Identification and Manipulation of the Caprazamycin Gene Cluster Lead to New Simplified Liponucleoside Antibiotics and Give Insights into the Biosynthetic Pathway

Leonard Kaysser, Liane Lutsch, Stefanie Siebenberg, Emmanuel Wemakor, Bernd Kammerer, and Bertolt Gust

From the 1Pharmazeutische Biologie, Pharmazeutisches Institut, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen and the 2Klinische Pharmakologie, Institut für Pharmakologie und Toxikologie, Universitätsklinikum Tübingen, Ottfried-Müller-Strasse 45, 72076 Tübingen, Germany

Caprazamycins are potent anti-mycobacterial liponucleoside antibiotics isolated from Streptomyces sp. MK730-62F2 and belong to the translocase I inhibitor family. Their complex structure is derived from 5′-O-aminoribosyl-glycyluridine and comprises a unique N-methyldeoxypentapeptide side chain in penetration of the bacterial cell (15, 16). Apparently, the acyl-caprazols (4) represent the most simplified anti-biotically active liponucleosides and a good starting point for further optimization of this class of potential therapeutics.

Although chemical synthesis and biological activity of CPZs and LPMs has been studied in some detail, their biosynthesis remains speculative and only few data exists about the formation of other translocase I inhibitors (17, 18). Nevertheless, we assume that the CPZ biosynthetic pathway is partially similar to that of LPMs, FR-90043 (6), and muraymycins (7) and presents a model for the comprehension and manipulation of liponucleoside formation. Considering the unique structural features of the CPZs we also expect some unusual biotransformations to be involved in the formation of, e.g., the (+)-caprazol.

Here we report the identification and analysis of the CPZ gene cluster, the first cluster of a translocase I inhibitor. A set of gene disruption experiments provide insights into the biosynthetic origin of the CPZs and moreover, heterologous expres-
Manipulation of the Caprazamycin Gene Cluster

ExPERIMENTAL PROCEDURES

Bacterial Strains and General Methods—Chemicals, microbiological, and molecular biological agents were purchased from standard commercial sources. Streptomyces sp. MK730-62F2 and Streptomyces coelicolor M512 (SCP1, SCP2, ΔactIIorf4, ΔredD) and their respective derivatives were maintained and grown on either MS agar (2% soy flour, 2% mannitol, 2% agar; components purchased from Carl Roth, Karlsruhe, Germany) or TSB medium (BD Biosciences). E. coli strains were cultivated in LB medium (components purchased from Carl Roth) supplemented with appropriate antibiotics. Mycobacterium phlei was cultured in nutrient agar (BD Biosciences) and used as an indicator strain in agar diffusion assays for the detection of bioactivity in culture extracts of Streptomyces sp. MK730-62F2, S. coelicolor M512, and their derivatives. DNA isolation and manipulations were carried out according to standard methods for E. coli (19) and Streptomyces (20).

Production, Extraction, and Detection of Caprazamycin Derivatives—50 ml of TSB media was inoculated with spore suspension of Streptomyces sp. MK730-62F2, S. coelicolor M512, or a derivative thereof. The cultures were incubated for 2 days at 30 °C and 200 rpm. For the production of CPZs, 1 ml of the pre-cultures were inoculated into 100 ml of a medium containing 1% soytone, 1% soluble starch, and 2% D-maltose adjusted to pH 6.7 (components purchased from BD Biosciences). The cultures were incubated for 7 days at 30 °C and 200 rpm. For rapid identification of CPZs, cells were harvested and extracted with ice-cold methanol. The extract was directly applied to LC-MS and agar diffusion assay. Partial purification of CPZs was achieved by the adjustment of the culture supernatant to pH 4 and its subsequent extraction with an equal volume of butanol. The organic phase was evaporated and extracts were resolved in 500 µl of methanol. LC-MS/MS analysis was performed on a Surveyor HPLC system equipped with a Reprosil-Pur Basic C18 (5 µm, 250 × 2 mm) column (Dr. Maisch, Ammerbuch, Germany) coupled to a Thermo Finnigan TSQ Quantum triple quadrupole mass spectrometer (heated capillary temperature, 320 °C; sheath gas, nitrogen). For sample separation, a linear gradient from 2 to 40% acetonitrile in aqueous formic acid (0.1%) over 4 min followed by a linear gradient from 40 to 100% acetonitrile in aqueous formic acid (0.1%) over 31 min was used; the flow rate was 0.2 liters min⁻¹ and detection at 262 nm. Positive electrospray ionization ((+)ESI) was performed with electrospray voltage of 3.8 kV and collision-induced dissociation spectra were recorded with collision
Manipulation of the Caprazamycin Gene Cluster

