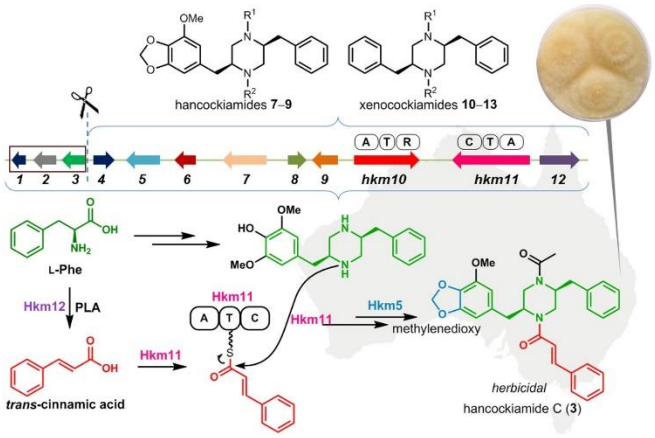


Hancockiamides: Phenylpropanoid Piperazines from *Aspergillus hancockii* are Biosynthesised by a Versatile Dual Single-module NRPS Pathway

Hang Li,[‡a] Alastair E. Lacey,[‡b] Si Shu,[a] John A. Kalaitzis,[c] Daniel Vuong,[b] Andrew Crombie,[b] Jinyu Hu,[a] Cameron L. M. Gilchrist,[a] Ernest Lacey,[bc] Andrew M. Piggott,*[c] and Yit-Heng Chooi*[a]

The hancockiamides are an unusual new family of *N*-cinnamoylated piperazines from the Australian soil fungus *Aspergillus hancockii*. Genomic analyses of *A. hancockii* identified a biosynthetic gene cluster (*hkm*) of 12 genes, including two single-module nonribosomal peptide synthetase (NRPS) genes. Heterologous expression of the *hkm* cluster in *A. nidulans* confirmed its role in the biosynthesis of the hancockiamides. We further demonstrated that a novel cytochrome P450, Hkm5, catalyses the methylenedioxy bridge formation, and that the PAL gene *hkm12* is dispensable, but contributes to increased production of the cinnamoylated hancockiamides. *In vitro* enzymatic assays and substrate feeding studies demonstrated that NRPS Hkm11 activates and transfers *trans*-cinnamate to the piperazine scaffold and has flexibility to accept bioisosteric thienyl and furyl analogues. This is the first reported cinnamate-activating fungal NRPS. Expression of a truncated cluster lacking the acetyltransferase gene led to seven additional congeners, including an unexpected family of 2,5-dibenzylpiperazines. These pleiotropic effects highlight the plasticity of the pathway and the power of this approach for accessing novel natural product scaffolds.

Introduction

Piperazine is well-known to be a privileged scaffold in drug discovery and development.[1, 2] Piperazine derivatives can be found in many pharmaceuticals, including the antihistamine cyclizine, the atypical antipsychotic aripiprazole, the cardiotonic vesnarinone and the HIV-1 protease inhibitor indinavir.[3] Piperazine alkaloids are also found in a range of plants and microorganisms, including fungi.[4] In our ongoing chemotaxonomic exploration of Australian Aspergilli,[5-10] we previously described a new species, *Aspergillus hancockii*, which was found to produce a diverse suite of over 60 secondary metabolites.[11] Among these metabolites, a group of novel piperazine alkaloids was discovered, which we named the hancockiamides.

Herein, we describe the detailed characterisation of the chemical structures, bioactivities and biosynthesis of hancockiamides A–F (**1**–**6**). Significantly, five of the six hancockiamides were found to contain *N*-cinnamoyl substitution. Cinnamic acid is commonly incorporated in plant natural products, such as terpenoids (denhaminols),[12] glucosides (polyanthosides),[13] chalcones (mallopenin)[14] and lignans (longipedlignans),[15] but is rare among fungal secondary metabolites. While cinnamate-derived benzoic acid has been shown to act as starter units for fungal polyketides, such as squalestatins[16], zaragozic acids[17] and strobirulins[18], the hancockiamides represent the first example where the phenylpropanoid pathway has shown to be recruited into the fungal NRPS pathway that synthesises the "privileged" piperazine scaffolds. Furthermore, the 4-hydroxy-3,5-dimethoxy and methylenedioxy benzylic substitutions on several of the hancockiamides are reminiscent of plant phenylpropanoids, such as syringaresinol[19] and etoposide[20]. The novel structural features of hancockiamides prompted us to investigate their biosynthesis by heterologous expression of the whole and truncated biosynthetic gene clusters in *A. nidulans*, and by substrate feeding and *in vitro* assay of the NRPS expressed in yeast.

Results and Discussion

Discovery and structure elucidation of the hancockiamides

Hancockiamides A–F (**1**–**6**) were extracted with acetone from 14-day stationary cultures of *A. hancockii* MST FP2241 on rice and were purified by silica gel chromatography and reversed-phase preparative HPLC (Supporting Information). HRESI(+)MS analysis of hancockiamide A (**1**) (*m/z* 495.2252, $[M + Na]^+$),

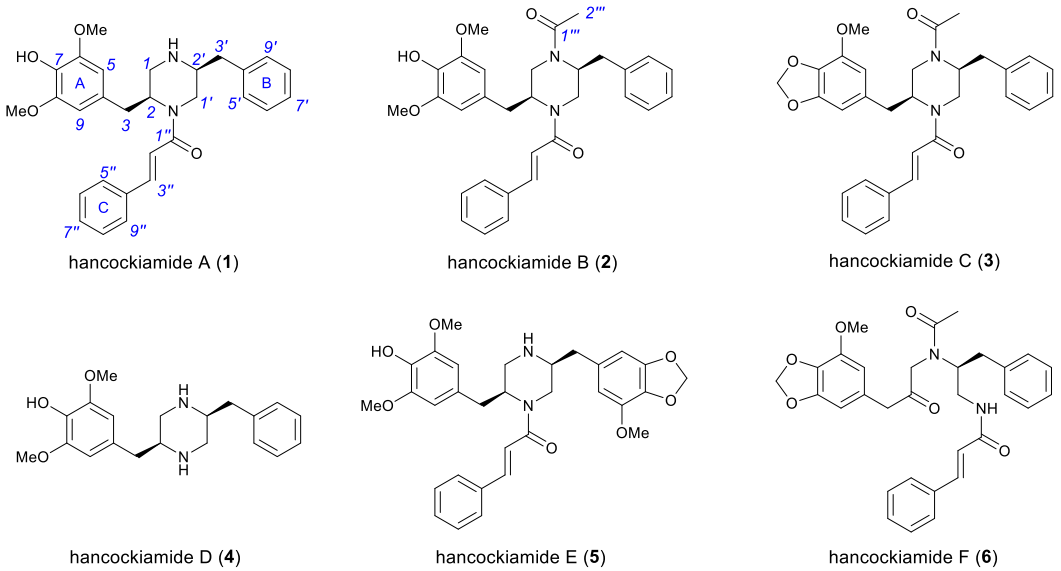


Fig. 1 Structures of the hancockiamides identified from *Aspergillus hancockii*.

