Biosynthesis

Chemical and Genetic Studies on the Formation of Pyrrolones During the Biosynthesis of Cytochalasans

Haili Zhang,[a, b] Verena Hantke,[a, b] Pia Bruhnke,[a, b] Elizabeth J. Skellam,*[a, b, c] and Russell J. Cox*[a, b]

Abstract: A key step during the biosynthesis of cytochalasans is a proposed Knoevenagel condensation to form the pyrrolone core, enabling the subsequent 4+2 cycloaddition reaction that results in the characteristic octahydroisoindolone motif of all cytochalasans. In this work, we investigate the role of the highly conserved α,β-hydrolase enzymes PyiE and ORFZ during the biosynthesis of pyrichalasin H and the ACE1 metabolite, respectively, using gene knockout and complementation techniques. Using synthetic aldehyde models we demonstrate that the Knoevenagel condensation proceeds spontaneously but results in the 1,3-dihydro-2*H*-pyrrol-2-one tautomer, rather than the required 1,5-dihydro-2*H*-pyrrol-2-one tautomer. Taken together our results suggest that the α,β-hydrolase enzymes are essential for first ring cyclisation, but the precise nature of the intermediates remains to be determined.

Introduction

Cytochalasans are secondary metabolites that have been isolated from a wide range of fungi and that possess numerous potent bioactivities extensively reviewed elsewhere.[1,2] As might be expected for metabolites that have been known since the mid-1960s,[3] their biosynthesis has been extensively studied,[4,5] and their intriguing structures have been the target for numerous total synthesis campaigns.[6]

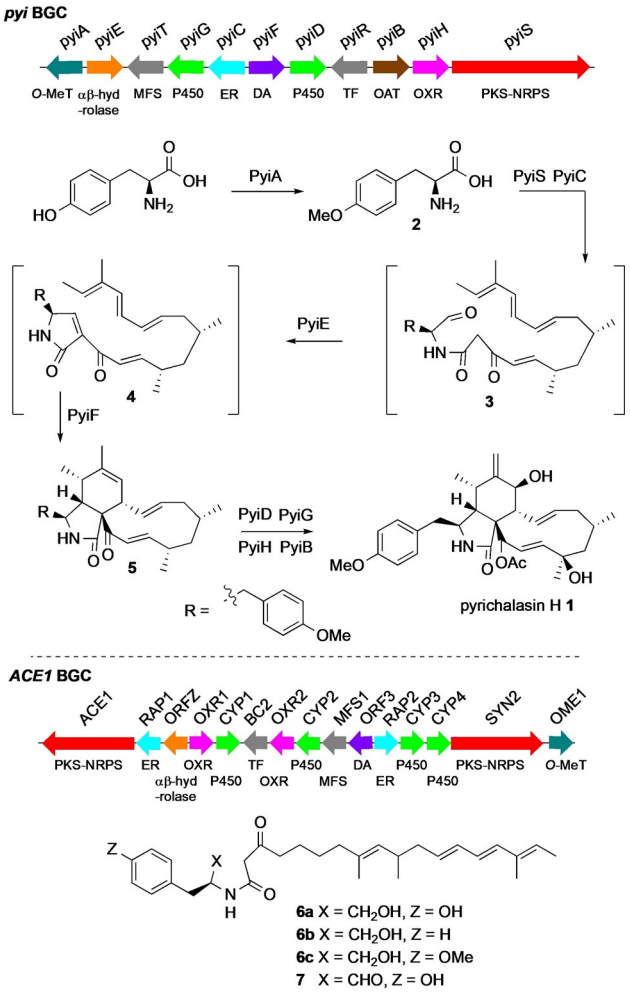
We have investigated the biosynthesis of pyrichalasin H **1** from the fungus *Magnaporthe grisea* (Scheme 1).[7] The results to-date are consistent with an overall pathway in which *O*-methyl tyrosine **2** is created by PyiA and used as the initial building-block by the polyketide synthase non-ribosomal peptide synthetase (PKS-NRPS) PyiS which acts in concert with the *trans*-enoyl reductase (ER) PyiC. PyiS has a reductive release system and it is hypothesised that an aldehyde intermediate

[a] Dr. H. Zhang, Dr. V. Hantke, P. Bruhnke, Dr. E. J. Skellam, Prof. Dr. R. J. Cox
Institute for Organic Chemistry, Leibniz Universität Hannover
Schneiderberg 1B, 30167, Hannover (Germany)
E-mail: elizabeth.skellam@unt.edu
russell.cox@oci.uni-hannover.de

[b] Dr. H. Zhang, Dr. V. Hantke, P. Bruhnke, Dr. E. J. Skellam, Prof. Dr. R. J. Cox
Biomolekulares Wirkstoff Zentrum (BMWZ), Leibniz Universität Hannover
Schneiderberg 38, 30167, Hannover (Germany)

[c] Dr. E. J. Skellam
Current Address: Department of Chemistry
University of North Texas, 1508 W Mulberry, 30167, Denton (Texas, USA)

Supporting information and the ORCID identification numbers for the authors of this article can be found under:
https://doi.org/10.1002/chem.202004444.



Scheme 1. Comparison of the *pyi* and *ACE1* BGC and proposed biosynthesis of metabolites encoded by these BGC.

such as **3** is released. Cyclisation of **3** by a Knoevenagel process is then proposed to give **4**,[8] and Diels–Alder (DA) reaction could then give the observed cytochalasan skeleton of **5**. The function of tailoring enzymes PyiBDGH has been proven using gene knockout (KO) and heterologous expression experiments during the formation of the mature metabolite **1**.[9] Meanwhile, *pyiR* encodes a transcription-factor which we have shown to be responsible for increasing the titre of **1**.[10] Recently we showed that PyiF is involved in the proposed DA reaction.[11,12]

These observations leave PyiE, which encodes a putative hydrolytic enzyme, to be proposed as the Knoevenagel catalyst *by default* since all other proteins encoded by the pyrichalasin biosynthetic gene cluster (*pyi* BGC) have proven roles. In addition, compounds **3** and **4** (or their analogues in other pathways) have never been directly observed, so uncertainty remains regarding the early biosynthetic steps.

We have also investigated the parallel ACE1 biosynthetic pathway[13] in the closely related fungus *Pyricularia oryzae* (Scheme 1). *P. oryzae* causes rice blast disease and the ACE1 metabolite is involved in avirulence signalling between the fungal pathogen and resistant rice plants. The *ACE1* BGC is only active in single cell appressoria for a matter of hours during initial fungal penetration of plant leaves, so the identity of the ACE1 metabolite is not known. It has also been impossible, for the same reasons, to perform productive KO experiments in the ACE1 system.[14] However, the high gene-for-gene homologies between the *ACE1* and *pyi* genes, and the fact that *ACE1* genes can complement knockouts in the *pyi* BGC mean that the ACE1 metabolite is very likely to be a cytochalasan derived from **2**. Heterologous expression experiments[15] have also reinforced this hypothesis. Expression of *ACE1* (PKS-NRPS) and *RAP1* (*trans*-ER) in *Aspergillus oryzae* lead to the production of various shunt compounds such as **6a** and **6b** in the absence of **2**,[15] and **6c** when **2** is supplied. It is assumed that these are derived by rapid reduction of putative aldehyde intermediates such as **7**.[16] Interestingly, putative intermediate **7** differs from the proposed pyrichalasin aldehyde intermediate **3** by being two carbons longer showing that the ACE1 PKS is differently programmed.

