Molecular Cloning and Analysis of the Ergopeptine Assembly System in the Ergot Fungus Claviceps purpurea

Telmo Correia,1,3 Nicolas Grammel,1,3,4 Ingo Ortel,2,3,4 Ulrich Keller,2,4 and Paul Tudzynski1,*

1Institut für Botanik Westfälische Wilhelms-Universität
Schlossgarten 3
D-48149 Münster
2Institut für Chemie Fachgebiet Biochemie Technische Universität Berlin
Franklinstrasse 29
D-10587 Berlin-Charlottenburg
Germany

Summary

Claviceps purpurea produces the pharmacological important ergopeptides, a class of cyclo-structured alkaloid peptides containing D-lysergic acid. These compounds are assembled from D-lysergic acid and three different amino acids by the nonribosomal peptide synthetase enzymes LPS1 and LPS2. Cloning of alkaloid biosynthesis genes from C. purpurea has revealed a gene cluster including two NRPS genes, cpps 1 and cpps 2. Protein sequence data had assigned earlier cpps1 to encode the trimodular LPS1 assembling the tripeptide portion of ergopeptides. Here, we show by transcriptional analysis, targeted inactivation, analysis of disruption mutants, and heterologous expression that cpps 2 encodes the monomodular LPS2 responsible for D-lysergic acid activation and incorporation into the ergopeptide backbone. The presence of two distinct NRPS subunits catalyzing formation of ergot peptides is the first example of a fungal NRPS system consisting of different NRPS subunits.

Introduction

D-lysergic acid (1,6,10,6-methyl-ergoline-8-carboxylic acid) is an important pharmacophore that in its amidated form exerts useful activities for therapy for the treatment of migraine, hypertension, or prolactine disorders [1, 2]. The specific effects of the various amides of D-lysergic acid depend on the amide substituents which confer to D-lysergic acid structural similarity to dopamine, serotonin, or dopamine, causing agonist or antagonist behavior against the corresponding neurotransmitter families [3]. Most prominent among the D-lysergic acid amides are the ergopeptides, in which the amide component is a tripeptide chain in the modified form of a bicyclic cyclo-lactam structure (Figure 1). Simpler D-lysergic acid derivatives are ergometrine or D-lysergic acid α-hydroxyethylamide (Figure 1), where D-lysergic acid is attached to small amino alcohols [4, 5].

Main producers of ergopeptides and the D-lysergic acid amides are members of the genus Claviceps [6]. Whereas Claviceps purpurea can produce both groups of D-lysergic acid derivatives, Claviceps paspali produces simple D-lysergic acid amides or paspalic acid, the immediate precursor of D-lysergic acid [7, 8]. Other species, notably Claviceps fusiformis, lack the ability to synthesize D-lysergic acid, producing clavine alkaloids instead, which have still simpler structures [9]. The main representatives of the clavines represent intermediates in the ergot alkaloid biosynthesis (for review see [4, 5]). Figure 2 shows the various steps of the ergoline pathway from the formation of dimethylallyl tryptophan (DMAT) to the conversion of paspalic acid to D-lysergic acid. DMAT synthase (DMATS) has been the first enzyme of the ergot alkaloid pathway that was purified to homogeneity [10, 11]. The DMATS-encoding gene (dmw) was the first cloned alkaloid biosynthesis gene from the clavine-producing species Claviceps fusiformis [12]. Importantly, an ortholog of dmw named cpd1 was shown to be present in ergopeptide-producing Claviceps purpurea, which proved the common biosynthetic origin of the clavines and the D-lysergic acid-derived group of alkaloids [13].

Previously, investigations of the assembly of D-lysergyl tripeptide lactams, the immediate precursor of ergopeptides, in C. purpurea have shown that they are formed by a large nonribosomal peptide synthetase (NRPS) multienzyme that, during purification, separated in two enzyme activities [14]. The two enzymes, D-lysergyl peptide synthetase 1 and 2 (LPS1 and 2) have sizes of 370 and 140 kDa, respectively. Synthesis starts with binding of D-lysergic acid to LPS2 as thioester followed by transfer to LPS1, where three successive condensations into the D-lysergyl mono-, di-, and tripeptide thioester intermediate take place, with final release of the end product D-lysergyl tripeptide lactam [15]. These data indicated that LPS1 and LPS2 would contain four D-lysergic acid modules, LPS2 would represent the D-lysergic acid module (Figure 2).

Sequencing the flanking regions of the recently cloned DMAT synthase gene cpd1 from C. purpurea led to the identification of further genes most probably involved in ergoline ring synthesis (Figure 3). Importantly, downstream to cpd1 the gene encoding the trimodular enzyme LPS1, cpps 1, was identified, indicating the presence of an ergot alkaloid gene cluster [17]. This finding suggested that further chromosome walking on the flanking regions of cpd1 would disclose the gene of LPSC, for which protein sequence data were not available. In this report, we describe the genetic organization of an as yet unidentified region of the ergot alkaloid.
cluster with a novel NRPS gene cpps2, which by genetic and biochemical characterization was shown to encode LPS2.

Results

Sequence Analysis of LPS1, Prediction of the Structure of LPS2, and Cloning of cpps2