indicated a molecular formula $C_{29}H_{32}N_2O_4$. Examination of the ^{1}H (Fig. S13) and ^{13}C (Fig. S14†) NMR spectra of **1** in DMSO-d_6 revealed two closely paired sets of signals in a 78:22 ratio. The ROESY spectrum (Fig. S18†) confirmed these two sets of signals were in chemical exchange, suggesting the presence of two slowly interconverting conformers. The NMR data for the major conformer consisted of resonances attributable to one symmetrically 1,3,4,5-tetrasubstituted aromatic ring (Ring A), two monosubstituted aromatic rings (Rings B and C) (Fig. 1), one isolated *trans*-alkene and one amide/ester carbonyl group (Table S7†). Two equivalent methoxy groups (6/8-OMe) and one hydroxy group (7-OH) were positioned on Ring A based on HMBC correlations (Fig. S16†). The COSY spectrum of **1** (Fig. S17†) revealed two additional spin systems – H_2-1, H-2, H_2-3 and H_2-1′, H-2′, H_2-3′ – with HMBC correlations from H_2-3 to C-5/9 and from H_2-3′ to C-5′/9′ positioning these two methylene groups at the benzylic positions of Rings A and B, respectively. Diagnostic HMBC correlations from H_2-1 to H-2′ and from H_2-1′ to H-2, along with deshielding of C-1, C-2, C-1′ and C-2′, suggested the two spin systems were linked by a central piperazine ring. Additional HMBC correlations from H-2, H_2-1′, H-2″ and H-3″ to C-1″ and from H-3″ to C-5″/9″ suggested the presence of a cinnamoyl group on 1′-N, while the final exchangeable proton (1-NH) suggested the second piperazine nitrogen was underivatised. The observed peak doubling in the NMR spectra of **1** was thus attributed to slowly interconverting amide conformers. Detailed analysis of the remaining 2D NMR correlations (Fig. 2) confirmed the structure of **1** as shown in Fig. 1. The 2,2′-*cis* relative configuration of the piperazine ring was established based on the presence of a 10.4 Hz *trans*-diaxial coupling between H-1′b and H-2 and a 2.3 Hz axial-equatorial coupling between H-1′a and H-2. We have tentatively assigned a 2*S*,2′*S* absolute configuration for **1** based on our proposed biosynthesis incorporating two molecules of L-phenylalanine (L-Phe) (Fig. 2).

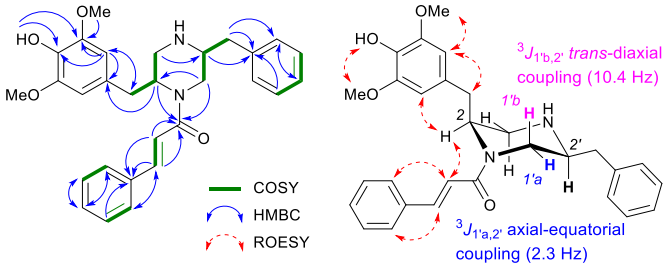


Fig. 2 Key 2D NMR correlations for hancockiamide A (**1**)

Five additional hancockiamide analogues were also isolated from the crude extract of *A. hancockii*, as shown in Fig. 1. Hancockiamide B (**2**) was identified as the N^1-acetyl analogue of **1**, while hancockiamide C (**3**) contains both an N^1-acetyl group and a methylenedioxy bridge between 7-OH and 8-OMe. Hancockiamide D (**4**) was found to be the $N^{1'}$-descinnamoyl analogue of **1**, while hancockiamide E (**5**) is similar to **1**, but with Ring B having the same methylenedioxy/methoxy substitution pattern as found on Ring A in **3**. Finally, hancockiamide F (**6**) was identified as a ring-opened analogue of **3**. Full details of the structure elucidation and characterisation of **2**–**6** are provided in the ESI†.

Identification of a two single-module NRPS gene cluster for hancockiamide biosynthesis

The unusual structures of **1**–**6** prompted us to investigate the molecular basis of their biosynthesis. Previous studies have demonstrated that fungal piperazine-type dipeptides are biosynthesised by a monomodular NRPS coupled with an NmrA-like reductase such as LnaA/LnaB[21] and HqlA/HqlB,[22] which are involved in the biosynthesis of an L-tyrosine-derived

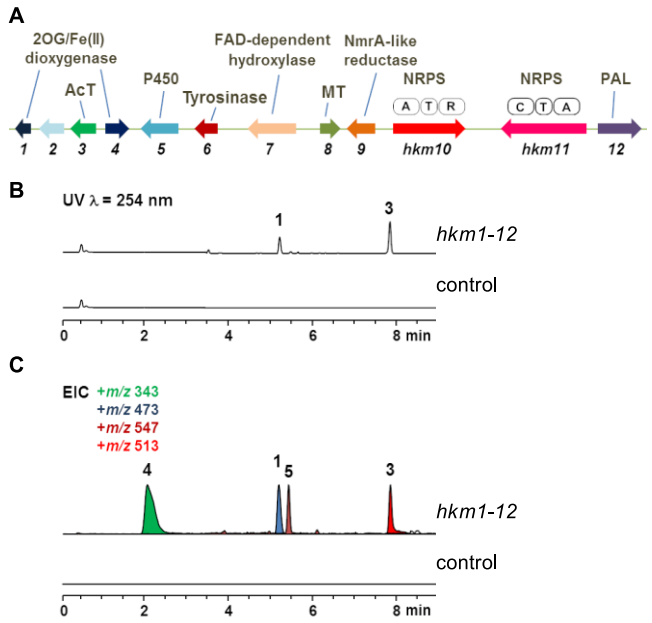


Fig. 3 Verification of the hkm gene cluster responsible for production of the hancockiamides. (A) The identified *hkm* gene cluster. AcT, acetyltransferase; MT, methyltransferase; PAL, phenylalanine ammonia lyase; A, adenylation; T, thiolation; C, condensation, R, reduction. (B) HPLC traces showing the production of **1** and **3** from *A. nidulans* expressing the entire *hkm* gene cluster. (C) Extracted ion chromatograms (EICs) showing the production of **1** and **3**–**5** from *A. nidulans* expressing the entire *hkm* gene cluster.

