MbtH and Adenylate-Forming Enzymes in the Biosynthesis of Aminocoumarin Antibiotics and Vancomycin

MbtH und Adenylat-Bildende Enzyme in der Biosynthese von Aminocoumarin-Antibiotika und Vancomycin

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Björn Boll
aus Göttingen

Tübingen
2011

Tag der mündlichen Qualifikation:	18.11.2011
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Lutz Heide
2. Berichterstatter:	Prof. Dr. Thilo Stehle

Table of contents

Publications and Presentations

1. Publications

Björn Boll, Tatjana Taubitz, Lutz Heide, **"MbtH-like proteins are required for the activation of tyrosine in the biosynthesis of aminocoumarin antibiotics and of vancomycin"** *The Journal of Biological Chemistry (2011), 286(42) 36281-90*

Björn Boll, Susanne Hennig, Chunsong Xie, Jae Kyong Sohng and Lutz Heide, **"Adenylate-forming enzymes of rubradirin biosynthesis: RubC1 is a bifunctional enzyme with aminocoumarin acyl ligase and tyrosine-activating domains"** *ChemBioChem (2011), 12, 1105-14*

Orwah Saleh, Bertolt Gust, Björn Boll, Hans-Peter Fiedler and Lutz Heide, **"Aromatic prenylation in phenazine biosynthesis: Dihydrophenazine-1-carboxylate dimethylallyltransferase from *Streptomyces anulatus*."** *The Journal of Biological Chemistry (2009), 284(21):14439-47*

2. Oral presentations

Björn Boll, Susanne Hennig and Lutz Heide. **"Enzymes of adenylate formation in the biosynthesis of rubradirin in *Streptomyces achromogenes* var. *rubradiris* NRRL 3061"** International VAAM-Workshop 2010: Biology of Bacterial Producing Natural Products. September 2010, Tübingen (Germany)

Björn Boll. **"Amide synthetases as tool for the generation of catechol-substituted aminocoumarins"** The Bacterial Cell Envelope: Structure, Function and Infection Interface, SFB 766 Symposium, March 2009, Freudenstadt (Germany)

Björn Boll. **"Amide synthetases as tool for the generation of catechol-substituted aminocoumarins"** 1st SFB 766 Symposium: The Bacterial Cell Envelope: Structure, Function and Infection Interface, March 2008, Freudenstadt (Germany)

3. Poster presentations

Björn Boll, Susanne Hennig and Lutz Heide. **"Investigation of adenylate-forming enzymes involved in rubradirin biosynthesis"**. Internationaler Workshop der Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM): Biology and Chemistry of Antibiotic-Producing Bacteria. September 2011, Bonn (Germany)

Björn Boll, Susanne Henning and Lutz Heide. **"Biochemical studies of adenylate forming enzymes of rubradirin biosynthesis"**. 2nd International SFB 766 Symposium. May 2011, Kloster Irsee / Kaufbeuren (Germany).

Björn Boll, Silke Alt, Ute Metzger and Lutz Heide. **"Biochemical and crystallization studies of amide synthetases for generation of catechol-substituted aminocoumarin antibiotics."** New Trends in Infectious Disease Research. 6th Joint Ph.D. students meeting of the collaborative research centers SFB 630, 766, 544. November 2010, Ellwangen (Germany).

Björn Boll, Silke Alt, Susanne Henning, Ute Metzger, Thilo Stehle and Lutz Heide. **"Biochemical and crystallization studies of amide synthetases"**. SFB 766 Symposium 2010. March 2010, Freudenstadt (Germany).

Björn Boll, Jae Kyong Sohng and Lutz Heide. **"Adenylate-forming enzymes in the biosynthesis of rubradirin in *Streptomyces achromogenes* var. *rubradiris* NRRL 3061"** 15th International Symposium on the Biology of Actinomycetes (ISBA). September 2009, Shanghai (PR China). Receipt of **award for the best poster**.

Björn Boll, Jae Kyong Sohng and Lutz Heide. **"Enzymes of aminocoumarin biosynthesis and of aminocoumarin ligation from the rubradirin biosynthetic gene cluster"** Internationaler Workshop der Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM): Biology and Chemistry of Antibiotic-Producing Bacteria. Oktober 2008, Berlin (Germany)

4. Grants and awards

Poster award for the poster **"Adenylate-forming enzymes in the biosynthesis of rubradirin in *Streptomyces achromogenes* var. *rubradiris* NRRL 3061"** in the session of "Combinatorial Biosynthesis & Pathway Engineering" on the 15th International Symposium on the Biology of Actinomycetes (ISBA). September 2009, Shanghai (PR China).

