Preparation of pyrrolo[2,3-b]indoles carrying a β-configured reverse C3-dimethylallyl moiety by using a recombinant prenyltransferase CdpC3PT†

Wen-Bing Yin,‡§[a] Xia Yu,‡[a] Xiu-Lan Xie[b] and Shu-Ming Li*[a]

Received 14th January 2010, Accepted 19th February 2010
First published as an Advance Article on the web 22nd March 2010
DOI: 10.1039/c000587h

Six β-configured reversely C3-prenylated pyrrolo[2,3-b]indoles were successfully prepared by using a recombinant prenyltransferase from *Neosartorya fischeri*. For this purpose, the putative prenyltransferase gene *NFIA_074280* (termed herewith *cdpC3PT*) was cloned into pQE60 and overexpressed in *Escherichia coli*. The overproduced His_6-CdpC3PT was purified to near homogeneity and incubated with five cyclic tryptophan-containing dipeptides in the presence of dimethylallyl diphosphate (DMAPP). All of the substrates were accepted by CdpC3PT and converted to reversely C3-prenylated pyrrolo[2,3-b]indoles. Using cyclo-L-Trp-L-Trp as substrate, both mono- and diprenylated derivatives were obtained. The structures of the enzymatic products were confirmed by HR-ESI-MS, ^{1}H- and ^{13}C-NMR analyses as well as by long-range ^{1}H-^{13}C connectivities in heteronuclear multiple-bond correlation (HMBC) spectra after preparative isolation. ^{1}H-^{1}H spatial correlations in nuclear overhauser effect spectroscopy (NOESY) were used for determination of absolute configuration. The K_M values were determined at about 1.5 mM for DMAPP and in the range from 0.22 to 5.5 mM for cyclic dipeptides. The turnover number k_{cat} were found in the range of 0.023 to 0.098 s^{-1} and specificity constants k_{cat}/K_M from 14.2 to 122.7 M^{-1} s^{-1}. In contrast to the products of AnaPT bearing α-configured C3-dimethylallyl residues, the C3-prenyl moieties in the products of CdpC3PT have a β-configuration. Discovery and characterisation of CdpC3PT expand the usage of the chemoenzymatic approach for stereospecific synthesis of C3-prenylated derivatives.

Introduction

A number of naturally occurring indoline alkaloids carry a reverse prenyl moiety at position C3 of the indoline ring and a five-membered ring system between the indoline and the diketopiperazine ring (Fig. 1). These compounds are mainly found in the genera *Penicillium* and *Aspergillus* of ascomycota.[1] All of the known natural products from this group have a *cis*-configuration between H-2 and C3-prenyl moiety. Both α- and β-configured C3-dimethylallyl residues have been identified (Fig. 1). 5-*N*-acetylardeemin,[2] aszonalenin and epiaszonalenin as well as their acetylated form[1,3,4] belong to the first subgroup. Examples of β-configured derivatives are roquefortines C and D,[5–7] fructigenines A and B (verrucofortine),[8,9] rugulosuvines A and B,[10] brevicompanines A, B and C,[11,12] as well as amauromine.[13] Both α- and β-configured C3-dimethylallyl residues are found in the structure of epiamauromine derived from two tryptophan molecules.[14] Until now, more β-configured derivatives have been isolated from fungal strains than the α-configured ones. C3-prenylated indolines have been reported to show diverse biological activities.[15]

By using two indole prenyltransferases AnaPT and CdpNPT, four aszonalenin stereoisomers were synthesised successfully from (R)- and (S)-benzodiazepinedione.[16] Seven C3-prenylated pyrrolo[2,3-b]indoles (**2a–2f** and **4**) have been obtained by enzymatic conversion of six tryptophan-containing cyclic dipeptides (**1a–1f**) with AnaPT (Scheme 1).[15] By using HPLC, only one stereoisomer each was detected from the incubation mixtures with the mentioned prenyltransferases. MS, CD and NMR analyses, including nuclear overhauser effect spectroscopy (NOESY), showed clearly that all of the isolated products of AnaPT reactions carried an α-configured C3-dimethylallyl moiety. To expand the usage of prenyltransferases for chemoenzymatic synthesis, we identified an additional **c**yclic **d**ipe**p**tide **C3-p**renyl**t**ransferase (CdpC3PT) gene from the genome sequence of *Neosartorya fischeri*. In contrast to those of AnaPT, the products of CdpC3PT bear β-configured C3-dimethylallyl residues. Here, we report the gene cloning, protein overproduction and biochemical investigation of CdpC3PT.

Results and discussion

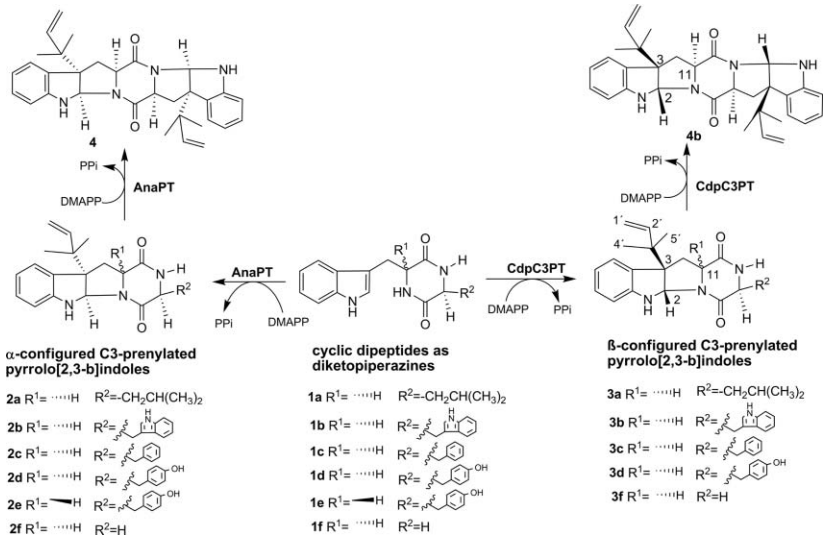
In the course of our investigation on indole prenyltransferases, one putative gene *NFIA_074280* from the genome sequence of *Neosartorya fischeri* NRRL181 raised our attention. *NFIA_074280* spans bp 569563-570866 of DS027696.1 in GenBank. Its deduced product, EAW17508, comprises 423 amino acids and showed significant sequence similarity to known indole prenyltransferases.[17,18] *NFIA_074280* and a putative

[a]*Philipps-Universität Marburg, Institut für Pharmazeutische Biologie, Deutschhausstrasse 17A, D-35037 Marburg, Germany. E-mail: shuming.li@Staff.uni-Marburg.de; Fax: +49-6421-2825365; Tel: +49-6421-2822461*
[b]*Philipps-Universität Marburg, Fachbereich Chemie, Hans-Meerwein-Strasse, 35032 Marburg, Germany*

† Electronic supplementary information (ESI) available: Spectra of compounds. See DOI: 10.1039/c000587h

‡ These authors contributed equally to this work

§ Present address: University of Wisconsin-Madison, Medical Microbiology and Immunology, 3455 Microbial Sciences Building, 1550 Linden Drive, Madison WI 53706, USA.