piperazine and herquline A, respectively. Thus, we hypothesised that the hancockiamides are also synthesised in a similar manner by a monomodular NRPS and an NmrA-like reductase that reduce the dihydropyrazine intermediate to pyrazine. Interestingly, **1**–**3** and **5**–**6** contain an unusual *N-trans*-cinnamoyl moiety, which we proposed to be derived from phenylalanine, and thus requires a phenylalanine ammonia lyase (PAL) to supply the cinnamate.[16-18, 23] Using these criteria and the genome mining strategy described previously,[24] we then mined the genome sequence of *A. hancockii* for a potential gene cluster for these novel piperazines using LnaA as a query sequence in BLASTP search. This resulted in the identification of the NRPS Hkm10, which shares the highest protein identity (56%) with LnaA. Furthermore, adjacent genes to *hkm10*, both up- and downstream, encode the expected NmrA-like reductase (Hkm9), a PAL (Hkm12), an acyltransferase (Hkm3), a tyrosinase (Hkm6), a methyltransferase (Hkm8), another monomodular NRPS (Hkm11) and several other tailoring enzymes (Fig. 3A and Table S2†), which match the proposed biosynthetic enzymes required for hancockiamide production. Thus, we proposed the *hkm* gene cluster as the candidate gene cluster responsible for biosynthesis of the hancockiamides.

To verify the involvement of the *hkm* gene cluster in hancockiamides biosynthesis, the entire 32.5 kb cluster was cloned as three fragments into the tripartite yeast–fungal artificial chromosome (YFAC) vectors with different auxotrophy markers for expression in a the *A. nidulans* LO7890[25] host (see Supporting Information). Given the gene cluster is not silent in *A. hancockii*, we hypothesised that promoter exchange is not required. Indeed, *A. nidulans* harbouring the entire *hkm* gene cluster accumulated hancockiamides A (**1**), C (**3**), D (**4**) and E (**5**) (Fig.3B and 3C), thus confirming the role of the *hkm* cluster in the biosynthesis of hancockiamides. Interestingly, the open-ring **6** isolated as a trace metabolite in the native producer *A. hancockii*[11] was not detected from the *A. nidulans* strain.

Truncation of *hkm* gene cluster boundaries provided insights into enzymes involved in the N-substitutions of hancockiamides

After verifying the identity of the *hkm* gene cluster, we aimed to characterise the biosynthetic steps unique to the hancockiamide pathway. Firstly, we attempted to characterise the enzymes involved in the *N*-substitutions, in particular the unusual *N-trans*-cinnamoylation. We hypothesised that the A-T-C monomodular NRPS encoded by *hkm11* could be involved in activating and transferring *trans*-cinnamic acid supplied by the PAL encoded by *hkm12* to the piperazine ring. Alternatively, an NCBI CD-Search[26] revealed that Hkm3 contains conserved domains corresponding to the pfam02458 transferase family and the PLN02663/PLN03157 hydroxycinnamoyl-transferases, and hence could be responsible for the *N-trans*-cinnamoylation. To test this hypothesis, we constructed *A. nidulans* strains expressing truncated versions of *hkm* cluster containing *hkm1–10*, *hkm1–11* and *hkm4–12* (see Supporting Information). Expression of the truncated gene cluster containing *hkm1–10* in *A. nidulans* resulted in the production of **4**, while *A. nidulans* expressing the *hkm1–11* genes clearly accumulated a small amount of **1** together with the major product **4** (Fig. 4A). These observations, together with the result from whole gene cluster (*hkm1–12*) expression (Fig. 3B), confirmed that the NRPS Hkm11 is required for the *N*-cinnamoylation of **4**, while PAL Hkm12 is dispensable but contributes to increase the production of cinnamoyl-containing **1**. The observation of a small amount of **1** in the absence *hkm12* could be due to the presence of endogenous PAL in *A. nidulans*, partially complementing the function of Hkm12. Indeed, a BLASTP search against the *A. nidulans* genome identified a putative PAL (AN6075.2) sharing 53% protein identity to Hkm12.

A. nidulans expressing the truncated cluster *hkm4–12* accumulated a new compound, hancockiamide G (**7**; Fig. 4B; *m/z* 471 $[M + H]^+$), with an almost identical UV spectrum to **3**. Large-scale cultivation of this strain afforded **7** in sufficient quantities for purification and structure elucidation. HRESI(+)MS of **7** (*m/z* 471.2273 $[M + H]^+$) indicated a molecular formula $C_{29}H_{30}N_2O_4$, corresponding to a decrease of C_2H_2O compared to **3**. Indeed, the NMR data for **7** (Table S15†) were very similar to those for **3**, with the main differences being the absence of signals for the N^1-acetyl group and the presence of an additional exchangeable proton. Detailed analysis of the remaining 2D NMR correlations (Fig. S64†) confirmed the structure of **7** as the N^1-desacetyl analogue of **1**, as shown in Fig. 4C. This suggests that Hkm3 is an *N*-acetyltransferase, which transfers an acetyl group to 1-N of piperazine. The production of **7** by *A. nidulans* expressing *hkm4-12* also

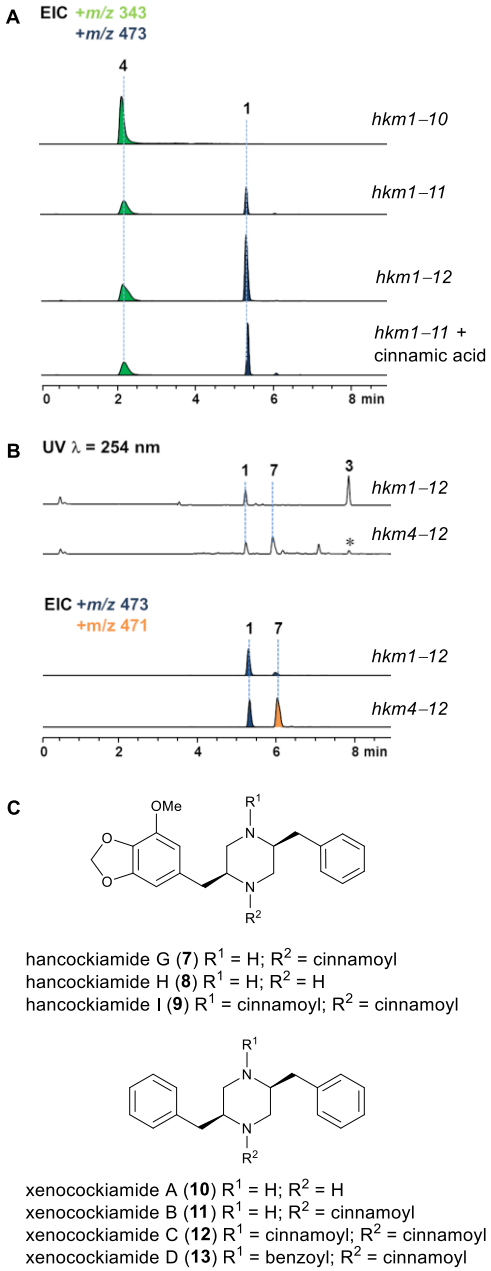


Fig. 4 Heterologous expression of the truncated *hkm* cluster in *A. nidulans*. (A) Extracted ion chromatograms (EICs) showing the production of **1** and **4** from *A. nidulans* transformants expressing truncated *hkm* gene clusters. (B) LCMS traces showing the production of **7** from *A. nidulans* expressing *hkm4–12*. Asterisk indicates unrelated peak. (C) Structures of hancockiamides G–I (**7**–**9**) and xenocockiamides A–D (**10**–**13**).