Scholarship from the Reinhold und Maria Teufel-foundation for participation at the 15th International Symposium on the Biology of Actinomycetes 2009.

Chapter 1 • *Summary*

Streptomycetes possess biosynthetic gene clusters that allow them to produce a multiplicity of secondary metabolites. These include pharmaceutically relevant compounds like prenylated phenazines as well as aminocoumarin- and glycopeptide antibiotics. For specific alterations in the structure of these substances an exact knowledge of the steps of their biosynthesis is of central importance. In this work, enzymes of the biosynthesis of different compounds were biochemically characterized. This allows structural modifications and creation of new derivatives in the future plus an enhancement of the production yield.

The first part of this work describes the identification and characterization of the prenyltransferase PpzP from *Streptomyces anulatus* which transfers a prenyl moiety to a phenazine. PpzP is encoded in a cluster with genes for the biosynthesis of phenazine-1-carboxylic acid (PCA) and the prenyl donor dimethylallyldiphosphate (DMAPP). Cloning and expression of the gene and purification of PpzP resulted in a 37 kDa soluble protein. Activity assays and mass spectrometric analyses confirmed the formation of a C-C bond between the C-1 atom of the isoprenoid substrate and the C-9 atom of the aromatic compound. In contrast to many other prenyltransferases the reaction of PpzP is independent of magnesium or other divalent cations. The K_m values for the substrates were determined as 116 µM for DMAPP and 35 µM for PCA with a turnover number k_{cat} of 0,435 s^{-1}. The sequence of PpzP shows clear homology to the family of aromatic ABBA prenyltransferases. Therefore PpzP broadens the spectrum of accepted substrates of this family, previously limited to phenolic compounds, to phenazine derivatives.

In the second part the adenylate-forming enzymes of the aminocoumarin biosynthetic gene cluster of rubradirin were identified and characterized. These enzymes catalyze the activation of L-tyrosine as a precursor of the aminocoumarin moiety, as well as the formation of an amide bond between an acyl moiety and this aminocoumarin ring. Interestingly, the cluster of rubradirin contains three genes coding for putative enzymes that may catalyze this reaction. Therefore, all three genes were cloned and expressed and the proteins purified for biochemical studies. The 55 kDa Orf4 was shown to be an active amide synthetase *in vitro*. However, the

56 kDa RubF6 was inactive despite its 88 % sequence identity to Orf4, but site directed mutagenesis of the ATP-binding loop converted it into an active enzyme. The third 138 kDa protein, RubC1, was shown to be a unique bifunctional enzyme. It is comprised of an amide synthetase domain as well as a domain for L-tyrosine adenylation with subsequent binding on a peptidyl carrier domain. This natural hybrid enzyme is singular among known proteins and presents a particularly effective machinery for aminocoumarin antibiotic biosynthesis.

The third part is concerned with MbtH-like proteins. Their effect on enzymes that catalyze the adenylation of amino acids was characterized biochemically. The MbtH-like proteins, comprised of approximately 70 amino acids, are encoded in gene clusters of non-ribosomal peptide synthetases. Their function in the biosynthesis was unknown at the beginning of this study but has recently been elucidated, with this study contributing to it. Investigation of the role of MbtH-like proteins in the biosynthesis of the aminocoumarin antibiotics novobiocin, clorobiocin and simocyclinone D8 as well as the glycopeptide antibiotic vancomycin proved that they influence the activity of tyrosine-adenylating enzymes. It could be shown that the tyrosine-activating enzymes CloH, SimH and Pcza361.18, involved in the biosynthesis of clorobiocin, simocyclinone D8 and vancomycin, respectively, require the presence of MbtH-like proteins in a molar ratio of 1:1. They form a heterotetramer consisting of two adenylating enzymes and two MbtH-like proteins. In contrast, NovH from novobiocin biosynthesis showed activity even in the absence of MbtH-like proteins, but could be stimulated by them. NovH and CloH share 83 % identity in their amino acid sequence, yet show a striking difference in their requirement for MbtH-like proteins. To further this phenomenon, 3D structure models were created and compared. This showed that one amino acid differs in the otherwise complete conserved active center. A site-directed mutagenesis of this amino acid in CloH (L383M) indeed resulted in an MbtH-independent mutant. All investigated tyrosine-adenylating enzymes exhibited remarkable promiscuity for MbtH-like proteins from different pathways and organisms. Additionally, the MbtH-like protein YbdZ from *E. coli* was found to co-purify with the heterologously expressed tyrosine-adenylating enzymes, leading to incorrect biochemical results. Therefore, a knock-out strain was created in which the corresponding gene was deleted. This was of central importance for a reliable biochemical characterization of the tyrosine-adenylating enzymes.