Scheme 1 Preparation of C3-prenylated pyrrolo[2,3-b]indoles by using recombinant prenyltransferases.

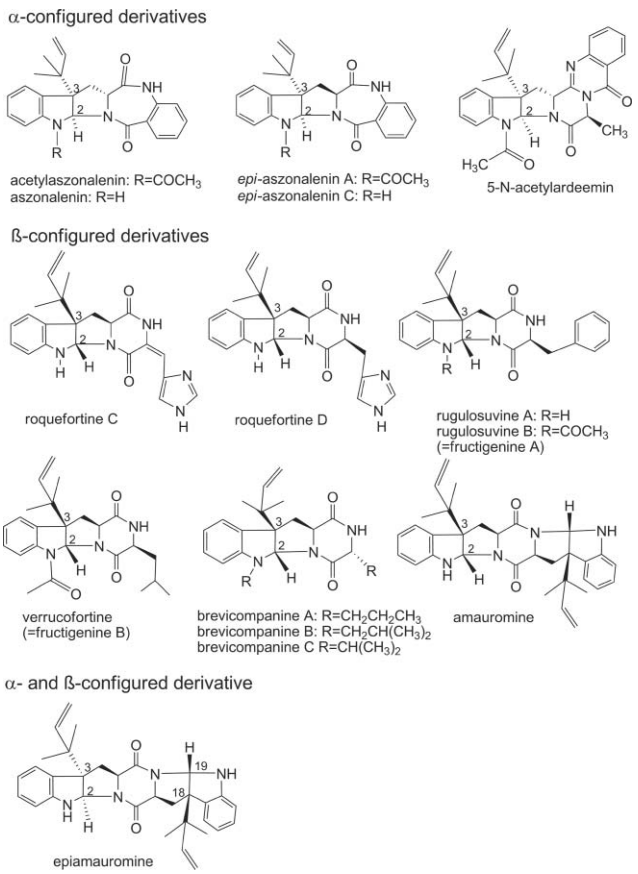


Fig. 1 Examples of C3-prenylated pyrrolo[2,3-b]indoles.

non-ribosomal peptide synthetase (NRPS) gene *NFIA_074300* are separated from each other by only 3.9 kb (Fig. 2) and could therefore belong to a same gene cluster for a secondary metabolite. The deduced product of *NFIA_074300*, EAW17510, showed a similar domain structure as the cyclic dipeptide synthetase FtmPT1 (brevianamide F synthetase) involved in the biosynthesis of fumitremorgins,[19] *i.e.* ATCATC (A: adenylation; T: thiolation and C: condensation), and could therefore be responsible for the formation of a cyclic dipeptide. Thus, we speculated that EAW17508 would function as a cyclic dipeptide prenyltransferase. By sequence comparison and analysis, it was neither possible to propose the natural substrate for the putative prenyltransferase, nor its prenylation pattern, *i.e.* regular or reverse, or prenylation position at the indole ring. The end product encoded by this cluster is also unknown. One putative cytochrome P450 oxidoreductase gene *NFIA_074290* is located between the putative prenyltransferase gene *NFIA_074280* and the putative NRPS gene *NFIA_074300*. Orthologous genes, *i.e. pc21g15430*, *pc21g15470* and *pc21g15480*, with sequence identities of 67, 60 and 59% on the amino acid level to *NFIA_074280*, *NFIA_074290* and *NFIA_074300*, respectively, have been identified in the genome sequence of *Penicillium chrysogenum* (Fig. 2).[20] Three additional genes *pc21g15440*, *pc21g15450* and *pc21g15460* are located between *pc21g15430* and *pc21g15470*, indicating that the product encoded by the cluster from *P. chrysogenum* is very likely different from that in *N. fischeri*.

To prove the function of EAW17508, *i.e.* CdpC3PT, the coding region of *NFIA_074280* was amplified by homologous recombinant DNA technique through three rounds PCR.[1,21] The PCR fragment consisting exclusively of the two exons was cloned *via* pGEM-T easy vector into pQE60 to create the expression construct pWY25 (see Experimental section). Soluble proteins were obtained from transformants of *E. coli* cells harbouring pWY25 after 6 h induction by 0.5 mM of IPTG at 37 °C. His_6-CdpC3PT was purified with Ni-NTA agarose to near homogeneity as judged by SDS-PAGE (Fig. 3) and a protein yield of 7 mg of purified His_6-tagged CdpC3PT per litre of cultures was obtained. A major protein band with migration near the 45 kDa size marker was observed, which corresponded well to the calculated value of 47.6 kDa for His_6-CdpC3PT.

In analogy to AnaPT,[15] the purified CdpC3PT (1.3 μM) was incubated with five tryptophan-containing cyclic dipeptides (**1a–1d** and **1f**) (1 mM) in the presence of DMAPP (2 mM) for different time. HPLC analysis of the incubation mixtures (Fig. 4) showed that all of the five substrates were accepted by

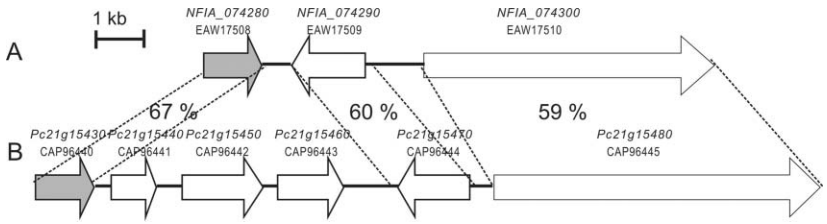


Fig. 2 Putative clusters containing *cdpC3PT* and its orthologue from *Neosartorya fischeri* (A) and *Penicillium chrysogenum* (B).

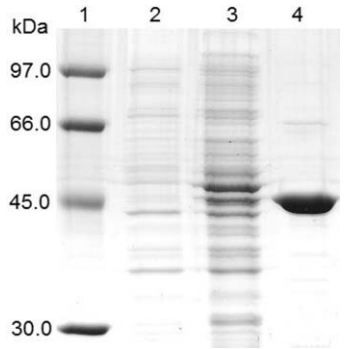


Fig. 3 Purification of CdpC3PT as a His_6-tagged protein. The 12% (w/v) SDS polyacrylamide gel was stained with Coomassie brilliant blue G-250. Lanes 1: molecular mass standard; 2: total protein before induction; 3: total protein after induction; 4: purified His_6- CdpC3PT.

