# Simultaneous C7- and N1-prenylation of cyclo-L-Trp-L-Trp catalyzed by a prenyltransferase from *Aspergillus oryzae*†

**Hui-Xi Zou,[a,b] Xiu-Lan Xie,[c] Uwe Linne,[c] Xiao-Dong Zheng[a] and Shu-Ming Li*[b]**

***Received 12th February 2010, Accepted 16th April 2010***
***First published as an Advance Article on the web 14th May 2010***
**DOI: 10.1039/c002850a**

A putative prenyltransferase gene *cTrpPT* was amplified from *Aspergillus oryzae* DSM1147, cloned into pQE70 and overexpressed in *Escherichia coli*. The overproduced $His_6$-CTrpPT was purified to near homogeneity and incubated with L-tryptophan or tryptophan-containing cyclic dipeptides in the presence of dimethylallyl diphosphate. The formation of the enzyme products was monitored with HPLC. It was shown that CTrpPT differed clearly from other known indole prenyltransferases in several aspects. This enzyme showed higher substrate specificity towards aromatic substrates, but lower regioselectivity regarding the prenylation position than other indole prenyltransferases. Cyclo-L-Trp-L-Trp was much better accepted than other cyclic dipeptides tested in this study. In comparison to other indole prenyltransferases with one dominant enzyme product, at least two product peaks were detected in the reaction mixtures of CTrpPT. $^1$H- and $^{13}$C-NMR analyses, including long-range $^1$H–$^{13}$C connectivities in Heteronuclear Multiple-Bond Correlation (HMBC) and Nuclear Overhauser Effect Spectroscopy (NOESY), proved the structures of the enzyme products as C7- and N1-prenylated derivatives with a ratio of 1 : 1.2 using cyclo-L-Trp-L-Trp as substrate. The $K_M$ values were determined at about 2.5 mM for dimethylallyl diphosphate and 0.3 mM for cyclo-L-Trp-L-Trp with a turnover number of 0.33 $s^{-1}$.



## Introduction

Prenyltransferases are found in all domains of life and are involved in the biosynthesis of primary and secondary metabolites.[1–4] They catalyze the transfer reaction of a prenyl moiety from a prenyl donor, usually as diphosphate to a terpenoid, serine residue of a protein or an aromatic nucleus.[2] The latter group are known as aromatic prenyltransferases, mainly found in plants, bacteria and fungi.[1,2,4] Significant progress has been achieved in the last years on the molecular biological, biochemical and structural biological investigation of aromatic prenyltransferases, including indole prenyltransferases from ascomycota, which catalyze prenylation of diverse substrates at different positions of the indole rings.[5–7]

The known indole prenyltransferases from fungi accepted only dimethylallyl diphosphate (DAMPP) as prenyl donor, but usually showed broad substrate specificities towards their aromatic substrates, such as simple tryptophan derivatives or tryptophan-containing cyclic dipeptides.[8–12] The second common feature of the known indole prenyltransferases is the high regioselectivity of the prenyl transfer reaction. In most cases, only one or one dominant product was detected in the reaction mixture. For example, FtmPT1 catalyzed the C2-prenylation of all of the tested tryptophan-containing cyclic dipeptides,[8] while AnaPT prenylated these compounds at position C3 of the indole ring.[12,13]

In the last years, several C4 prenylating enzymes of tryptophan have been identified, *e.g.* FgaPT2 from *Aspergillus fumigatus*,[14] DmaW from a Clavicipitalean fungus[15] and MaPT from *Malbranchea aurantiaca*.[16] It has been shown that all of these enzymes accepted a series of simple tryptophan derivatives and converted them to C4-prenylated derivatives. FgaPT2 accepted even tryptophan-containing cyclic dipeptides. The prenylation position remained C4 in this case.[17] The high regiospecificity of the indole prenyltransferases was also observed with 7-DMATS from *Aspergillus fumigatus*,[18] which prenylated diverse simple tryptophan derivatives at position C7 of the indole ring.[11]

In the course of our investigation of indole prenyltransferases, one putative prenyltransferase gene *AO090120000023* from the genome sequence of *Aspergillus oryzae* RIB40[19] raised our attention. *AO090120000023* is located on chromosome 5 directly next to *AO090120000024* (Fig. 1), a putative non-ribosomal peptide synthetase (NRPS) gene. Homologous genes of *AO090120000023* and *AO090120000024*, *i.e.AFLA_090190* and *AFLA_090200* with sequence identity of 98% on the amino acid level, have also been identified in the genome of *Aspergillus flavus* (GenBank). This indicates that the putative prenyltransferase and NRPS genes very likely belong to a cluster for the biosynthesis of a secondary metabolite, which could not be determined by sequence analysis. It can, however, be expected that the products of *AO090120000023* and *AFLA_090190*, as well as those of *AO090120000024* and *AFLA_090200*, catalyze the same reaction, respectively. Interestingly, the deduced dimodular NRPS protein BAE62663 of *AO090120000024* in *A. oryzae*, or its orthologue EED48939 in *A. flavus*, consists of an unusual domain structure of ATCTC

[a]*Zhejiang University, Department of Food Science and Nutrition, 310029 Hangzhou, Zhejiang, People's Republic of China*
[b]*Philipps-Universität Marburg, Institut für Pharmazeutische Biologie und Biotechnologie, Deutschhausstrasse 17A, D-35037 Marburg, Germany. E-mail: shuming.li@Staff.uni-Marburg.de; Fax: 0049-6421-2825365; Tel: 0049-6421-2822461*
[c]*Philipps-Universität Marburg, Fachbereich Chemie, Hans-Meerwein-Strasse, 35032 Marburg, Germany*
† Electronic supplementary information (ESI) available: LC-ESI-TOF-MS of $His_6$-CTrpPT, and NMR data and spectra. See DOI: 10.1039/c002850a

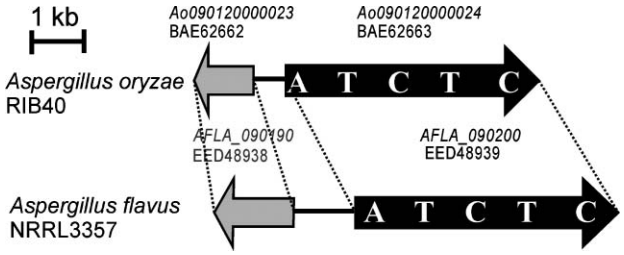



**Fig. 1** Genetic organisation of the putative prenyltransferase and NRPS genes in *Aspergillus oryzae* and *Aspergillus flavus*.