suggests that *hkm1* and *hkm2*, encoding putative phytanoyl-CoA dioxygenase and cellobiose dehydrogenase-like cytochrome protein, respectively (Table S2†), are not essential for hancockiamide biosynthesis. We also noted that *hkm1* contains a nonsense mutation at the N-terminus of the translated protein sequence, which may render the gene non-functional (Fig. S1†). Indeed, *A. nidulans* expressing *hkm3–12* is capable producing **1** and **3** in similar yield to *hkm1–12* (Fig. S2†), confirming that *hkm1* and *hkm2* are not required for hancockiamide biosynthesis.

When we set up the large-scale culture of *A. nidulans* expressing *hkm4–12* for the purification of **7**, we isolated two additional hancockiamides H–I (**8**–**9**), and a second parallel family of related metabolites, xenocockiamides A–D (**10**–**13**) (Fig. S3-4†). HRESI(+)MS and detailed analysis of the 2D NMR correlations (Fig. S70†) confirmed the structure of **8** to be the 7,8-methylenedioxy analogue of **4** (Fig. 4C). HRESI(+)MS of **9** (*m/z* 601.2695 $[M + H]^+$) suggested an increase of C_9H_6O compared to **7**. The ^1H NMR signals for **9** (Fig. S73†) were broadened to such an extent that few individual peaks could be resolved. However, analysis of the HSQC spectrum (Fig. S75 †) revealed signals consistent with one methoxy group, one methylenedioxy group and two cinnamoyl groups. Therefore, **9** was tentatively assigned to be the N^1-cinnamoyl analogue of **7** (Fig. 4C).

Detailed analysis of the spectroscopic data for **10**–**12** (see ESI†) suggested these compounds were the non-oxygenated Ring A analogues of **4**/**8**, **1**/**7** and **9**, respectively (Fig. 4C). The structure of **10** was confirmed by comparison of spectroscopic data (Fig. S76-S79†) and HPLC retention time (Fig. S9†) with an authentic synthetic sample of 2,5-dibenzylpiperazine. Likewise, the structures of **11**–**12** were confirmed following mono- and di-derivatisation of synthetic **10** with cinnamoyl chloride and comparisons to the isolated compounds (Fig. S10-S11† and S90-S96†). HRESI(+)MS of **13** (*m/z* 501.2539 $[M + H]^+$) indicated a decrease of C_2H_4 compared to **12**. The ^1H NMR signals for **13** (Fig. S97 †) were also significantly broadened. Based on HRESI(+)MS and HSQC analysis (Fig. S97†), we hypothesised that **13** may be the corresponding N^1-benzoyl analogue, which was confirmed by comparing to synthetic **13** obtained by derivatisation of synthetic **11** with benzoyl chloride (see ESI†).

Simple truncation of the *hkm* cluster has surprisingly generated more molecular diversity than expected. The observation of the additional metabolites **9**–**13** in *A. nidulans* expressing *hkm4–12* is particularly interesting. While production of desacetyl analogue **7** was expected with the lack of *hkm3* acetyltranferase gene, and **8** can be explained as the methylenedioxy version of **4**, the observation of dicinnamoylated **9** was unexpected. This suggests that NRPS Hkm11 has the flexibility to tolerate the bulky **7** as a substrate and the absence of the acetyltransferase Hkm3 opens up the opportunity for Hkm11 to introduce a second *N*-cinnamoyl moiety. Even more curiously, the *A. nidulans hkm4–12* strain produced a parallel family of unsubstituted 2,5-dibenzylpiperazines, xenocockiamides **10**–**13**, which were not observed in either the *A. nidulans* strain expressing the full *hkm* cluster or in the native producer *A. hancockii*. This suggests that the acetylation of 1-N could be important in controlling the biosynthetic efficacy of the hancockiamide pathway. Alternatively, either removal of *hkm1–3* has an epistatic effect on other cluster genes or optimal activities of the 4-hydroxy-3,5-dimethoxy- and methylenedioxy-forming enzymes are somehow linked to the presence of the acetyltransferase Hkm3, perhaps via protein-protein interactions, as we observed in elsinochrome biosynthesis.[27] These induced pleiotropic effects highlight the strategies employed by Nature for diversifying and shaping metabolite

distributions arising from a single biosynthetic pathway and how heterologous partial gene cluster expression can be used not only for pathway elucidation, but also as a powerful tool for the rapid generation of new natural product analogues.

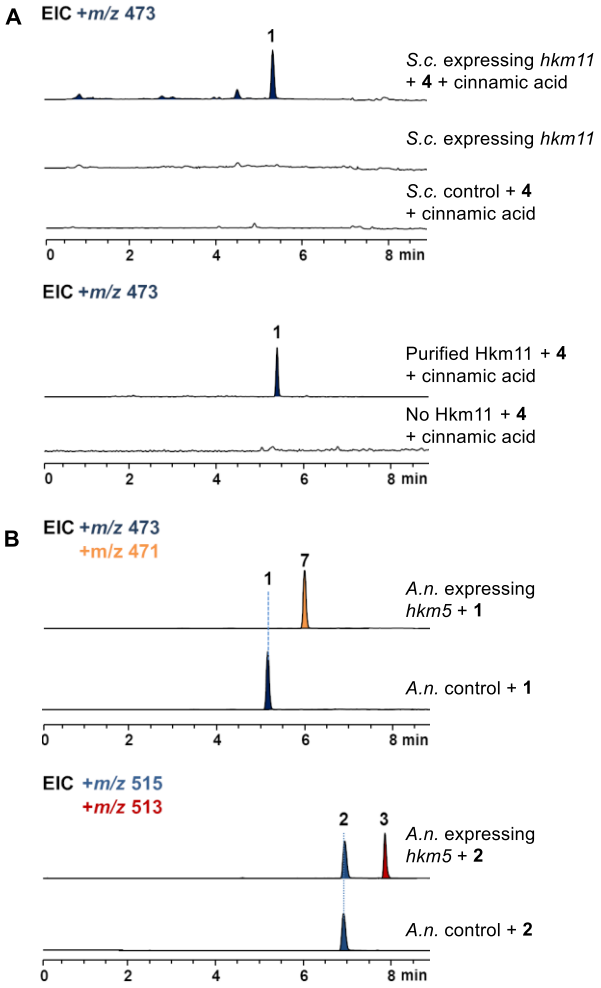


Fig. 5 *In vivo* and *in vitro* verification of the functions of Hkm11 and Hkm5. (A) LCMS analysis showing *in vivo* and *in vitro* verification of Hkm11 using *Saccharomyces cerevisiae* (*S.c.*) expressing *hkm11* and the purified Hkm11 (B) EICs showing *A. nidulans* (*A.n.*) expressing *hkm5* converted **1** and **2** into **7** and **3**, respectively.