Detailed analysis of the amino acid sequence deduced from the gene sequence of ccpps1 had shown that LPS1, as expected from previous biochemical work, contains three A and three PCP domains (also called T domains) (Figure 3A). However, the sequence also revealed that LPS1 has only two C domains, i.e., in module 2 and module 3. No such C domain was seen aminoterminally to the first A domain in module 1. This suggested that the C domain for formation of the peptide bond between D-lysergic acid and the first amino acid of the cyclol peptide chain was located either on LPS2 or on another as yet unknown protein component of the LPS multienzyme complex. Another interesting feature of the deduced amino acid sequence of ccpps1 is that LPS1 has a carboxy-terminal domain with no similarity to the typical release domain of NRPS, such as the thioesterase (Te) domain [18, 19]. Instead, this domain has some similarity to C domains as well as to heterocyclization domains (Cy domains) of NRPS [16]. However, it is ~70 amino acids shorter than the typical C and Cy domains and shows significant differences in the signature sequences (C1 to C6 and Cy1 to Cy6, respectively [16]) of C and Cy domains. In particular, the C3 motif of the LPS1 carboxy-terminal domain, Q-R-A-Q-D-G-V-S, differs from both the C3 consensus H-H-x-I-S-D-G-W and Cy3 consensus D-x-x-x-x-D-x-x-S of the C and Cy domains, respectively (Figure 3B). The double H of the C3 motif as well as the first D of the Cy3 motif are missing, which suggests a different mechanism of this domain. Therefore, this domain cannot be regarded as a regular NRPS C domain. In view of the release of the D-lysergyl tripeptide chain from LPS1 as an acyl-diketopiperazine, it is most likely involved in product release catalyzing the lactam formation step from D-lysergyl tripeptide-PCP thioester. We designated this domain as the cyclization domain (Cyc domain).

In order to identify the gene of LPS2, which from the biochemical data should be composed at least of an A and a PCP domain, a chromosome walking was performed downstream of the gene cpox2. During this, several ORFs could be detected [4]; sequence analyses led to the identification of a putative P450 mono-oxygenase gene (cpP450-1), a putative catalase gene (cpcat2), a putative oxidase gene (cpox3), and, interestingly enough, a gene showing significant homology to fungal NRPS, which was named ccpps2 (Figure 3A). This gene (GenBank accession number AJ439610) has a coding sequence of 3924 bp interrupted by an intron of 65 bp, leading to a putative protein of 1308 amino acids with a calculated Mr of 140 kDa.

Since ccpps2, according to the size of the encoded protein and its vicinity to other ergot alkaloid biosynthesis genes in C. purpurea strain P1, was a possible candidate gene for LPS2, its deduced amino acid sequence was analyzed in respect to the overall structure and domain composition. Alignments of Ccpps2 with various NRPS sequences in the database revealed that it is composed of four domains (Figure 3A). An A domain in the protein sequence commences at a distance of 280 amino acids from the amino-terminal end. The A domain is followed by a typical PCP domain (T) of about 100 amino acid residues, whereas in the carboxy-terminal portion of Ccpps2 a domain of 500 amino acids is present, showing high similarity to the regular C domains of peptide synthetases. Interestingly, the amino-terminal domain of Ccpps2 with about 280 amino acids has no similarity to known proteins except for a stretch of ~100 amino acids between the amino acid positions 40 to about 140 with weak homology to the carboxy-terminal half of C domains. Based on these findings, ccpps2 was considered the gene of LPS2, particularly because of the presence of the carboxy-terminal C domain that exactly matches what would be expected from the biochemical investigations of the LPS system and the structure and size of LPS1.

Expression Analysis of ccpps2 in C. purpurea

In order to show the significance of Ccpps2 for D-lysergyl peptide synthesis in C. purpurea, Northern blot analyses of total RNA using the ccpps2 gene were performed. It
is known that ergot alkaloid synthesis is regulated by the presence of phosphate and that repression by high phosphate concentrations can be relieved by tryptophan, an inducer of ergot alkaloid synthesis [20, 21]. Indeed, we could show recently that several genes of the ergot alkaloid cluster (Figure 3A), including the gene encoding DMATS (cpd1) and the oxidase genes (cpox1, cpP450-1), are coregulated in terms of their dependence of expression on the phosphate concentrations in the medium, which confirms their potential role in alkaloid biosynthesis (Y. Lübke, T.C., and P.T., unpublished data). Expression analysis of the cpss2 gene therefore was performed with mycelia grown under low- and high-phosphate conditions (see Experimental Procedures). The data presented in Figure 4 clearly show a transcript of cpss2 of the expected size which is upregulated under low phosphate condition, comparable to the behavior of cpd1 and cpss1 as controls. This regulation pattern correlated with alkaloid production of the cultures (data not shown).

Functional Analysis of cpss2
To show the direct involvement of cpss2 in ergopeptine assembly, gene inactivation was performed by a gene replacement approach: cpss2 was interrupted by a phleomycin-resistance cassette as shown in Figure 5. From the 5’ and 3’ genomic region of cpss2, two fragments were cloned upstream and downstream, respectively, of the phleomycin-resistance cassette of vector pAN8-1UM. The ApaI/BamHI fragment carrying the whole replacement construct was used to transform C. purpurea strain P1. Phleomycin-resistant transformants obtained were individually checked for the presence of the replacement fragment in the genomes (generated by double crossover events) by PCR as described in Experimental Procedures (Figure 5A). From a total of 98 primary transformants, 26 showed the correct “diagnostic” PCR fragments of 1692 bp (generated with primers p2HUX and P2HLX) and 1829 bp (primers P2HUCP/ P2HLCP). This is a rather high frequency (27%) when compared to the previous values of 1%-2% observed in
Chemistry & Biology

Figure 3. Map of the Ergot Alkaloid Biosynthesis Gene Cluster in C. purpurea Strain P1 and Section of an Amino Acid Alignment of Cyc, C, and Cy Domain Regions of NRPS
(A) Location of cpps1 and cpps2 in the ergot alkaloid biosynthesis gene cluster and the deduced structures of LPS 1 and LPS 2. Cpcat2, cpP450, cpox3, and cpps2 have been sequenced during this study, whereas the other genes were sequenced previously [17]. The narrow white boxes in cpps1 and 2 denote the introns in these genes.
(B) Conserved motif HHxxxDxxS from regular C domains of LPS1, the corresponding Cyc domain motif QRAQFDxxS from the carboxy-terminal domain of LPS1, and the corresponding Cy motif DxxxxDxxS from heterocyclization domains of NRPS. Sequences are from the two internal C domains of LPS1 (GenBank accession number CAB39315), the Cy domains of pyochelin synthetase (PchF1) (GenBank AAD55801), of bleomycin synthetase (BlmIV) (GenBank AAG02364), and from the Cyc domain of LPS1 (GenBank CAB39315).

