the course of studies on the biosynthesis of maleidrides,¹⁵ we examined extracts of the cornexistin producing fungus *Paecilomyces variotii*.¹⁶ These extracts contained mainly the



Scheme 1 Pathways emanating from emodin 9 and related compounds.

^aSchool of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK. E-mail: tom.simpson@bristol.ac.uk

^bSyngenta, Jealott's Hill International Research Centre, Bracknell, RG42 6EY, UK

^cSchool of Biological Sciences, 24 Tyndall Avenue, Bristol, BS8 1TQ, UK

^dInstitute for Organic Chemistry, Leibniz University of Hannover, 30167, Germany

^eBMWZ, Leibniz University of Hannover, 30167, Germany

† Electronic supplementary information (ESI) available: All experimental and characterisation details. CCDC 1839028 and 1839029. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c8sc03778g

benzophenone monodictyphenone **7** and other unknown aromatic metabolites. Here we determine the structures of the unknowns, the associated biosynthetic gene cluster and reveal key new redox steps on the fungal pathway to chrysophanol **6**.

Isolation of metabolites

Extracts of *P. variotii* K5103 contained several major aromatic metabolites (Fig. 1) which were isolated by semi-preparative HPLC. The material eluting at 11.4 min (typical yield of 20 mg L⁻¹ after 15 days fermentation) had ¹H and ¹³C NMR (see ESI†) and UV spectra which were identical to those for monodictyphenone **7**.¹⁷ Closer inspection of the ¹H NMR spectrum revealed the presence (ca. 25%) of a co-eluting isomer whose NMR spectra (see ESI†) matched those of the previously reported cephalone F **10** from *Graphiopsis chlorocephala*.¹⁸ Monodictyphenone **7** is believed to be formed by a Baeyer-Villiger monooxygenase (BVMO) cleavage of the C-10/C-10a bond of chrysophanol **6** and the formation of cephalone F **10** can be rationalized by alternative cleavage at the C-4a/C-10 bond. The reported ¹H NMR spectrum of cephalone F **10** contained signals due to a minor component (ca. 5%) clearly corresponding to those of monodictyphenone **7**, suggesting that the putative BVMO in *G. chlorocephala* has a complementary regioselectivity to that in *P. variotii*.

The peak eluting at 13.6 minutes (typical yields of 15 mg L⁻¹) had the same molecular formula as **7**, C₁₅H₁₂O₆ (calc. 311.0532, measured 311.0536 for [M + Na]⁺), but whereas the main low resolution mass fragment ion in monodictyphenone **7** occurred at M-44 (-CO₂), indicating that a carboxylic acid was present, the main fragment ion now occurred at M-62 (-CO₂, -H₂O). The NMR spectra also had similarities to monodictyphenone **7**, e.g. a methyl group (δ_{H} 2.05), a highly conjugated ketone (δ_{C} 181.7)

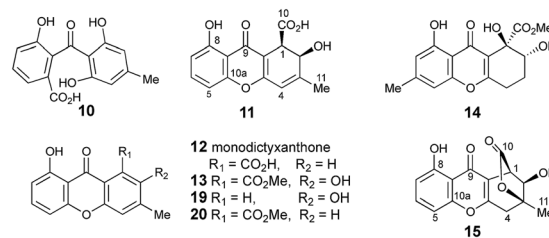


Fig. 3 Cephalone F **10** and xanthonones discussed in the text.

and three adjacent aromatic hydrogens (δ_{H} 6.91, 7.55 and 6.75) confirmed by their characteristic *ortho* couplings and COSY correlations.

