Journal Name

ARTICLE

Genetic and Chemical Characterisation of the Cornexistin Pathway Provides Further Insight into Maleidride Biosynthesis

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

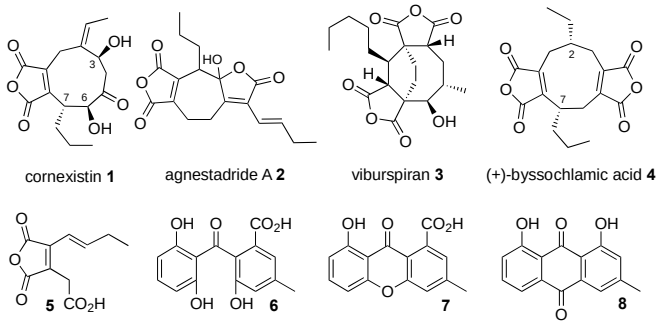
www.rsc.org/

Katherine Williams,[a] Agnieszka. J. Szwalbe,[b] Claire Dickson,[b] Tim R. Desson,[c] Nicholas. P. Mulholland,[c] Jason. L. Vincent,[c] John. M. Clough,[c] Andrew. M. Bailey,[d] Craig. P. Butts,[b] Christine. L. Willis,[b] Thomas. J. Simpson[b] and Russell. J. Cox[a,b*]

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.[1] 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.[2] Cornexistin **1** is a maleidride: these are 7-, 8- and 9-membered carbocyclic rings fused to one or (usually) two maleic anhydrides.[3] Related compounds include agnestadride A **2** (heptadride),[3] viburspiran **3**[4] (octadride) and byssochlamic acid **4** (nonadride).[3,5] Maleic anhydrides have been associated with a wide range of useful biological properties.[6] We recently elucidated the complete biosynthetic pathway of byssochlamic acid **4**.[5] The first step of the pathway is catalysed by a fungal highly reducing polyketide synthase (hrPKS),[7] 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[8] and Tang and coworkers have shown that the biosynthesis of rubratoxins follows the same pathway.[9] 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.[10] 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[11] was compared to the previously identified byssochlamic acid BGC,[5] 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).[5] 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

[a] Institute for Organic Chemistry, and BMWZ, Leibniz University of Hannover, Schneiderberg 1B, 30167, Germany
[b] School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK
[c] Syngenta, Jealott's Hill, Bracknell, Berkshire, RG42 6EY, UK
[d] School of Biological Sciences, Bristol Life Sciences Building, University of Bristol, 24 Tyndall Ave, Bristol BS8 1TH, UK

Electronic Supplementary Information (ESI) available: including all experimental and characterisation details. See DOI: 10.1039/x0xx00000x

putative oxidase encoding genes and genes of unknown function in the putative cornexistin BGC.

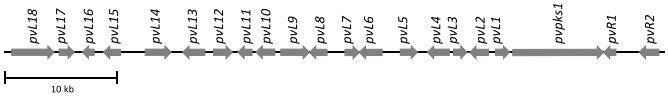


Figure 1. Cornexistin biosynthetic gene cluster.

Gene	Putative function	Gene	Putative function
pv-pks1	Polyketide synthase (PKS)	***pvL11***	Dienelactone hydrolase
pvL1	Hydrolase	***pvL12***	MFS transporter
pvL2	2-methylcitrate dehydratase (2MCD)	***pvL13***	Cytochrome P450
pvL3	Ketosteroid isomerase-like (KI-like)	***pvL14***	Transketolase
pvL4	MFS transporter	***pvL15***	Aldo/keto reductase
pvL5	Non-heme iron dioxygenase	***pvL16***	Hypothetical protein
pvL6	Citrate synthase-like (CS-like)	***pvL17***	Fumarylacetoacetate hydrolase
pvL7	Gluconolactonase	***pvL18***	NRPS-like enzyme
pvL8	C6 transcription factor	***pvR1***	Phosphatidylethanolamine-binding protein-like (PEBP-like)
pvL9	Serine carboxypeptidase	***pvR2***	Phosphotransferase
pvL10	MFS transporter		

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 **6**, monodictyxanthone **7**, chrysophanol **8**, and three unidentified congeners **A**, **B** and **C** (Figure 2A).[12] 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**.