CdpC3PT. In all the cases, product formation has been clearly observed already after incubation for 1 h (Fig. 4) and was dependent on the presence of active protein and DMAPP (data not shown). Increasing of product formation with extension of incubation time has also been observed. After incubation for 24 h, conversion rates of 31, 12, 6, 30 and 37% were determined for cyclo-L-Trp-L-Leu (**1a**), cyclo-L-Trp-L-Trp (**1b**), cyclo-L-Trp-L-Phe (**1c**), cyclo-L-Trp-L-Tyr (**1d**) and cyclo-L-Trp-Gly (**1f**), respectively, which were calculated from the ^{1}H-NMR signal intensity of the enzymatic products and remained respective substrate in the reaction mixtures (see Experimental section). One product peak each was detected for **1a**, **1c**, **1d** and **1f** (Fig. 4). In the HPLC chromatogram of the incubation mixture of **1b** (Fig. 4), two product peaks have been observed.

For structure elucidation, the enzymatic products were isolated on a preparative scale and substances in the range of 0.2–1 mg were subjected to high resolution ESI-MS and NMR analyses. HR-ESI-MS (Table 1) confirmed the presence of one dimethylallyl moiety in the structures of **3a–3d** and **3f** by detection of $[M+1]^+$ or $[M+Na]^+$ ions, which are 68 daltons larger than those of the respective substrates. The assignments of the NMR data listed in Table 2 were

Table 1 HR-ESI-MS data of the enzymatic products

Comp.	Chemical Formula	HR-ESI-MS data		Deviation
		Calculated	Measured	ppm
3a	$C_{22}H_{29}N_3O_2$	368.2338 $[M+1]^+$	368.2362	6.5
3b	$C_{27}H_{28}N_4O_2$	441.2291 $[M+1]^+$	441.2308	3.9
3c	$C_{25}H_{27}N_3O_2$	402.2182 $[M+1]^+$	402.2216	8.4
3d	$C_{25}H_{27}N_3O_3$	418.2131 $[M+1]^+$	418.2112	4.5
3f	$C_{18}H_{21}N_3O_2$	334.1531 $[M+Na]^+$	334.1540	2.7
4b	$C_{32}H_{36}N_4O_2$	509.2917 $[M+1]^+$	509.2920	0.6

verified by analysing the ^{1}H-^{13}C heteronuclear single-quantum correlation (HSQC) and heteronuclear multiple bond correlation (HMBC) spectra (Figures S1–S6, ESI†). Comparison of the ^{1}H-NMR data of the isolated products **3a–3d** and **3f** (Table 2) with those of the respective substrates (data not shown) revealed clearly the presence of signals for a reverse dimethylallyl moiety in the spectra of the isolated compounds at 5.08–5.14 (d or dd) for H-1′, 5.95–5.98 (dd) for H-2′, 1.00–1.01 ppm (s) for H-4′ and 1.11–1.13 (s) for H-5′, respectively. The signals of H-2 were strongly upfield shifted from approximately 7.0–7.2 in the spectra of the substrates to 5.50–5.56 ppm in the spectra of the enzymatic products (Figures S1–S6, ESI†). This indicated that the prenylation has very likely taken place at position C3 of the cyclic dipeptides and pyrrolo[2,3-b]indole derivatives were formed during the enzymatic reactions. Cross peaks in the HSQC spectra of **3a–3d** and **3f** revealed that the singlets of H-2 at 5.5 ppm correlated with the signals of C-2 at 77.5–77.6 ppm, proving the disappearance of the double bonds between C-2 and C-3 of the indole rings in the structures of **3a–3d** and **3f**. HMBC spectra showed clear connectivity from H-2 to C-11, which is conducted by the formation of a chemical bond between C-2 and N-12 of the diketopiperazine ring. Furthermore, HMBC spectra showed strong connectivities from H2 to C3′, H4′ to C3, and H5′ to C3. These results proved that the structures of **3a–3d** and **3f** are indeed C3-prenylated indolines with a fused five-membered ring between the indoline and diketopiperazine ring. This means that CdpC3PT catalysed, in analogy to AnaPT,[15] the reverse C3-prenylation of tryptophan-containing cyclic dipeptides and meanwhile the formation of a pyrrolo[2,3-b]indole structure.

The ^{1}H-NMR spectra and the assigned chemical shifts of **3a–3d** and **3f** (Table 2) are similar to those of **2a–2d** and **2f** (Scheme 1) reported previously.[15] Without considering the NOESY spectra, however, clear differences of chemical shifts were found for H-2 and H-10_{syn}. The chemical shifts of H-2 in **3a–3d** and **3f** were detected at approximately 5.5 ppm, while those of **2a–2d** and **2f** with α-configured C3-prenyl moieties at approximately 5.4 ppm.[15] The chemical shifts of H-10_{syn} in **3a–3d** and **3f** are found from 2.52 to 2.56 ppm, about 0.3 ppm upfield shifted in comparison to those of **2a–2d** and **2f** at 2.79 to 2.84 ppm.[15] These data indicated that the structures of **3a–3d** and **3f** differ from those of **2a–2d** and **2f** very likely by different configurations at C-2 and C-3 of the indoline rings.

Unambiguous proof of the stereochemistry was provided by NOESY experiments for **3a–3d** and **3f** (Table 3). Strong NOE correlations between H-2 and the protons of the prenyl moiety, *i.e.* H-2′, H-4′and H-5′, proved the *cis*-configuration between H-2 and C3-prenyl moieties of the indoline rings.

For compounds **3a–3d** and **3f**, only very weak NOE correlation was observed between H-2 and H-11. In the cases of **2a–2d** and **2f**, the NOE correlation for these protons was determined as