The methyl signal showed long range couplings and COSY correlations to signals at δ_{H} 6.08 and 4.86, the latter being further coupled to a doublet (8.5 Hz) at 4.07 ppm. COSY and HMBC correlations (Fig. 2), were consistent with the dihydroxanthone structure **11**. Slow crystallization from methanol gave crystals which allowed a high quality X-ray crystal structure (see ESI†) to be obtained (Flack parameter -0.01(4)) which confirmed structure **11**, here named agnestin A, and the (1*R*,2*R*) absolute stereochemistry.¹⁹ It is relatively stable due to the *gauche* relationship of the carboxyl and hydroxyl substituents although slow dehydration does occur to give the known monodictyxanthone **12**²⁰ whose presence was also detected in older (>25 days) cultures. The related dehydro-compound **13** (Fig. 3) is known from fungal endophytes of *Picea glauca*,²¹ and the similar tetrahydro-xanthone analogue, α -diversonolic ester **14** is known from *Penicillium diversum*.²²

The peak eluting at 14.3 minutes (typical yields of 5–10 mg L⁻¹, Fig. 1) was a further structural isomer with the molecular formula C₁₅H₁₂O₆ (HRMS calc. 311.0532, measured 311.0521 for [M + Na]⁺). The ¹H NMR spectrum indicated a close resemblance to agnestin A **11**: three adjacent mutually coupled aromatic hydrogens were clearly present at δ_{H} 6.92, 7.58 and 6.78 ppm. The main difference was the presence of two methine singlets (δ_{H} 3.87 and 4.11 ppm) and a closely coupled methylene AB signal centred at 3.15 ppm. The two singlets showed long-range COSY correlations to a methyl group (1.53 ppm), and the C-3 resonance at δ_{C} 158.1 ppm in agnestin A **11** was replaced by one at 86.8 ppm. These observations, combined with HMBC correlations (Fig. 2) were consistent with the bicyclic lactone-containing structure **15**, which we have named agnestin B.

Consistent with this, samples of agnestin A **11** were observed (HPLC) to slowly convert to agnestin B **15** and give colourless crystals on slow evaporation from acetonitrile. X-ray analysis (see ESI†) confirmed the structure of agnestin B **15** and also explained the lack of observable coupling between H-1 and H-2 due to the observed dihedral angle of 77° (see ESI†). The 1*R*,2*R* absolute configuration of agnestin B **15** was assigned on the basis of its formation from rearrangement of agnestin A **11**.

To examine if either agnestin A **11** or agnestin B **15** was merely an extraction-artefact formed by acid-catalysed rearrangement of the other, a comparison of extraction conditions was made. One culture was worked-up under the normal acidified conditions, while another was extracted under neutral

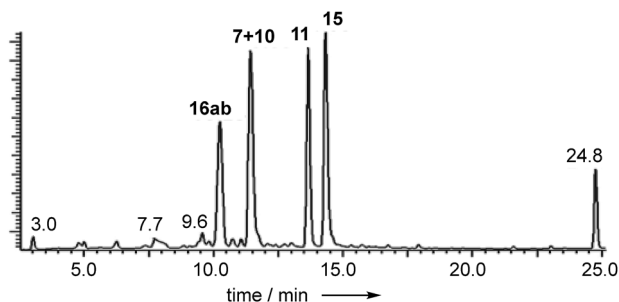


Fig. 1 Typical DAD chromatogram of wild type *P. variotii* K5103 culture extract. Bold, compound numbers; other numbers retention times.

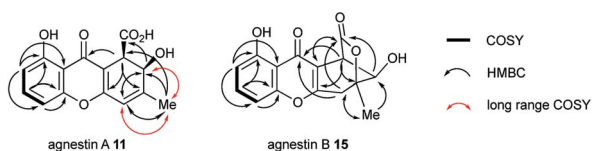


Fig. 2 2D correlations (H to C and H to H) observed for agnestin A and agnestin B.



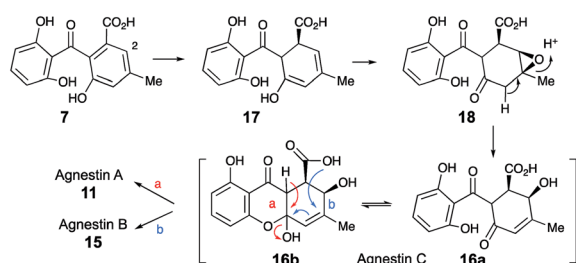
conditions. In each case both agnestins were observed, suggesting that both are true *in vivo* products.