Hydrolytic enzymes similar to PyiE, and its ACE1 homolog ORFZ, are encoded in the BGC of all known cytochalasans,[17] as well as in other BGC which encode the biosynthesis of compounds such as fusarin C where the PKS-NRPS has a reductive release mechanism.[18] However, the precise function of these hydrolytic enzymes remains unknown. As part of ongoing studies into the biosynthesis of cytochalasans we therefore decided to investigate PyiE from *M. grisea* and ORFZ from *P. oryzae* in more detail.

Results

Knockout of *pyiE*

Disruption of *pyiE* in *M. grisea* was achieved by insertion of a hygromycin resistance cassette using the bipartite method of Neilsen and co-workers.[19] Examination of organic extracts from *M. grisea ΔpyiE* showed that, in comparison to extracts from wild-type (WT) *M. grisea*, production of **1** was significantly reduced, concomitantly with an increase in production of known pyriculols.[20] We then re-introduced a copy of either *pyiE* itself or its homolog *ORFZ*, under the control of the *Aspergillus oryzae amyB* promoter (P_{amyB}), into the KO strain using previously reported methods.[21] In previous work we had shown that P_{amyB} is functional in *M. grisea* and we therefore expected that biosynthesis of **1** should be restored to WT titres.[7,11] However, LCMS analysis showed that **1** was not produced in significantly increased titre for either complementation experiment (Figure 1 and Figures S6.1–S6.6).

The LCMS profiles obtained from these strains were similar to those we had observed during KO of *pyiA*.[7] Since the target gene *pyiE* is adjacent to *pyiA* (Scheme 1) we considered it possible that either *pyiA* itself, or the putative bidirectional promoter region between *pyiE* and *pyiA*, may have been damaged by the *pyiE* engineering. We tested this possibility by feeding *O*-methyl tyrosine **2** to cultures of *M. griseaΔpyiE*, *M. griseaΔpyiE:*P_{amyB}*·pyiE* and *M. griseaΔpyiE:*P_{amyB}*·ORFZ*. In the case of the *M. griseaΔpyiE:*P_{amyB}*·pyiE* and *M. griseaΔpyiE:*P_{amyB}*·ORFZ* strains production of **1** was substantially restored (Figure 1 and Figures S6.1 to S6.6).

In the *ΔpyiE* + **2** experiments a major new peak eluting at 5.1 minutes was also observed (Figure 2). Mass data corresponded to the pre-Diels–Alder intermediates or shunt metabolites. Mass spectrometric analysis indicated a mass of 519 Da, also consistent with compounds related to **1** (e.g. **10**, *m/z* 521). This compound was detected in both WT and the *ΔpyiE* mutant, but in much lower amounts (Figure S6.9). The major isomer eluting at 5.1 min was purified by mass-directed preparative reverse-phase chromatography. HRMS indicated a formula of $C_{29}H_{45}NO_7$ (observed 542.3096, calculated 542.3094 for $C_{29}H_{45}NNaO_7$) consistent with the 29 main-chain carbons of **1**. Full NMR analysis (Supporting Information, section S7) showed the compound to be **9**, confirming it as a pre-cyclisation shunt compound. Reduction of the expected C-21 aldehyde to a primary alcohol is consistent with the known propensity of fungi to reduce free aldehydes.[15,16a]

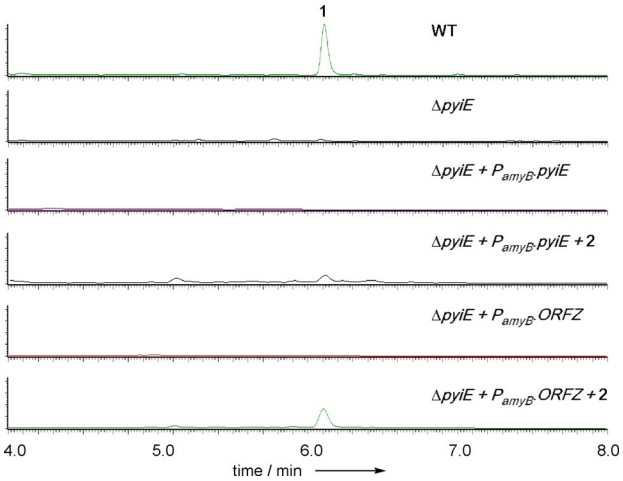


Figure 1. Production of pyrichalasin H **1** in WT *M. grisea* NI980 and selected mutants. UV at 275 nm.

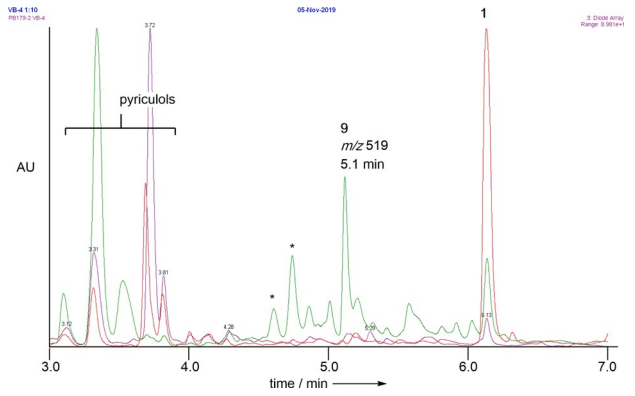


Figure 2. Red trace, WT *M. grisea* supplemented with *O*-methyl tyrosine **2**; Green trace, *M. grisea* Δ*pyiE* supplemented with *O*-methyl tyrosine **2**; Purple trace, *M. grisea* Δ*pyiE* alone. * unrelated metabolites.

Dihydroxylation of the skeleton of **9** at C-15/C-16 and hydroxylation at C-9 are consistent with modifications previously observed in shunt metabolites (e.g. **10**) in *M. grisea*, for example during KO studies of the Diels–Alderase gene *pyiF*.[7] The structural assignment was further confirmed by observation of a major mass fragment at *m/z* 364 corresponding to a facile homo-allylic fragmentation[22] of the polyketide triol sidechain (Figure 3).