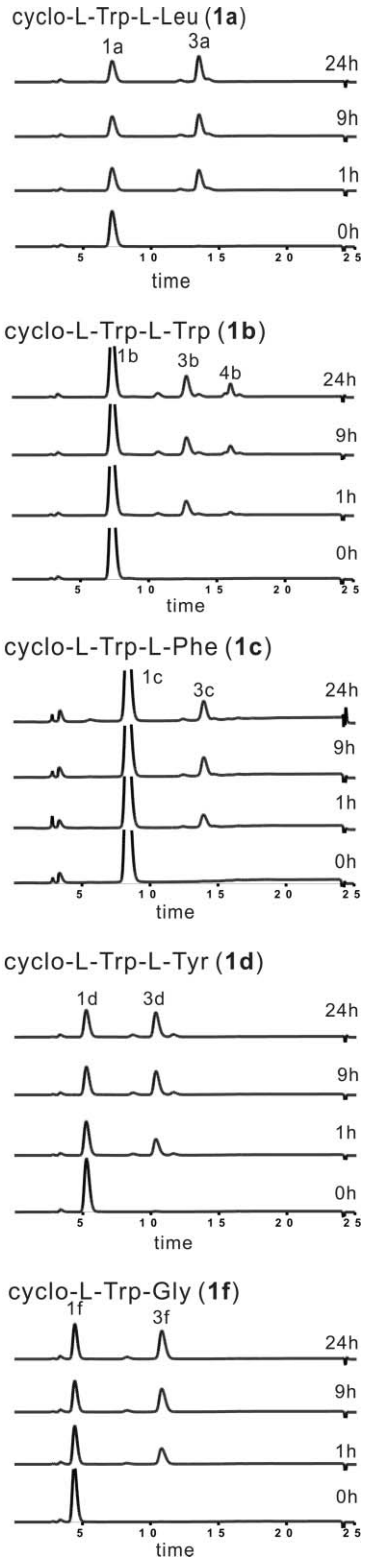


Fig. 4 HPLC chromatograms of incubation mixtures of five tryptophan–containing cyclic dipeptides with recombinant CdpC3PT. Detection was carried with a diode array detector and illustrated for absorption at 254 nm.

medium. Moreover, H-10$_{anti}$ in **3a–3d** and **3f** instead of H-10$_{syn}$ in **2a–2d** and **2f**,[15] showed medium or strong NOE correlations with protons of the prenyl moiety (Table 3). H-10$_{syn}$ in **3a–3d** and **3f** as well as **2a–2d** and **2f** showed strong correlation with H-11. The correlations observed for **2a–2d** and **2f** between H-11 and those of the prenyl group were not detected in **3a–3d** and **3f**. From the results of NOE correlations, we concluded that H-2 and C3-prenyl moiety must be substituted on the opposite side to H-11. This demonstrated that, in contrast to **2a–2d** and **2f**, **3a–3d** and **3f** carry β-configured C3-dimethylallyl moieties as illustrated in Table 2 and Scheme 1. The steric hindrance of the β-configured prenyl residue makes the indoline ring bend in opposite direction and causes a spatial proximity between H4 and H10$_{syn}$. This is confirmed by the observation of strong NOE correlation between these two protons (Table 3).

The NMR data of **3a** and **3c** corresponded also well to those reported for brevicompanine B[22] and rugulosuvine A,[23] respectively.

The second product peak **4b** of cyclo-L-Trp-L-Trp (**1b**) (Fig. 4) was also isolated by repeated chromatography and subjected to MS analysis. The positive HR-ESI-MS of **4b** showed an ion at *m/z* 509.2920 (Table 1), which could be interpreted as $[M+1]^+$ of a diprenylated derivative with a molecular mass 136 daltons larger than that of the substrate.

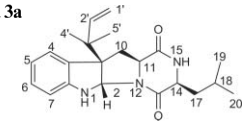
The ^{1}H-NMR spectrum of **4b** showed signals of two identical tryptophanyl moieties, which are reversely prenylated at position C3 of the indoline rings. C3-diprenylated derivatives of cyclo-L-Trp-L-Trp containing two pyrrolo[2,3-b]indole systems have been reported, *e.g.* amauromine (Fig. 1) from *Amauroascus sp*[13] and epiamauromine (Fig. 1) from *Aspergillus ochraceus*.[14] ^{1}H-NMR data of **4b** (Table 2) corresponded perfectly to those of amauromine,[13] indicating that **4b** has an identical structure as amauromine or both compounds are enantiomers. The absolute configuration of **4b** was not determined in this study. However, it can be expected that the symmetrical **4b** was formed by a second prenylation and cyclisation of **3b** under the catalysis of CdpC3PT (Scheme 1). Therefore, both tryptophanyl moieties must have the same configuration as in **3b**. Therefore, **4b** and amauromine have same configuration as illustrated in Fig. 1 and Scheme 1.

The results provided in this study proved that both AnaPT and CdpC3PT catalyse the formation of indoline derivatives carrying fused five-membered rings with a *cis*-configuration. However, they introduced the ring system from opposite sides. Similar as observed with AnaPT,[15,16] the reaction catalysed by CdpC3PT includes at least three steps, *i.e.* attachment of a reverse prenyl moiety to C3 of the indole ring, breaking of the double bond between C2 and C3 and the formation of a C–N bond between C2 and N12. The detailed mechanism of this reaction is unknown. As proposed for AnaPT,[15] it can be speculated that attacking of dimethylallyl cation *via* its C3′ to C3 of the indole ring of the cyclic dipeptides would result in formation of an intermediate with a positive charge at C2. Formation of a C–N bond between C2 and N12 would lead to an intermediate with positive charge at N12, which is to be converted to the enzymatic products by releasing of a proton.

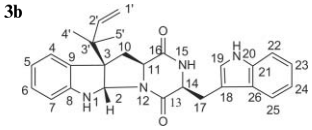
After structure elucidation of the enzymatic products, CdpC3PT was characterised biochemically. For determination of the dependence of the CdpC3PT activity on metal ions, incubations of **1a** with DMAPP were carried out in the presence of different metal ions at a final concentration of 5 mM. Incubations with EDTA at a final concentration of 5 mM and without additives were used as controls. Our results showed that product

Table 2 ^{1}H-NMR and ^{13}C-NMR data of enzymatic products ($CDCl_3$)