Extracts of younger cultures of *P. variotii* (4–5 days) show the presence of a major peak at 10.2 minutes with a molecular formula of $C_{15}H_{14}O_7$ (calc. 305.0667, measured 305.0660 for $[M - H]^-$) corresponding to formal addition of H_2O to the agnestins. However, on attempted isolation, the compound rearranged to give mainly agnestin A **11** with some agnestin B **15** (see ESI†). Initial NMR studies of a freeze-dried HPLC-purified fraction showed the presence of two major components, one of which appeared to be a benzophenone due to the presence of the characteristic *ortho* coupled one-proton triplet and two proton doublet associated with the “symmetrical” 2,6-dihydroxyphenolic ring. The other component has the triplet, doublet, doublet pattern more typical of constraining this ring into a xanthone derivative. Other signals are present at δ_H 5.8, 4.7, 3.8, 3.2 and 2.1 ppm. Similar signals were observed in a spectrum obtained directly from an HPLC-purified fraction (WET 1H NMR with solvent suppression)²³ using a 600 MHz cryo-probed spectrometer. Monitoring the 1H NMR spectrum over time showed that the mixture gradually converted to

mainly agnestin A **11** after 15 hours (see ESI†). Simple biosynthetic rationale requires the 2-OH to be introduced by an oxidation at C-2 of monodictyphenone **7**, and the molecular formula requires a further reduction to give the dihydroxanthone core.

We propose that this intermediate, agnestin C, is an equilibrating *ca.* 50 : 50 mixture of structures **16a** and **16b** which could be formed *via* epoxidation of the dihydro-derivative **17** of monodictyphenone **7** to give epoxide **18** followed by rearrangement to give the allylic alcohols **16a/b**. In the absence of evidence to the contrary it is of course possible that epoxidation could precede reduction. Elimination of water could give agnestin A **11** or agnestin B **15** by either of the mechanisms shown in Scheme 2.

In addition to monodictyxanthone **12**, older (*ca.* 25 days) cultures contained two further minor xanthone derivatives. These were readily identified by comparison of their spectroscopic characteristics with a compound **19** reported²⁴ by Tarek *et al.* and a natural product **20** isolated by Li *et al.* from co-cultures of two South China coast fungal strains (Fig. 3).²⁵ The former is presumably formed from agnestin A by decarboxylation and oxidation, whereas the latter corresponds to the methyl ester of monodictyxanthone **12**. Interestingly, these extracts also contained detectable amounts of the anthraquinones emodin **9** and chrysophanol **6**.



Scheme 2 Postulated formation of agnestins A **11** and B **15** via agnestin C **16ab**.

Genetic analysis

A draft genome sequence of *P. variotii* K5103 was previously obtained.¹⁶ Identification of a putative biosynthetic gene cluster (BGC) responsible for the biosynthesis of the agnestins (*agn*) was aided by comparison with the known *mdp* **7** and *ged* **8** clusters from *A. nidulans* (Table 1).^{9,10}

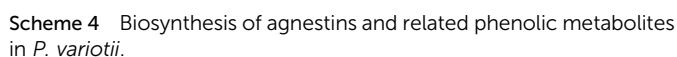
Table 1 Analysis of the agnestin biosynthetic gene cluster from *P. variotii* K5103. Proteins homologous with those encoded by the *mdp* and *ged* BGC are shaded. Mdp, monodictyphenone BGC, Ged, geodin BGC, nr-PKS, non-reducing polyketide synthase