(A: adenylation; T: thiolation and C: condensation). The two T and C domains are expected to be responsible for the thioester formation of two amino acids and their subsequent condensation to a cyclic dipeptide, as demonstrated by brevianamide F synthetase in the biosynthesis of fumitremorgins.[20] Given the functionality of this protein, the two amino acids of the putative cyclic dipeptide must be activated by one single adenylation domain. It seems, therefore, that one amino acid is adenylated by the single A-domain and then transferred to the two T-domains. However, it cannot be excluded that two different amino acids would be activated by the same domain. We speculated that BAE62663 and EED48939 would be responsible for the formation of a cyclic dipeptide consisting of two identical amino acids, *e.g.* two tryptophan residues, which is then prenylated by the putative prenyltransferase BAE62662 in *A. oryzae* or EED48938 in *A. flavus*. Here, we report the cloning and heterologous expression of an orthologue of the putative prenyltransferase gene *AO090120000023*, termed *cTrpPT* in this study, from *A. oryzae* DSM1147, identification of the overproduced enzyme as a prenyltransferase with high preference for cyclo-L-Trp-L-Trp as aromatic substrate. This enzyme catalyzed prenylation at position N1 or C7 of the indole ring.

## Results and discussion

According to genome annotation of *A. oryzae* RIB40, *AO090120000023* spans bp 47468 to 48655 of AP007166 in GenBank and consists exclusively of one exon with a length of 1188 bps. The deduced protein BAE62662 comprises 395 amino acids. By sequence analysis using FGENESH (see Experimental Section) and comparison with known indole prenyltransferases, we predicted that the putative prenyltransferase gene has an additional small exon at its 3′-end, which spans bp 47308 to 47407 of AP007166. The two exons with a length of 1184 and 100 bps, respectively, are interrupted by an intron of 64 bps. As a consequence, the deduced polypeptide comprises 427 amino acids, 32 amino acids longer than that predicted by genome annotation. To distinguish with the annotated sequence in GenBank, we use the name *cTrpPT* for the putative prenyltransferase gene in this study.

The deduced gene product CTrpPT with a calculated molecular mass of 49.0 kDa showed clear sequence similarity to known prenyltransferases of the dimethylallyl tryptophan synthase superfamily including indole prenyltransferases. For example, by using the program "BLAST 2 SEQUENCES" (see Experimental section), CTrpPT shares an identity of 35% with FgaPT2[14] and 28% with 7-DMATS,[18] respectively. It could be therefore expected that CTrpPT catalyzes a prenyl transfer reaction onto the indole ring of tryptophan or derivatives thereof. To prove the function of CTrpPT, we cloned the coding sequence of CTrpPT from *A. oryzae* DSM1147 for heterologous expression in *E. coli*.

The initial cloning attempt was carried out by PCR amplification of the entire coding region using cDNA as template, which had been obtained from mRNA by reverse transcription. This strategy failed, because the cloned PCR product still contained the predicted intron sequence, although the mRNA was obtained in high quality (data not shown). Expression of the first exon as given in GenBank (coding sequence for BAE62662) with the expression vector pQE70 in *E. coli* did not result in a detectable amount of overproduced protein under various conditions including different *E. coli* strains, IPTG concentrations, cultivation media and induction temperature (data not shown). Therefore, we changed our strategy and tried to clone the entire coding region by PCR amplification of the two exons from the genomic DNA and fusion of the two PCR fragments with the help of a short overlapped region by two further PCR amplifications (see Experimental Section). The successfully amplified PCR product was cloned into pGEM-T easy vector to create the cloning construct pHX19. Sequencing of pHX19 showed that thymine at position 891 in the sequence of genome reference strain RIB40 was replaced by cytosine. This difference was very likely caused by the PCR amplification, because the PCR products obtained from mRNA and for expression of the first exon of the initial attempt had no change at this position. However, this change had no effect on the amino acid sequence.

For gene expression, the coding sequence of *cTrpPT* was released from pHX19 and cloned into the expression vector pQE70, resulting in the plasmid pHX21. Overexpression of pHX21 in *E. coli* resulted in clear production of soluble CTrpPT, which was subsequently purified to near homogeneity with Ni-NTA agarose as judged by SDS-PAGE (Fig. 2), and a protein yield of 5 mg of purified $His_6$-tagged CTrpPT per litre of culture has been obtained. The observed molecular mass was about 44 kDa, and is somehow smaller than the calculated value of 50254.5 Da for $His_6$-CTrpPT. To ensure that the purified protein was not a truncated derivative, the protein sample was sent for MS analysis. LC-ESI-TOF mass spectroscopy detected a dominant peak at $m/z$ 50124.0 Da (Fig. S1, ESI†). This value corresponds very well to that of $His_6$-CTrpPT after removal of the methionine residue at

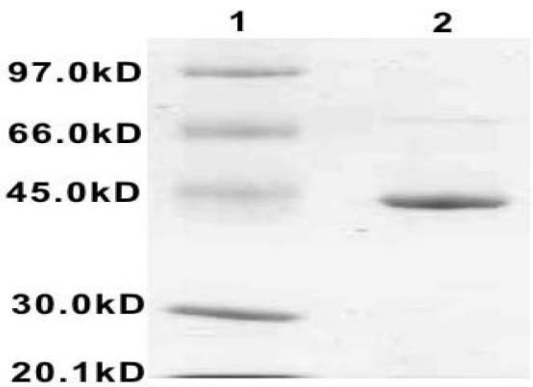



**Fig. 2** Purification of CTrpPT as a $His_6$-tagged protein. The 12% (w/v) SDS polyacrylamide gel was stained with Coomassie Brilliant Blue G-250. Lane 1: Molecular mass standard; 2: Purified $His_6$- CTrpPT.