Figure 4. Northern Analysis of cpps2 in High- and Low-Phosphate Conditions and Coregulation of Other Genes of the Ergot Alkaloid Cluster
Wild-type mycelia were grown under low- and high-phosphate conditions (see Experimental Procedures), RNA was extracted and, after separation in a agarose gel, was blotted on a nylon membrane. Probes used for hybridization were as follows: 4.0 kb HindIII fragment (cpps1, 17), a 0.8 kb SalI fragment (cppd1, 17), and a 3.5 kb SalI fragment (cpps2, Figure 3). A plus sign denotes RNA under low-phosphate conditions, and a minus sign indicates high-phosphate conditions.

knockout experiments using the pathogenic field isolate Claviceps purpurea 20.1 [17, 22, 23]. The identity of the PCR fragments was confirmed by sequence analysis (data not shown). It is known that C. purpurea P1 and its parent strain ATCC 20102 are polykaryotic, harboring more than 10 nuclei per cell compartment [24]. From this, it is clear that single integration events always lead to heterokaryons. In fact, all of our positive primary transformants also showed the wild-type fragment. To obtain homokaryotic segregants, a positive transformant was propagated from which single-cell units had to be prepared by protoplast subculturing and hyphal tip isolation because of the inability of P1 to produce conidia (see Experimental Procedures). From a total of 600 subcultures thus generated, one homokaryotic mutant could finally be isolated. This transformant was lacking the wild-type PCR fragment containing the diagnostic band only (shown for the left flank of disruption construct in Figure 5B); it was named Δcpps2-1. Southern analyses confirmed the PCR data (Figure 5C) Final proof for the inactivation of cpps2 in the Δcpps2-1 mutant came from Northern analysis which showed that no cpps2 transcript could be detected in the mutant under conditions that stimulate expression of cpps2 in strain P1 (Figure 5D).

Analysis of the Alkaloid Spectrum of Δcpps2-1
Since strain Δcpps2-1 represents a bona fide deletion mutant, it was analyzed for its ability to form ergopep-
The Ergopeptine Assembly System in *C. purpurea*

**Figure 5. Functional Analysis of ccpps2**  
(A) Gene replacement strategy; the design of the replacement fragment (upper part) and the genomic situation after a successful gene replacement by double crossover (lower part) are presented. Primer binding sites are indicated by arrow heads (for details see Experimental Procedures). The narrow white box indicates the intron in *ccpps2*.

(B) PCR analysis of a Δccpps2 mutant and the wild-type strain (P1) using diagnostic primers for the left flank (HI) and wild-type primers (WT). Left panel, wild-type primers; right panel, with "left flank" primers. For details, see Experimental Procedures.

(C) Southern analysis of a replacement transformant: genomic DNA of the wild strain and mutant Δccpps2-1 was digested with ClaI, transferred to a nylon membrane, and hybridized to a 3.5 kb SalI fragment of the *ccpps2* region (see Figure 3). The labels a, b, and c denote the hybridizing ClaI fragments expected to be present in Δccpps2-1 and wild-type P1, respectively.

(D) Northern analysis: total RNA from wild-type and mutant Δccpps2-1 was extracted from mycelia grown under alkaloid production conditions, blotted on a nylon membrane, and hybridized to the *ccpps2* probe (see C).

The Ergopeptine Assembly System in *C. purpurea*  

Tines and other alkaloids. Cultivation of the mutant, along with an ectopic integration mutant and the recipient strain P1 as controls, was in standard alkaloid production medium with limiting phosphate concentrations. Analysis of the individual components of the alkaloid mixtures was performed by TLC, HPLC, or electrospray mass spectrometry. Figure 6 shows that strain P1 produces ergotamine, ergocryptine, and ergosine as main alkaloids. Moreover, from the clavines, elymoclavine was detected in significant amounts (data not shown). By contrast, the Δccpps2-1 mutant did not produce any ergopeptine on solid or in liquid medium but accumulated D-lysergic acid, which is absent from strain P1, indicating accumulation of that compound due to blockage of ergopeptine synthesis. Radioisotope labeling of ergopeptines using radioactive amino acids in short-term incubations of washed mycelium and TLC separations of the radioactive alkaloid mixtures in P1 or the Δccpps2-1 mutant confirmed these results (data not shown). The fact that strain Δccpps2-1 accumulated D-lysergic acid instead of forming ergopeptines clearly indicated that ergoline ring synthesis was unaffected by gene disruption in *ccpps2* (Figure 6).

**Enzymatic Analysis of Δccpps2-1**

Protein fractions of the Δccpps2-1 mutant along with protein fractions from the parent strain P1 were tested for the presence of LPS1 and LPS2 by measuring the ability to form enzyme thioester with 14C-valine or 14C-phenylalanine and 3H-dihydrolysergic acid, respectively. To this end, protein from extracts of broken cells of the two strains was fractionated by ammonium sulfate precipitation and subjected to gel filtration on Ultrogel AcA34. Testing fractions from the gel filtrations from both strain P1 and the mutant gave a peak of valine or phenylalanine thioester formation shortly behind the void volume of the column, which represented LPS1 activity (Figure 7). The peak of activity of dihydrolysergyl enzyme thioester of LPS2 in strain P1 could be detected separately from that of LPS1 in the fractionation range of about Mr 150,000. In SDS-gel electrophoretic separations of the fraction containing dihydrolysergyl thioester-forming activity, a ~140 kDa band could be visualized by autoradiography after incubation with radioactive dihydrolysergic acid and ATP (Figure 7). By contrast, no dihydrolysergic acid thioester formation was detected in protein extracts of the Δccpps2-1 mutant, which con-
Figure 6. Mass Spectrometric Analysis of Alkaloid Extracts from Strain C. purpurea Strain P1 and the Δcpps2 Mutant

Extracts from each strain grown on solid or liquid media were subjected to electrospray mass spectrometry as described in Experimental Procedures.