Confirmation of monomodular NRPS Hkm11 as the *N-trans*-cinnamoyl transferase and Hkm5 as the methylenedioxy-forming cytochrome P450

To further corroborate the function of NRPS Hkm11, we expressed the intron-free *hkm11* under the *ADH2* promoter in *Saccharomyces cerevisiae* BJ5464-NpgA[28]. Upon feeding **4** and *trans*-cinnamic acid to this strain, the production of **1** was clearly observed (Fig. 5A), which suggested that Hkm11 was successfully expressed in the yeast strain and can transfer cinnamic acid onto **4** to form **1**. We then purified the recombinant C-terminal His-tagged Hkm11 from *S. cerevisiae* (Fig. S5†) and assayed its activity *in vitro*. In the presence of Hkm11, **4** and *trans*-cinnamic acid were converted into **1** (Fig. 5A), thus confirming that Hkm11 activates and transfers cinnamic acid onto **4** to form **1**. Thus, Hkm11 represents the first example of a fungal NRPS that can activate and transfer a *trans*-cinnamic acid moiety.

Next, we attempted to determine the enzyme that catalyses the methylenedioxy bridge formation in **3**. The 1,3-benzodioxole moiety is commonly found in plant lignans[29] and alkaloids[30]. Previous studies have demonstrated that the formation of the methylenedioxy bridge in plant natural products is catalysed by cytochrome P450 enzymes.[31, 32] Given that the *hkm* cluster also encodes a cytochrome P450 (Hkm5), we then proposed that the methylenedioxy bridge in hancockiamides could be formed by Hkm5. To interrogate the function of Hkm5, we heterologously expressed *hkm5* under the alcohol-inducible promoter *PalcA* in *A. nidulans*. Compared to the control, when *A. nidulans* expressing *hkm5* was fed **1**, the production of **7** was clearly observed (Fig. 5B). We further showed that *A. nidulans* expressing *hkm5* can partially convert **2** and **4** into their corresponding methylenedioxy-containing products **3** and **8**, respectively. (Fig. 5B and Fig. S6†). This clearly demonstrated that the formation methylenedioxy bridge was catalysed by Hkm5. These results also suggest that, *in vivo*, the biosynthesis of **3** likely proceeds via a branched pathway.

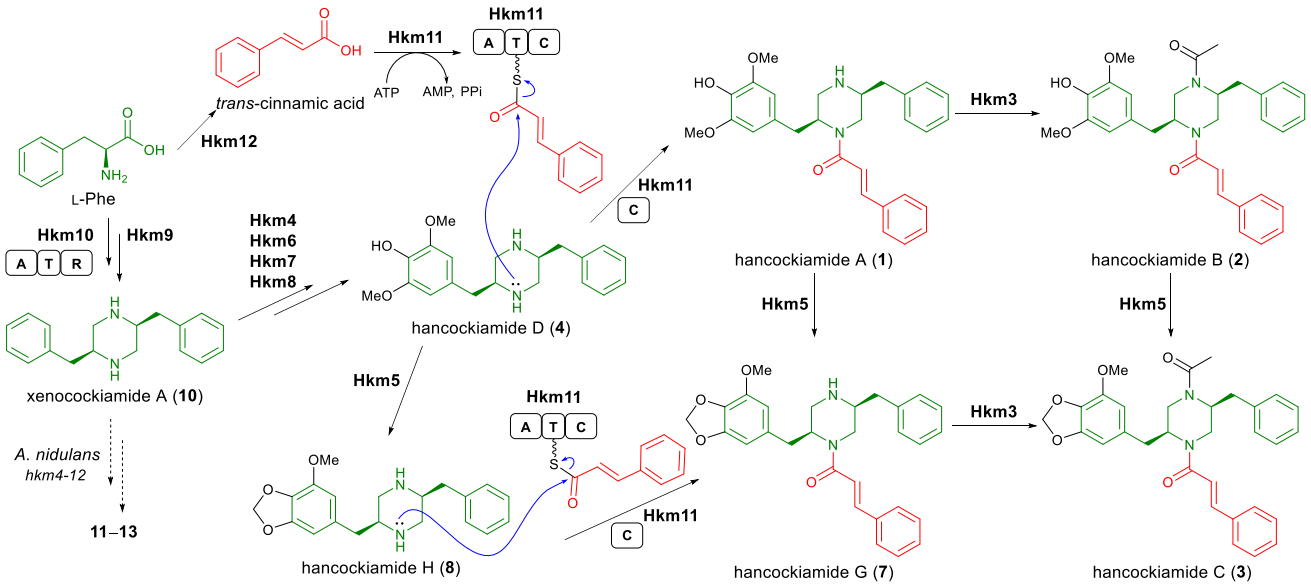
Based on the above results, we then proposed a biosynthetic pathway to the hancockiamides. Previous studies have shown that fungal monomodular NRPSs catalyse piperazine formation by first activating and reducing two molecules of L-tyrosine (L-Tyr) to the amino aldehydes and then forming the coupled 2,5-dihydropyrazine, which can be reduced to the piperazine by an NmrA-like reductase.[21, 22] Accordingly, in the *hkm* cluster, the NRPS Hkm10 and the NmrA-like reductase Hkm9 are proposed to convert two molecules of L-Phe to the corresponding piperazine **10** (Scheme 1). Notably, the discovery of **10**–**13** from *A. nidulans* expressing *hkm4*–*12* further supported the hypothesis that the hancockiamides are biosynthesised from two molecules of L-Phe.

Proceeding from **10** to **4** should involve a series of hydroxylations and *O*-methylations. Scanning the remaining biosynthetic genes in *hkm* gene cluster, we proposed that a tyrosinase encoded by *hkm6* should catalyse an aromatic hydroxylation, then a 2-oxoglutarate (2OG)-dependent Fe(II) dioxygenase encoded by *hkm4* and an FAD-dependent phenol hydroxylase encoded by *hkm7* could catalyse consecutive hydroxylations to install two more hydroxy groups, and a methyltransferase encoded by *hkm8* could catalyse two methylations using two molecules of *S*-adenosyl-L-methionine (SAM). Subsequently, **4** can be enzymatically converted into **1**–**3** by Hkm3, Hkm5, Hkm11 and Hkm12 (Scheme 1).