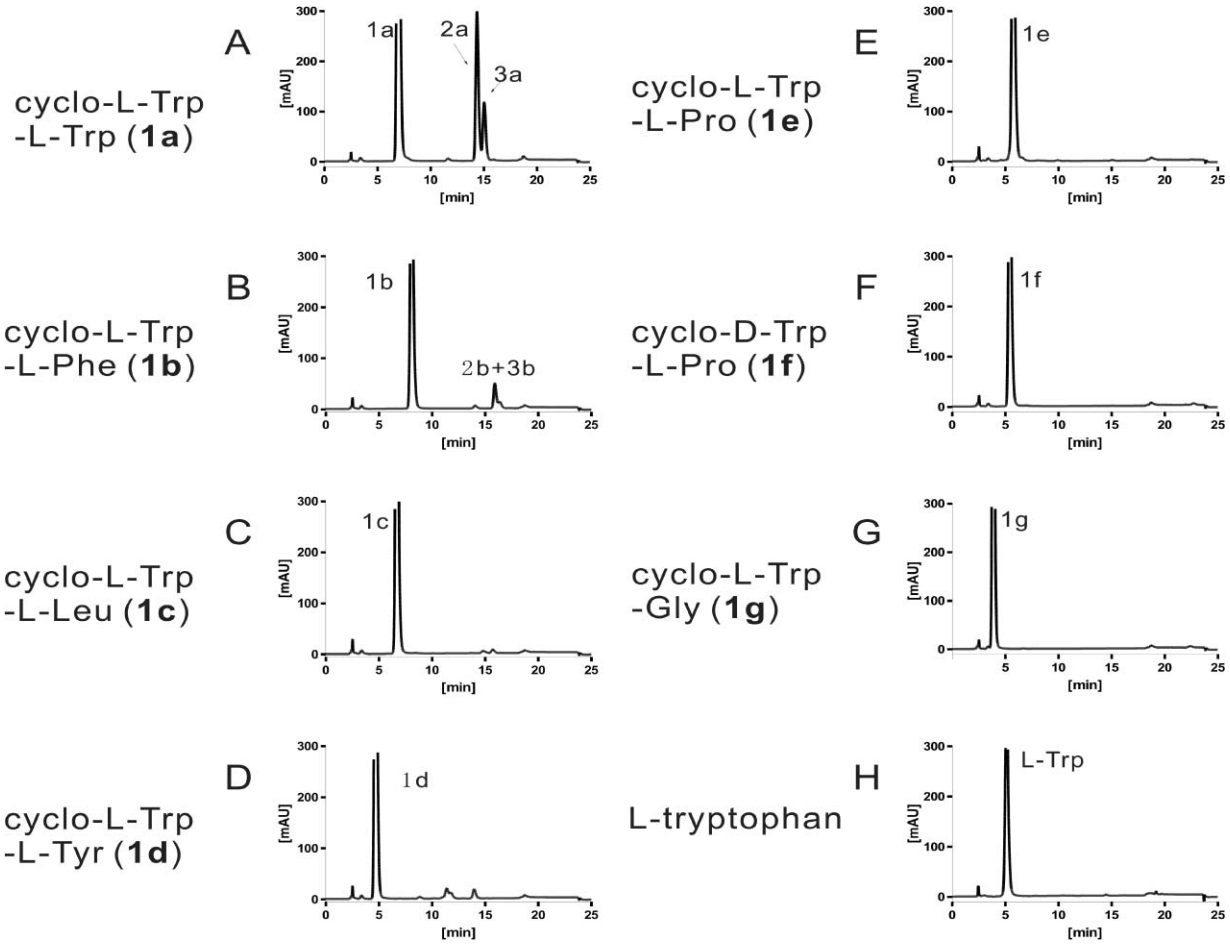



**Fig. 3** HPLC chromatograms of incubation mixtures of L-tryptophan and tryptophan-containing cyclic dipeptides with recombinant CTrpPT. Detection was carried with a Diode Array detector and illustrated for absorption at 277 nm.

the N-terminus by *E. coli* methionine aminopeptidase[21,22] with a mass of 50123.3 Da.

The purified CTrpPT (2 μM) was then incubated with L-tryptophan and seven tryptophan-containing cyclic dipeptides (1 mM) in the presence of 1 mM DMAPP for 16 h and HPLC analysis of the incubation mixtures was used to monitor the formation of enzyme products. As shown in Fig. 3, cyclo-L-Trp-L-Trp (**1a**) was accepted by CTrpPT as the best substrate. For **1a**, a conversion rate of 67% has been calculated by NMR analysis of the reaction mixture (see Experimental Section). Cyclo-L-Trp-L-Phe (**1b**) was accepted with a conversion rate of about 10%. Cyclo-L-Trp-L-Leu (**1c**) and cyclo-L-Trp-L-Tyr (**1d**) were also accepted by CTrpPT, but with much lower relative activity than **1a** and **1b**. At least two product peaks each were detected in the reaction mixtures of **1a–d**. These peaks were absent in the control assays with heat-inactivated enzyme (data not shown), demonstrating that they are enzyme products or derivatives thereof. No detectable product formation has been observed with cyclo-L-Trp-L-Pro (**1e**), cyclo-D-Trp-L-Pro (**1f**), cyclo-L-Trp-Gly (**1g**) or L-tryptophan.