The genes were replaced in \textit{E. coli} BW25113/pIJ790/cpzLK09 by using the PCR targeting system (24). Resulting cosmids were confirmed by restriction analysis. Excision of the cassette was performed in \textit{E. coli} BT340 taking advantage of the FLP/FRT recognition sites adjacent to the resistance cassette (25). Positive cosmids were screened for their apramycin sensitivity and verified by restriction analysis and PCR using primers \textit{B1_test_rv, B1_test_fw, B2_test_fw, B4_test_fw, B5_test_rv, B6_test_rv, and B7_test_rv} (supplemental Table S1). Cosmids cpzWP01 (\textit{Δcpz33–34}), cpzWP02 (\textit{Δcpz32–34}), cpzWP05 (\textit{Δcpz1–4}), cpzWP06 (\textit{Δcpz1–5}), and cpzWP07 (\textit{Δcpz1–6}) were transferred into \textit{E. coli} ET12567 (26) and introduced into \textit{S. coelicolor} M512 by triparental intergeneric conjugation with the help of \textit{E. coli} ET12567/pUB307 (27). Kanamycin resistance clones were selected, confirmed by PCR, and designated as \textit{S. coelicolor} M512/cpzWP01(1–3), \textit{S. coelicolor} M512/cpzWP02(1–3), \textit{S. coelicolor} M512/cpzWP05(1–3), \textit{S. coelicolor} M512/cpzWP06(1–3), and \textit{S. coelicolor} M512/cpzWP07(1–3).

Generation of \textit{Δcpz21} and \textit{Δcpz23} Mutants in \textit{S. coelicolor} M512—Deletion mutants were generated in accordance with the generation of the mutants for the detection of the cluster boundaries. Primer pairs \textit{cpz21_rv/cpz21_fw} and \textit{cpz23_rv/cpz23_fw} were used to amplify the apramycin resistance cassette (supplemental Table S1). The resulting mutants were designated \textit{S. coelicolor} M512/cpzLL06(1–3) (\textit{Δcpz21}) and \textit{S. coelicolor} M512/cpzLL07(1–3) (\textit{Δcpz23}).

To generate the expression plasmids for mutant complementation, \textit{cpz21} and \textit{cpz23} were amplified from cosmid cpzLK09 using primer pairs \textit{cpz21Eco_fw/cpz21Hind_rv} and \textit{cpz23Eco_fw/cpz23Hind_rv} and cloned into the vector pGEM-T (Promega). The genes were subcloned into the EcoRI/Spel sites of expression vector pUWL201 (28) under the control of the \textit{ermE} promoter. This resulted in plasmids pL06 (cpz21) and pL07 (cpz22), respectively. DNA sequencing of these plasmids confirmed the correct sequence of all constructs. For protoplast transformation, the two plasmids were transferred into the non-methylating \textit{E. coli} strain ET12567 and DNA was isolated by standard procedures. Transformation of the \textit{S. coelicolor} mutant strains by polyethylene glycol-mediated protoplast transformation (20) finally generated strains \textit{S. coelicolor} M512/cpzLL06/pL06 and \textit{S. coelicolor} M512/cpzLL07/pL07.

RESULTS AND DISCUSSION

Identification and Cloning of the Caprazamycin Gene Cluster—The unusual structure of CPZs and the lack of information about the biosynthetic origin make it difficult to select genetic probes for the identification of the gene cluster. However, the formation of the permethylated 1-rhamnose moiety is known from other antibiotics like elloramycin (29) and spinosyn (30). Oligonucleotides deduced from a multiple sequence alignment based on the elloramycin methyltransferase \textit{elm}M1 led to the amplification of a partial sequence of \textit{cpz28} with high similarity to sugar O-methyltransferases. Primer walking revealed two adjacent genes \textit{cpz29} and \textit{cpz30} to be homologous to other O-methyltransferases suggesting the presence of the CPZ gene cluster. To our knowledge, this is the first study demonstrating the successful application of degenerated primers for O-meth-
Manipulation of the Caprazamycin Gene Cluster

FIGURE 2. Organization of the caprazamycin gene cluster. The putative assignment of the genes to different steps in the biosynthesis is indicated. Bars above the cluster mark the gene deletions performed in this study. – indicates that the deletion of the respective region led to an abolishment of CPZ production. + indicates that CPZ production was not influenced.

yltransferases for probe development. Commonly, methyltransferases are considered to be too diverse on the nucleotide sequence level and too widely distributed in bacterial metabolism to be useful in the identification of a specific gene cluster.

Perfect matching primers were applied to a genomic library of *Streptomyces* sp. MK730-62F2 constructed in a SuperCos1 vector. Eight positive cosmids of 3000 could be identified and proven to contain overlapping DNA by restriction mapping. Cosmid 31C2 was finally selected for complete shotgun sequencing (nucleotide sequence of the gene cluster has been deposited at GenBank, accession number FJ490409).

Sequence Analysis of the Caprazamycin Gene Cluster—A contiguous 42.3-kb region could be assembled with an average GC content of 70.2%, a typical value for *Streptomyces* DNA. In silico sequence analysis guided by BLAST homology searches (21), conserved protein domain searches (31), and the GC frame plot method (32) revealed 34 candidate genes. A total of 23 open reading frames, designated cpz9–31, were assigned to the CPZ gene cluster putatively encoding for biosynthesis, resistance, transport, and regulatory functions (Fig. 2). Table 1 summarizes the orthologous and proposed functions of the annotated genes. Notably, most of the putative gene products did not show homology to proteins found previously in other secondary metabolite gene clusters that reflects the unusual structure of the CPZs.