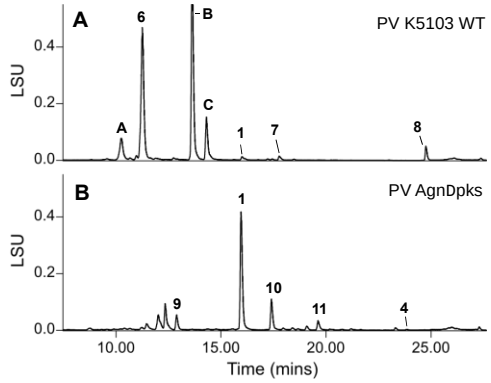


Figure 2. ELSD chromatograms showing the comparison of crude extracts from 11 day old cultures of *P. variotii*: **A**, WT strain; **B**, AgnΔpks strain. See ESI for detailed detection of **4** using HR-LCMS.

The xanthone nrPKS was identified by comparison with other known xanthone pks genes (such as *mdpG* from *Aspergillus nidulans,*[13] and *claG* from *Cladosporium fulvum*)[14] and then disrupted using a geneticin resistance cassette to produce strain PVAgnΔpks. 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 3*S*,6*S*,7*R* for the first time (Figure 3).

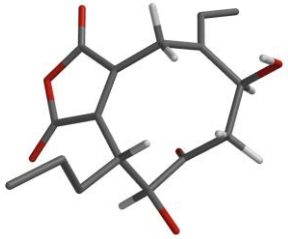
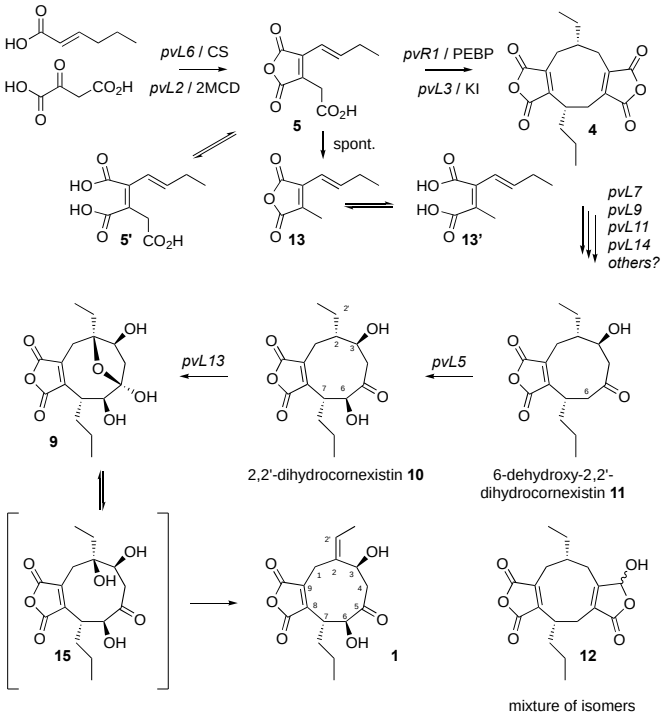


Figure 3. X-ray structure and absolute configuration of cornexistin **1**.

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.[15] Byssochlamic acid **4** itself was also detected by HR-LCMS analysis (see ESI).



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 2*S* epimer.

The *pvpks1* 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).

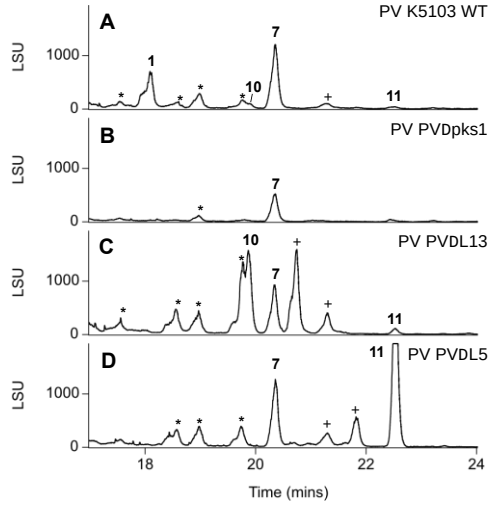


Figure 4. ELSD comparison of extracts from WT *P. variotii* with extracts from the putative late gene disruptions and *P. variotii* PVΔpks1. *, xanthone related; +, unrelated.