To prove the relationship of the two product peaks **2a** and **3a** (Fig. 3) in the incubation mixture with cyclo-L-Trp-L-Trp, we carried out incubations with various enzyme amounts and for different times. Fig. 4 shows clearly that the formation of both peaks increased parallel to protein amounts or incubation time, with a ratio of 1 : 1.2 for **2a** : **3a**. These results have been reproduced with different protein batches and fractions after Ni-NTA agarose purification. This demonstrated that both compounds are enzyme products of cyclo-L-Trp-L-Trp (**1a**) and are very likely formed independently of each other.

For structure elucidation, the enzyme products of **1a** and **1b** were isolated on a preparative scale and subjected to ESI-MS and NMR analyses. ESI-MS data confirmed the presence of one dimethylallyl moiety each in the structures of **2a** and **2b**, as well as of **3a** and **3b** by detection of $[M + 1]^+$ ions, which are 68 Daltons larger than those of the respective substrates.

For unambiguous assignments of the $^1$H- and $^{13}$C-NMR signals given in Table S1 (ESI†), we took $^1$H–$^{13}$C heteronuclear single-quantum correlation (HSQC), heteronuclear multiple bond correlation (HMBC) and Nuclear Overhauser Effect Spectroscopy (NOESY) spectra (Figs. S2.1–S5.4, ESI†).

Comparison of the $^1$H-NMR data of **2a** and **2b** (Table S1, ESI†) with those of the substrates clearly revealed the presence of signals for a regular dimethylallyl moiety at 3.5 ppm (d) for H-1′, 5.3 ppm (t) for H-2′, 1.78-1.79 ppm (s) for H-4′ and 1.71 ppm (s) for H-5′. Inspection of the $^1$H-NMR spectra of **2a** and **2b** also provided evidence for the presence of signals for three coupling protons on the benzene ring instead of four in those of the substrates (data not shown), indicating a prenylation at position C4 or C7 of the indole

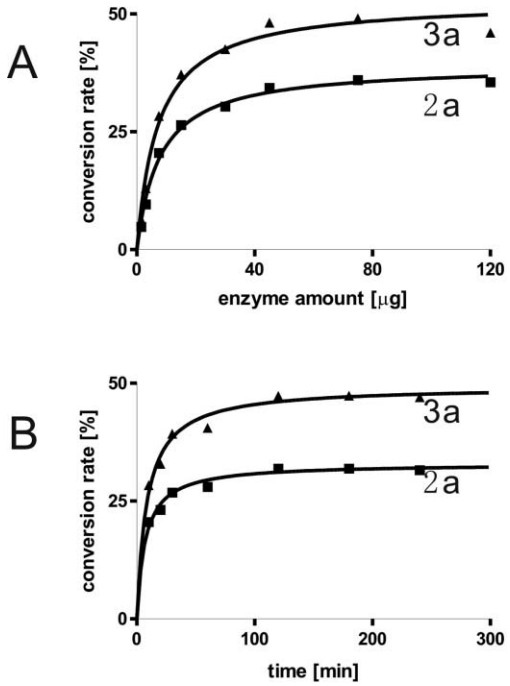



**Fig. 4** Dependence of product formation of **1a** on protein amount after incubation for 30 min (A) and incubation time with a protein amount of 15 µg (1 µM) (B).

ring. Unambiguous proof of the C7-prenyl moiety in **2a** and **2b** was provided by HMBC and NOESY correlation. A cross peak in the HSQC spectrum of **2a** revealed that the doublet of H-1′ at 3.52 ppm correlated with the signal of C-1′ at 30.4 ppm, which showed clear connectivity to H-6 of the indole ring in the HMBC spectrum (Fig. 5). In addition, H-1′ of **2a** correlated obviously with C6, C7 and C8 of the indole ring in the HMBC spectrum. Correlations between H-1′ at 3.54 ppm of the dimethylallyl moiety with C6, C7 and C8 of the indole ring were also clearly observed in the HMBC spectrum of **2b**. NOESY confirmed furthermore the interaction between H-1′ of the dimethylallyl moiety with H-6 of the indole ring, as well as NH-1 with H-1′ and H-2′ of the dimethylallyl moiety (Fig. 5).

In contrast to those of **2a** and **2b**, the $^1$H-NMR spectra of **3a** and **3b** clearly showed the presence of signals for a reverse dimethylallyl moiety at 5.2 ppm for H-1′, 6.1 ppm for H-2′ and two singlets at 1.7 ppm (s) for H-4′ and H-5′. The HMBC spectrum of **3a** (ESI†) showed connectivity of H-2 of the indole ring to C-3′ of the dimethylallyl moiety, proving unequivocally the prenylation at N1 of the indole ring. This conclusion was also confirmed by interaction of H-2 and H-7 of the indole ring with the two methyl groups of the dimethylallyl moiety in the NOESY of **3a**. These NOE contacts were also observed in NOESY of **3b** (ESI†).

In summary, CTrpPT catalyzed the simultaneous prenylation of cyclo-L-Trp-L-Trp (**1a**) and cyclo-L-Trp-L-Phe (**1b**) at positions C7 and N1. In addition to monoprenylation, AnaPT also catalyzed the diprenylation at position C3 of the both tryptophanyl units in cyclo-L-Trp-L-Trp.[12] In the incubation mixture of cyclo-L-Trp-L-Trp with CTrpPT, no additional substances to **2a** and **3a** were identified. This means that no or only a minor amount of N1, C7-diprenylated, N1, N20-diprenylated or C7, C22-diprenylated was produced by CTrpPT. The detailed mechanism of the CTrpPT reaction is unknown. We speculate that generating of a dimethylallyl cation would initiate the enzyme reaction (Scheme 1), as in the case of other prenyltransferases, *e. g.* FgaPT2.[23] Attacking of this intermediate *via* its C1 to C7 of the indole ring would result in formation of the intermediate **4** with a positive charge at C8. Abstraction of a proton and reformation of the double bond between C7 and C8 would result in **2**. The dimethylallyl cation can also attack N1 of the indole ring, resulting in formation of the intermediate **5**, which is then to be converted to **3** by releasing a proton.