As proposed, the CPZ gene cluster would start with cpz9, which encodes for a putative regulator of the AraC family. Most members of this family are positive transcriptional activators containing a helix-turn-helix motif. They are known from sugar degradation and other pathways but are rarely found in gene clusters of secondary metabolism (33). The predicted gene product of cpz22 shows homology to ABC-transporters. Similar proteins can be found in many antibiotic gene clusters and are usually involved in self-resistance and export (34). cpz12 and cpz27 are two putative sugar kinase genes similar to tunicamycin resistance proteins, e.g. TmRd from *Deinococcus radiodurans*, which structure has been reported recently (35). The 2’-, 3’-, and 5’-hydroxy groups of the uridine have been suggested as potential targets for phosphorylation by TmRd, resulting in inactivation of the nucleoside antibiotic tunicamycin.

Cpz10 exhibits similarity to the Fe(II)/2-oxoglutarate-dependent oxygenase family (36). Cpz11 and Cpz26 are two putative methyltransferases that contain conserved S-adenosylmethionine-binding domains (cd02440) but share low overall homology to each other (10% identity/18% similarity). Interestingly, both genes are translationally coupled to possible resistance genes cpz12 and cpz27 by overlap of start and stop codons. cpz13 shows weak homology to aminotransferases and is most likely translationally coupled to the predicted serine hydroxymethyltransferase gene cpz14.

The genes cpz15–23 seem to be co-translated as indicated by the overlap of start and stop codons. This subcluster would encode for Cpz15, another hypothetical Fe(II)/2-oxoglutarate-dependent oxygenase, a putative nucleotidyltransferase Cpz16 and Cpz17, which shows similarity to the glycosyltransferases. Cpz18 seems to belong to the class III aminotransferases, whereas Cpz19 resembles pyrimidine-nucleoside phosphorylases. Cpz20 and Cpz21 are similar to a putative acyl-CoA synthase (TMCL4) and a carboxysterase (TMCL1) from the tautomycetin gene cluster (37). Another possible esterase is encoded by cpz23 the last gene in the proposed subcluster. The deduced gene product of cpz25 is a hypothetical alcohol dehydrogenase.

cpz28, cpz29, cpz30, and cpz31 apparently constitute an operon for the attachment and methylation of a deoxysugar as indicated by probable translational coupling of these genes. They show strong similarity to O-methyltransferases and glycosyltransferases from known antibiotic gene clusters in particular to proteins participating in the formation of efloramicin (29), 38) and spinosyn (39). Both compounds contain the same permethylated L-rhamnose moiety as found in the CPZs.

Interestingly, we could not identify genes for the dTDP-L-rhamnose biosynthesis on the cosmid. This was initially surprising, because all genes for the production of a bacterial secondary metabolite are usually clustered. However, neither the gene cluster of efloramicin (40) nor of spinosyn (41), steffymycin (42), or arranciamycin (43) contain genes for dTDP-L-rhamnose formation. We therefore suggest the genes for the CPZs deoxysugar biosynthesis to be located elsewhere on the genome of the natural producer.

Heterologous Expression of the Caprazamycin Gene Cluster—To investigate whether the genes found on cosmid 31C2 were sufficient for biosynthesis of CPZs we intended to express the cosmid heterologously. For this purpose, the betalactamase gene on the backbone of 31C2 was replaced with an integration cassette of pIJ787 (23, 44) containing the attP attachment site and the integrase gene (int) of phage FCS1, a tetra-cycline resistance gene (tet) and an origin of transfer (oriT) using λ-Red recombination. The generated cosmid cpzLK09 was introduced into *S. coelicolor* M512 by polyethylene glycol-mediated protoplast transformation (20) and three kanamycin
Manipulation of the Caprazamycin Gene Cluster