To investigate if byssochlamic acid **4** is an intermediate on the cornexistin pathway we fed the PVΔpks1 strain of *P. variotii* with byssochlamic acid (+)-**4** purified from WT *B. fulva*. After 10 days the *P. variotii* mycelia and media were extracted with acidified EtOAc and the concentrated organic extract examined by LCMS. No cornexistin could be detected, but a mixture of the previously reported[3] C-10 epimers of dihydrobyssochlamic acid **12** was isolated (see ESI for details).

The AgnΔpks strain was used for all further gene disruptions to investigate the pathway, as the increased nonadride titre made the analysis more sensitive. Disruption of either the hydrolase (*pvL1*) or the citrate synthase (*pvL6*), to produce strains AgnΔpks-PVΔL1 and AgnΔpks-PVL6 respectively, led to loss of production of cornexistin **1** (Figures 5BC) but no accumulation of any potential intermediates.

In contrast, disruption of the PEBP gene (*pvR1*) to produce strains AgnΔpks-PVΔR1 led to a threefold reduction in titre of cornexistin **1** and the production of three previously reported[3,5] metabolites (Figure 5D). Comparison of chromatographic and spectroscopic propertied confirmed that these were the monomer **5**, its hydrolysis product **5'** and **13** / **13'** which arise by spontaneous decarboxylation of **5**.[3]

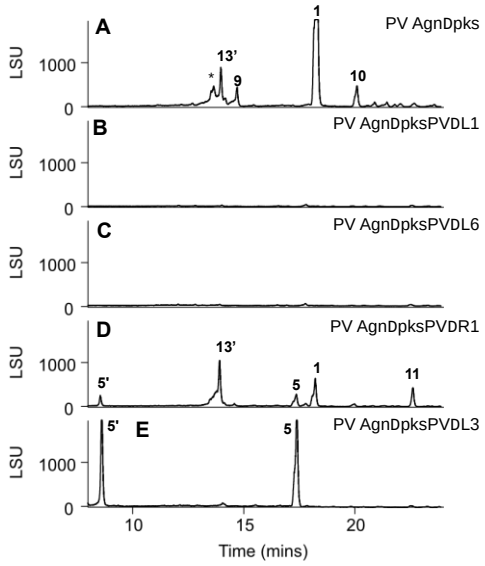


Figure 5. ELSD comparison of *P. variotii* AgnΔpks with extracts from the putative early gene disruptions. *Unrelated compounds. Note the ratios of **5**, **5'** and **13'** depend on fermentation and extraction conditions. See references 3 and 5 for more details.

Disruption of the KI-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 KI genes in the byssochlamic acid BGC.[5] Knockout of *pvL11*, encoding a putative dienelactone 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 KI-like proteins; and a pair of PEBP-like proteins. In previous work[5] we showed that both PEBP and both KI encoding genes must be expressed for high production of **4**. In the case of cornexistin **1**, however, only one copy each of the KI and PEBP genes is present and required for biosynthesis. The KI-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]

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 *P. variotii* (see ESI), but it could not be isolated and further characterised. *In lieu* of this we fed **4** isolated from WT *B. fulva* to the maleidride PKS KO strain of *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 *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 (TropI) is implicated in maleic anhydride hydrolysis during stipitatonic acid biosynthesis in *Talaromyces stipitatu*s and its KO shows a similar reduction in titre.[18] Proteins encoded by *pvL7*, *pvL9* and *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 *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 *pvL13* and *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 (*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 *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 *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 removal of 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/F066104/1) 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. *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