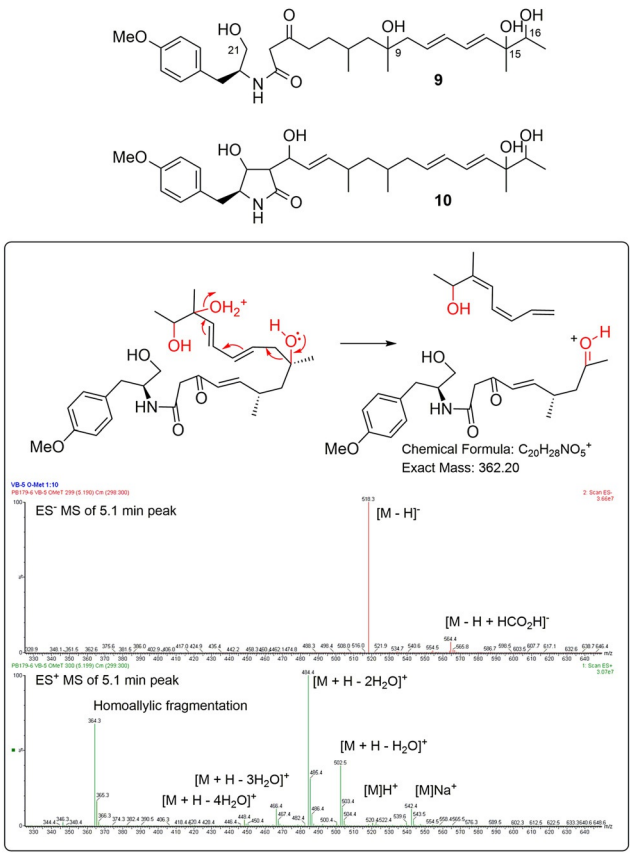


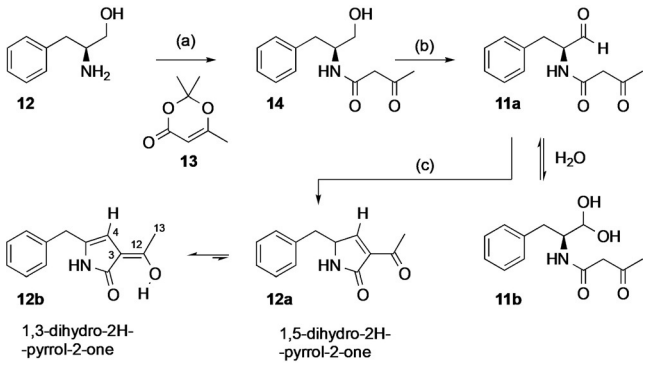
Figure 3. Structure of Δ*pyiE* shunt metabolite **9** isolated in this study and shunt metabolite **10** isolated in the Δ*pyiF* study.[11]

Heterologous expression of *ACE1*, *RAP1* and *ORFZ* in *Aspergillus oryzae*

Previously we had observed that when we co-expressed *ACE1* + *RAP1* in *A. oryzae*, linear alcohol shunt products **6a–c** were observed rather than expected aldehyde **7** (Scheme 1).[15] To examine whether the inclusion of the corresponding α,β-hydrolase ORFZ prevented reduction and enabled biosynthesis of the expected aldehyde **7** and ring-closed congeners analogous to **4**, we co-expressed *ACE1*, *RAP1* and *ORFZ* in *A. oryzae* using previously reported methods.[15,21] However, the same shunt intermediates **6a**,**b** were observed in the absence of exogenous **2** (Figure S6.7), while **6c** was produced in the presence of **2** (Figure S6.8). To investigate the possibility that the Diels–Alderase, ORF3,[11] is also required, we co-expressed *ACE1*, *RAP1*, *ORFZ* and *ORF3* in *A. oryzae*. However, the same shunt intermediates **6a–c** were again observed (Figure S6.7). PCR and RT-PCR analyses showed that all transformed genes were being expressed in *A. oryzae*, and previous complementation studies had shown that ORFZ and ORF3 were correctly cloned. We therefore concluded that aldehyde reduction out-competes any ring-closing in this host. Realising that we could not investigate the function of PyiE or ORFZ in vivo, we decided to investigate the Knoevenagel reaction in vitro.

Chemical investigation of Knoevenagel step

In order to investigate the chemistry of the proposed Knoevenagel step we synthesised the model compound **11a** which features a phenylalanine-derived aldehyde similar to proposed intermediate **3**, but with a truncated β-keto polyketide designed for simpler synthesis and analysis. Thus, amino-alcohol 2*S*-**12** was treated with acetone diketene adduct **13**, to give **14** and oxidised under mild Pfitzner-Moffatt conditions[23] to give the required aldehyde **11a. 11a** was treated with 1 M NaOH in an attempt to generate **12a** (Scheme 2). However, in aqueous solution **12a** rapidly equilibrates to its 1,3-dihydro tautomer **12b**. This structure is supported by: chemical shift data, especially for the C-13 methyl (δ_H 2.15 ppm); the chemical shift of H-4 (δ_H 5.54 ppm); the chemical shifts of C-12 and C-3 (δ_C



Scheme 2. Synthesis of model pyrrol-2-one **12**. Reagents and conditions: (a) toluene, 100 °C; (b) EDCI, Cl_2CHCO_2H, DMSO:toluene 1:1, RT, 16 h; (c) 1 M aq. NaOH, 0 °C, 5 min.

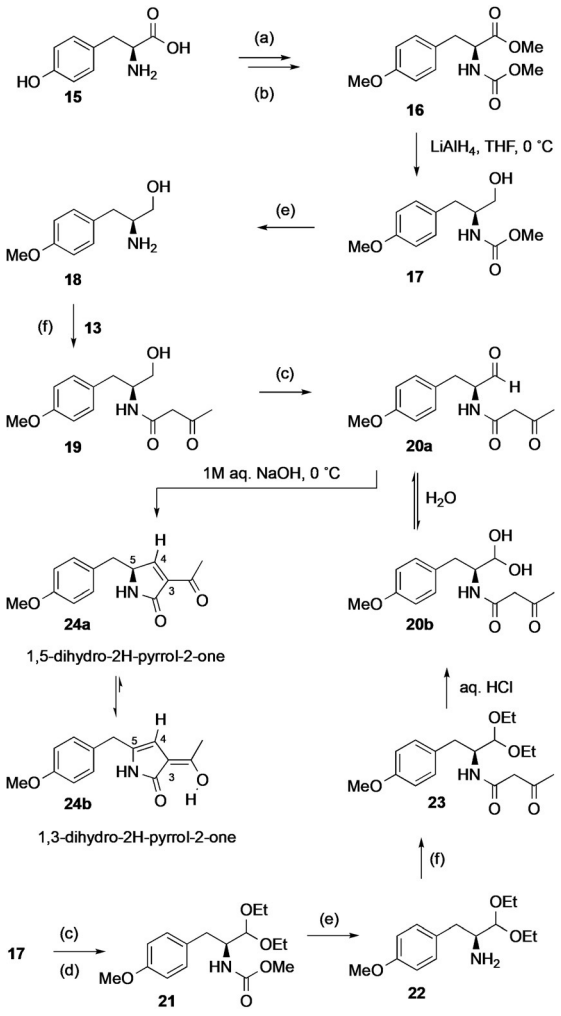
172.3 and 106.9 ppm, respectively); the observation of a distinct UV absorption (329 nm) corresponding to the aromatic heterocycle; and loss of optical rotation (e.g. for **20a** $[\alpha]_D = -12.0$ ($c = 0.9$, CH_2Cl_2); for **23**, $[\alpha]_D = -16.9$ ($c = 0.3$, MeOH); for **12** $[\alpha]_D = 0$).