TABLE 1
Deduced functions of genes in the caprazamycin gene cluster

| Gene | AA | Protein homolog                                                                 |
|------|----|----------------------------------------------------------------------------------|
| cpz1 | 477 | SchA30, Streptomyces sp. SCC 2136                                                |
| cpz2 | 151 | SSEG 002365, Streptomyces sviceus ATCC 29083                                     |
| cpz3 | 278 | SchA31, Streptomyces sp. SCC 2136                                                |
| cpz4 | 513 | SACE 7046, Saccharopolyspora erythraea                                            |
| cpz5 | 419 | hmgA, Streptomyces sp. CL 190                                                   |
| cpz6 | 349 | Naipyra, Streptomyces antibioticus                                                |
| cpz7 | 260 | SACE 3529, S. erythraea                                                         |
| cpz8 | 212 | SACE 5947, S. erythraea                                                          |
| cpz9 | 348 | SSEG 03352, S. sviceus ATCC 29083                                                |
| cpz10| 182 | RS290926, 0489, Roseibacter sp. SK209-2-6                                        |
| cpz11| 210 | SCO 1731, S. coelicolor A3(2)                                                    |
| cpz12| 189 | SACE 10045, S. sviceus ATCC 29083                                                |
| cpz13| 441 | SACE 4299, S. coelicolor                                                        |
| cpz14| 424 | Orf-4, Streptomyces atroolivaceus                                                 |
| cpz15| 274 | Xccc100 2413, Xanthomonas campestris                                              |
| cpz16| 233 | Smed 4814, Sinorhizobium medica                                                   |
| cpz17| 377 | AprG2, Streptomyces tenebrarius                                                   |
| cpz18| 424 | CetH, Actionomyces sp. LU 9419                                                   |
| cpz19| 460 | GK 2312, Geobacillus kaustophilus                                                 |
| cpz20| 354 | TMC4L, Streptomyces sp. CK 4412                                                  |
| cpz21| 500 | TMCL1, Streptomyces sp. CK 4412                                                  |
| cpz22| 1238| SAV 5299, Streptomyces avermitilis MA 6480                                      |
| cpz23| 346 | Avx9X, Streptomyces viridochromogenanes                                             |
| cpz24| 598 | Cja 1569, Cellvibrio japonicus                                                   |
| cpz25| 341 | SAV 2980, S. avermitilis MA 6480                                                 |
| cpz26| 417 | SSDG 05270, S. pristinaespiralis ATCC 25486                                      |
| cpz27| 192 | SSEG 10045, S. sviceus ATCC 29083                                                |
| cpz28| 396 | spnC, Saccharopolyspora spinoa                                                    |
| cpz29| 265 | elmM3, Streptomyces oliveace                                                     |
| cpz30| 419 | elmM1, S. oliveace                                                                |
| cpz31| 396 | elmGT, S. olivace                                                                  |
| cpz32| 543 | Tfu 2432, Thermobifida fusca                                                      |
| cpz33| 556 | SchA32, Streptomyces sp. SCC 2136                                                |
| cpz34| 544 | SchA33, Streptomyces sp. SCC 2136                                                |

Accession number: CAH10130, YP002007296, CAH10131, YP001105278, YP001108055, YP002206529, YP001903823, YP001313540, YP001106493, AAN85510, YP001903823, YP148165, AB194375, ABR9378, NP926476, AAK83171, YP001982049, NP824156, YP002195839, YP002205959, AAG23272, CAD57141, CAD57139, YP169413, YP290488, CAH10132, CAH10133, AAK83171.

Identity/similarity: 84/90, 90/95, 76/84, 57/65, 43/60, 27/41, 26/40, 23/34, 16/30, 36/53, 33/49, 12/14, 44/58, 35/46.

Proposed function: Feruloyl-CoA synthase, Acyl dehydratase, Transcriptional regulator, PAPS 3-phosphatase, Hypothetical protein, HMG-CoA synthase, Type III polyketide synthase, β-Hydroxylase, Methyltransferase, Kinase, Aminotransferase, Hydroxymethyltransferase, Aminotransferase, Pyrimidine phosphorylase.

* Overall homology (%).
* Gene lacks an appropriate stop codon and is considered to be incomplete on cosmid 31C2 cpz. Bold genes are proposed to be essential for caprazamycin production.
* PAPS, adenosine 3'-phosphate, 5'-phosphosulfate.

resistance clones were selected, referred to as S. coelicolor M512/cpzLK09(1), -2, and -3. Extracts of cultures of the wild-type and mutant strains were applied to HPLC and ESI-MS/MS. In the wild-type strain the known CPZs A–G, whose different fatty acid side chains result in three different masses were detected readily as depicted in the selected ion monitoring chromatograms of Fig. 3A for CPZ E and F at Rt = 21.28 min, m/z 1113 [M + H]+ for CPZ E and F at Rt = 21.28 min, m/z 1113 [M + H]+ for CPZ C, D, and G at Rt = 22.39 min, and m/z 1147 [M + H]+ for CPZ A and B at Rt = 24.37 min. Characteristic MS/MS fragmentation patterns were observed by collision-induced dissociation of the methylated rhamnose moiety. Fragments of m/z 814, 701, and 569 were obtained from the parent ion m/z 945 [M + H]+ (Fig. 3B).

Analogous fragments with a mass shift of +14 Da and +14 Da resulted from m/z 931 [M + H]+ and m/z 959 [M + H]+ representing the CPZ derivatives with fatty acids of different chain length (supplemental Fig. S1). Molecular ions of m/z 558, 427, and 315, assigned to components of the caprazol structure, were found in all three spectra. S. coelicolor M512 without the gene cluster did not produce any of these new substances (supplemental Figs. S3 and S4). The analytical data strongly implicates the production of non-glycosylated CPZs, which are similar to the type-(III) LPMs (Fig. 1, 3A and supplemental Fig. S4), prominent mass peaks for the CPZ aglyca were observed in S. coelicolor M512/cpzLK09 (selected ion monitoring chromatograms in Fig. 3B). For the CPZ E and F aglyca m/z 931 [M + H]+ at Rt = 18.34 min; for CPZ C, D, and G aglyca m/z 945 [M + H]+ at Rt = 19.51 min; and for CPZ A and B aglyca m/z 959 [M + H]+ at Rt = 20.98 min. Displaying the more hydrophilic character of the free carboxylic acid group, the aglyca elute 3 min earlier from the reversed phase HPLC column than the corresponding intact CPZs. Positive ESI-collision-induced dissociation fragmentation of the new compounds was identical to CPZs except for the absence of the L-rhamnose moiety. Fragments of m/z 814, 701, and 569 were obtained from the parent ion m/z 945 [M + H]+ (Fig. 3B).