1. E. C. Oerke, *J. Agric. Sci.*, 2006, **144**, 31-43.
2. M. Nakajima, K. Itoi, Y. Takamatsu, S. Sato, Y. Furukawa, K. Furuya, T. Honma, J. Kadotani, M. Kozasa and T. Haneishi, *J. Antibiot.*, 1991, **44**, 1065–1072; S. C. Fields, L. Mireles-Lo and B. C. Gerwick, *J. Nat. Prod.*, 1996, **59**, 698-700.
3. A. J. Szwalbe, 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, **51**, 17088-17091.
4. M. Saleem, H. Hussain, I. Ahmed, S. Draeger, B. Schulz, K. Meier, M. Steinert, G. Pescitelli, T. Kurtán, U. Flörke and K. Krohn, *Eur. J. Org. Chem.*, 2010, **2011**, 808–812.
5. 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.
6. X. Chen, Y. Zheng and Y. Shen, *Chem. Rev.*, 2007, **107**, 1777–1830.
7. R. J. Cox, *Org. Biomol. Chem.*, 2007, **5**, 2010–2026; Y.-H. Chooi and Y. Tang, *J. Org. Chem.*, 2012, **77**, 9933–9953.
8. R. Fujii, Y. Matsu, A. Minami, S. Nagamine, I. Takeuchi, K. Gomi and H. Oikawa, *Org. Lett.*, 2015, **17**, 5658–5661.
9. J. Bai, D. Yan, T. Zhang, Y. Guo, Y. Liu, Y. Zou, M. Tang, B. Liu, Q. Wu, S. Yu, Y. Tang and Y. Hu, *Angew. Chem.*, 2017, **129**, 4860–4864.
10. J. S. Clark, J. M. Northall, F. Marlin, B. Nay, C. Wilson, A. J. Blake and M. J. Waring, *Org. Biomol. Chem.*, 2008, **6**, 4012–4025; J. S. Clark, F. Marlin, B. Nay and C. Wilson, *Org. Lett.*, 2003, **5**, 89–92.
11. Full cornexistin BGC uploaded to genbank. Accession: MF197864
12. Y. M. Chiang, E. Szewczyk, A. D. Davidson, R. Entwistle, N. P. Keller, C. C. C. Wang and B. R. Oakley, *Appl. Environ. Microbiol.*, 2010, **76**, 2067-2074; A. Krick, S. Kehraus, C. Gerhaeuser, K. Klimo, M. Nieger, A. Maier, H.-H. Fiebig, I. Atodiresei, G. Raabe, J. Fleischhauer and G. M. Koenig, *J. Nat. Prod.*, 2007, **70**, 353-360.
13. 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.
14. S. Griffiths, C. H. Mesarich, B. Saccomanno, A. Vaisberg, P. J. G. M. de Wit, R. J. Cox and J. Collemare, *Proc. Nat. Acad. Sci. USA*, 2016, **113**, 6851-6856.
15. S. G. Smith and J. M. Goodman, *J. Am. Chem. Soc.,* 2010, **132**, 12946–12959.
16. M. L. Nielsen, L. Albertsen, G. Lettier, J. B. Nielsen, U. H. Mortensen, *Fung. Genet. Biol.,* 2006, **43**, 54-64.
17 B. Bonsch, V. Belt, C. Bartel, N. Duensing, M. Koziol, C. M. Lazarus, A. M. Bailey, T. J. Simpson and R. J. Cox, *Chem. Commun.*, 2016, **52**, 6777–6780.
18. A. al Fahad, A. Abood, T. J. Simpson and R. J. Cox, *Angew. Chem. Int. ed.*, 2014, **53**, 7519–7523.
19. J. Barriuso, D.T. Nguyen, J. W.-H Li, J. N. Roberts, G. MacNevin, J. L. Chaytor, S. L. Marcus, J.C. Vederas and D-K. Ro, *J. Am. Chem. Soc.*, 2011, 133, 8078-8081
20. Y. Matsuda, T. Awakawa, T. Wakimoto and I. Abe, *J. Am. Chem. Soc.*, 2013, **135**, 10962-10965.