Bioactivity of hancockiamides and xenocockiamide

Equipped with a range of hancockiamide analogues, we tested **1**–**13** for *in vitro* antibacterial, antifungal, antiparasitic, cytotoxic and herbicidal activities (Table S22†). Notably, **1** and **4** displayed potent cytotoxic activity against murine myeloma NS-1 cells (MIC at 1.6 and 3.1 μg/mL) but were inactive against neonatal foreskin fibroblasts, suggesting a potential

Scheme 1 Proposed biosynthesis of hancockiamides and xenocockiamides.



antitumour application. More interestingly, compound **3**, the likely end metabolite of *hkm* pathway, showed potent *Arabidopsis thaliana* seed anti-germination activity (MIC at 6.3 μg/mL), but was inactive against the monocot *Eragrostis tef* seed, suggesting that **3** could be a herbicidal lead targeting monocots. We propose that the herbicidal activity of **3** could be due to its phenylpropanoid-like structural features, which may act on the plant lignan pathways, and hence warrants further investigations.

Production of novel hancockiamides by exploiting the substrate promiscuity of Hkm11

Motivated by the bioactivities of the hancockiamides and to take advantage of the biosynthesis, we then attempted to test the substrate flexibility of Hkm11 and to explore the possibilities for accessing more unusual hancockiamides by adding unnatural cinnamic acid congeners. We used *S. cerevisiae* expressing *hkm11* to perform precursor feeding assays via feeding **4** and *trans*-cinnamic acid congener. As a result, we have tested 14 cinnamic acid congeners (Table S6†) and found that Hkm11 is capable of activating and transferring (*E*)-3-(2-thienyl)acrylic acid (Fig. 6A), (*E*)-3-(2-furyl)acrylic acid (Fig. 6B) and (*E*)-3-(3-thienyl)acrylic acid (Fig. 6C) onto **4** to generate the corresponding novel hancockiamides congeners **14**–**16**, respectively. Interestingly, while Hkm11 can activate and transfer acids smaller than cinnamic acid, larger molecules, such as caffeic acid, 3,4,5-trimethoxycinnamic acid, and 2,5-dimethoxycinnamic acid could not be incorporated (Table S6†). The production of unnatural hancockiamides **14**–**16** as new congeners expands the post-modification diversity of **4**, and the discovery of substrate tolerance for Hkm11 suggests that it could be an efficient tool for enzymatic synthesis of additional hancockiamides through directed evolution of Hkm11.

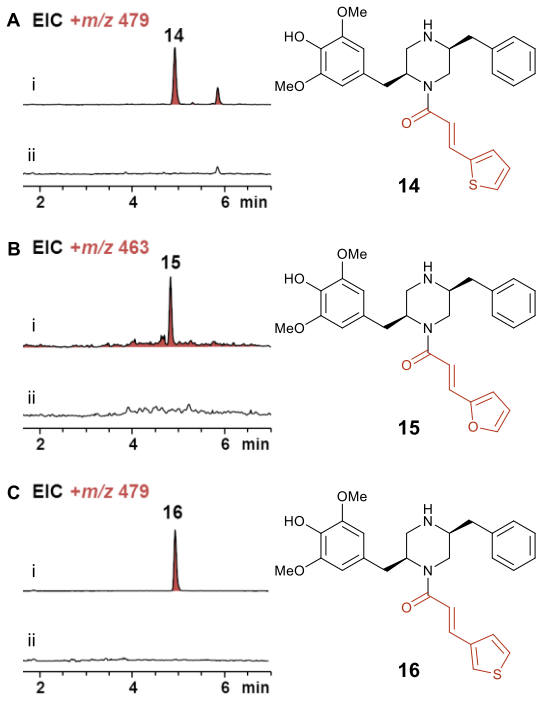


Fig. 6 Production of unnatural hancockiamide congeners by precursor feeding experiments. (A) Production of **14**. i) *S. cerevisiae* expressing *hkm11* fed with **4** and (*E*)-3-(2-thienyl)acrylic acid; ii) *S. cerevisiae* control fed with **4** and (*E*)-3-(2-thienyl)acrylic acid. (B) Production of **15**. i) *S. cerevisiae* expressing *hkm11* fed with **4** and (*E*)-3-(2-furyl)acrylic acid; ii) *S. cerevisiae* control fed with **4** and (*E*)-3-(2-furyl)acrylic acid. (C) Production of **16**. i) *S. cerevisiae* expressing *hkm11* fed with **4** and (*E*)-3-(3-thienyl)acrylic acid; ii) *S. cerevisiae* control fed with **4** and (*E*)-3-(3-thienyl)acrylic acid.

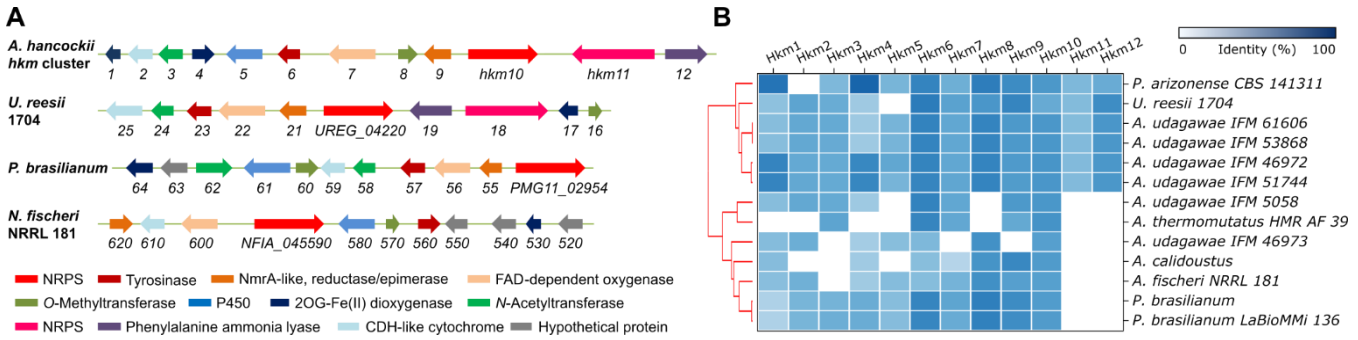


Fig. 7 Homologous biosynthetic gene clusters conserved in other fungi. (A) Selected homologous gene clusters; homologues are in same colours (see Table S5†). (B) Additional homologous gene clusters identified by cblaster in NCBI GenBank (see full results in Fig. S5†).