Although the masses for the CPZs could only be detected in the natural producer Streptomyces sp. MK730-62F2 (Fig. 3A and supplemental Fig. S4), prominent mass peaks for the CPZ aglyca were observed in S. coelicolor M512/cpzLK09 (selected ion monitoring chromatograms in Fig. 3B) for the CPZ E and F aglyca m/z 931 [M + H]+ at Rt = 18.34 min; for CPZ C, D, and G aglyca m/z 945 [M + H]+ at Rt = 19.51 min; and for CPZ A and B aglyca m/z 959 [M + H]+ at Rt = 20.98 min. Displaying the more hydrophilic character of the free carboxylic acid group, the aglyca elute 3 min earlier from the reversed phase HPLC column than the corresponding intact CPZs. Positive ESI-collision-induced dissociation fragmentation of the new compounds was identical to CPZs except for the absence of the L-rhamnose moiety. Fragments of m/z 814, 701, and 569 were obtained from the parent ion m/z 945 [M + H]+ (Fig. 3B).

Analogous fragments with a mass shift of +14 Da and +14 Da resulted from m/z 931 [M + H]+ and m/z 959 [M + H]+ representing the CPZ derivatives with fatty acids of different chain length (supplemental Fig. S1). Molecular ions of m/z 558, 427, and 315, assigned to components of the caprazol structure, were found in all three spectra. S. coelicolor M512 without the gene cluster did not produce any of these new substances (supplemental Figs. S3 and S4). The analytical data strongly implicates the production of non-glycosylated CPZs, which are similar to the type-(III) LPMs (Fig. 1, 3A) isolated previously (46). These compounds have been reported to show excellent activity against Mycobacteria. Thus, culture extracts from the heterologous producer were applied to an agar diffusion assay against M. phlei (Fig. 4). An inhibition zone of similar size could be observed with extracts from the wild-type and mutant strains, whereas extracts from S. coelicolor M512 without the gene cluster showed no bioactivity in this assay.
The production of the non-glycosylated CPZ derivatives correlates with the absence of genes for the dTDP-L-rhamnose biosynthesis on the cosmid. Apparently, *S. coelicolor* M512 is unable to provide the dTDP-L-rhamnose in compensation as the corresponding enzymes are not encoded on the genome. Both, analytical and biological data verify that the genes identified on cosmid 31C2 indeed represent the CPZ biosynthetic gene cluster.

**Validation of the Cluster Boundaries**

A set of gene disruption experiments was carried out to determine the cluster boundaries. Sequence analysis of overlapping cosmids suggested the CPZ gene cluster to be inserted into a genomic region apparently conserved in several *Streptomyces* strains. *cpz1* and *cpz3* at the one end of the cluster and *cpz33* and *cpz34* at the other are almost identical with a continuous part of sequence from *Streptomyces* sp. SCC 2136 (47). Primer walking and terminal sequencing of overlapping cosmids showed that this similarity extends further upstream of *cpz1*.

Although *cpz4* encodes for a hypothetical protein with unknown function, *cpz5* showed homology to 3-hydroxymethylglutaryl (HMG)-CoA synthases. 3-Hydroxymethylglutaryl-CoA synthases catalyze the aldol addition of acetyl-CoA onto acetoacetyl-CoA and usually participate in the mevalonate pathway (48). A biosynthetic route to the uncommon 3-MG moiety was proposed involving a 3-hydroxymethylglutaryl-CoA synthase, a dehydratase, and a hydrogenase. No function in CPZ formation could be assigned to a putative type III polyketide synthase encoded by *cpz6* and the possibly co-transcribed genes *cpz7* and *cpz8*.

To validate the left border of the cluster *cpz1*, *cpz2*, *cpz3*, and *cpz4* were deleted in cpzLK09 to generate cosmid cpzWP05 and *cpz5* was additionally deleted to generate cpzWP06. By inactivation of the suggested biosynthetic pathway to 3-MG in...
We hoped to produce compounds similar to the highly bioactive type (IV) LPMs (Fig. 1). Cpz6 was inactivated in addition to cpz1–cpz5 in cosmid cpzWP07. At the right end of the cluster cpz33 and cpz34 encoding for a hypothetical protein and a metallophosphoesterase were deleted in cosmid cpzWP01. A possible nucleotidyltransferases encoded by cpz32 was additionally deleted in cosmid cpzWP02.

After introducing the modified cosmids into S. coelicolor M512, positive candidates were selected by their kanamycin resistance and verified by PCR. Cultivation and analysis by HPLC and ESI-MS/MS revealed production of CPZ aglyca in all mutants (data not shown). In addition, bioassays of the culture extracts against M. phlei did not show any difference in inhibitory activity compared with S. coelicolor M512/cpzLK09 containing the intact gene cluster. In the case of /H9004 cpz4, /H9004 cpz5, and /H9004 cpz32 complementation by host genes seems unlikely as the S. coelicolor genome contains no homologues. Therefore we concluded cpz1–6 and cpz32–34 to be non-essential in CPZ biosynthesis. Given that cpz6 is most likely co-transcribed with its downstream positioned genes, a functional knock-out of cpz7 and cpz8 can be assumed in cosmid cpzWP07. Consequently, the biosynthetic gene cluster for CPZs is predicted to span from cpz9 to cpz31 (Fig. 2).