Gene	Putative function	Swissprot homolog (% id.)	Mdp homolog (% id.)	Ged homolog (% id.)	Ref.
<i>agnL12</i>	MFS transporter	lepC, B8NJG7 (39)	—	—	27
<i>agnL11</i>	Transcription factor	lepB, B8NJG5 (31)	—	—	—
<i>agnL10</i>	Transcription factor	MdpE, Q5BH32 (33)	MdpE, (33)	GedR, (34)	28
<i>agnL9</i>	Regulation	MdpA, (43)	MdpA, (43)	GedD, (40)	—
<i>agnL8</i>	Dehydratase	SCD1, Q00455 (63)	MdpB, (57)	—	29
<i>agnL7</i>	Hydrolase	MdpF, Q5BH31 (69)	MdpF, (69)	GedB, (65)	30
<i>agnL6</i>	Reductase	CPUR_05429, M1W270 (74)	MpdC, (72)	—	31
<i>agnL5</i>	Unknown	PtaG, A0A067XNI6 (43)	—	—	—
<i>agnL4</i>	Oxidoreductase	TpcG, Q4WQZ1 (59)	MdpK, (59)	GedF, (61)	32
<i>agnL3</i>	BVMO	CPUR_05427, M1WG92 (51)	MdpL, (42)	GedK, (43)	33
<i>agnL2</i>	Anthrone oxidase	GedH, P0DOB2 (46)	MdpH2, (43)	GedH, (46)	34
<i>agnL1</i>	Decarboxylase	TpcK, Q4WQY7 (72)	MdpH1, (59)	GedI, (67)	—
<i>agnPKS</i>	nr-PKS	MdpG, Q5BH30 (66)	MdpG, (66)	GedC, (65)	35
<i>agnR1</i>	Oxidoreductase	—	—	—	—



Our results are not consistent with either of these possibilities. Knock-out of *agnL4* results in accumulation of emodin **9**, but only traces of chrysophanol **6** (see ESI†). Müller has shown³⁸ that the initial “ketoreduction” step during the conversion of **9** to **6** *in vitro* requires prior chemical reduction of **9** to the hydroquinone (shown by NMR to exist in the tautomeric forms **21a** and **21b**) before MdpC (=AgnL6) mediated reduction to

Based on these observations we now propose a unifying pathway to monodictyphenone and the agnestins (Scheme 4) which is consistent with all available evidence. AgnPKS (=MdpG) assembles and cyclises an octaketide bound to the PKS as a thiolester. This is then hydrolysed by AgnL7 (=MdpF) to give atochryson carboxylic acid **30** as the first enzyme-free intermediate. AgnL1 then catalyses the concerted decarboxylation-elimination required to convert atochryson carboxylic acid **30** to emodin anthrone **31** which is then oxidized to emodin **9** by



AgnL2 (=MdpH1). Emodin **9** then undergoes reduction catalysed by AgnL4 (=MdpK) to give the dihydroquinone tautomer **21b**. This is the substrate for AgnL6 (=MdpC) reduction to **22**, followed by AgnL8 (=MdpB) dehydration and likely spontaneous autoxidation to chrysophanol **6**. Baeyer–Villiger oxidation by AgnL3 (=MdpL) gives monodictyphenone **7** along with some cephalone **F 10**. Formation of **7** presumably occurs *via* hydrolysis of the lactone **32** which has not been previously reported as a natural product. Close examination of the NMR spectra of isolated samples of monodictyxanthone **12** (see ESI†), show the presence of a co-eluting structural isomer. The chemical shifts are fully consistent with values calculated for **32** which we name monodictylactone.

Further conversion to agnestins **A 11** and **B 15** (Scheme 4), requires reduction to dihydro-monodictyphenone **17**, oxidation to agnestin **C 16ab** probably *via* **18**, and rearrangement to either agnestin **A 11** or agnestin **B 15** directly, although we have demonstrated that **11** and **15** also interconvert. Examination of the *agn* cluster, reveals *agnR1* as the only unassigned oxidoreductase encoding gene present which could be involved in this conversion. KO of *agnR1*, however, revealed it is not involved in the pathway (see ESI†), and thus genes involved in the proposed oxidation/reduction may be located elsewhere on the genome of *P. variotii*. Such split BGCs have been observed before in related systems, for example prenyl transferases involved during the biosynthesis of the shamixanthones in *A. nidulans* are not encoded within the *mdp* BGC itself.¹¹ The remaining metabolites **12**, **19** and **20** are probably formed by spontaneous decarboxylations, dehydrations and methanolysis reactions (Scheme 4).