Genome mining of *hkm*-like biosynthetic gene clusters

Considering the bioactivities exhibited by some of the hancockiamides, we attempted to search for homologous biosynthetic gene clusters in other fungi. Using the antiSMASH[33] ClusterBlast algorithm, we discovered three homologous gene clusters in *Uncinocarpus reesii* 1704, *Penicillium brasilianum*, and *Neosartorya fischeri* NRRL 181 (Fig. 7). The core genes that are conserved across all species are *hkm10*, *hkm9*, *hkm4*, *hkm6*, *hkm7* and *hkm8*, which are responsible for the production of hancockiamide D (**4**). The *U. reesii* genome also encodes homologues of Hkm11, Hkm12 and Hkm3, while lacking the P450 Hkm5 required for methylenedioxy bridge formation. This suggests that *U. reesii* should have the potential to encode the production of **2**-like *N-trans*-cinnamoyl- and *N*-acetyl-substituted piperazines, but without the methylenedioxy bridge. By contrast, the homologous gene clusters in *P. brasilianum* and *N. fischeri* lack homologues of the PAL Hkm12 and the NRPS Hkm11, which are responsible for installation of the *N-trans*-cinnamoyl group. Instead, *P. brasilianum* encodes two acetyltransferases, which suggests that *P. brasilianum* may produce **4**-like compounds that contain two *N*-acetyl groups. Indeed, *P. brasilianum* has been reported to produce brasiliamides,[34, 35] thus it is likely that the *PMG11_02954* cluster identified here is the candidate gene cluster for brasiliamide. More recently, during the preparation of this manuscript, the biosynthetic gene cluster responsible for brasiliamide biosynthesis has been reported from *P. brasilianum* WZXY-M122-9.[36] That study showed that a clavaminic acid synthetase (CAS)–like Fe(II)/2OG oxygenase BrsJ catalyses the piperazine ring cleavage, although the homologue of BrsJ was not found in the *hkm* cluster. This could explain the presence of the open-ring **6** as a minor product in *A. hancockii* culture. Likewise, it is reasonable to predict that *N. fischeri* may encode the **8**-like methylenedioxy-containing piperazines as *N. fischeri* encodes the homologue of P450 Hkm5. Indeed, the methylenedioxy-containing brasiliamides G[37] and H[38] have been reported from *N. pseudofischeri* and *N. hiratsukae*, respectively.

Using the cblaster[39] tool for deep genome mining, we found additional homologous gene clusters in the NCBI GenBank database, e.g. *A. udagawae* and *Penicillium arizonense* (Fig. 7 and S7†). Based on the biosynthesis for hancockiamides established in this study and earlier works,[21, 22] we propose that the previously reported fungal phenylpropanoids nigerazines[40, 41], helvamide[42] and chrysosporazines[43, 44] (Fig. S8†) should be encoded by similar biosynthetic gene cluster and biosynthesised in a similar manner to the hancockiamides.

Conclusions

In conclusion, we have discovered a group of six phenylpropanoid-like piperazine alkaloids, hancockiamides A–F (**1**–**6**), together with seven heterologously produced piperazines (**7**–**13**). The *N-trans*-cinnamoyl moiety, which is rare among fungal natural products, was shown to be supplied by a dedicated PAL encoded in the *hkm* cluster and transferred to the piperazine scaffold by an unprecedented monomodular NRPS Hkm11. This represents a unique example of the merging of the phenylpropanoid and NRPS pathways. Furthermore, a novel fungal P450 was shown to form the methylenedioxy bridge from a phenolic methoxy and the *ortho*-phenolic hydroxy groups, analogous to the plant methylenedioxy-forming P450 in the biosynthesis of sesamin.[31, 32] Importantly, acetyltransferase gene *hkm3* proved to be important for pathway metabolite distribution, with its omission leading to the production of a parallel family of 2,5-dibenzylpiperazines, the xenocockiamides. This study provides biosynthetic insights into the expanding family of *N*-cinnamoyl/benzoyl-substituted piperazine metabolites, and paves way for uncovering additional such piperazines via genome mining.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Dr Atul Bhatnagar and Mr Kymren Bolling-McDougall (MQ) for assistance with acquisition of spectroscopic data. This research was funded, in part, by the Cooperative Research Centres Projects scheme (CRCPFIVE000119) and the Australian Research Council (FT160100233, FT130100142).