Chapter 2 • *Zusammenfassung*

Streptomyceten besitzen Biosynthese-Gencluster, die sie befähigen eine Vielzahl von sekundären Metaboliten zu bilden. Darunter befinden sich auch pharmazeutisch relevante Verbindungen wie prenylierte Phenazine sowie Aminocoumarin- und Glycopeptid-Antibiotika. Für die gezielte Veränderung dieser Substanzen ist ein genaues Verständnis der Abläufe ihrer Biosynthese unerlässlich. In dieser Arbeit wurden Enzyme der Biosynthese von verschiedenen Verbindungen biochemisch charakterisiert. Dies erlaubt in Zukunft Modifikationen der chemischen Strukturen und die Herstellung neuer Derivate sowie eine Verbesserung der Produktausbeute.

Im ersten Teil der Arbeit wurde die Prenyltransferase PpzP aus *Streptomyces anulatus* identifiziert und charakterisiert, welche eine Prenyleinheit auf ein Phenazin überträgt. *ppzP* liegt in einem Cluster mit den Genen für die Biosynthese von Phenazin-1-Carbonsäure (PCA) und dem Prenyldonor Dimethylallyldiphosphat (DMAPP). Klonierung und Expression des Gens und Reinigung von PpzP führte zu einem 37 kDa großen, löslichen Protein. Aktivitätstest und massenspektrometrische Untersuchungen bestätigten die Bildung einer C-C Bindung zwischen dem C1-Atom des isoprenoiden Substrats und dem C9-Atom des Aromaten. Dabei ist die Reaktion von PpzP im Gegensatz zu vielen anderen Prenyltransferasen unabhängig von Magnesium oder anderen zweiwertigen Kationen. Die K_m-Werte für die Substrate wurden mit 116 µM für DMAPP und 35 µM für PCA bestimmt mit einer Wechselzahl k_{cat} von 0,435 s^{-1}. Die Sequenz von PpzP zeigt deutliche Homologie zu der Familie von aromatischen ABBA-Prenyltransferasen. Dadurch erweitert PpzP den Bereich der von dieser Familie akzeptierten Substrate, der bisher nur phenolische Verbindungen enthielt, um Phenazin-Derivate.

Der zweite Teil beschreibt die Identifizierung und Charakterisierung der adenylat-bildenden Enzyme aus dem Gencluster des Aminocoumarin-Antibiotikums Rubradirin. Diese Enzyme werden für die Aktivierung von L-Tyrosin als Vorstufe des Aminocoumarin-Rings benötigt, weiterhin für die Bildung einer Amidbindung zwischen einer Acyl-Einheit und jenem Aminocoumarin Ring. Überraschenderweise enthält das Biosynthese-Gencluster von Rubradirin drei putative Gene, die diese Reaktion katalysieren könnten. Daher wurden alle drei Gene kloniert und exprimiert und die Proteine gereinigt und biochemisch charakterisiert. Für das 55 kDa große Protein Orf4 konnte eine *in vitro* Amidsynthetaseaktivität nachgewiesen werden. Das 56 kDa große Protein RubF6 hingegen war trotz einer 88 %igen Sequenzidentität zu

Orf4 inaktiv, konnte jedoch durch zielgerichtete Mutagenese der ATP-bindenden Schleife in ein aktives Enzym umgewandelt werden. Das dritte Protein, RubC1 mit 138 kDa, wurde als bifunktionales Enzym identifiziert. Es umfasst sowohl eine Amidsynthetase-Domäne als auch eine Domäne für die Adenylierung von L-Tyrosin und anschließende Bindung an eine Peptidyl-Carrier-Domäne. Solch eine hybride Zusammensetzung dieser Domänen ist bislang einzigartig und zeigt einen sehr effizienten Mechanismus für die Biosynthese von Aminocoumarin Antibiotika auf.