As pyrichalasin H **1** is derived from *O*-methyl tyrosine **2** we also synthesised a *p*-methoxy derived aldehyde **20a** (Scheme 3). L-Tyrosine **15** was protected as its *N*-methyl carbamate (MOC) and then methylated to give **16** (Scheme 3). $LiAlH_4$ reduction led to **17** and saponification then gave amino-alcohol **18**. Treatment with acetone diketene adduct **13** under thermolytic conditions gave **19**. Modified Pfitzner–Moffatt oxidation of **19** resulted in the desired compound **20a**, which once again equilibrated with its hydrate form **20b** in water. Aldehyde **20a**, however, was not easy to purify by flash chromatography, although mg amounts could be purified by reverse-phase HPLC. In an alternative route, Moffatt oxidation of **17** earlier in the synthesis was followed by formation of the corresponding diethyl acetal to give **21**. Then basic removal of the MOC group to **22** and attachment of the *N*-acetoacetyl group using acetone diketene adduct **13** gave **23** which could be prepared and purified at larger scale. This material could be deprotected under mild aqueous acid conditions, followed by neutralisation for convenient production of **20a** in situ.

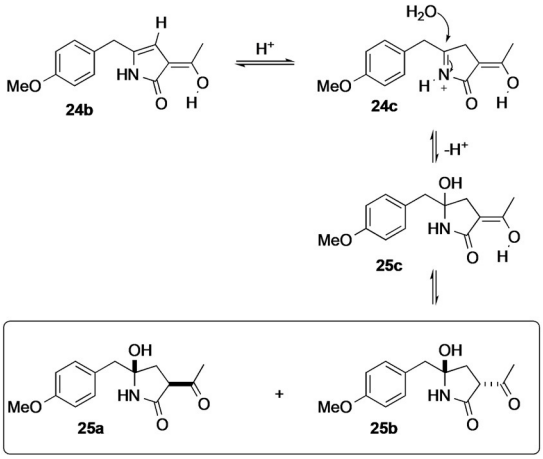
Once again, brief base treatment achieved cyclisation to give the 1,3-dihydro-2*H*-pyrrol-2-one tautomer **24b**. This compound proved significantly less stable than the phenylalanine analogue **12b** and it was difficult to purify and obtain full NMR data. In aqueous solution **24b** rapidly reacted further to give a co-eluting mixture of two new compounds **25** with *m/z* 263 which were more stable and which could be purified and fully characterised. Full NMR analysis showed these compounds to be a pair of diastereomers based on the 5-hydroxy-pyrrolid-2-one skeleton, presumably derived by addition of water to **24b** (Scheme 4).

In order to obtain further evidence that the obtained heterocycles exist predominantly as their 1,3-dihydro-2*H*-pyrrol-2-one tautomers we synthesised a model compound unable to tautomerise in this way. Briefly, *p*-methoxyaniline **26** was diazotized at low temperature, and the resulting diazonium salt **27** was treated with methallyl alcohol **28** in the presence of iron sulfate to give diazo compound **29** (Scheme 5). A Zn/HCl reduction then gave the amino alcohol **30**, and the usual sequence of acetoacetylation to **31**, Moffatt oxidation to **32** and aqueous base cyclization gave the expected 1,2-dihydro-2*H*-pyrrol-2-one tautomer **33**. This differed from the non-methylated compound **12b** in lacking a distinct UV absorption for the pyrrolone (e.g. λ_{max} shift from 329 to 275 nm); having a δ_H for H-4 of 7.71 rather than 5.54 ppm; different carbon chemical shifts for C-12 and C-3 (δ_C 194.3 and 135.2 ppm); and a more obvious methyl-13 ketone resonance (δ_H 2.51 ppm).

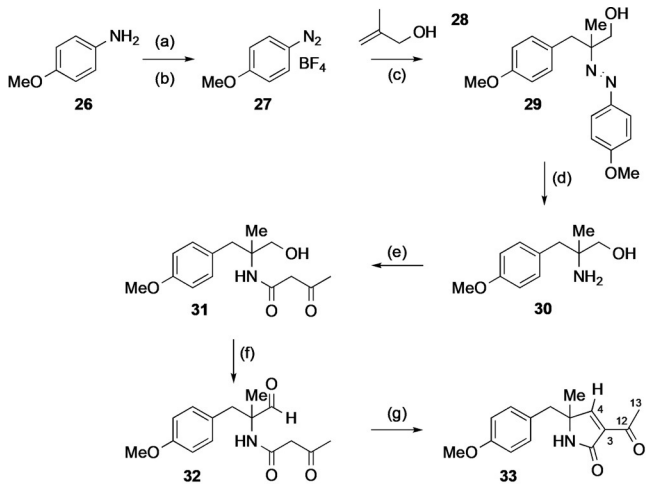
With synthetic aldehydes in hand we next set out to study their behaviour under aqueous conditions, and in the presence of preparations of the putative Knoevenagel catalyst ORFZ. The native *M. oryzae ORFZ* gene was cloned into the expression vector pET28a(+), confirmed by sequencing. The plasmid was transformed into *E. coli* BL21 (DE3) and ORFZ was expressed by addition of IPTG to a final concentration of 0.1-



Scheme 3. Synthesis of *p*-methoxy model compounds. Reagents and conditions: (a) MOCCl, THF, H_2O, $NaHCO_3$; (b) MeI, K_2CO_3, DMF; (c) EDCI, Cl_2CHCO_2H, DMSO:toluene 1:1, RT, 16 h; (d) $(EtO)_3CH$, EtOH, *p*-TsOH, 50 °C, 16 h; (e) 25% KOH, MeOH, H_2O, 50 °C, 12 h; (f) **13**, toluene 100 °C, 2 h.



Scheme 4. Formation of hemiaminal diastereomers **25ab**.



Scheme 5. Synthesis of α-methyl model compounds. Reagents and conditions: (a) HBF_4; (b) $NaNO_2$, $<5\,^{\circ}C$; (c) $FeSO_4.7H_2O$; (d) Zn, HCl, MeOH, RT, 16 h; (e) **13**, toluene, 100 °C, 2 h; (f) EDCI, Cl_2CHCO_2H, DMSO:toluene 1:1, RT, 16 h.

1 mM (Figure S3.1). His_6-ORFZ was isolated from the soluble protein fraction using Ni-NTA chromatography and the purified protein was analysed by ESI Q-TOF to confirm the correct sequence.