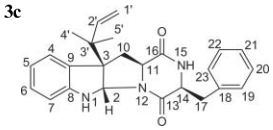
Compound **3a**



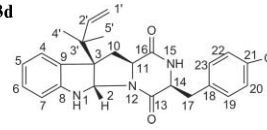
3b



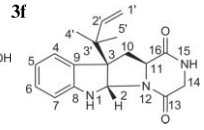
3c



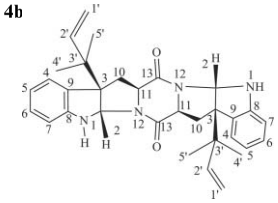
3d



3f



4b



Position	δ_C	δ_H, multi., J in Hz	δ_C	δ_H, multi., J in Hz	δ_C	δ_H, multi., J in Hz	δ_C	δ_H, multi., J in Hz	δ_C	δ_H, multi., J in Hz	δ_H, multi., J in Hz
2	77.6	5.50, s	77.5	5.55, s	77.6	5.54, s	77.5	5.54, s	77.5	5.56, s	5.44, s
3	61.3	—	61.3	—	61.4	—	61.4	—	61.2	—	—
4	124.9	7.16, d, 7.4	124.9	7.14[a]	124.9	7.14, d, 7.4	125.0	7.14, d, 7.4	124.9	7.16, d, 7.2	7.09, dd, 7.6, 1.1
5	118.7	6.76, t, 7.5	118.8	6.75, td, 7.5, 0.8	118.8	6.76, td, 7.5, 0.9	118.8	6.76, t, 7.7	118.8	6.77, t, 7.5	6.70, td, 7.4, 1.0
6	128.7	7.10, t, 7.6	128.8	7.11[a]	128.8	7.10, td, 7.6, 1.0	128.8	7.11, t, 7.5	128.8	7.11, t, 8.1	7.02, td, 7.6, 1.2
7	109.1	6.58, d, 7.8	109.0	6.61, d, 7.8	109.0	6.60, d, 7.6	109.0	6.60, d, 7.9	109.0	6.60, d, 7.9	6.48, d, 7.9
8	149.8	—	149.7	—	149.7	—	149.6	—	149.7	—	—
9	128.9	—	128.7	—	128.7	—	128.6	—	128.6	—	—
10_{syn}[b]	35.7	2.55, dd, 12.7, 6.4	35.7	2.52, dd, 12.6, 6.3	35.9	2.52, dd, 12.6, 6.2	35.9	2.52, dd, 12.6, 6.1	36.2	2.56, dd, 12.6, 6.2	2.50, dd, 12.8, 6.6
10_{anti}[b]		2.47, t, 11.9		2.42, t, 11.9		2.40, t, 11.9		2.38, t, 11.9		2.45, t, 12.0	2.45, t, 11.9
11	58.6	3.95, m	58.8	3.92, ddd, 11.1,6.2, 1.6	58.7	3.93, ddd, 11.2, 6.2, 1.6	58.7	3.93,ddd,10.7,6.1, 2.0	58.0	3.97, ddd, 11.0, 5.9, 2.1	3.88, dd, 10.5, 6.4
13	166.3	—	165.8	—	165.3	—	165.1	—	163.4	—	—
14	53.3	3.95,m	54.4	4.31, ddd, 10.9, 3.6, 1.8	55.9	4.21, ddd, 10.4, 3.3, 1.5	56.0	4.15, dd, 8.9, 2.2	46.3	4.05, dd, 17.0, 1.9	—
	—	—	—	—	—	—	—	—	—	3.90, dd, 17.1, 4.0	—
15	—	5.67, s	—	5.69, s	—	5.54, s	—	5.64, s	—	6.24, br.s	—
16	169.3	—	168.8	—	168.8	—	168.8	—	169.5	—	—
17	38.6	2.02, ddd, 14.2, 10.2, 3.8	26.8	3.74, ddd, 14.7, 3.3, 0.8	36.7	3.59, dd, 14.4, 3.5	35.9	3.47, dd, 16.9, 2.5	—	—	—
		1.57, ddd, 14.5, 9.6, 5.0		2.98, dd, 15.0, 11.0		2.79, dd, 14.5, 10.6		2.76, dd, 14.3, 10.3		—	—
18	24.4	1.68, m	109.5	—	135.3	—	127.1	—	—	—	—
19	23.0	0.99, d, 6.6	123.1	7.11, s	128.8	7.19, d, 7.1	130.2	7.06, d, 8.4	—	—	—
20	20.8	0.92, d, 6.6	—	8.13, s	129.1	7.34, t, 7.3	116.0	6.79, d, 8.4	—	—	—
21	—	—	136.3	—	127.5	7.29, t, 7.4	154.8	—	—	—	—
22	—	—	111.3	7.38, d, 8.2	129.1	7.34, t, 7.3	116.0	6.79, d, 8.4	—	—	—
23	—	—	122.7	7.22, td, 7.7, 0.9	128.8	7.19, d, 7.1	130.2	7.06, d, 8.4	—	—	—
24	—	—	119.9	7.13[a]	—	—	—	—	—	—	—
25	—	—	118.2	7.55, d, 8.1	—	—	—	—	—	—	—
26	—	—	126.3	—	—	—	—	—	—	—	—
1′	114.4	5.12, d, 10.8	114.3	5.12, dd, 10.8, 1.0	114.3	5.13, dd, 10.8, 1.0	114.3	5.13, d, 10.8	114.4	5.14, dd, 10.8, 0.7	5.13, dd, 10.9, 1.1
		5.08, d, 17.3		5.08, dd, 17.4, 1.0		5.08, dd, 17.4, 0.9		5.08, d, 17.6		5.09, dd, 17.4, 0.6	5.08, dd, 17.4, 1.1
2′	143.3	5.97, dd, 17.3, 10.8	143.3	5.97, dd, 17.4, 10.8	143.3	5.96, dd, 17.3, 10.8	143.3	5.95, dd, 17.5, 10.8	143.2	5.98, dd, 17.4, 10.8	5.99, dd, 17.4, 10.8
3′	40.5	—	40.6	—	40.6	—	40.6	—	40.7	—	—
4′	22.7	1.01, s	22.6	1.01, s	22.6	1.01, s	22.7	1.00, s	22.7	1.01, s	1.01, s
5′	22.2	1.12, s	22.2	1.11, s	22.2	1.11, s	22.2	1.11, s	22.3	1.13, s	1.11, s

[a] Overlapping signals; [b] H-10_{syn} has a *cis*-configuration to H-11, H-10_{anti} a *trans*-configuration.

Table 3 NOE results of C3-prenylated pyrrolo[2,3-b]indoles (with an exception for **4b**)

Protons	Strength
H-2 to H-1′	Weak
H-2 to H-2′	Strong
H-2 to H-4′	Strong
H-2 to H-5′	Strong
H-2 to H-11	Weak
H-10$_{anti}$ to H-1′	Weak
H-10$_{anti}$ to H-2′	Medium
H-10$_{anti}$ to H-4′	Strong
H-10$_{anti}$ to H-5′	Strong
H-10$_{anti}$ to H-11	Medium
H-10$_{syn}$ to H-11	Strong
H-10$_{syn}$ to H-4	Strong
H-11 to H-2′	Not observed
H-11 to H-4′	Not observed
H-11 to H-5′	Not observed

Table 4 Preliminary parameters of the tested substrates[a]

Substrate	K_M/mM	k_{cat}/s^{-1}	k_{cat}/K_M/M^{-1} s^{-1}
1a	2.1	0.085	40.5
1b	0.35	0.023	65.7
1c	0.22	0.027	122.7
1d	1.5	0.06	40.0
1f	5.5	0.078	14.2
DMAPP[b]	1.4	0.098	70.0
DMAPP[c]	1.6	0.035	21.9

[a] Due to low solubility, the aromatic substrates were only tested up to 1 mM (**1b–1d**) or 5 mM (**1a** and **1f**). [b] **1a** as aromatic substrate. [c] **1b** as aromatic substrate.

formation was independent of the presence of metal ions. Even in the presence of the chelating agent EDTA, no decreasing of the enzyme activity was detected, in comparison to that of incubation without additives. This finding corresponded well to the behaviour of other known prenyltransferases.[17,18] The enhancing effect of Ca^{2+} on the activity of CdpC3PT seems stronger than other prenyltransferases. Addition of Ca^{2+}, Mg^{2+} and Mn^{2+} to the reaction mixtures increased the enzyme activity to 400, 240 and 220% of that without additives, respectively.