(A and B) HPLC separations of extracts from strain P1 and the Δcpps2 mutant, respectively.

(C–E) Mass peaks of the indicated peaks from (A). (C), ergosin m/z = 548 [M + H]; (D), ergotamin m/z = 582 [M + H]; (E), ergocryptin m/z = 576 [M + H].

(F) Mass peaks of compounds accumulated in (B) (D-lysergic acid m/z = 269 [M + H]).

Obtained solely LPS1 at a level comparable to strain P1 (Figure 7). This represents good evidence that cpps2 is encoding LPS2.

Heterologous Expression of cpps2 in E. coli
To characterize the gene product of cpps2 further, cpps2 was engineered for expression as a hexaHis fusion protein in E. coli. After verification of the intron of cpps2 by RT-PCR, which revealed the correctness of the intron-exon junctions inferred from analysis of the genomic DNA sequence (data not shown), a cDNA was constructed via PCR amplification of fragments of the exon-encoding regions of cpps2, allowing the two exons to join via an EcoRV restriction site (see Experimental
Figure 7. Comparative Functional Analysis of LPS2 in Protein Fractions from C. purpurea Strains P1 and Δcpps2-1 and from E. coli (A) Analysis of protein extracts of C. purpurea strains P1, Δcpps2 mutant, and E. coli carrying plasmid cpps2_pNG5 (cpps2) for the presence of LPS2 activity. I and II show gel filtration on Ultragel AcA 34 of an ammonium sulfate-fractionated protein extract from C. purpurea strain P1 (I) producing ergopeptines and C. purpurea Δcpps2 (II). Each fraction was assayed for the presence of LPS1 and LPS2 by the thioester formation assay with 14C-phenylalanine (or 14C-valine) (circles) and 3H-dihydrolysergic acid (squares), respectively. Protein is denoted by diamonds. The inset in I shows an autofluorograph of an SDS polyacrylamide gel electrophoretic separation of the peak fraction containing LPS2 activity after incubation enzyme with 3H-dihydrolysergic acid and ATP. No LPS2 activity is present in C. purpurea Δcpps2. III shows gel filtration on Superdex 200 of recombinant LPS2-32 partially purified from E. coli after expression from plasmid (cpps2_pNG5). Fractions were tested for the ATP-pyrophosphate exchange dependent on D-lysergic acid and ATP. No LPS2 activity is present in C. purpurea Δcpps2. (B) Alignment of substrate specificity determining amino acid residues in binding pockets of LPS2 and several NRPS adenylation domains with known substrates. Shown are the extracted residues from the LPS2 A domain and the A domains of NRPS A domains and Aryl-AMP-ligases (arylcarboxylate activating enzymes). D-LSA, D-lysergic acid. Numbers refer to the corresponding residues in the PheA domain of gramicidin synthetase 1 [34, 35]. CepB, chloroeremomycin synthetase (GenBank accession number CAI11795); Cda, CDA peptide synthetase I (GenBank CAB38518); SafAI, saframycin Mx1 synthetase A (GenBank AAC41129); SnbA, pristinamycin synthetase I (GenBank CAA67140); YbtE, yersinia-bactin synthetase (GenBank AAC69591); ACMS I, actinomycin synthetase I (GenBank AAD30111); EntE, enterobactin synthetase (GenBank AAN79156); HPG, dihydroxyphenylglycine; 3h4mPhe, 3-hydroxy-4-methyl-phenylalanine; HPk, 3-hydroxypicolinic acid; 4-MHA, 4-methyl-3-hydroxyanthranilic acid; DHB, 2,3-dihydroxybenzoic acid, Sal, salicylic acid. 

Procedures. Since the intron is located in the region of the gene encoding the amino-terminal domain, the resultant change of the amino acid sequence (LF→DI) was not expected to disturb the activity of the A domain. Two constructs were prepared on the basis of expression vectors pOE 32 (LPS2-32) and pOE 70 (LPS2-70) with a hexaHis-encoding sequence at the 5’ and the 3’ ends, respectively. Expression was in E. coli M15.
Attempts to purify the soluble fraction of either enzyme on Ni-NTA matrix failed due to lack of binding to the affinity matrix. Using the D-lysergic acid-dependent ATP-pyrophosphate exchange as assay, LPS2-32 was partially purified by a classical purification protocol [25] involving ammonium sulfate precipitation and several steps of gel filtration and anion exchange chromatography (Figure 7). In each separation step, there was a peak of D-lysergic acid activation which was absent from the nontransformed strain (data not shown). Like the native LPS2 enzyme, LPS2-32 did not activate any of the amino acids present in the peptide chain of D-lysergyl peptide lactam or tryptophan (which, like D-lysergic acid, contains an indole ring as characteristic structural element). Attempts to demonstrate dihydroxypergacyclic acid enzymethioester formation with LPS2-32 failed. Most probably, LPS2-32 does not contain a 4’-phosphopantetheine cofactor necessary for covalent binding of the substrate as thioester. Incubations of the protein with two different 4’-phosphopantetheine transferases such as Sfp [26] or OrfC [27] in the presence or absence of coenzyme A with 3H-dihydroxypergacyclic acid did not show any increase in bound radioactivity.