Thus we have identified the protein responsible for the first essential reductive step in the aromatic deoxygenation of anthraquinones, *e.g.* emodin **9** to chrysophanol **6**. The chemical requirement for this had been elegantly demonstrated by Müller and co-workers using an initial chemical reduction step. We have now firmly established the genetic and biochemical basis for this important process.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank BBSRC (BB/J006289/1 and Bristol Centre for Synthetic Biology BB/L01386X/1) and Syngenta for funding. LCMS were provided by EPSRC (EP/F066104/1) and DFG (INST 187/621). 500 MHz NMR (EP/L011999/1) was provided by EPSRC. *P. variotii* was a gift from Syngenta and was sequenced at The Genome Analysis Centre (Norwich, UK) under contract to Syngenta through Genome Enterprises Ltd.

Notes and references

- 1 K.-S. Masters and S. Brase, *Chem. Rev.*, 2012, **112**, 3717–3776; T. Wezeman, S. Brase and K. S. Masters, *Nat. Prod. Rep.*, 2015, **32**, 6–28.

- 2 K. M. Henry and C. A. Townsend, *J. Am. Chem. Soc.*, 2005, **127**, 3724–3733.
- 3 A. J. Birch, J. Baldas, J. R. Hlubucek, T. J. Simpson and P. W. Westerman, *J. Chem. Soc., Perkin Trans. 1*, 1976, 898–904.
- 4 S. A. Ahmed, E. Bardshiri, C. R. McIntyre and T. J. Simpson, *Aust. J. Chem.*, 1992, **45**, 249–274.
- 5 B. V. Franck and F. H. Münster, *Fortschr. Chem. Org. Naturst.*, 1973, **30**, 151–206.
- 6 P. S. Steyn, *Tetrahedron*, 1970, **26**, 51–57.
- 7 A. Rhodes, G. A. Somerfield and M. P. McGonagle, *Biochem. J.*, 1963, **88**, 49–357.
- 8 I. Fujii, Y. Ebizuka and U. Sankawa, *J. Biochem.*, 1988, **103**, 878–883.
- 9 M. L. Nielsen, J. B. Nielsen, C. Rank, M. L. Klejnstrup, D. M. K. Holm, K. H. Broggard, B. J. Hansen, J. C. Frisvad, T. O. Larsen and U. H. Mortensen, *FEMS Microbiol. Lett.*, 2011, **321**, 157–166.
- 10 Y.-M. Chiang, E. Szewczyk, A. D. Davidson, R. Entwistle, N. P. Keller, C. C. C. Wang and B. R. Oakley, *Appl. Environ. Microbiol.*, 2011, **76**, 2067–2074.
- 11 J. F. Sanchez, R. Entwistle, J.-H. Hung, J. Yaegashi, S. Jain, Y.-M. Chiang, C. C. C. Wang and B. R. Oakley, *J. Am. Chem. Soc.*, 2011, **133**, 4010–4017.
- 12 M. T. Nielsen, J. B. Nielsen, D. C. Anyaogu, D. K. Holm, K. F. Nielsen, T. O. Larsen and U. H. Mortensen, *PLoS One*, 2013, **8**, e72871.
- 13 R. Schor and R. J. Cox, *Nat. Prod. Rep.*, 2018, **35**, 230–256.
- 14 S. Griffiths, C. Mesarich, B. Saccomanno, A. Vaisberg, P. Wit, R. Cox and J. Collemare, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 6851–6856.
- 15 K. Williams, A. J. Szwalbe, N. P. Mulholland, J. L. Vincent, A. M. Bailey, C. L. Willis, T. J. Simpson and R. J. Cox, *Angew. Chem., Int. Ed.*, 2016, **55**, 6784–6788.
- 16 A. J. Wozniak, K. Williams, D. E. O'Flynn, A. M. Bailey, N. P. Mulholland, J. L. Vincent, C. L. Willis, R. J. Cox and T. J. Simpson, *Chem. Commun.*, 2015, **15**, 17088–17091.
- 17 A. Krick, S. Kehraus, C. Gerhäuser, K. Klimo, M. Nier, A. Maier, H.-H. Fiebig, I. Atodiresei, G. Raabe, J. Fleischhauer and G. M. König, *J. Nat. Prod.*, 2007, **70**, 353–360.
- 18 T. Asai, S. Otsuki, H. Sakurai, K. Yamashita, T. Ozeki and Y. Oshima, *Org. Lett.*, 2013, **15**, 2058–2061.
- 19 Crystallographic coordinates for **11** and **15** have been deposited with the CCDC with accession numbers 1839028 and 1839029 respectively.
- 20 A. Krick, S. Kehraus, C. Gerhäuser, K. Klimo, M. Nier, A. Maier, H.-H. Fiebig, I. Atodiresei, G. Raabe, J. Fleischhauer and G. König, *J. Nat. Prod.*, 2007, **70**, 353–360.
- 21 M. W. Sumarah, E. Puniani, B. A. Blackwell and J. D. Miller, *J. Nat. Prod.*, 2008, **71**, 1393–1398.
- 22 J. S. E. Holker, E. O'Brien and T. J. Simpson, *J. Chem. Soc., Perkin Trans. 1*, 1983, 1365–1368; K. C. Nicolau and A. Li, *Angew. Chem., Int. Ed.*, 2008, **47**, 6579–6582.
- 23 R. J. Ogg, P. B. Kingsley and J. S. Taylor, *J. Magn. Reson., Ser. B*, 1994, **104**, 1–10.
- 24 S. Tarcz, X. Xie and S.-M. Li, *RSC Adv.*, 2014, **4**, 17986–17992.