Deletion of cpz21 and cpz23 and Production of Hydroxyacylcaprazols—Because cpz5 seems not to be required for CPZ formation and the corresponding mutant S. coelicolor M512/cpzWP06 did not accumulate the desired β-hydroxyacylcaprazols (Fig. 5), we searched for possible acyltransferases within the gene cluster. Two acyl moieties, the 3-MG and β-hydroxy fatty acids have to be attached during CPZ biosynthesis and the two putative hydrolases Cpz21 and Cpz23 could be considered for these transfer reactions. Cpz21 is predicted to contain a typical α/β-hydrolase fold, the catalytic triad Ser<sup>166</sup>–Glu<sup>226</sup>–His<sup>409</sup>, and a GXGS motif (49). The overall homology of Cpz21 is strongest with TMCL1 from Streptomyces sp. CK4445 (37). TMCL1, also named TmcC, is assigned to the esterification of a dialkyglycerol moiety to the linear polyketide during tautomycetin formation.

The amino acid sequence deduced from cpz23 shows highest overall homology to several hypothetical proteins from Streptomyces including AviX9 of the avilamycin gene cluster from Streptomyces viridochromogenes (50). According to the conserved protein domain search the C terminus of Cpz23 is similar to SGNH-hydrolases, a diverse family of lipases and carboxyesterases (51).

Both genes were individually deleted from cosmid cpzLK09 using A-Red-mediated recombination. To create in-frame deletions, the disruption cassette from pIJ773 was subsequently removed by the use of FLP-recombinase (25) generating cosmids cpzLL06 (Δcpz21) and cpzLL07 (Δcpz23). After introduction of the cosmids in S. coelicolor M512 kanamycin-resistant mutants were cultivated and extracts were analyzed by LC-MS/MS. Production of the CPZ aglyca was abolished in S. coelicolor M512/cpzLL06 and S. coelicolor M512/cpzLL07 (data not shown). This proves that both Cpz21 and Cpz23 play an essential role in CPZ biosynthesis. Metabolites were only identified in extracts of S. coelicolor M512/cpzLL06 (Δcpz21) with m/z 803 [M + H]<sup>+</sup> at Rt = 16.10 min; m/z 817 [M + H]<sup>+</sup> at Rt = 17.20 min; and m/z 831 [M + H]<sup>+</sup> at Rt = 18.56 min by LC-ESI-MS mass scan in positive mode (selected ion monitoring...
chromatograms in Figs. 3C and supplemental S4, and UV chromatograms in supplemental Fig. S5). Collision-induced fragmentation corresponds to the CPZ aglyca by sequential loss of the aminoribose (−131 Da), uracil (−111 Da), and ribose (−132 Da) but m/z values of product ions indicate the absence of the 3-methylglutarate (129 Da) (see also supplemental Fig. S1). Overall, fragments matched exactly the predicted characteristics of β-hydroxycyclopiazacrols, structurally minimized liponucleosides antibiotics. In a bioassay against M. phlei (Fig. 4) extracts of a S. coelicolor M512/cpzLL09 and S. coelicolor M512/cpzLL06 (Δcpz21) cultivation broths showed both similar inhibiting activities, whereas no growth inhibition could be observed with S. coelicolor M512 and S. coelicolor M512/cpzLL07 (data not shown) extracts. Co-expression of intact copies of the deleted genes under the constitutive ermF* promoter in the respective mutants restored the production of CPZ aglyca (data not shown).

Although similar to the type (IV) LPMS (4) the β-hydroxycyclopiazacrols are expected to be slightly more hydrophilic due to the additional OH-group at the 3a-C position. Thus, they represent interesting novel compounds for further investigations, e.g. in structure/activity relationships. Moreover, the data indicated that Cpz21 is most likely involved in attachment of the 3-MG moiety. Beside Cpz21 two other enzymes with homology to para-nitrobenzyl esterases are known from bacterial secondary metabolism. Both of these enzymes, TmcC and TtmK from the tautomycin gene cluster (52), were proposed to catalyze the attachment of an acyl group. However, inactivation of the corresponding genes in the gene cluster did not lead to the identification of an accumulated intermediate. Therefore the data presented here provides the first functional evidence that these family of enzymes indeed act as acyltransferases.

A Model for Caprazamycin Biosynthesis—Sequence analysis of the gene cluster combined with analytical data from heterologous expression and gene inactivation experiments may allow a first proposal of the CPZ biosynthetic pathway (Fig. 5) although many of the suggested reactions remain speculative at present.