**Discussion**

The results presented here show that the ergot alkaloid gene cluster from *Claviceps purpurea* harbors two NRPS genes, *cpps1* and *cpps2*. *Cpps1* had been identified earlier as the gene encoding the nonribosomal peptide synthetase LPS1 (370 kDa), which is responsible for to reside at the amino-terminal end of LPS1 necessary earlier as the gene encoding the nonribosomal peptide to a role attributed to the C domain originally believed

The analyses showed that LPS2 is a stand-alone initiation module analogous to bacterial NRPS systems such as the phenylalanine recruiting module for initiation of gramicidin S or tyrocidine in *Bacillus brevis* [16] or the aryl carboxylic acid recruiting composite EntE/EntB in the initiation of siderophore synthesis in *E. coli*, *Versinia pestis*, or in *Vibrio cholerae* [29]. Such initiation modules or units consist at least of an A and a PCP domain, and they normally interact in trans with their cognate C domain located on the subunit harboring the next module of the biosynthetic sequence catalyzing the elongation step following activation of the starter residue. In bacterial NRPS systems, even such C domains can occur as autonomous proteins, e.g., the stand-alone VibH, which condenses enzyme thioester-activated aryl carboxylic acids with a free amine during the biosynthesis of the siderophore vibriobactin in *V. cholerae* [30]. However, the analysis of the deduced amino acid sequence of *cpps2* revealed that in contrast to the integral or, even more, to the composite bacterial initiation modules, LPS2, in addition to its A and PCP domain, does contain an integral “regular” C domain at its carboxyl terminus in the arrangement (A-PCP-C). From the deduced domain arrangement and calculated size (140 kDa, 1308 aa), the *cpps2* gene product meets all criteria required to fulfill the role of LPS2 [14]. In particular, the presence of the C domain at the carboxyl terminus of LPS2 readily points to a role attributed to the C domain originally believed to reside at the amino-terminal end of LPS1 necessary for condensation of D-lysergic acid with the amino acid activated by the first module of LPS1. The domain arrangement (A-PCP-C) of the *cpps2* gene product is nonetheless unusual in NRPS systems, because C domains are usually cis-acting domains in respect to their downstream A and PCP domains [19]. This may be a consequence of the observed substrate specificity of C domains, which has been reported to be high for the substrate bound to the downstream PCP-bound substrate and low for the substrate from the upstream module [32]. Accordingly, elongation subunits in NRPS very often commence at their amino-terminal ends with a C domain, which reflects the observed functional intimacy between C domains and their downstream A domains. In the LPS system, the D-lysergic acid amidating C domain would therefore act in trans to its downstream A domain. The significance of the carboxy-terminal trans-acting C domain of LPS2 is not clear. Whether it is a mere exception to the rule or may have significance in the light of the very high substrate specificity of LPS2 for D-lysergic acid [14] remains to be seen in the future, when swapping of that C domain between LPS2 and LPS1 will become possible. The only known example of a stand-alone module having a carboxy-terminal C domain is the monomodular VibF of vibriobactin synthesis in *Vibrio cholerae* [33]. Like *Cpps2*, *VibF* (domain arrangement Cy1-Cy2-A-C1-PCP-C2) has the carboxy-terminal C domain (termed C2), which is the only catalytically active one (in contrast to C1) in acylating a free small diamine, dihydroxybenzoyl-norspermidine, at both its primary and secondary amino groups with a small peptide gener-
ated by the Cy1-Cy2-A-(C1)-PCP domain assembly upstream of C2. However, the difference to LPS2 lies in the fact that VibF acts in trans acylating a free substrate, whereas LPS2 acts in transacylating the amino acid covalently tethered to the first module of LPS1, which may require specific protein-protein interactions. Comparison of the C domain sequence of Cpps2 (LPS2) with the internal C domain sequences of LPS1 gave no hint for characteristic peptide sequences possibly involved in protein-protein interactions (data not shown).

Direct and unequivocal evidence for the identity of Cpps2 and LPS2 was obtained by gene replacement of ccpps2 in the C. purpurea genome by a phleomycin-resistance cassette that created phenotype accumulating D-lysergic acid instead of ergotamine and other ergopeptines. This clearly indicated a blockade in the assembly of D-lysergyl tripeptide lactam. Comparative enzymatic studies of protein extracts derived from the Δcpps2 mutant and its parent strain P1 showed that the ergopeptine-producing strain P contained both LPS1 and LPS2, whereas extracts from the Δcpps2 mutant were devoid of any LPS2-related activity (Figure 7). Moreover, Northern analysis of RNA obtained from cultures grown in low and high concentrations of phosphate indicated the same regulation pattern for ccpps2 as for cpps1 and other genes located in the gene cluster that show growth-linked repression by phosphate (e.g., the DMAT synthase gene cdpe1) concomitantly with lowered alkaloid production in cultures of Claviceps sp. [20] (Y. Lübbe, T.C., and P.T., unpublished data). These findings indicate a clear involvement of the gene product of ccpps2 in D-lysergyl peptide assembly and also coregulation with expression of cpps1 by phosphate. This is particularly important in view of the fact that the known fungal NRPS consist of single polypeptide chains, which warrants equimolarity of the various modules in these multienzymes. In order to warrant such conditions of balanced levels of the different modules for the LPS system, coordinate transcription of cpps1 and ccpps2 is obviously necessary.

Final proof for the identity of Cpps2 with LPS came from expression experiments with ccpps2 cDNA in E. coli which yielded a protein of the expected size of 140 kDa that specifically catalyzed the D-lysergic acid- (and dihydrolysergic acid-)dependent ATP-pyrophosphate exchange (Figure 7). Like wild-type enzyme, the enzyme did not activate any of the amino acids of ergopeptines or tryptophan. In contrast to its ability to catalyze the ATP-pyrophosphate exchange dependent on D-lysergic acid, the LPS2 from E. coli did not catalyze the enzyme-thioester formation with dihydrolysergic acid, as did the wild-type enzyme. We argue that this is probably due to the lack of 4′-phosphopantetheine cofactor. Possibly, the LPS2-apoenzyme may be a specific substrate for a Ppant transferase from the ergot fungus C. purpurea that is not yet known.