Der dritte Teil befasst sich mit MbtH-ähnlichen Proteinen. Dabei wurde ihre Wirkung auf Enzyme, welche die Adenylierung von Aminosäuren katalysieren, biochemisch untersucht. Die rund 70 Aminosäuren umfassenden MbtH-ähnlichen Proteine sind in Biosynthese-Genclustern codiert, die Peptide über nicht-ribosomale Peptidsynthetasen bilden. Für die Untersuchung der Funktion von MbtH-ähnlichen Proteinen, die zu Beginn dieser Studie unbekannt war, wurden verschiedene Antibiotika-Biosynthese-Gencluster hinzugezogen. Im speziellen waren dies die Cluster der Aminocoumarin-Antibiotika Novobiocin, Clorobiocin und Simocyclinon D8 sowie des Glycopeptid-Antibiotikums Vancomycin. Dabei zeigte sich, dass MbtH-ähnliche Proteine die Aktivität der Tyrosin-adenylierenden Enzyme beeinflussen. Die Tyrosin-aktivierenden Enzyme der Clorobiocin, Simocyclinon D8 und Vancomycin Biosynthese, CLoH, SimH und Pcza361.18 benötigen ein MbtH-ähnliches Protein in einem molaren Verhältnis von 1:1 für optimale Aktivität. Dabei formen die Proteine ein Heterotetramer aus je zwei adenylierenden Enzymen und je zwei MbtH-ähnlichen Proteinen. Im Gegensatz dazu benötigt NovH aus der Novobiocin-Biosynthese kein MbtH-ähnliches Protein, wird jedoch durch dessen Zugabe in seiner Aktivität gesteigert. Um den Unterschied in der MbtH-Abhängigkeit zwischen den zu 83 % identischen Proteinen NovH und CloH zu untersuchen, wurden 3D-Strukturmodelle erstellt und verglichen. Dabei war nur eine Aminosäure im ansonsten identischen aktiven Zentrum verschieden. Eine zielgerichtete Mutagenese dieser Aminosäure in CloH (L383M) führte zu einer MbtH-Unabhängigkeit dieser Mutante. Die Abhängigkeit der Tyrosin-aktivierenden Enzyme beschränkte sich nicht auf ein spezifisches MbtH Protein, sondern ließ sich auch mit Homologen aus andern Clustern stimulieren. Insbesondere das MbtH-ähnliche Protein YbdZ aus *E. coli* führte durch die Co-Reinigung der heterolog exprimierten Tyrosin-adenylierenden Enzyme zu einer Verfälschung der Aktivitätsmessungen. Daher wurde ein knock-out Stamm erstellt, in dem das entsprechende Gen deletiert wurde. Dies war von zentraler Wichtigkeit für eine zuverlässige Charakterisierung der Tyrosin-aktivierenden Enzyme.

Chapter 3 • *Introduction*

1. Streptomycetes and their secondary metabolites

Streptomycetes are Gram-positive, filamentous bacteria and belong to the order Actinomycetales. Their DNA has a high G+C content (> 70 %) and their genome spans over 8 – 10 Mb (Bentley *et al.* 2004; Bibb *et al.* 1984; Ikeda and Nakagawa 2003). They are found predominantly in soil and represent an important part of the microflora, noted for their distinct odor which results from production of a volatile metabolite, geosmin. They form a mycelium which consists of a network of branching filaments of cells, similar in appearance to the mycelium of some fungi. Under certain conditions, caused by not completely understood stimuli emanating from the environment or from cellular metabolism, the mycelium differentiate from aerial filaments into spores (Chater 1998; Chater 2001).