In vitro assays

Aldehyde **20a** was prepared at a concentration of *ca* 0.5 mM in water and incubated at 28 °C. The composition of the solution was monitored by analytical LCMS. Under these conditions **20** appeared to be relatively stable as the hydrate form **20b** (Figure 4), but cyclised slowly to form the previously characterised alcohols **25ab** (*ca* 10% after 17 h). Aldehyde **20** was then tested in buffer solutions. Buffers such as Tris which contain a free amine are unsuitable for use with aldehydes so phosphate buffer was initially tested at pH values between 6.0 and 8.0. Under these conditions **20** behaved similarly to its reactions in water (Figure 5). However a small amount of an additional new compound **34** with *m/z* 522 was observed after 23 h. In phosphate at pH 7.0 a much higher proportion of alcohols **25** were

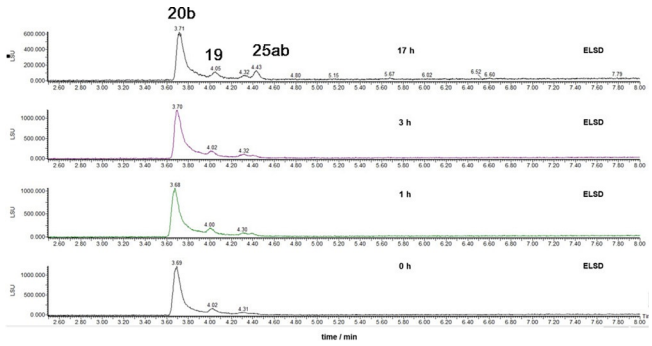


Figure 4. Model **20** in H_2O at 28 °C over time. Alcohol **19** is a residual synthetic intermediate (see Scheme 3) and not formed in the incubation.

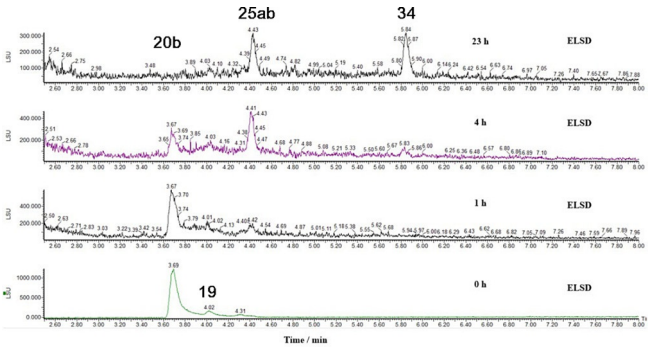


Figure 5. Model **20** in potassium phosphate buffer (pH 7.0, 75 mM) at 28 °C under air over time. Alcohol **19** is a residual synthetic intermediate (see Scheme 3) and not formed in the incubation.

observed and the new compound with *m/z* 522 also increased significantly (Figure 5).

At pH 8.0 in phosphate buffer aldehyde **20** appeared to be quantitatively converted to the *m/z* 522 compound after 24 h (Figure 6). A mass of 522 corresponds to a potential dimerization and oxidation of **25**. The experiments were therefore repeated under an atmosphere of N_2. Under these conditions the *m/z* 522 compound was no longer observed, but at pH 8.0, for example, complete conversion to the alcohols **25** was achieved (see Supporting Information, section S8).

New compound **34** was purified. Full NMR analysis confirmed this to be an oxidised dimer of the alcohol **25**. A series of intra-unit HMBC correlations (Figure 7) confirmed the presence of two different monomers, and a series of inter-unit

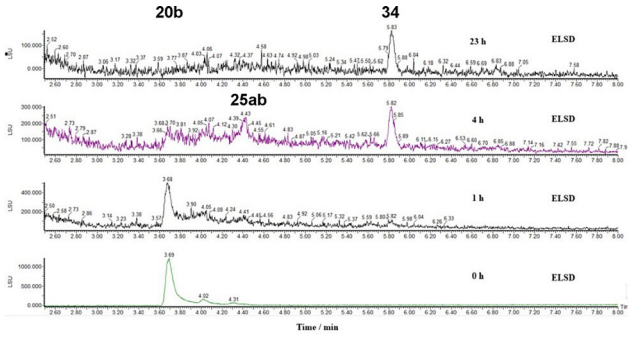


Figure 6. Model **20** in potassium phosphate buffer (pH 8.0, 75 mM) at 28 °C under air over time.

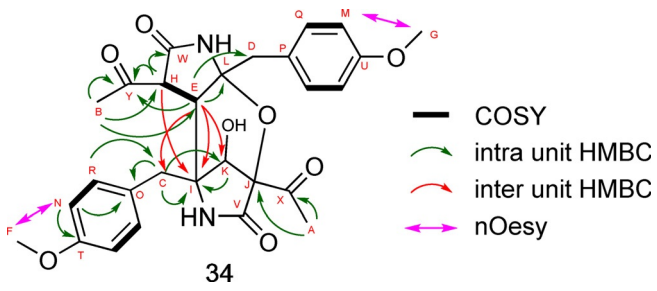


Figure 7. Elucidated structure of oxidised dimer of **34**.

HMBC correlations confirmed the carbon-carbon linkage shown in Figure 7. Chemical shift considerations suggest that carbons L and J (Figure 7) are linked via oxygen rather than the amide nitrogen.

Further investigations into buffers indicated that HEPES buffer induced the minimum formation of **25** and dimer **34**. Therefore this buffer was used in subsequent in vitro assays with enzymes. Attempts were then made to observe the effect of adding either purified ORFZ or cell-free extracts of *E. coli* expressing ORFZ, to **20** (e.g. Figures S8.1–S8.11). However, despite numerous attempts, no catalysis could be observed.