Due to their low amounts, it was not possible to get enough enzyme products from incubation mixtures with **1c** and **1d** for an unequivocal structure elucidation. However, the ESI-MS data confirmed the monoprenylation in the enzyme products and NMR spectra provided indications for the presence of N1- and C7-prenylated derivatives in both cases (data not shown).

To evaluate the effect of metal ions on the activity of CTrpPT, enzyme assays containing 1 µM of CTrpPT, 1 mM of cyclo-L-Trp-L-Trp (**1a**) and 1 mM of DMAPP were carried out in the presence of metal ions or EDTA. Consistent with most known indole prenyltransferases,[5] addition of EDTA to the reaction mixture did not reduce the enzyme activity. In contrast, a relative activity of 236% was observed in the reaction mixture with 5 mM of EDTA in comparison to the incubation without additives. Addition of $Ca^{2+}$, $Mg^{2+}$ and $Mn^{2+}$ to a final concentration of 5 mM to the reaction mixtures increased the enzyme activity to about 400, 140 and 220% of that without additives, respectively.

$K_M$ values of 0.3 and 0.6 mM have been determined for **1a** and **1b**, respectively (Table 1). Significant difference in the reaction velocity was observed for **1a** and **1b**. The catalytic efficiency $k_{cat}/K_M$ for **1a** is 130-fold higher than that of **1b**. It can be

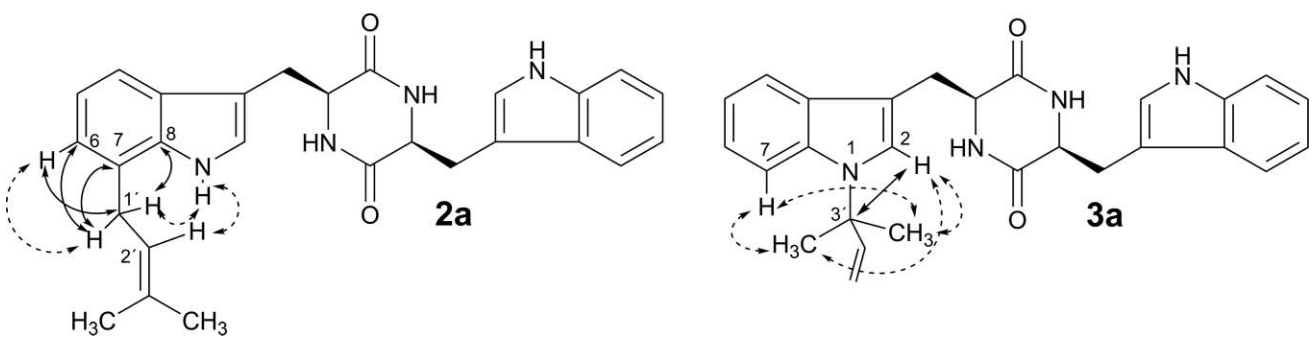
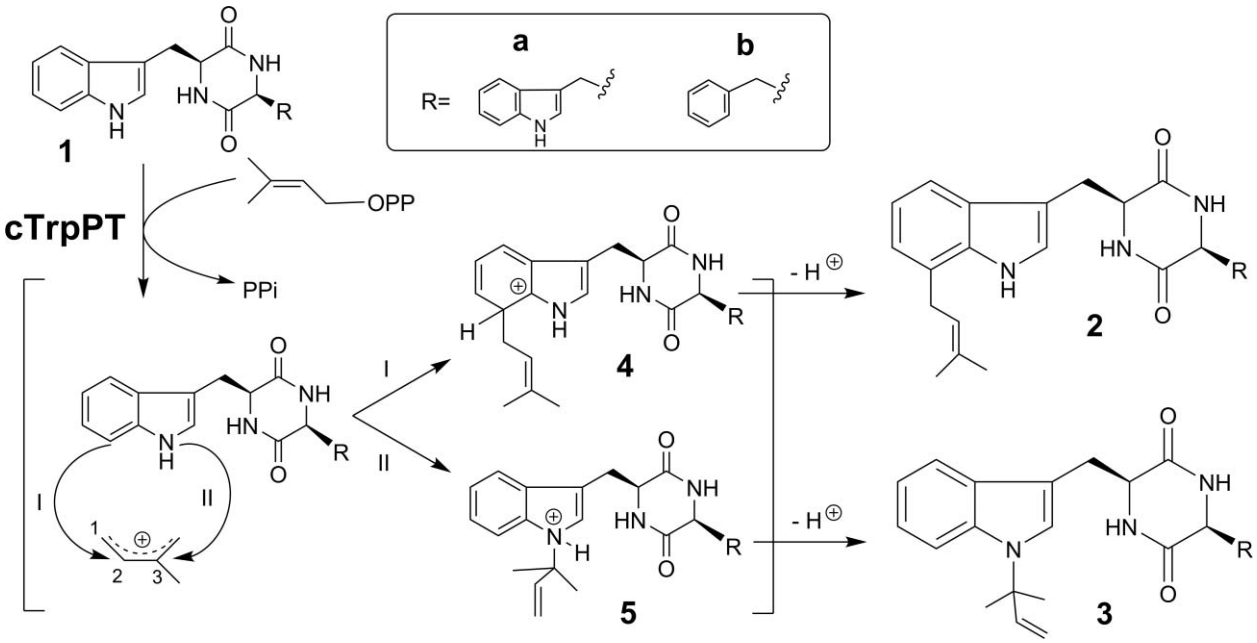



**Fig. 5** HMBC (solid lines) and NOE correlations (dashed lines) of enzyme products **2a** and **3a**.




**Scheme 1** Proposed reaction mechanism for CTrpPT reaction.

**Table 1** Preliminary parameters of CTrpPT

| Substrate | $K_M$/mM | $k_{cat}$/$s^{-1}$ | $k_{cat}/K_M$ [$s^{-1}$ $M^{-1}$] | Ratio of $k_{cat}/K_M$ |
|---|---|---|---|---|
| Cyclo-L-Trp-L-Trp (**1a**) | 0.30 | 0.33 | 1100 | 100 |
| Cyclo-L-Trp-L-Phe (**1b**) | 0.60 | 0.0048 | 8 | 0.72 |
| DMAPP[a] | 2.50 | 0.39 | | |