For comparison of the behaviour of CdpC3PT towards the cyclic dipeptides, kinetic parameters were determined for DMAPP and all of the five aromatic substrates by incubation with CdpC3PT at 37 °C for 30 min. For this purpose, the dependence of the product formation on incubation time had been proven and found to be linear up to 30 min for **1a**, nearly linear up to 30 min for **1b** and 45 min for **1c**, **1d** and **1f** (see please Figure S7, ESI†). Michaelis–Menten constants (K_M) as well as the turnover numbers (k_{cat}) were determined by Hanes-Woolf analysis and are given in Table 4. The obtained values were also confirmed by Lineweaver–Burk and Eadie-Hofstee analyses.

Using **1a** and **1b** as aromatic substrates, comparable K_M values for DMAPP were determined at 1.4 and 1.6 mM, respectively, which are much higher than the K_M values of other cyclic dipeptide prenyltransferases for DMAPP, *e.g.* FtmPT1 at 56 μM[24] and AnaPT at 156 μM.[1] However, the natural substrates were used for determination of kinetic parameters of FtmPT1 and AnaPT. In the case of CdpC3PT, the natural substrate is still unknown. **1b** and **1c** were found to have similar K_M values of 0.35 and 0.22 mM and turnover numbers of 0.023 and 0.027 s^{-1}, respectively (Table 4). **1a**, **1d** and **1f** with K_M values of 2.1, 1.5 and 5.5 mM, respectively, showed much lower affinity towards CdpC3PT than **1b** and **1c**. Interestingly, the turnover numbers of these three substrates were two- to three-fold of those of **1b** and **1c**. Consequently, CdpC3PT showed similar catalytic efficiency towards **1a**, **1b** and **1d** with specificity constants k_{cat}/K_M of 40.5, 65.7 and 40 M^{-1} s^{-1}, respectively. Using **1c** as substrate, CdpC3PT was found to have the best catalytic efficiency with a specificity constant of 122.7 M^{-1} s^{-1}. This value is however only 1.9% of that of AnaPT at 6522 M^{-1} s^{-1} determined by using its natural substrates DMAPP and (R)-benzodiazepinedinone.[1]

Conclusions

In this study, we described the cloning and biochemical investigation of a new indole prenyltransferase CdpC3PT, which catalysed the reverse prenylation of cyclic dipeptides at position C3 of the indole ring. Similar to AnaPT from the biosynthetic gene cluster of acetylaszonalenin[1] from the same fungus, *i.e.* *N. fischeri* NRRL181, CdpC3PT showed also broad substrate specificity and accepted all of the five tryptophan-containing cyclic dipeptides tested. Both AnaPT and CdpC3PT catalysed the formation of a five-membered ring between the original indole and diketopiperazine rings with a *cis*-configuration between H-2 and C3-dimethylallyl moiety. However, they introduce the prenyl moiety from different sides, *i.e.* AnaPT from behind and CdpC3PT from the front, so that two compounds with different stereochemistry at position C2 and C3 could be obtained from one substrate. They are therefore complement to each other regarding the prenylation and expand their potential to be used for chemoenzymatic synthesis.

Experimental section

Chemicals

Dimethylallyl diphosphate was prepared according to the method described for geranyl diphosphate by Woodside.[25] Cyclo-L-Trp-L-Leu (**1a**), cyclo-L-Trp-L-Trp (**1b**), cyclo-L-Trp-L-Phe (**1c**), cyclo-L-Trp-L-Tyr (**1d**) and cyclo-L-Trp-Gly (**1f**) were purchased from Bachem (Bubendorf, Switzerland).

Bacterial strains, plasmids and cultural conditions

pGEMT easy vector and pQE60 were obtained from Promega (Mannheim, Germany) and Qiagen (Hilden, Germany), respectively.

Escherichia coli XL1 Blue MRF′ (Stratagene) was used for cloning and over expression experiments and grown in liquid or on solid Luria-Bertani medium with 1.5% (w/v) agar at 37° C.[26] Carbenicillin (50 μg mL^{-1}) was used for selection of recombinant *E. coli* strains.

Neosartorya fischeri NRRL181 was kindly provided by ARS Culture Collection (Peoria, Illinois USA).

Cultivation of *N. fischeri* and DNA isolation

For DNA isolation, mycelia of *N. fischeri* from plates were inoculated into 300 mL Erlenmeyer flask containing 100 mL YES

media consisting of yeast extract (0.6% (w/v)), sucrose 0.2% (w/v) (pH 5.8) and cultivated at 30 °C and 170 rpm for 48 h. DNA isolation from *N. fischeri* was carried out according to the protocol described by Ausubel *et al.*[27]

DNA isolation, PCR amplification and gene cloning

Standard procedures for DNA isolation and manipulation were performed as described.[26]

PCR amplification was carried out on an iCycler from BioRad (Munich, Germany). The entire coding sequence of *cdpC3PT* was obtained after three rounds PCR amplification by using genomic DNA as template and Expand High Fidelity Kit (Roche Diagnostics GmbH, Mannheim, Germany). The first round PCR is used for amplification of the two exons. The primers for the first exon were CdpC3PT_a1 (5′-TCGTCTGATAAACCTCATCCT-3′) at the 5′-end and CdpC3PT_a2 (5′- **TCTGTGGCAATAGC-CAAGTCCTCGCCAGGG**TAATATGATTG CAGATTCTCC-3′) at the 3′-end. The primers for the second exon were CdpC3PT_b1 (5′- **GGGTCTACCCGGAGAATCTGCAAT-CATATTA**CCCTGGCGAG GACTTGGCT-3′) at the 5′-end and CdpC3PT_b2 (5′- AACAGGAAGGCACTAATAAGC-3′) at the 3′-end. Bold letters represent overlapping region of the two exons. The PCR products of the first and second exon were mixed in a molar ratio of 1 : 1 and used as template for a second round of PCR to get a fragment consisting of the two exons with help of the overlapping region. A third round PCR was then carried out by using a nested primer pair and the PCR product from the second round as template. The nested primers are CdpC3PT_for (5′-**ATTCC**ATG**G**CAGTGTCGTCGACCG-3′) at the 5′-end and CdpC3PT_rev (5′-**GAAGATCT**GTGGTAGTACATGGTCA-3′) at the 3′-end of the gene. Bold letters represent mutations inserted in comparison to the original genome sequence to give the underlined restriction sites NcoI located in the start codon in CdpC3PT_for and BglII located in the predicted stop codon in CdpC3PT_rev, respectively. A PCR fragment of 1286 bp containing the entire coding sequence of *cdpC3PT* could be amplified after the third round PCR. The PCR fragment was cloned into pGEMT easy vector resulting in plasmid pWY24, which was subsequently sequenced (Eurofins MWG Operon, Ebersberg, Germany) to confirm the sequence. To create the expression vector pWY25, pWY24 was digested with NcoI and BglII and the resulted NcoI-BglII fragment of 1278 bp was ligated into pQE60, which had been digested with the same enzymes, previously.