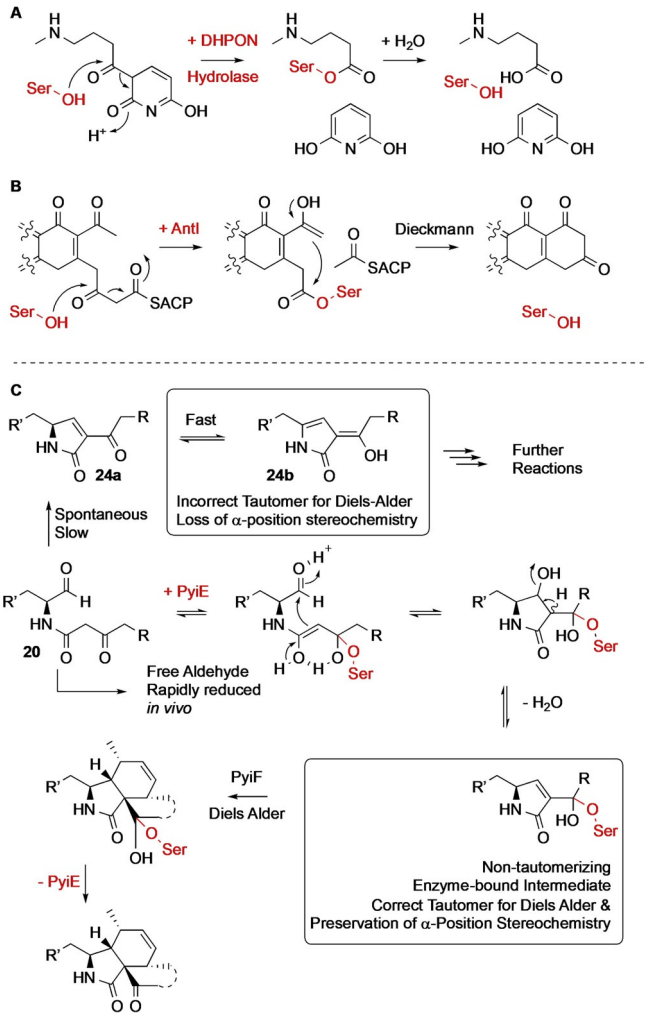
Sequence analysis

α,β-Hydrolases related to PyiE and ORFZ are commonly found encoded in fungal PKS-NRPS biosynthetic gene clusters, especially those with a reductive release mechanism.[17] For example, the cytochalasin E (CcsE),[5a] fusarin (Fus2),[18a,24] pseurotin (PsoG),[18c] and phomacin (SnoG)[25] gene clusters all encode homologous α,β-hydrolases. No structural data for orthologs of ORFZ/PyiE has been reported, but database (ESTHER)[26] and structural prediction (Phyre-2,[27] Swiss-Model)[28] searches indicate that well-studied enzymes such as acylaminoacyl peptidase (3fnb), 2,6-dihydroxy-pseudo-oxynicotine (DHPON, 2jbw) hydrolase,[29,30] and AntI[31] (6hxa) possess the same overall protein fold. DHPON hydrolase catalyses a retro-aldol C–C bond cleavage in *Arthrobacter nicotinovorans* (Scheme 6A), while AntI catalyses retro-Claisen reaction followed by a ring-closing cyclisation during the biosynthesis of polyketide-derived anthraquinones in *Photorhabdus luminescens* (Scheme 6B).

Sequence alignment of DHPON hydrolase, AntI, PyiE, ORFZ, CcsE, Fus2, PsoG and SnoG was performed using COBALT,[32] and a model structure of ORFZ was constructed in Swiss-Model using 3fnb as a template (see Supporting Information) to give an adequate resulting structure (QMEAN −3.48). The ORFZ model aligns with AntI with an RMSD of 3.5 Å over 344 residues and comparison of the sequence and structural alignments gives useful information (Figures S9.2).

AntI has a nucleophilic serine (S245) located at a 'nucleophilic elbow' in the highly conserved motif GXSXGG.[31] This motif and the catalytic serine is conserved in DHPON-hydrolase and all the PyiE/ORFZ family members. In AntI S245 forms part of a catalytic triad with D326 and essential H355. H355 has been shown to protonate the retro-aldol nucleofuge.[31] Structural homologs of D326 and H355 are present in DHPON-hydrolase (D300 and H329 respectively) which also catalyses C–C bond cleavage, but these catalytic residues are absent from PyiE, ORFZ and the other cytochalasan hydrolases.

After the retro-Claisen reaction, AntI also catalyses a ring-closing reaction (Scheme 6B). This has been shown to involve enolisation followed by a Dieckmann cyclisation. Residues D175, S176, E181 and R24 of AntI are involved in catalysis of this step, while Y20 is thought to be involved in substrate orientation.[31] These residues are conserved in DHPON-hydrolase and in the PyiE family of hydrolases (e.g. D174, S175, E178, R17 and F13 respectively in PyiE and D165, E169, S243, F12 and R16 in ORFZ, Figure S9.2).



Scheme 6. Suggested alternative 'Hydrolase'/'Diels–Alderase' route to the cytochalasan skeleton. **A**, reactions catalysed by the DHPON hydrolase; **B**, reactions catalysed by AntI; **C**, suggested analogous pathway to Cytochalasans. Note that DHPON-hydrolase may also use water directly as a nucleophile.[29,34]

Discussion

Knock-out of *pyiE* in *M. grisea* significantly reduced the titres of **1**, but initial complementation of *pyiE* did not restore **1** to WT levels. When exogenous *O*-methyltyrosine **2** was fed to the *pyiE* complemented strain, production of **1** was restored, and new compounds were detected in the *ΔpyiE* strain (Figure 2). This is consistent with damage to the promoter region of *pyiA* which encodes the essential tyrosine *O*-methyl transferase. The major new compound from the *ΔpyiE* strain chemically complemented with **2** was isolated, purified and structurally characterised as the linear *O*-methyltyrosine octaketide **9**. This compound has not undergone Knoevenagel condensation and instead the expected aldehyde has been reduced to an alcohol. The shunt compound has also been oxidized at C-9, C-15 and C-16 (Figure 3) through addition of hydroxyl groups. These modifications are highly similar to the shunt pyrichalasin precursor **10** observed when pyiF was disrupted in *M. grisea*.[7] This KO result is similar to the results of KO experiments ob-

served in the cases of fusarin (*fus2*)[18a] and pseurotin (*psoG*)[18a,c] which also demonstrated that these genes are essential for biosynthesis, although intermediates or precursors were not identified.

Heterologous complementation of *ΔpyiE* strains with *ORFZ* from the ACE1 BGC shows that PyiE and ORFZ have the same function. This reinforces previous observations in which we showed that other enzymes from the ACE1 cluster can complement the *pyi* pathway: for example ORF3 can complement PyiF (the Diels–Alderase); and late stage tailoring enzymes from the ACE1 pathway such as CYP1 and CYP3 are also functional in the pyrichalasin system.[7]

In vitro investigations using synthetic model compounds **11** and **20** show that they do not cyclise rapidly. In water or phosphate buffer at pH 7.0 a very slow spontaneous Knoevenagel cyclisation is observed, followed by further addition of water at the β-position to give **25**, and these reactions are accelerated at higher pH. In strong base the cyclisation is fast, but further water addition does not occur. Detailed studies of the major pyrrolinone tautomer show conclusively that the 1,3-dihydro-form **24b** is formed: this tautomer cannot undergo the required Diels–Alder cyclisation. Formation of the 1,3-dihydro tautomer **24b** also removes the α-stereocentre which is conserved in the cytochalasans, strongly suggesting that non-enzymatic equilibration of 1,3- and 1,5 dihydro tautomers does not occur during cytochalasan biosynthesis. In phosphate buffer in the presence of atmospheric oxygen further oxidation and dimerisation of **24b** takes place to give **34**, emphasising the inherent reactivity of these intermediates. Our attempts to show in vitro activity of expressed and purified ORFZ with model compound **20** were not successful. It may be that ORFZ is incorrectly folded, as observed for many fungal protein targets of *E. coli* expression;[33] or that **20** is not a sufficiently elaborate substrate for the protein. Alternatively, ORFZ may require interaction with other protein partners (vide infra) or the PKS-NRPS itself. No doubt the lack of soluble and active ORFZ/PyiE family members has limited progress in this area.