[a] **1a** as aromatic substrate.

expected that other cyclic dipeptides tested in this study, *i.e.* **1c–g** (Fig. 3) are much poorer substrates for CTrpPT. As mentioned in the **Introduction**, both the natural prenyl donor and acceptor of CTrpPT are still unknown. Based on the domain structure of the NRPS BAE62263, we speculated that cyclo-L-Trp-L-Trp could be the natural aromatic substrate of the prenyltransferase. The $K_M$ value of CTrpPT at 0.3 mM for cyclo-L-Trp-L-Trp is comparable to that of the cyclic dipeptide prenyltransferase AnaPT at 0.23 mM for its natural substrate (*R*)-benzodiazepinedione,[13] but much higher than that of FtmPT1 at 0.057 mM for its natural substrate cyclo-L-Trp-L-Pro.[8] The $k_{cat}/K_M$ value of CTrpPT for cyclo-L-Trp-L-Trp is only 17% of that of AnaPT[13] and 1.0% of that of FtmPT1.[8] These data can neither confirm, nor exclude the possibility that cyclo-L-Trp-L-Trp is the natural substrate of CTrpPT. Inactivation of *cTrpPT* in *A. oryzae* or *A. flavus* and analysis of the secondary metabolite production would provide direct evidence of its function. Prerequisite for the success of this approach is the expression of the gene cluster of interest in the fungal strain. Expression of the NRPS gene *AO090120000024* from *A. oryzae* or *AFLA_090200* from *A. flavus* in a fungal strain and identification of the accumulated dipeptide would provide an alternative approach.

Unusually, CTrpPT showed a $K_M$ value of 2.5 mM for DMAPP, much higher than other known prenyltransferases in the range of 4–190 μM.[8,13,14,18,24,25] However, no product formation was observed with CTrpPT, when geranyl diphosphate instead of DMAPP was used as a prenyl donor for cyclo-L-Trp-L-Trp (data not shown). An additional interesting biochemical feature of CTrpPT is the substrate inhibition by both DMAPP (over 7.5 mM) and aromatic substrates (over 2 mM).

## Conclusions

In this study, we identified a new indole prenyltransferase CTrpPT, which accepted cyclo-L-Trp-L-Trp as substrate much better than other cyclic dipeptides and catalyzed simultaneous prenylation at different positions of the indole ring, *i.e.* the regular prenylation at position C7 and the reverse prenylation at position N1. This is valid at least for cyclo-L-Trp-L-Trp and cyclo-L-Trp-L-Phe, which had not been observed for indole prenyltransferases before.

## Experimental section

### Chemicals

DMAPP was prepared according to the method described for geranyl diphosphate by Woodside *et al*.[26] Cyclic dipeptides were obtained from Bachem (Bubendorf, Switzerland).

### Computer-assisted sequence analysis

FGENESH (Softberry; http://www.softberry.com/berry.phtml) and the DNASIS software package (version 2.1; Hitachi Software Engineering) were used for intron prediction and sequence analysis, respectively. Sequence similarities were obtained by alignments of amino acid sequences using the BLAST2 sequences program (release 2.9.9; http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi).

### Bacterial and fungal strains, plasmids and culture conditions

*A. oryzae* strain DSM1147 was purchased from DSMZ (Braunschweig, Germany), and used to obtain mRNA and genomic DNA as templates for PCR amplification. For DNA isolation, mycelia of *A. oryzae* from plates were inoculated into a 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

pGEM-T easy (Promega; Mannheim, Germany) and pQE70 (Qiagen; Hilden, Germany) were used as cloning and expression vectors, respectively. *E. coli* XL1 Blue MRF' (Stratagene) was used for cloning and expression experiments, and was grown in liquid Luria–Bertani (LB) medium, or on solid LB medium with 1.5% (w/v) agar, at 37 °C.[27] Addition of 50 μg $mL^{-1}$ carbenicillin was used for selection of recombinant *E. coli* strains.

### DNA isolation, PCR amplification and cloning

Standard procedures for DNA isolation and manipulation in *E. coli* were performed, as described.[27] DNA isolation from *A. oryzae* was carried out according to the protocol described by Ausubel *et al.*[28]

PCR amplification was carried out on an iCycler from Bio-Rad. Using the Expand High Fidelity kit (Roche Diagnostics), the entire coding sequence of *cTrpPT* was amplified from genomic DNA as the template after three rounds of PCR. In the first round PCR, two exons were amplified separately. Primers CTRPPTEXP7 (5′-CCCAAGCTCCATAATCACCTTATCTTGTATCGGCTATG-ACTACCTATACG-3′) at the 5′- end and CTRPPTEXP16 (5′-***CGGTGCAGTCATTCAGGTCCATATGTGGT***ACATAAGAA -GCCAAATTGTC-3′) at the 3′- end were used for the first exon, and CTRPPTEXP9 (5′-***CATACATTGACAATTTGG-CTTCTTATGT***ACCACATATGGACCTGAATG-3′) at the 5′-end and CTRPPTEXP18 (5′-TTACTACCCCTAGCTAGTGA-TAATAAACAGTAATATATGGCCCCGTA-3′) at the 3′- end for the second exon. Bold italic letters represent overlapping region of the two exons. PCR products of the first and the second exon were then mixed in a molar ratio of 1 : 1 and used as the template for a second round of PCR to get a fragment consisting of the two exons with the help of the overlapping region. A nested primer pair, CTRPPTEXP1 (5′-GG***GCATGC***CTA-CCTATACGCTGTC-3′) at the 5′-end and CTRPPTEXP6 (5′-GCCCC***GGATCC***GTGATAATAAACA -3′) at the 3′-end, and the PCR products from the second round were used in the final round PCR. The bold letters in the last two primers represent mutations inserted to give the italic restriction sites SphI, located at the start codon in CTRPPTEXPT1, and BamHI, located at the predicted stop codon in CTRPPTEXPT. The 1298 bp PCR fragment obtained was cloned into pGEM-T easy vector, resulting in plasmid pHX19, which was subsequently sequenced (MWG-Biotech, Germany). To create the expression vector pHX21, pHX19 was digested with SphI and BamHI. The resulted SphI-BamHI fragment of 1284 bp was ligated into pQE70, which had been digested with the same enzymes.