In addition to their particularly large genome they contain giant linear plasmids (Kinashi 2011). This plasmids and the linear genome contain genes for the biosynthesis of secondary metabolites organized in clusters, *i.e.*, cohered parts on the genome. They possess all genes required for biosynthesis, regulation and resistance, *e.g.* for a specific antibiotic compound. Since Waksman discovered streptomycin as the first therapeutically useful *Streptomyces* antibiotic in the year 1943, it has been found that Streptomycetes synthesize an amazing variety of chemically distinct inhibitors of many different cellular processes (Chater *et al.* 2010). So far, more than 7000 different secondary metabolites have been discovered in Streptomyces isolates (Bérdy 2005) and more than 3000 bioactive compounds could be isolated, *e.g.* important antibiotics like tetracyclines, vancomycin and erythromycin. A mathematical model in 2001 estimated that this genus is capable of producing a total number of antimicrobials in the order of a 100,000 (Watve *et al.* 2001). Furthermore, substances effective as cytostatics, antifungal agents, virostatics, immunosuppressants and herbicides have been found (von Döhren and Gräfe 2008). This makes the Streptomycetes an extraordinary interesting class from the pharmaceutical point of view.

The development of molecular microbiology and recombinant DNA technology led to an increased knowledge about DNA sequences and a detailed comprehension of the genetic and biochemical principals of secondary metabolites. The increasing

sequencing of genomes as well as the isolation and characterization of gene clusters involved in the biosynthesis of secondary metabolites and the elucidation of the biochemical steps improved the understanding of the biosynthetic machineries and regulatory networks of a targeted molecule and its related metabolic fluxes. The development of tools and methodologies to efficiently carry out genetic manipulations in the producers has strengthened the effectiveness of combinatorial biosynthesis and metabolic engineering (Chen *et al.* 2010).

2. Aminocoumarin antibiotics

Aminocoumarin antibiotics are produced by different *Streptomyces* strains. There are three structural related aminocoumarin antibiotics, novobiocin produced by *S. niveus (*syn. *S. spheroides* NCIMB 11891) (Hoeksema *et al.* 1955), clorobiocin from *S. roseochromogenes* var. *oscitans* DC 12.976 (Mancy *et al.* 1974) and coumermycin A1 from *S. rishiriensis* (Berger and Batcho 1978). Two more aminocoumarin antibiotics have been discovered, namely simocyclinone D8 from *S. antibioticus* Tü 6040 (Holzenkämpfer *et al.* 2002a; Schimana *et al.* 2000) and rubradirin produced by *S. achromogenes* var. *rubradiris* NRRL3061 (Bhuyan *et al.* 1965; Sohng *et al.* 1997). The structural moiety of the aminocoumarin antibiotics that is giving them their name is a 3-amino-4,7-dihydroxycoumarin moiety, which is linked *via* an amide bond to an acyl moiety. While all aminocoumarin share this feature, they differ in the acyl moiety. In novobiocin, clorobiocin and coumermycin A1, the aminocoumarin ring is linked to the deoxysugar noviose by a glycosidic bond (Heide 2009a).

The biological effects of clorobiocin, novobiocin and simocyclinone D8 have been studied extensively (Edwards *et al.* 2009; Maxwell and Lawson 2003). The aminocoumarins inhibit the DNA gyrase by binding to this target with extremely high affinity, which is two orders of magnitude lower than modern fluoroquinolones (Maxwell and Lawson 2003). Since prokaryotic gyrases and topoisomerases IV from the type II topoisomerases are different from eukaryotic topoisomerases (Champoux 2001), they are a promising anti-infective drug target. Crystallographic studies of clorobiocin in complex with gyrase from *E. coli* or *S. aureus* showed that only the aminocoumarin ring and noviose are essential for the interactions with gyrase (Lewis *et al.* 1996; Tsai *et al.* 1997).