The inspection of the specificity-determining region of the A domain (amino acid binding pocket) of LPS2 by comparison with the amino acid pocket of the PheA domain of tyrosidine synthetase I according to previously published procedures [34, 35] indicated uniqueness of the D-lysergic acid binding pocket (Figure 7B). It showed similarity to some A domains activating hydrophobic amino acids (best 50% score with the p-hydroxyphenylglycine-activating module of chloroeremomycin synthetase [36]). Furthermore, no convincing similarity of the LPS2 A pocket residues was seen with the corresponding residues characterizing the A domain pockets of aryl-AMP ligases [37]. Strikingly, there was also no evident similarity with the specificity-determining residues of tryptophan-activating A domains of various NRPS, suggesting that the indole portion may not play a role in the recognition of D-lysergic acid by LPS2 (Figure 7B). Since the best matches were observed with the pocket residues of A domains catalyzing hydrophobic amino acids, it may be suggested that the LPS2 A domain substrate binding pocket resembles typical A domains binding amino acids (e.g., the ones of LPS1) rather than the binding pockets of A domains activating aryl carboxylic acids. Possibly, therefore, ring C and D of D-lysergic acid, which have no aromatic character, would contribute to recognition by the A domain (Figure 1). The indole ring system of D-lysergic (ring A and B) as recognition element would play a minor role, being too distant from the carboxyl and the methylamino group in ring D. Unfortunately, no substrate analogs of D-lysergic acid were available to test this hypothesis, and one must therefore await structure data of the LPS2 A domain in the future that will give insight into the molecular details of D-lysergic acid recognition and binding.

The fact that LPS1 and 2 are distinct enzymes raises the question of whether this exception to the hitherto described fungal NRPS systems is of vital importance for Claviceps purpurea and ergot fungi in general. It is known that D-lysergic acid is constituent not only of ergopeptines but also of the simpler D-lysergic acid peptides and amides such as ergosecaline or ergometrine and D-lysergic acid α-hydroxyethylamide, respectively [4]. A number of C. purpurea strains can produce both ergopeptines and ergometrine simultaneously [6], which indicates the presence in one strain of two different biosynthesis systems, both using D-lysergic acid as substrate. Ergometrine is derived from D-lysergyl alanine [38], for which a D-lysergic acid module like LPS2 and an alanine module with an additional releasing domain would be required. In strains producing both ergopeptines and simple D-lysergic acid amides, LPS2 could play a central role by interacting either with LPS1-type enzymes or alanine modules in some kind of natural combinatorial biosynthesis. This could at least partly explain why the D-lysergic acid module stands alone and is not part of a multimodular NRPS, because its standing alone would easily permit its integration and participation in different enzyme systems. Moreover, the presence or absence of such genes of LPS2- and LPS1-like enzymes may facilitate the understanding of the basis of diversity in product formation in various species of ergot fungi.

**Significance**

The results presented here demonstrate unequivocally that alkaloid peptides are assembled by an NRPS system consisting of two distinct multienzymes (LPS1 and LPS2) harboring three and one modules, respectively.
This is in contrast to known fungal NRPS systems which harbor all their modules on one polypeptide chain. Cloning of the ergot alkaloid gene cluster of *C. purpurea* revealed the presence of the corresponding genes, *cpps1* and *cpps2*, in an alkaloid biosynthesis gene cluster. Analysis of *cpps2* along with *cpps1* by transcription analysis and targeted inactivation in *C. purpurea* revealed coordinate expression with other genes in the cluster upon induction of alkaloid synthesis in cultures. Enzymatic analysis of a Δcpps2 mutant and of the heterologously expressed protein gave clear evidence that *cpps2* encodes LPS2, the D-lysergic acid-activating module which, together with LPS1, assembles the D-lysergyl tripeptide precursors of ergopeptides. The existence of LPS2 as a monomodular enzyme points to its role in other D-lysergyl peptide assemblies in ergot fungi such as the D-lysergyl-alkylamides which may determine the spectrum of products formed. Thus, the observed accumulation of D-lysergic acid in Δcpps2 mutants points to the possibility that the ergot alkaloid gene cluster in different ergot fungi may differ from each other by the presence or absence of genes involved in steps of ergoline synthesis or peptide assemblies. The availability of the independent D-lysergic acid module LPS2 and its gene *cpps2* opens new and interesting perspectives for the understanding of the biosynthesis of the wide spectrum of naturally occurring D-lysergic acid amides and for design of new pharmacological lysergic acid derivatives by combinatorial biosynthesis.

**Experimental Procedures**

**Strains and Cultures**

*C. purpurea* strain P1 (1029/NS) was described previously [21]. It produces mainly ergotamine, with minor amounts of ergocryptine and ergosine. Maintenance and culture conditions for strain P1 and the replacement mutant Δcpps2-1 (see Results) were as described [14, 17].

**Chemicals and Radiochemicals**

[9,10-3H]-9,10-dihydroergocryptine (17.5 Ci/mmol) was from Hartmann Analytics, Braunschweig (Germany). L-[U-14C]-phenylalanine (443 mCi/mmol), L-[U-14C]-alanine (148 mCi/mmol), and L-[U-14C]-valine (260 mCi/mmol) were from Amersham International. [9,10-3H]-9,10-dihydrolysergic acid was prepared from [9,10-3H]-9,10-dihydrolysergoline [14]. All other materials were of the highest purity commercially available.

**Methods of Analysis**

Radioactive determinations were as described [14]. Protein concentrations were determined according to [39]. SDS-polyacrylamide gel electrophoresis was performed according to [40]. Autoradiography of radiolabeled samples in SDS-polyacrylamide gels was as described [14]. Immunoblot analyses were performed by standard procedures. Enzyme assays were based on the ATP-pyrophosphate reaction dependent on D-lysergic acid and amino acids or on enzyme-thioester formation from dihydrolysergic acid or the relevant amino acid substrates of LPS1 [14]. TLC for separation of alkaloid mixtures was performed on silica gel plates using ethyl acetate:methanol:water (75:15:5, by volume) or ethyl acetate:dimethylformamide:ethanol (75:15:5, by volume) [14].