### Overproduction and purification of $His_6$-CTrpPT protein

For gene expression, *E. coli* XL1 Blue MRF′ cells harbouring the plasmid pHX21 were cultivated in 300 mL Erlenmeyer flasks containing 100 mL liquid LB medium, supplemented with carbenicillin (50 μg $mL^{-1}$), and grown at 37 °C to an absorption at 600 nm of 0.6. For induction, IPTG was added to a final concentration of 0.1 mM, and the cells were cultivated for a further 16 h at 30 °C before harvest. The bacterial cultures were centrifuged, and the pellets were resuspended in lysis buffer (10 mM imidazole, 50 mM $NaH_2PO_4$ and 300 mM NaCl, pH 8.0, 2–5 mL per g wet weight). After addition of 1 mg lysozyme per mL, and incubation on ice for 30 min, the cells were sonicated six times, 10 s each time at 200 W. To separate the cellular debris from the soluble proteins, the lysate was centrifuged at 14 000 *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) was carried out according to the manufacturer's instructions. The protein was eluted with 250 mM imidazole in 50 mM $NaH_2PO_4$ and 300 mM NaCl, pH 8.0. In order to change the buffer, the protein fraction was passed through a NAP-5 column (GE Healthcare), which had been equilibrated with 50 mM Tris/HCl and 15% (v/v) glycerol, pH 7.5. $His_6$-CTrpPT was eluted with the same buffer as for column equilibration and stored at –80 °C for enzyme assays.

### Protein analysis and determination of molecular mass of active $His_6$-CTrpPT

Protein monomers were analyzed by SDS-PAGE according to the method of Laemmli[29] and stained with Coomassie Brilliant Blue G-250.

The molecular mass of the recombinant active $His_6$-CTrpPT was determined by gel filtration 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), conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa) and ribonuclease A (13.7 kDa) (GE Health Care; Freiburg, Germany). The protein was eluted with 20 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl. The molecular mass of the recombinant active $His_6$-CTrpPT was determined as 144 kDa, proving that CTrpPT acts as a homotrimer.

### Assays for CTrpPT

All of the enzyme assays contained 50 mM Tris-HCl, pH 7.5, 0.1–4.5% (v/v) of glycerol and 10 mM $CaCl_2$ and were carried out in duplicate. The reaction mixtures were incubated at 37 °C and the reactions were terminated by addition of 100 μl methanol per 100 μl reaction mixtures. After removal of the protein by centrifugation (14 000 × *g*, 10 min, 4 °C), the enzyme products were analyzed on an HPLC system as described below. The standard assays for determination of the enzyme activity contained 2 μM of CTrpPT, 1 mM of DMAPP and 1 mM of cyclic dipeptides. The assays for determination of the kinetic parameters (100 μl) of DMAPP contained 0.08 μM of CTrpPT, 2 mM of cyclo-L-Trp-L-Trp and DMAPP at final concentrations of 0, 0.1, 0.2, 0.5, 1.0, 2.0, 5.0, 7.5, 10 and 15 mM. For determination of the kinetic parameters of cyclo-L-Trp-L-Trp, 5 mM DMAPP, 0.08 μM of CTrpPT and cyclo-L-Trp-L-Trp at final concentrations of 0, 0.05, 0.1, 0.2, 0.5, 1.0, 2.0, 3.0 or 4.0 mM were used. The incubation time was 30 min. For determination of the kinetic parameters of

cyclo-L-Trp-L-Phe, 2 μM of CTrpPT, 2 mM DMAPP and cyclo-L-Trp-L-Phe at final concentrations of 0, 0.05, 0.1, 0.2, 0.5, 1.0, 2.0 and 5.0 mM were used. The incubation time for these assays was 60 min.

### Determination of the conversion rate

The incubation mixtures contained enzyme products and the remaining substrate were initially analyzed by $^1$H-NMR after being extracted with ethyl acetate and evaporated to dryness. The ratio of products to substrate was calculated by comparison of the integrals of one or more corresponding protons. The conversion rate was defined as the ratio of products in mol to the sum of products and substrate. The incubation mixtures after NMR analysis were subsequently analyzed on an HPLC system as mentioned above to determine the respective absorption coefficient.

### HPLC conditions for analysis and isolation of enzyme products of CTrpPT

The enzyme products of the incubation mixtures of CTrpPT were analyzed by HPLC on an Agilent series 1200 by using a Multospher 120 RP 18–5μ column (250 × 4 mm, 5 μm, C+S Chromatographie Service, Langenfeld, Germany) at a flow rate of 1 mL min$^{-1}$. Double distilled water (solvent A) and methanol (solvent B) were used as solvents. For analysis of enzyme products, a linear gradient of 50–80% (v/v) solvent B in 15 min was 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 photodiode array detector and illustrated at 277 nm in Fig. 3.

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. For isolation of the enzyme products of cyclo-L-Trp-L-Trp, cyclo-L-Trp-L-Tyr and cyclo-L-Trp-L-Leu, a linear gradient of 70–85% (v/v) solvent B in A in 15 min at a flow rate of 2.5 mL min$^{-1}$ was used. The column was then washed with 100% solvent B for 5 min and equilibrated with 50% (v/v) solvent B for 10 min. For isolation of the enzyme products of cyclo-L-Trp-L-Phe, a linear gradient of 60–90% (v/v) solvent B in 22.5 min at a flow rate of 2.5 mL min$^{-1}$ was used, then followed a linear gradient of 90–100% (v/v) solvent B in 7.5 min at a flow rate of 1.0 mL min$^{-1}$. The column was then washed with 100% solvent B for 5 min and equilibrated with 50% (v/v) solvent B for 10 min at a flow rate of 2.5 mL min$^{-1}$.