A key question in the biosynthesis of CPZs and translocase I inhibitors of the same class is the origin of the glycuridine (Fig. 5, 9). Metabolic labeling studies have shown that uridine is incorporated directly into related uridyl antibiotics such as tunicamycins (53). A pathway to the tunicamycins has been proposed to start with the oxidation of uridine to form uridine 5'-aldehyde (8) (17). We suggest a similar reaction for CPZ biosynthesis, which may be catalyzed by the putative alcohol dehydrogenase Cpz25. The resulting product (8) could undergo a subsequent aldol addition with a pyridoxal phosphate-glycine adduct to generate 9. This mechanism would be very similar to that of the well studied serine hydroxymethyltransferases, which are known to produce β-hydroxy α-amino acids from glycine and various aldehydes (54). Cpz14, with significant sequence similarity to serine hydroxymethyltransferases, is an obvious candidate for the catalysis of this reaction. The next step would be the transfer of a 3-amino-3-carboxypropyl group to the 5''-amino group of 9 to form 10. A corresponding reaction occurs in the nocardiadin biosynthesis (55). In this pathway, the gene product Nat utilizes S-adenosylmethionine to transfer the 3-amino-3-carboxypropyl moiety to a nucleophilic acceptor (56). Nat shows conserved domains of S-adenosylmethionine-dependent methyltransferases, which are also found in Cpz11 and Cpz26, but overall sequence similarity is low with 16% to Cpz26 and 15% to Cpz11. Hence, both genes may be candidates for a 3-amino-3-carboxypropyl transfer in CPZ biosynthesis, although we rather consider them to be involved in the two N-methylation steps discussed below. We further speculate that 10 could be a common intermediate in the caprazamycin (1), liposidomycin (2), FR900493 (6), and the muraymycin (7) biosynthesis. β-Hydroxylation of the 3-amino-3-carboxypropyl group of 10 would lead to 11 and could be catalyzed by either Cpz10 or Cpz15. Both proteins show homology to oxygenases.

Subsequent biosynthetic steps, including formation and transfer of the aminoribose, cyclization, and N-methylation of the diazepanone ring and attachment of the fatty acid would finally lead to 4 (Fig. 5). Compounds of this structure were accumulated in the Δcpz21 mutant strain and are probable intermediates of the CPZ pathway. Reasonable candidate genes for these biosynthetic steps can be found in the cluster. However, the sequence of these reactions, described in the following paragraphs, is speculative at present.

Cyclization of 11 by amide bond formation between the carboxyl group and the secondary amino group would immediately result in the characteristic diazepanone ring. For this reaction, a previous activation of the carboxyl group, e.g. in the form of an acyl adenylate, a coenzyme A ester, or an acyl phosphate would be required. The hypothetical acyl-CoA synthase Cpz20 or the putative kinases Cpz12 and Cpz27 may be involved in this reaction.

Interestingly, a contiguous set of genes, cpz16–19, was found in the CPZ cluster, which can be assigned to all steps required for the generation and attachment of the aminoribosyl moiety. This reaction sequence may start from a second molecule of 8, derived from uridine by a Cpz25-mediated oxidation as described above. Subsequently, the 5-aldehyde group could undergo an aminotransfer reaction, yielding a 5-amminated nucleoside possibly catalyzed by the hypothetical aminotransferase Cpz18. CetH, an orthologue of Cpz18, has recently been assigned to the aminotransferase reaction in biosynthesis of the aminocyclitol cetonicyclette (57). 5-Amino-ribose-1-phosphate and uracil would be generated from the aminated nucleoside by Cpz19, a putative pyrimidine-nucleoside phosphorylase. A similar reaction has been shown in fluorotheonine biosynthesis where a 1- phosphoribosyl derivative is formed under catalysis of the pyrimidine phosphorylase FlB (58). Subsequently, the potential nucleotidyldtransferase Cpz16 may convert the 5-amino-ribose-1-phosphate to dNTP-5-amino ribose. Then, the putative glycosyltransferase Cpz17 could transfer the aminoribose moiety forming a glycosidic bond.

Generally, ribosyl moieties are attached by phosphoribosyltransferases (59) using 5-phosphoribosyl-1-diphosphate as a donor to generate a 5'- phosphoribosylated product. Then, the 5'-phospho group is removed by a phosphatase. Similar reactions have recently been shown to lead to the ribosyl moiety in butirosin biosynthesis involving BtrL and BtrP (60). However, no orthologues to BtrL and BtrP were found in the CPZ gene.
cluster, making the pathway described above a more likely alternative.

The fatty acid moieties of LPMs and CPZs are probably derived from primary metabolism, as feeding studies with labeled palmitic acid in *Streptomyces griseus* showed the direct incorporation into LPMs (61). Hydroxylation of the fatty acids could either occur within primary metabolism or by oxygenases Cpz10 or Cpz15. Cpz23 may be involved in the attachment of the hydroxy fatty acids, due to its homology to lipases.

In the diazepanone ring, both nitrogens are methylated. The N-methylation reactions are likely to be catalyzed by Cpz11 and/or Cpz26. Notably, Cpz11 shows sequence similarity (55%) to one of the few characterized N-methyltransferases AtM1 from the gene cluster of AT2433 (62).

The biosynthetic origin of the 3-MG moiety remains elusive. By our inactivation experiments we could exclude an involvement of the putative 3-hydroxyacyl-CoA synthase Cpz25. Therefore, this moiety is likely generated by enzymes encoded outside the cluster, probably in the form of a coenzyme A ester. We assign the catalysis of the subsequent acyltransfer to Cpz21 as indicated by functional investigations in this study.