**Mass Spectrometry**

Mass spectrometric detection was carried out using an Esquire3000+ ion trap instrument (Bruker Daltonik GmbH, Germany) equipped with a API-ESI interface. Ions were scanned with a scan speed of 13,000 Da/s. Using ESI in positive ion mode, mass spectra were acquired from m/z = 100 to 700. For spectra acquisition, a total of 10 scans were summarized. The HPLC system consisted of a HP series 1100 solvent delivery system (Waldbronn, Germany). HPLC separation was carried out on a C18 column (100 x 2 mm, phenomenex LUNA 3 μ) at a flow rate of 300 μl/min, operated at 30°C. Compounds were eluted using a gradient (A = H2O, 0.05% formic acid, B = acetonitril, 10 to 100% B in 30 min).

**Nucleic Acid Extraction and Analysis**

Standard recombinaton DNA methods were performed according to [41]. The plasmids used for subcloning fragments were pUC19, pT728 U (Amersham Pharmacma Biotech), or pSP72 (Promega). Expression plasmids were pQE32 and pQE70 (Qiagen). *Escherichia coli* strains used for cloning were *TOP10F* (invitrogen) and, for expression, M15 (Qiagen). All of these strains were grown according to standard protocols. Extraction of genomic DNA and of RNA, Southern and Northern analyses, and DNA sequencing were performed as described [23]. Sequence comparisons, multiple sequence alignments, and identity scores were computed with Megalign (DNASTAR), BESTFIT (HUSAR), or CLUSTAL X [42]. For PCR analyses, REDTaq (Sigma-Aldrich), Bio Therm (Genecraft), and Vent DNA polymerase (Biolabs) were used according to the manufacturer’s instructions.

**Cloning of *cpps2* and Generation of a Replacement Vector**

A genomic λEMBL3 library of *C. purpurea* strain P1 (1029/NS) was probed with a 1.8 kb fragment of the regulation plasmid pCSS7Re (pHindIII containing the 3’ end of the *cpps2* gene using the plaque filter hybridization technique [41]). Of 69,000 screened plaques, 4 hybridizing λ clones were obtained and purified. One of them, λ63, was digested with several restriction enzymes (BamHI, EcoRI, HindIII, SalI, and XbaI), and the resulting fragments were further cloned into the pUC19 cloning vector. The overlapping subclones were sequenced to obtain the complete coding sequence of *cpps2*. For construction of the *cpps2* replacement vector, a 1.35 kb CiaI/PstI fragment of plasmid p6353mal6.4 (containing the whole coding region of *cpps2*) was subcloned into the pAN81 UM [43] to yield p.*cpps2*.*cpps2*. Digestion with Apal and NotI gained a 4.6 kb fragment including the 5’ region of *cpps2* and the phleomycin-resistance cassette (phl). The 3’ region of *cpps2* located between 2589 bp to 3794 bp (respective to the start codon) was amplified by PCR using primers PS2U and PS2L and cloned into the PC2R1 TOP vector to yield p.*cpps2*.*cpps2*. This plasmid construction was linearized with Apal and NotI and ligated with the Apal/NotI fragment of p.*cpps2*.*cpps2*, resulting in the circular replacement vector p.*cpps2*.*cpps2*. The replacement fragment was excised with Apal/BamHI and used to transform *C. purpurea* (see Figure 3).

**Transformation of *C. purpurea* and Molecular Characterization of Transformants**

Transformation of *C. purpurea* strain P1 was performed as follows: a mycelial suspension of strain P1 was spread on the surface of cellophane sheets on BII agar [44] and grown for 2–3 days at 28°C. The cellophane sheets were placed on petri dishes containing 15 ml of lytic solution corresponding to SMac buffer (0.2 M potassium maleate [pH 5.2]) containing 15 mg/ml β-D-Glucanase Enzyme (Interspec, USA), so that the mycelia faced downwards. The mycelia could then be easily detached and were incubated for about 2 hr at 80 rpm and 28°C. The protoplasts were filtered through sterile filter funnels (POR1 and POR2). They were further pelleted by centrifugation 10 min at 4500 g, washed twice with STC buffer (0.85 M sorbitol, 10 mM Tris, 50 mM CaCl2), and brought to a final concentration of 10° protoplasts/ml STC. 10° protoplasts were transformed as described [22].

After 8 hr of regeneration at 28°C, protoplasts were overlaid with 10° BII/6 containing 100 μg/ml phleomycin (Cayla, final concentration 33 μg/ml). Positive transformants were screened by transferring them to 100 μg/ml bleomycin selection plates. Due to the inability of the production strain P1 to sporulate, the purification of the mutants was performed by protoplasting (see above) or hyphal tip dissection. Phleomycin-resistant transformants carrying a homoplasmy integration of the replacement construct were identified by PCR using primers P2HX (5’-TGAATGCTCGTGAAACCCAAAAATAA), P2HLX (5’-AAAGGCAAGAAAGAACTCGC-3’) (shown in Figure 4).
performed in order to verify the homologous integration of the re-lia were used in each case, and the amount of starting material was
and transferred to a nylon filter (HybondN

hybridized with a 3.5 kb SalI fragment (p63SalI3,5; cpps2) and a 0.8 kb Sall fragment (p2SalII0,8; cd1)

5A) for the right flank, P2HLCP (5′-CAAGGCGGCTGTCGAAAT−3′), P2HLCP (5′-GCCGACTGCTCCCCTATGG-3′) for the left flank, and 10−100 ng of genomic DNA. The predicted 1692 bp (right) and 1829 bp (left) fragments were amplified from DNA of strain Δcpps2-1. The lack of the wild-type gene copy in the Δcpps2 mutants was checked using the primers P2WX (5′-ACGCGCAAGATGCGCTTGC-3′) and P2HLX (5′-AAAGGCGCAGAAATACGC-3′), which gave rise to a 1922 bp fragment with the wild-type strain. Southern analysis was performed in order to verify the homologous integration of the re-placement vector. Genomic DNA from knockout mutants and wild-type was restricted with ClaI, separated via gel electrophoresis (1%), Ultrogel AcA 34 step was sufficient to show the presence or absence

placement vector. Genomic DNA from knockout mutants and wild-30 g wet weight [14]. For the purpose of the present work, the first