Overproduction and purification of His_6-CdpC3PT

For *cdpC3PT* expression, *E. coli* XL1 Blue MRF′ cells harbouring the plasmid pWY25 were cultivated in 300 mL Erlenmeyer flasks containing 100 mL liquid Luria-Bertani medium supplemented with carbenicillin (50 μg mL^{-1}) and grown at 37 °C to an absorption at 600 nm of 0.6. For induction, isopropyl thiogalactoside (IPTG) was added to a final concentration of 0.5 mM and the cells were cultivated for further 6 h at 37 °C before harvest. The bacterial cultures were centrifuged and the pellets were resuspended in lysis buffer (10 mM imidazole, 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0) at 2–5 mL per gram wet weight. After addition of 1 mg mL^{-1} lysozyme and incubation on ice for 30 min, the cells were sonicated 6 times for 10 s each at 200 W. To separate the cellular debris from the soluble proteins, the lysate was centrifuged at 13,000 x *g* for 30 min at 4 °C. One-step purification of the recombinant His_6-tagged fusion protein by affinity chromatography with Ni-NTA agarose resin (Qiagen, Hilden, Germany) was carried out according to the manufacturer's instructions. The protein was eluted with 250 mM imidazole in 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0. In order to change the buffer, the protein fraction was passed through a NAP-5 column (GE Healthcare, Freiburg, Germany), which had been equilibrated with 50 mM Tris-HCl, 15% (v/v) of glycerol, pH 7.5, previously. CdpC3PT was eluted with the same buffer and stored frozen at –80 °C for enzyme assays.

Protein analysis and determination of molecular mass of active His_6-CdpC3PT

Proteins were analysed by SDS-PAGE according to the method of Laemmli[28] and stained with Coomassie brilliant blue G-250.

The molecular mass of the recombinant His_6-CdpC3PT was determined by size exclusion chromatography on a HiLoad 16/60 Superdex 200 column (GE Health Care, Freiburg, Germany), which had been equilibrated with 20 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl. The column was calibrated with dextran blue 2000 (2000 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), carbonic anhydrase (29 kDa) and ribonuclease A (13.7 kDa) (GE Health Care, Freiburg, Germany). The proteins were eluted with 20 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl. The molecular mass of the recombinant His_6-CdpC3PT was determined as 130 kDa. This indicated that CdpC3PT acts likely as a homotrimer.

Assay for CdpC3PT activity

For quantitative determination of the enzyme activity, the reaction mixture (100 μL) contained 50 mM Tris-HCl (pH 7.5), 5 mM $CaCl_2$, 1 mM aromatic substrates, 1 mM DMAPP, 1.5% (v/v) glycerol, and 0.07 μM of purified recombinant CdpC3PT. The reaction mixtures were incubated at 37 °C and the reactions were terminated by addition of 100 μl methanol per 100 μl reaction mixtures. The protein was removed by centrifugation at 13,000 x g for 20 min. The enzymatic products were analysed by HPLC under conditions described below. For quantitative measurement of the enzyme activity, duplicate values were determined routinely. For determination of kinetic parameters of DMAPP, cyclic dipeptides at 1 mM, DMAPP at final concentrations of 0.0, 0.16, 0.40, 0.81, 1.62 and 4.05 mM were used as substrates. For determination of the kinetic parameters of cyclic dipeptides, DMAPP at a final concentration of 5 mM was used. Due to low solubility, concentrations of **1a** and **1f** up to 5.0 mM and **1b–1d** up to 1 mM were used.

Quantification of the enzymatic products

For quantification of the enzymatic products, CdpC3PT was incubated in a large scale (10 mL) containing each of the five cyclic dipeptides (1 mM), DMAPP (2 mM), $CaCl_2$ (5 mM), Tris-HCl (50 mM, pH 7.5), glycerol 1.5% (v/v) and CdpC3PT (2 mg). The reaction mixtures were extracted after incubation at 37 °C for 24 h with ethyl acetate. After evaporation of the solvent, the residues of

the ethyl acetate phase containing both enzymatic products and substrates were subjected to ^{1}H-NMR analysis. The conversion rate of a given substrate was determined by comparison of the integrals of the enzymatic product and the remained substrate in ^{1}H-NMR spectra. The absorption coefficients of the enzymatic products were then calculated by HPLC analysis of the samples after NMR analysis.

Preparative synthesis of enzymatic products for structure elucidation

The NMR samples for quantification of the enzymatic products were then purified on HPLC under the conditions described below for structure elucidation. The isolated products were subjected to ^{1}H-NMR, ^{13}C-NMR, ^{1}H-^{1}H COSY, ^{1}H-^{13}C HSQC and HMBC as well as high resolution electrospray ionization mass spectrometry (HR-ESI-MS).

HPLC conditions for analysis and isolation of enzymatic products of CdpC3PT

The enzymatic products of the incubation mixtures of CdpC3PT were analysed by HPLC on an Agilent series 1200 by using a LiChrospher RP 18-5 column (125 × 4 mm, 5 µm, Agilent) at a flow rate of 1 mL min^{-1}. Water (solvent A) and methanol (solvent B) were used as solvents. For analysis of enzymatic products, linear gradients of 50–80% (v/v) solvent B in 10 min and then of 80–100% (v/v) solvent B in 5 min were used. The column was then washed with 100% solvent B for 5 min and equilibrated with 50% (v/v) solvent B for 5 min. Detection was carried out by a Photo Diode Array detector and illustrated at 254 nm in the figures in this paper.

For isolation, the same HPLC equipment with a Multospher 120 RP-18 column (250 × 10 mm, 5 µm, C+S Chromatographie Service, Langenfeld, Germany) was used. Linear gradients of 50–80% (v/v) solvent B in 10 min and then of 80–100% (v/v) solvent B in 5 min at a flow rate of 2.5 mL min^{-1} were used. The column was then washed with 100% solvent B for 5 min and equilibrated with 50% (v/v) solvent B for 5 min.