Analogous to the biosynthesis of elloramycin, 1-rhamnose would be synthesized from enzymes encoded elsewhere on the genome (40). The dNDP-1-rhamnose probably constitutes the substrate for a transfer reaction to the CPZ aglycon catalyzed by the putative rhamnosyltransferase Cpz31. Sequential methylation of the deoxysugar moiety is likely catalyzed by the hypothetical sugar O-methyltransferases Cpz28, Cpz29, and Cpz30.

The identification and analysis of the caprazamycin gene cluster provides the first molecular basis for the proposal of a translocase I inhibitor biosynthetic pathway. Because the formation of intermediate 10 can be speculated to be similar for other structurally related compounds, this work may help in the development of probes for the discovery of gene clusters of other uridyl antibiotics. As proposed, several biosynthetic steps to the caprazamycins seem to be distinctive and unique in bacterial secondary metabolism. Apparently they represent intriguing subjects for further functional investigations. A detailed understanding of the caprazamycins biosynthetic pathway combined with the successful establishment of a heterologous expression system sets the basis for genetic and metabolic engineering toward the production of new liponucleoside antibiotics with improved properties.

Acknowledgments—We are grateful to Prof. Lutz Heide (Tübingen University, Germany) for comments on the manuscript and Prof. Tim Bugg (University of Warwick, UK) for valuable discussions. We thank the students of the "Wahlpflichtfach Pharmazeutische Bio und Gen- technologie 2007" of Tübingen University for the generation of mutant cosmids cpzWP01-07 and Dr. Manuel Wolpert for kind supervision. We also thank Stefanie Khartulayari for help with the ESI-MS.

REFERENCES

1. Igarashi, M., Nakagawa, N., Doi, N., Hattori, S., Naganawa, H., and Hamada, M. (2003) *J. Antibiot.* 56, 580–583
2. Igarashi, M., Takahashi, Y., Shitara, T., Nakamura, H., Naganawa, H., Miyake, T., and Akamatsu, Y. (2005) *J. Antibiot.* 58, 327–337
3. Takeuchi, T., Igarashi, M., Naganawa, H., and Hamada, M. (May 9, 2002) *J. Antibiot.* 55, 1099–1104
4. Hirano, S., Ichikawa, S., and Matsuda, A. (2005) *Angew. Chem. Int. Ed. Engl.* 44, 1854–1856
5. Price, N. P., and Tsvetanova, B. (2007) *J. Antibiot.* 60, 485–491
6. Ohnuki, T., Muramatsu, Y., Miyakoshi, S., Takatsu, T., and Inukai, M. (2003) *J. Antibiot.* 56, 268–279
7. Marchler, G. H., Mullokandov, M., Song, J. S., Thanki, N., Yamashita, Ke, Z., Krylov, D., Lanczycki, C. J., Liebert, C. A., Liu, C., Lu, F., Lu, S., and Marchler-Bauer, A., Anderson, J. B., Derbyshire, M. K., DeWeese-Scott, C., Gonzalez, N. R., Gwadz, M., Hao, L., He, S., Hurwitz, D. I., Jackson, D. J., Ke, Z., Krylov, D., Lanczycki, C. J., Liebert, C. A., Liu, C., Fu, F., Lu, S., Marchler, G. H., Mullikovand, M., Song, J. S., Thanki, N., Yamashita, R. A., Yin, J. I., Zhang, D., and Bryant, S. H. (2007) *Nucleic Acids Res.* 35, 237–240
8. Doumith, M., Weingarten, P., Wehmeier, U. F., Salah-Bey, M., Capdevila, C., Michel, J. M., Piepersberg, W., and Raynal, M. C. (2000) *Mol. Gen. Genet.* 264, 477–485
9. Patallo, E. P., Blanco, G., Fischer, C., Brana, A. F., Rohr, J., Mendez, C., and Salas, J. A. (2001) *J. Biol. Chem.* 276, 18765–18774
10. Waldron, C., Matsushima, P., Rostek, P. R., Jr., Broughton, M. C., Turner, J., Madduri, K., Crawford, K. P., Merlo, D. J., and Balz, R. H. (2001) *Chem. Biol.* 8, 487–499
11. Marchler-Bauer, A., Anderson, J. B., Derbyshire, M. K., DeWeese-Scott, C., Gonzalez, N. R., Gwadz, M., Hao, L., He, S., Hurwitz, D. I., Jackson, D. J., Ke, Z., Krylov, D., Lanczycki, C. J., Liebert, C. A., Liu, C., Fu, F., Lu, S., Marchler, G. H., Mullikovand, M., Song, J. S., Thanki, N., Yamashita, R. A., Yin, J. I., Zhang, D., and Bryant, S. H. (2007) *Nucleic Acids Res.* 35, 237–240
12. Bibb, M. J., Findlay, P. R., and Johnson, M. W. (1984) *Gene* 30, 157–166
13. Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K., and Ramos, J. L. (1997) *Microbiol. Mol. Biol. Rev.* 61, 393–410
14. Rodriguez, A. M., Olano, C., Vilches, C., Méndez, C., and Salas, J. A. (1993) *Mol. Microbiol.* 8, 571–582
Manipulation of the Caprazamycin Gene Cluster