using the primers P2HLX (5′-CTTCCTCGATATCATTACTTGGCG-3′) and 5′-GACATGGCATGTCGACG-3′ (reverse). Next, for completion of ccpps2_pNG2 region was cloned. The loids by a strain of Claviceps paspali Stevens and Hall in

engineered via PCR with genomic cpps2 as template to carry a single Sphi site encompassing the ATG-start codon and an EcoRV site encompassing the 3′ end of the exon. The resulting Sphi-EcoRV gene fragment was ligated for subcloning into plasmid pTZ18 (Promega), resulting in plasmid ccpps2_pNG1. Primers were 5′-TCGGATCATGCAAGAGACG-3′ and 5′-CCGACCTCTTCTAATTATG-3′ (forward) and 5′-GCCGACATGTCGACG-3′ (reverse). The Sphi-EcoRV gene fragment was cloned into plasmid pTZ18 for sequencing and later cloned into ccpps2_pNG1, resulting in ccpps2_pNG2. Primers were 5′-GATGATATTATGGTCGACG-3′ (forward) and 5′-GTCGAATACGCGGCAACCC-3′ (reverse). Next, for completion of ccpps2, ccpps2_pNG2 was linearized with PstI at nt position 1747. Into this site, a natural genomic 3.7 kb PstI fragment comprising the rest of ccpps2 ranging from nt 1747 to 4534 together with ~1 kb of its 3′ region was cloned. The resultant plasmid was named ccpps2_pNG3. Trimming the 3′ region of ccpps2 in the plasmid ccpps2_pNG3 was done by introducing a HindIII cleavage site 219 bp downstream of the stop codon. This was done by synthesizing via PCR an ~300 bp fragment using ccpps2_pNG3 as template with primers TTTGGTAAAGAGATGCAAGAG-3′ (forward) and 5′-CTTCCTCAGATACGTTGTCGACG-3′ (reverse). This Sphi-HindIII cleaved expression vector pQ32 (Quiagen), which resulted in plasmid ccpps2_pNG4. Into the Sphi site of ccpps2_pNG4 the 3855 bp Sphi fragment obtained from Sphi cleavage of ccpps2_pNG3. This gave ccpps2_pNG5 with the engineered ccpps2 gene carrying 5′ hexaHis encoding sequence under the control of the lacZ promoter stemmng from the pQE32 vector. Engineering of ccpps2 for expression in E. coli as a carboxy-terminal hexaHis fusion protein (LPS2-70) involved introduction of a BglII site into the ccpps2 sequence of ccpps2_pNG2 encompassing the last two codons of the gene, including the stop codon. This was done by PCR using the primers 5′-CCG TGG CCT GAT GAA TAA CAC TGC-3′ (forward) and 5′-GTC AGG ATC TAG CAT GAG GAG ATT TGA-3′ (reverse) with pNG5 as template. Several rounds of subcloning with various fragments derived from ccpps2_pNG5 resulted in ccpps2_pIO4 carring the engineered ccpps2 gene with a 3′-hexaHis-encoding sequence in the original stop codon of Cpp2 under the control of the lacZ promoter of pQE70.

Heterologous Expression of ccpps2

Expression of ccpps2 carried by plasmids ccpps2_pNG5 (LPS2-32) or ccpps2_pIO4 (LPS2-70) was in E. coli strain M15. Cultures of M15/

pNG5 (1.6 liters of 2X YT medium, 100 μg/ml ampicillin, 25 μg/ml kanamycin) were grown at 30°C to an A600 of 0.9 and then induced with 0.1 mM isopropylthiogalactoside. Cells were harvested after a further 14 hr of incubation at 30°C.

Enzyme Purifications

The procedure to isolate LPS1 and LPS2 from extracts of Claviceps purpurea strains was as described previously. Seven-day-old myce-
lia were used in each case, and the amount of starting material was 30 g wet weight [14]. For the purpose of the present work, the first

Ultrogo Aca 34 step was sufficient to show the presence or absence of LPS2 in extracts of strain P1 and the Δcpps2-1 mutant, respec-
tively [14]. Purification of recombinant LPS2 from E. coli by Ni-NTA matrix was not possible with the hexaHis-LPS2 versions of LPS2 encoded by ccpps2_pNG5 or ccpps2_pIO4. Therefore, LPS2-30 was chosen for partial purification from E. coli, following some of the steps of purification of the protein from C. purpurea. The various steps involved cell disintegration by rupture in the French press and ammonium sulfate precipitation of the polyim P-clearance supemat-
tant of cell extract (60% saturation). This was followed by gel filtra-

tion on Ultrogo Aca 34 (detection of enzyme activity by the D-lysergic acid-dependent ATP-zyrophophate exchange). Gel filtra-
tion was followed by two rounds of anion exchange chromatography on resource Q and Mono Q (Pharmacia). Finally, enzyme was sub-

jected to gel filtration on Superdex 200 (Figure 7). Buffer conditions were as described previously [14]. Enzyme activity in all separations eluted as a single peak.

Acknowledgments

We thank Ariane Zwintscher for excellent help in the mass spectro-
metric analyses. This work was supported by the Deutsche Forschungsgemeinschaft (Tu 50/7-3; Ke 452/1-12).

Received: August 21, 2003

Revised: October 8, 2003

Accepted: October 14, 2003

Published: December 19, 2003

References


ogy of ergot alkaloids and related compounds. In Ergot Alka-


vatives in medical chemistry and therapy. In Natural Prod-

ucts and Drug Development, P. Krosgard-Larson, C.H. Chris-


vatives: receptor affinity and selectivity. Farmaco 54, 288–

296.


kota X, H.D. Osiewacz, ed. (Heidelberg, Germany: Springer-Ver-


7. Kobel, H., Schreier, E., and Rutschmann, J. (1964). 6-Methyl-1β-


The Ergopeptine Assembly System in C. purpurea


Accession Numbers

The sequence for ccpps2 has been deposited in GenBank under the accession number AJ439610.