NMR experiments

Small amount (less than 1 mg) of each sample was dissolved in 0.2 mL of $CDCl_3$. Samples were filled into Wilmad 3 mm tubes from Rototec Spintec. Spectra were recorded at room temperature on a Bruker Avance 600 MHz spectrometer equipped with an inverse probe with z-gradient. The HSQC and HMBC spectra were recorded with standard methods.[29] Gradient-selected NOESY experiment[30] was performed in phase-sensitive mode. For all two-dimensional spectra, 32 to 64 transients were used. For NOESY spectra, a mixing time of 1.5 s and a relaxation delay of 3.0 s. ^{1}H spectra were acquired with 65 536 data points, while 2D spectra were collected using 4096 points in the F_2 dimension and 512 increments in the F_1 dimension. Typical experiment time for the HMBC and NOESY measurements was about 12 h. Chemical shifts were referenced to $CDCl_3$. All spectra were processed with Bruker TOPSPIN 2.1.

Mass spectrometry

The isolated products were analysed by HR-ESI-MS with a Q-Trap Quantum (Applied Biosystems).Positive HR-ESI-MS data of the enzymatic products are given in Table 1.

Nucleotide sequence accession number

The nucleotide sequence of the genomic DNA from *Neosartorya fischeri* NRRL181 reported in this study is available at GenBank under accession number DS027696. The coding sequence of *cdpC3PT* is available at GenBank under the name *NFIA_074280*.

Acknowledgements

This work was supported by a grant from the LOEWE program des Landes Hessen (SynMikro to S.-M. Li). Xie acknowledges the Deutsche Forschungsgemeinschaft for funding the Bruker AVANCE 600 spectrometer. Xia Yu is a recipient of a fellowship from China Scholarship Council.

References

1 W.-B. Yin, A. Grundmann, J. Cheng and S.-M. Li, *J. Biol. Chem.*, 2008, **284**, 100–109.
2 J. P. Karwowski, M. Jackson, R. R. Rasmussen, P. E. Humphrey, J. B. Poddig, W. L. Kohl, M. H. Scherr, S. Kadam and J. B. McAlpine, *J. Antibiot. (Tokyo)*, 1993, **46**, 374–379.
3 Y. Kimura, T. Hamasaki, H. Nakajima and A. Isogai, *Tetrahedron Lett.*, 1982, **23**, 225–228.
4 C. Rank, R. K. Phipps, P. Harris, J. C. Frisvad, C. H. Gotfredsen and T. O. Larsen, *Tetrahedron Lett.*, 2006, **47**, 6099–6102.
5 P. M. Scott and B. P. C. Kennedy, *J. Agric. Food Chem.*, 1976, **24**, 865–868.
6 T. Rundberget, I. Skaar and A. Flaoyen, *Int. J. Food Microbiol.*, 2004, **90**, 181–188.
7 C. Finoli, A. Vecchio, A. Galli and I. Dragoni, *J. Food Prot.*, 2001, **64**, 246–251.
8 R. P. Hodge, C. M. Harris and T. M. Harris, *J. Nat. Prod.*, 1988, **51**, 66–73.
9 K. Arai, K. Kimura, T. Mushiroda and Y. Yamamoto, *Chem. Pharm. Bull.*, 1989, **37**, 2937–2939.
10 A. G. Kozlovsky, V. M. Adanin, H. M. Dahse and U. Gräfe, *Appl. Biochem. Microbiol.*, 2001, **37**, 253–256.
11 M. Kusano, G. Sotoma, H. Koshino, J. Uzawa, M. Chijimatsu, S. Fujioka, T. Kawano and Y. Kimura, *J. Chem. Soc., Perkin Trans. 1*, 1998, 2823–2826.
12 K. Sprogoe, S. Manniche, T. O. Larsen and C. Christophersen, *Tetrahedron*, 2005, **61**, 8718–8721.
13 S. Takase, Y. Kawai, I. Uchida, H. Tanaka and H. Aoki, *Tetrahedron Lett.*, 1984, **25**, 4673–4676.
14 F. S. De Guzman and J. B. Glober, *J. Nat. Prod.*, 1992, **55**, 931–939.
15 W.-B. Yin, X.-L. Xie, M. Matuschek and S.-M. Li, *Org. Biomol. Chem.*, 2010, **8**, 1133–1141.
16 W.-B. Yin, J. Cheng and S.-M. Li, *Org. Biomol. Chem.*, 2009, **7**, 2202–2207.
17 N. Steffan, A. Grundmann, W.-B. Yin, A. Kremer and S.-M. Li, *Curr. Med. Chem.*, 2009, **16**, 218–231.
18 S.-M. Li, *Phytochemistry*, 2009, **70**, 1746–1757.
19 S. Maiya, A. Grundmann, S.-M. Li and G. Turner, *ChemBioChem*, 2006, **7**, 1062–1069.
20 M. A. van den Berg, R. Albang, K. Albermann, J. H. Badger, J. M. Daran, A. J. Driessen, C. Garcia-Estrada, N. D. Fedorova, D. M. Harris, W. H. Heijne, V. Joardar, J. A. Kiel, A. Kovalchuk, J. F. Martin, W. C. Nierman, J. G. Nijland, J. T. Pronk, J. A. Roubos, d. K. van, I. N. N. van Peij, M. Veenhuis, H. von Dohren, C. Wagner, J. Wortman and R. A. Bovenberg, *Nat. Biotechnol.*, 2008, **26**, 1161–1168.
21 J. H. Yu, Z. Hamari, K. H. Han, J. A. Seo, Y. Reyes-Dominguez and C. Scazzocchio, *Fungal Genet. Biol.*, 2004, **41**, 973–981.

22 L. Du, X. Yang, T. Zhu, F. Wang, X. Xiao, H. Park and Q. Gu, *Chem. Pharm. Bull.*, 2009, **57**, 873–876.
23 J. H. Chang and H. Moon, *Biotechnol. Bioprocess Eng.*, 2004, **9**, 59–61.
24 A. Grundmann and S.-M. Li, *Microbiology*, 2005, **151**, 2199–2207.
25 A. B. Woodside, Z. Huang and C. D. Poulter, *Org. Synth.*, 1988, **66**, 211–215.
26 J. Sambrook, and D. W. Russell, *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 2001.
27 F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, *Current Protocols in Molecular Biology*, John Wiley and Sons Inc, New York, 1996.
28 U. K. Laemmli, *Nature*, 1970, **227**, 680–685.
29 S. Berger, and S. Braun, *200 and More NMR Experiments. A Practical Course*, Weiley-VCH, Weinheim, Germany, 2004.
30 T.-L. Hwang and A. J. Shaka, *J. Am. Chem. Soc.*, 1992, **114**, 3157–3159.