### NMR spectroscopic analysis

A 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, HMBC, and NOESY spectra were recorded with standard methods.[30] The ROESY experiment[31] was performed in phase-sensitive mode using State-TPPI technique.[32] For all two-dimensional spectra, 32 to 64 transients were used. For ROESY spectra, a mixing time of 300 ms and a relaxation delay of 3.0 s were used. $^1$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. The typical experiment time for the HMBC and ROESY measurements was about 12 h. Chemical shifts were referenced to $CDCl_3$. All spectra were processed with Bruker TOPSPIN 2.1.

### MS data of the isolated compounds

The isolated products were analyzed by electrospray ionization (ESI) mass spectrometry (MS) with a Q-Trap Quantum (Applied Biosystems).

**2a**, Molecular formula: $C_{27}H_{28}N_4O_2$; HR-ESI-MS: $[M + H]^+$ (Cal. : 441.2291, found: 441.2257)

**3a**, Molecular formula $C_{27}H_{28}N_4O_2$; HR-ESI-MS: $[M + H]^+$ (Cal. : 441.2291, found: 441.2297)

**2b**, Molecular formula $C_{25}H_{27}N_3O_2$; ESI-MS: $[M + H]^+$ (Cal. : 402.2182, found 402.2212)

**3b**, Molecular formula $C_{25}H_{27}N_3O_2$; ESI-MS: $[M + H]^+$ (Cal. : 402.2182, found 402.2188)

### Nucleotide sequence accession number

The nucleotide sequence of the genomic DNA from *Aspergillus oryzae* RIB40 reported in this study is available at GenBank under accession number AP007166. The sequence of *cTrpPT* from *Aspergillus oryzae* DSM1147 is available at GenBank under accession GU722589.

## Acknowledgements

This work was supported by grants from the Deutsche Forschungsgemeinschaft (S.-M. Li), China Natural Science Foundation (30972051/C200103) as well as DAAD and CSC (PPP program). Xie acknowledges the Deutsche Forschungsgemeinschaft for funding the Bruker AVANCE 600 spectrometer.

## References

1 L. Heide, *Curr. Opin. Chem. Biol.*, 2009, **13**, 171–179.
2 S. -M. Li, *Appl. Microbiol. Biotechnol.*, 2009, **84**, 631–639.
3 P. H. Liang, *Biochemistry*, 2009, **48**, 6562–6570.
4 K. Yazaki, K. Sasaki and Y. Tsurumaru, *Phytochemistry*, 2009, **70**, 1739–1745.
5 N. Steffan, *et al.*, *Curr. Med. Chem.*, 2009, **16**, 218–231.
6 S. -M. Li, *Phytochemistry*, 2009, **70**, 1746–1757.
7 S. -M. Li, *Nat. Prod. Rep.*, 2010, **27**, 57–78.
8 A. Grundmann and S. -M. Li, *Microbiology*, 2005, **151**, 2199–2207.
9 N. Steffan, I. A. Unsöld and S. -M. Li, *ChemBioChem*, 2007, **8**, 1298–1307.
10 W. -B. Yin, *et al.*, *ChemBioChem*, 2007, **8**, 1154–1161.
11 A. Kremer and S. -M. Li, *Appl. Microbiol. Biotechnol.*, 2008, **79**, 951–961.
12 W. -B. Yin, *et al.*, *Org. Biomol. Chem.*, 2010, **8**, 1133–1141.
13 W. -B. Yin, *et al.*, *J. Biol. Chem.*, 2009, **284**, 100–109.
14 I. A. Unsöld and S. -M. Li, *Microbiology*, 2005, **151**, 1499–1505.
15 A. Markert, *et al.*, *Plant Physiol.*, 2008, **147**, 296–305.
16 Y. Ding, R. M. Williams and D. H. Sherman, *J. Biol. Chem.*, 2008, **283**, 16068–16076.
17 N. Steffan and S. -M. Li, *Arch. Microbiol.*, 2009, **191**, 461–466.
18 A. Kremer, L. Westrich and S. -M. Li, *Microbiology*, 2007, **153**, 3409–3416.
19 M. Machida, *et al.*, *Nature*, 2005, **438**, 1157–1161.
20 S. Maiya, *et al.*, *ChemBioChem*, 2006, **7**, 1062–1069.
21 A. Ben-Bassat, *et al.*, *J. Bacteriol.*, 1987, **169**, 751–757.
22 T. Meinnel, Y. Mechulam and S. Blanquet, *Biochimie*, 1993, **75**, 1061–1075.

23 L. Y. Luk and M. E. Tanner, *J. Am. Chem. Soc.*, 2009, **131**, 13932–13933.
24 I. A. Unsöld and S. -M. Li, *ChemBioChem*, 2006, **7**, 158–164.
25 A. Grundmann, *et al.*, *ChemBioChem*, 2008, **9**, 2059–2063.
26 A. B. Woodside, Z. Huang and C. D. Poulter, *Org. Synth.*, 1988, **66**, 211–215.
27 J. Sambrook, and D. W. Russell, *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 2001.
28 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.
29 U. K. Laemmli, *Nature*, 1970, **227**, 680–685.
30 S. Berger, and S. Braun, *200 and More NMR Experiments. A Practical Course*, Wiley-VCH, Weinheim, Germany, 2004.
31 T. -L. Hwang and A. J. Shaka, *J. Am. Chem. Soc.*, 1992, **114**, 3157–3159.
32 D. Marion, *et al.*, *J. Magn. Reson.*, 1989, **85**, 393–399.