Notes and references

1. A. F. Brito, L. K. S. Moreira, R. Menegatti and E. A. Costa, *Fundam. Clin. Pharmacol.*, 2019, **33**, 13-24.
2. R. V. Patel and S. W. Park, *Mini Rev. Med. Chem.*, 2013, **13**, 1579-1601.
3. A. K. Rathi, R. Syed, H.-S. Shin and R. V. Patel, *Expert Opin. Ther. Pat.*, 2016, **26**, 777-797.
4. V. J. Ram, A. Sethi, M. Nath and R. Pratap, in *The Chemistry of Heterocycles*, eds. V. Ji Ram, A. Sethi, M. Nath and R. Pratap, Elsevier, 2019, pp. 3-391.
5. H. J. Lacey, D. Vuong, J. I. Pitt, E. Lacey and A. M. Piggott, *Aust. J. Chem.*, 2016, **69**, 152-160.
6. N. K. Chaudhary, J. I. Pitt, E. Lacey, A. Crombie, D. Vuong, A. M. Piggott and P. Karuso, *J. Nat. Prod.*, 2018, **81**, 1517-1526.
7. H. Li, C. L. M. Gilchrist, H. J. Lacey, A. Crombie, D. Vuong, J. I. Pitt, E. Lacey, Y.-H. Chooi and A. M. Piggott, *Org. Lett.*, 2019, **21**, 1287-1291.
8. H. J. Lacey, C. L. M. Gilchrist, A. Crombie, J. A. Kalaitzis, D. Vuong, P. J. Rutledge, P. Turner, J. I. Pitt, E. Lacey, Y.-H. Chooi and A. M. Piggott, *Beilstein J. Org. Chem.*, 2019, **15**, 2631-2643.
9. H. Li, C. L. M. Gilchrist, C.-S. Phan, H. J. Lacey, D. Vuong, S. A. Moggach, E. Lacey, A. M. Piggott and Y.-H. Chooi, *J. Am. Chem. Soc.*, 2020, **142**, 7145-7152.
10. C. L. M. Gilchrist, H. J. Lacey, D. Vuong, J. I. Pitt, L. Lange, E. Lacey, B. Pilgaard, Y.-H. Chooi and A. M. Piggott, *Fungal Genet. Biol.*, 2020, **143**, 103435.
11. J. I. Pitt, L. Lange, A. E. Lacey, D. Vuong, D. J. Midgley, P. Greenfield, M. I. Bradbury, E. Lacey, P. K. Busk, B. Pilgaard, Y.-H. Chooi and A. M. Piggott, *PLoS One*, 2017, **12**, e0170254.
12. C. Levrier, M. C. Sadowski, C. C. Nelson, P. C. Healy and R. A. Davis, *J. Nat. Prod.*, 2015, **78**, 111-119.
13. H. Nguyen Ngoc, M. Alilou, M. Stonig, D. T. Nghiem, L. T. Kim, J. M. Gostner, H. Stuppner and M. Ganzera, *J. Nat. Prod.*, 2019, **82**, 2941-2952.
14. S. Cheenpracha, S. G. Pyne, B. O. Patrick, R. J. Andersen, W. Maneerat and S. Laphookhieo, *J. Nat. Prod.*, 2019, **82**, 2174-2180.
15. J. Liu, P. Pandey, X. Wang, X. Qi, J. Chen, H. Sun, P. Zhang, Y. Ding, D. Ferreira, R. J. Doerksen, M. T. Hamann and S. Li, *J. Nat. Prod.*, 2018, **81**, 846-857.
16. 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.
17. N. Liu, Y.-S. Hung, S.-S. Gao, L. Hang, Y. Zou, Y.-H. Chooi and Y. Tang, *Org. Lett.*, 2017, **19**, 3560-3563.
18. R. Nofiani, K. de Mattos-Shipley, K. E. Lebe, L.-C. Han, Z. Iqbal, A. M. Bailey, C. L. Willis, T. J. Simpson and R. J. Cox, *Nat. Commun.*, 2018, **9**, 3940.
19. M. Janvier, P.-H. Ducrot and F. Allais, *ACS Sustain. Chem. Eng.*, 2017, **5**, 8648-8656.
20. B. J. Schultz, S. Y. Kim, W. Lau and E. S. Sattely, *J. Am. Chem. Soc.*, 2019, **141**, 19231-19235.
21. R. R. Forseth, S. Amaike, D. Schwenk, K. J. Affeldt, D. Hoffmeister, F. C. Schroeder and N. P. Keller, *Angew. Chem. Int. Ed.*, 2013, **52**, 1590-1594.
22. X. Yu, F. Liu, Y. Zou, M.-C. Tang, L. Hang, K. N. Houk and Y. Tang, *J. Am. Chem. Soc.*, 2016, **138**, 13529-13532.
23. M. W. Hyun, Y. H. Yun, J. Y. Kim and S. H. Kim, *Mycobiology*, 2011, **39**, 257-265.
24. R. A. Cacho, Y. Tang and Y.-H. Chooi, *Front. Microbiol.*, 2015, **5**, 774.
25. Y.-M. Chiang, M. Ahuja, C. E. Oakley, R. Entwistle, A. Asokan, C. Zutz, C. C. C. Wang and B. R. Oakley, *Angew. Chem. Int. Ed.*, 2016, **55**, 1662-1665.
26. A. Marchler-Bauer, S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng and S. H. Bryant, *Nucleic Acids Res.*, 2010, **39**, D225-D229.
27. J. Hu, F. Sarrami, H. Li, G. Zhang, K. A. Stubbs, E. Lacey, S. G. Stewart, A. Karton, Andrew M. Piggott and Y.-H. Chooi, *Chem. Sci.*, 2019, **10**, 1457-1465.
28. S. M. Ma, J. W.-H. Li, J. W. Choi, H. Zhou, K. K. M. Lee, V. A. Moorthie, X. Xie, J. T. Kealey, N. A. Da Silva, J. C. Vederas and Y. Tang, *Science*, 2009, **326**, 589-592.
29. H. Satake, E. Ono and J. Murata, *J. Agric. Food Chem.*, 2013, **61**, 11721-11729.
30. P. C. G. Rida, D. LiVecche, A. Ogden, J. Zhou and R. Aneja, *Med. Res. Rev.*, 2015, **35**, 1072-1096.
31. E. Ono, M. Nakai, Y. Fukui, N. Tomimori, M. Fukuchi-Mizutani, M. Saito, H. Satake, T. Tanaka, M. Katsuta, T. Umezawa and Y. Tanaka, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 10116-10121.
32. E. Ono, J. Murata, H. Toyonaga, M. Nakayasu, M. Mizutani, M. P. Yamamoto, T. Umezawa and M. Horikawa, *Plant Cell Physiol.*, 2018, **59**, 2278-2287.
33. K. Blin, T. Wolf, M. G. Chevrette, X. Lu, C. J. Schwalen, S. A. Kautsar, H. G. Suarez Duran, Emmanuel L. C. de los Santos, H. U. Kim, M. Nave, J. S. Dickschat, D. A. Mitchell, E. Shelest, R. Breitling, E. Takano, S. Y. Lee, T. Weber and M. H. Medema, *Nucleic Acids Res.*, 2017, **45**, W36-W41.
34. T. Fujita, D. Makishima, K. Akiyama and H. Hayashi, *Biosci. Biotechnol. Biochem.*, 2002, **66**, 1697-1705.
35. T. Fujita and H. Hayashi, *Biosci. Biotechnol. Biochem.*, 2004, **68**, 820-826.
36. B. Yuan, D. Liu, X. Guan, Y. Yan, J. Zhang, Y. Zhang, D. Yang, M. Ma and W. Lin, *Appl. Microbiol. Biotechnol.*, 2020, **104**, 6149-6159.
37. J. Paluka, K. Kanokmedhakul, M. Soytong, K. Soytong and S. Kanokmedhakul, *Fitoterapia*, 2019, **137**, 104257.
38. J. Paluka, K. Kanokmedhakul, M. Soytong, K. Soytong, J. Yahuafai, P. Siripong and S. Kanokmedhakul, *Fitoterapia*, 2020, **142**, 104485.
39. C. L. M. Gilchrist, http://doi.org/10.5281/zenodo.3953322, 2020.
40. T. Iwamoto, A. Hirota, S. Shima, H. Sakai and A. Isogai, *Agric. Biol. Chem.*, 1985, **49**, 3323-3325.
41. T. Iwamoto, S. Shima, A. Hirota, I. Akira and H. Sakai, *Agric. Biol. Chem.*, 1983, **47**, 739-743.
42. T. Fukuda, T. Furukawa, K. Kobayashi, K. Nagai, R. Uchida and H. Tomoda, *J Antibiot*, 2019, **72**, 8-14.
43. A. H. Elbanna, Z. G. Khalil, P. V. Bernhardt and R. J. Capon, *Org. Lett.*, 2019, **21**, 8097-8100.

44. O. G. Mohamed, A. A. Salim, Z. G. Khalil, A. H. Elbanna, P. V. Bernhardt and R. J. Capon, *J. Nat. Prod.*, 2020, **83**, 497-504.