Conclusions

Taking the above observations into account we can make the following conclusions. First, Knoevenagel reaction of putative intermediate **3** (Scheme 1) requires catalysis by PyiE/ORFZ at physiological pH. The model studies show the uncatalysed reaction is slow at pH 7 and the *pyiE* KO studies and *A. oryzae* heterologous expression studies show that intermediate aldehydes such as **3** are quickly reduced in vivo. This agrees with observations from other heterologous expression studies in our group[11,15] and from Oikawa and co-workers,[4e,16a] which show that expression of early genes from cytochalasan pathways in *A. oryzae* always leads to alcohol shunt intermediates. Corroborating observations from Lebrun and co-workers show that ORFZ (hydrolase) and ORF3 (Diels–Alderase) are the most highly expressed genes from the ACE1 BGC (up to 7-fold higher than the ACE1 PKS-NRPS itself), suggesting a high concentration of these catalysts is required to avoid the reductive shunt pathways.[13]

Second, the model studies show that rapid tautomerisation of **24a** to **24b** creates a substrate which is incompetent for further Diels–Alder reaction because the alkene is no longer present at C-3,C-4. Tautomer **24b** is itself reactive and undergoes other reactions leading it away from the cytochalasan pathway. Interestingly the observed hemi-aminal functionality of **25ab** mirrors the same functionality found in in fusarin C[18a] and talaroconvolutin B (Figure 8),[35] for example.

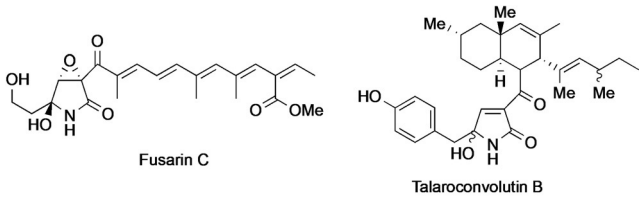


Figure 8. Molecular structure of fusarin C and talaroconvolutin B.

These conclusions show that the previously widely held view that Knoevenagel cyclisation leads to the formation of the 1,5-dydro-2*H*-pyrrol-2-one tautomer **4** prior to Diels–Alder catalysis is an over-simplification. We have shown that in vivo, Knoevenagel catalysis must be rapid enough to out-compete aldehyde shunt reduction, and that Diels–Alder reaction must also occur very rapidly before (non-enzymatic) tautomerisation and loss of chirality.

It is intriguing that the PyiE enzyme family preserve the nucleophilic active site serine at a 'nucleophilic elbow' position found in DHPON hydrolase and AntI, but lack the acid/base residues required for retro-Claisen reaction. However the acid/base residues shown by Bode and co-workers to be involved in cyclisation in AntI are conserved in PyiE and ORFZ.[31] This suggests that an alternative possible mechanism for PyiE and its homologues could involve addition of the serine nucleophile of PyiE/ORFZ to the β-ketone of **20** during reaction (Scheme 6C) to give an enzyme-bound tetrahedral intermediate, without retro Claisen cleavage because of the lack of the catalytic triad H and D residues. The formation of a tetrahedral centre at the β-position would prevent incorrect tautomerisation, providing the correct (enzyme-bound) substrate for the subsequent Diels–Alder reaction. This in-turn suggests that the hydrolase (e.g. PyiE/ORFZ) and Diels–Alderase (e.g. PyiF/ORF3) catalysts may work closely together or as a hetero-dimer, or conceivably with the PKS-NRPS itself. However, further detailed in vitro experiments will be required to test this idea, and more reliable methods for the production of soluble and active PyiE/ORFZ and PyiF/ORF3 proteins will be required to probe this hypothesis further. However synthetic and analytic progress made here sets the scene for future advances in this area.

Acknowledgements

The authors would like to thank Dr. Jennifer Senkler and Prof. Dr. Hans-Peter Braun from the Institute of Plant Genetics, Leibniz University of Hanover for protein sequence analysis by ESI

Q-TOF. Inken Hertrampf (LUH) and Franck Siacku (LUH) are thanked for technical assistance. DFG is thanked for the provision of NMR and LCMS equipment (INST 187/621-1, INST 187/686-1). V.H. was funded by DFG (CO 1328/2-1) and H.Z. was funded by the China Scholarship Council (CSC 201506200065). Open access funding enabled and organized by Projekt DEAL.

Conflict of interest

The authors declare no conflict of interest.