- 25 C. Li, J. Zhang, C. Shao, W. Ding, Z. She and Y. Lin, *Chem. Nat. Compd.*, 2011, **47**, 382–384.
- 26 K. Throckmorton, P. Wiemann and N. Keller, *Toxins*, 2015, **7**, 3572–3607.
- 27 P. Reihl and J. Stolz, *J. Biol. Chem.*, 2005, **280**, 39809–39817.
- 28 C. P. Woloshuk, K. R. Foutz, J. F. Brewer, D. Bhatnagar, T. E. Cleveland and G. A. Payne, *Appl. Environ. Microbiol.*, 1994, **60**, 2408–2414.
- 29 Y. Kubo, Y. Takano, N. Endo, N. Yasuda, S. Tajima and I. Furusawa, *Appl. Environ. Microbiol.*, 1996, **62**, 4340–4344.
- 30 V. Tiranti, C. Viscomi, T. Hildebrandt, I. Di Meo, R. Mineri, C. Tiveron, M. D. Levitt, A. Prella, G. Fagioliari, M. Rimoldi and M. Zeviani, *Nat. Med.*, 2009, **15**, 200–205.
- 31 N. P. Keller, N. J. Kantz and T. H. Adams, *Appl. Environ. Microbiol.*, 1994, **60**, 1444–1450.
- 32 J. W. Cary, K. C. Ehrlich, J. M. Bland and B. G. Montalbano, *Appl. Environ. Microbiol.*, 2006, **72**, 1096–1101.
- 33 K. C. Ehrlich, B. Montalbano, S. M. Boue and D. Bhatnagar, *Appl. Environ. Microbiol.*, 2005, **71**, 8963–8965.
- 34 K. C. Ehrlich, P. Li, L. Scharfenstein and P.-K. Chang, *Appl. Environ. Microbiol.*, 2010, **76**, 3374–3377.
- 35 A. Watanabe, I. Fujii, U. Sankawa, M. E. Mayorga, W. E. Timberlake and Y. Ebizuka, *Tetrahedron Lett.*, 1999, **40**, 91–94.
- 36 M. L. Nielsen, L. Albertsen, G. Lettier, J. B. Nielsen and U. H. Mortensen, *Fungal Genet. Biol.*, 2006, **43**, 54–64.
- 37 T. J. Simpson, *ChemBioChem*, 2012, **13**, 1680–1688.
- 38 M. A. Schätzle, S. M. Husain, S. Ferlino and M. Müller, *J. Am. Chem. Soc.*, 2012, **134**, 14742.
- 39 D. Conradt, M. A. Schätzle, J. Haas, C. A. Townsend and M. Müller, *J. Am. Chem. Soc.*, 2015, **137**, 10867–10869.