Keywords: biosynthesis · cytochalasan · hydrolase · Knoevenagel · pyrrolone

[1] K. Scherlach, D. Boettger, N. Remme, C. Hertweck, *Nat. Prod. Rep.* **2010**, *27*, 869 – 886.
[2] E. J. Skellam, *Nat. Prod. Rep.* **2017**, *34*, 1252 – 1263.
[3] D. C. Aldridge, J. J. Armstrong, R. N. Speake, W. B. Turner, *J. Chem. Soc. C* **1967**, 1667 – 1676.
[4] a) A. Probst, C. Tamm, *Helv. Chim. Acta* **1981**, *64*, 2065 – 2077; b) J. L. Robert, C. Tamm, *Helv. Chim. Acta* **1975**, *58*, 2501 – 2504; c) M. Binder, J. R. Kiechel, C. Tamm, *Helv. Chim. Acta* **1970**, *53*, 1797 – 1812; d) J. C. Vederas, C. Tamm, *Helv. Chim. Acta* **1976**, *59*, 558 – 566; e) H. Oikawa, Y. Murakami, A. Ichihara, *J. Chem. Soc. Perkin Trans. 1* **1992**, 2955 – 2959; f) R. Wyss, C. Tamm, J. C. Vederas, *Helv. Chim. Acta* **1980**, *63*, 1538 – 1541.
[5] a) K. Qiao, Y.-H. Chooi, Y. Tang, *Metab. Eng.* **2011**, *13*, 723 – 732; b) Y. Hu, D. Dietrich, W. Xu, A. Patel, J. A. J. Thuss, J. Wang, W.-B. Yin, K. Qiao, K. N. Houk, J. C. Vederas, Y. Tang, *Nat. Chem. Biol.* **2014**, *10*, 552 – 554; c) M. L. Nielsen, T. Isbrandt, L. M. Petersen, U. H. Mortensen, M. R. Andersen, J. B. Hoof, T. O. Larsen, *PLoS One* **2016**, *11*, e0161199; d) K. Ishiuchi, T. Nakazawa, F. Yagishita, T. Mino, H. Noguchi, K. Hotta, K. Watanabe, *J. Am. Chem. Soc.* **2013**, *135*, 7371 – 7377.
[6] a) G. Stork, Y. Nakahara, Y. Nakahara, W. J. Greenlee, *J. Am. Chem. Soc.* **1978**, *100*, 7775 – 7777; b) E. J. Thomas, J. W. F. Whitehead, *J. Chem. Soc. Perkin Trans. 1* **1989**, 507 – 518; c) M. Zaghouani, C. Kunz, L. Guedon, F. Blanchard, B. Nay, *Chem. Eur. J.* **2016**, *22*, 15257 – 15260; d) J. R. Reyes, N. Winter, L. Spessert, D. Trauner, *Angew. Chem. Int. Ed.* **2018**, *57*, 15587 – 15591; *Angew. Chem.* **2018**, *130*, 15813 – 15817.
[7] C. Wang, V. Hantke, R. J. Cox, E. J. Skellam, *Org. Lett.* **2019**, *21*, 4163 – 4167.
[8] J. Schümann, C. Hertweck, *J. Am. Chem. Soc.* **2007**, *129*, 9564 – 9565.
[9] C. Wang, K. Becker, S. Pfütze, E. Kuhnert, M. Stadler, R. J. Cox, E. J. Skellam, *Org. Lett.* **2019**, *21*, 8756 – 8760.
[10] V. Hantke, C. Wang, E. J. Skellam, R. J. Cox, *RSC Adv.* **2019**, *9*, 35797 – 35802.
[11] V. Hantke, E. J. Skellam, R. J. Cox, *Chem. Commun.* **2020**, *56*, 2925 – 2928.
[12] E. Vedejs, J. B. Campbell, Jr., R. C. Gadwood, J. D. Rodgers, K. L. Spear, Y. Watanabe, *J. Org. Chem.* **1982**, *47*, 1534 – 1546.
[13] J. Collemare, M. Pianfetti, A.-E. Houlle, D. Morin, L. Camborde, M.-J. Gagey, C. Barbisan, I. Fudal, M.-H. Lebrun, H. U. Böhnert, *New Phytol.* **2008**, *179*, 196 – 208.
[14] H. U. Böhnert, I. Fudal, W. Dioh, D. Tharreau, J.-L. Notteghem, M.-H. Lebrun, *Plant Cell* **2004**, *16*, 2499 – 2513.
[15] Z. Song, W. Bakeer, J. W. Marshall, A. A. Yakasai, R. M. Khalid, J. Collemare, E. Skellam, D. Tharreau, M.-H. Lebrun, C. M. Lazarus, A. M. Bailey, T. J. Simpson, R. J. Cox, *Chem. Sci.* **2015**, *6*, 4837 – 4845.
[16] a) R. Fujii, A. Minami, K. Gomi, H. Oikawa, *Tetrahedron Lett.* **2013**, *54*, 2999 – 3002; b) R. Fujii, T. Ugai, H. Ichinose, M. Hatakeyama, T. Kosai, K. Gomi, I. Fujii, A. Minami, H. Oikawa, *Biosci. Biotechnol. Biochem.* **2016**, *80*, 426 – 431.
[17] A. Minami, T. Ugai, T. Ozaki, H. Oikawa, *Sci. Rep.* **2020**, *10*, 13556.
[18] a) E.-M. Niehaus, K. Kleigrewe, P. Wiemann, L. Studt, C. M. K. Sieber, L. R. Connolly, M. Freitag, U. Güldener, B. Tudzynski, H.-U. Humpf, *Chem. Biol.* **2013**, *20*, 1055 – 1066; b) P. Wiemann, C.-J. Guo, J. M. Palmer, R. Sekonyela, C. C. C. Wang, N. P. Keller, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17065 – 17070; c) Y. Tsunematsu, M. Fukutomi, T. Saruwatari, H. Noguchi, K. Hotta, Y. Tang, K. Watanabe, *Angew. Chem. Int. Ed.* **2014**, *53*, 8475 – 8479; *Angew. Chem.* **2014**, *126*, 8615 – 8619; d) L. Li, M.-C. Tang, S. Tang, S. Gao, S. Soliman, L. Hang, W. Xu, T. Ye, K. Watanabe, Y. Tang, *J. Am. Chem. Soc.* **2018**, *140*, 2067 – 2071; e) B. G. Donzelli, S. B. Krasnoff, A. C. Churchill, J. D. Vandenberg, D. M. Gibson, *Curr. Genet.* **2010**, *56*, 151 – 162; f) M. Gressler, C. Zaehle, K. Scherlach, C. Hertweck, M. Brock, *Chem. Biol.* **2011**, *18*, 198 – 209.
[19] M. Nielsen, L. Albertsen, G. Lettier, J. Nielsen, U. Mortensen, *Fungal Genet. Biol.* **2006**, *43*, 54 – 64,.
[20] S. Jacob, T. Grötsch, A. J. Foster, A. Schüffler, P. H. Rieger, L. P. Sandjo, J. C. Liermann, T. Opatz, E. Thines, *Microbiology* **2017**, *163*, 541 – 553.
[21] K. A. Pahirulzaman, K. Williams, C. M. Lazarus, *Methods Enzymol.* **2012**, *517*, 241 – 260.
[22] J. Yadav, H. Chawla, S. Dev, *Tetrahedron Lett.* **1977**, *18*, 201 – 204.
[23] K. E. Pfitzner, J. G. Moffatt, *J. Am. Chem. Soc.* **1965**, *87*, 5661 – 5670.
[24] Z. Song, R. J. Cox, C. M. Lazarus, T. J. Simpson, *ChemBioChem* **2004**, *5*, 1196 – 1203; D. O. Rees, N. Bushby, R. J. Cox, J. R. Harding, T. J. Simpson, C. L. Willis, *ChemBioChem* **2007**, *8*, 46 – 50.
[25] H. Li, H. Wei, J. Hu, E. Lacey, A. N. Sobolev, K. A. Stubbs, P. S. Solomon, Y.-H. Chooi, *ACS Chem. Biol.* **2020**, *15*, 226 – 233.
[26] N. Lenfant, T. Hotelier, E. Velluet, Y. Bourne, P. Marchot, A. Chatonnet, *Nucleic Acids Res.* **2013**, *41*, D423 – D429.
[27] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. Sternberg, *Nat. Protoc.* **2015**, *10*, 845 – 858.
[28] T. Schwede, J. Kopp, N. Guex, M. C. Peitsch, *Nucleic Acids Res.* **2003**, *31*, 3381 – 3385.
[29] C. Schleberger, P. Sachelaru, R. Brandsch, G. E. Schulz, *J. Mol. Biol.* **2007**, *367*, 409 – 418.
[30] P. Sachelaru, E. Schiltz, G. L. Igloi, R. Brandsch, *J. Bacteriol.* **2005**, *187*, 8516 – 8519.
[31] Q. Zhou, A. Bräuer, H. Adihou, M. Schmalhofer, P. Saura, G. L. C. Grammbitter, V. R. I. Kaila, M. Groll, H. B. Bode, *Chem. Sci.* **2019**, *10*, 6341 – 6349.
[32] J. S. Papadopoulos, R. Agarwala, *Bioinformatics* **2007**, *23*, 1073 – 1079.
[33] J. Hu, H. Li, Y.-H. Chooi, *J. Am. Chem. Soc.* **2019**, *141*, 8068 – 8072.
[34] G. Dunn, M. G. Montgomery, F. Mohammed, A. Coker, J. B. Cooper, T. Robertson, J.-L. Garcia, T. D. H. Bugg, S. P. Wood, *J. Mol. Biol.* **2005**, *346*, 253 – 265.
[35] S. Suzuki, T. Hosoe, K. Nozawa, K. Kawai, T. Yaguchi, S. Udagawa, *J. Nat. Prod.* **2000**, *63*, 768 – 772.

Manuscript received: October 4, 2020
Accepted manuscript online: November 4, 2020
Version of record online: January 14, 2021