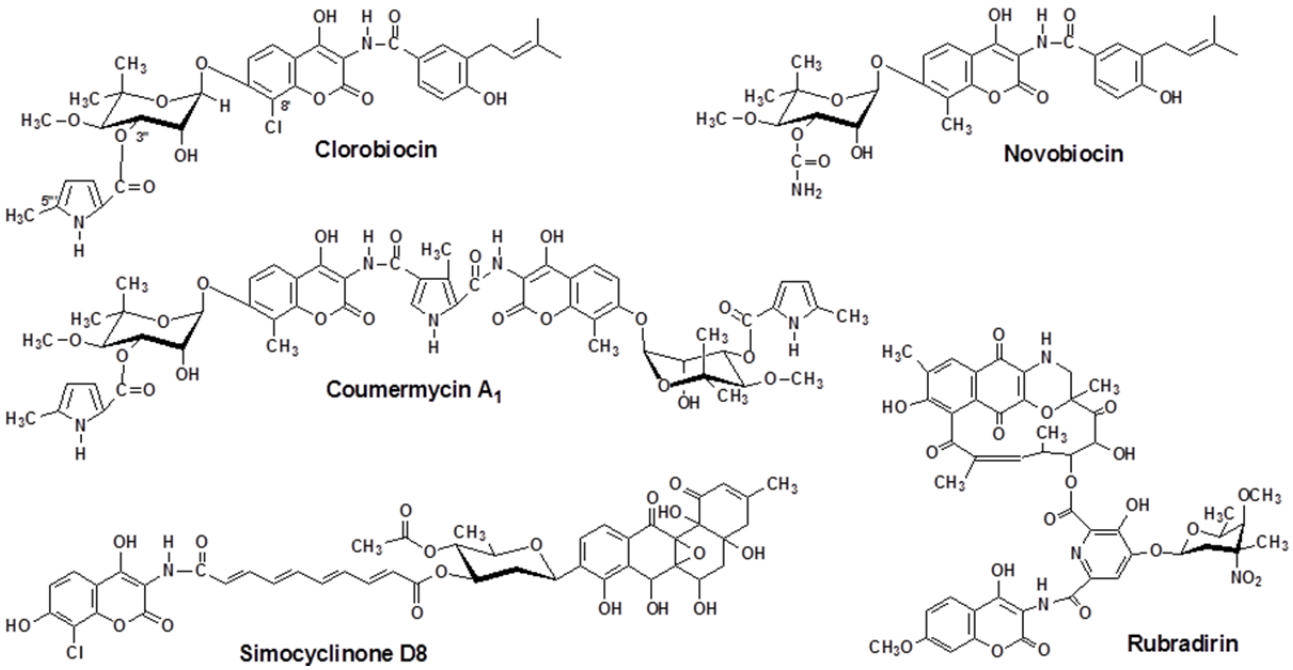


Figure 1 Chemical structures of the aminocoumarin antibiotics.

In contrast, simocyclinone D8 inhibits DNA gyrase by a completely new mode of action that has been shown in recent biochemical and X-ray crystallographic studies. Simocyclinone D8 interacts with two separate pockets of the enzyme and thereby prevents its binding to DNA (Edwards *et al.* 2009; Flatman *et al.* 2005; Oppegard *et al.* 2009).

A completely different mode of action can be observed for rubradirin. It exhibits activity against many Gram-positive bacteria, including *Staphylococcus aureus* strains with multiple antibiotic resistance (Bhuyan *et al.* 1965). Whereas rubradirin inhibits translation at the bacterial ribosomes (Reusser 1973), the rubradirin aglycone has been shown to inhibit bacterial RNA polymerase by a different mechanism from those of ansamycins such as rifamycins, streptovaricins, and tolypomycins (Wherli 1977). Further biological effects of the rubradirin aglycone have also been described (Reusser 1979; Reusser 1988).

Of all aminocoumarin antibiotics, only novobiocin has been licensed in the United States in 1964 (Albamycin®, Pharmacia & Upjohn, now Pfizer) for the treatment of human infections with multi resistant bacteria such as *Staphylococcus aureus* and *Staphylococcus epidermidis*. However, clinical use of Albamycin® remained restricted due to its toxicity in eukaryotes, poor solubility in water, quick development of drug resistance and low activity against Gram-negative bacteria (resulting from poor permeability). Since it was no longer produced and sold, Pfizer

requested a withdrawal that was approved by the United States Food and Drug Administration (FDA) in 2009.

In addition to their antimicrobial activity, novobiocin and its derivatives showed an synergistic effect with antitumor drugs to reduce drug resistance (Duan and You 2009; Rappa *et al.* 2000; Shiozawa *et al.* 2004; Su *et al.* 2007). Furthermore, interactions between novobiocin and its derivatives with the eukaryotic heat shock protein 90 (Hsp90) have been observed (Marcu *et al.* 2000). Hsp90 is a ubiquitous heat shock protein in eukaryotes, plays a key role in protein folding, signaling systems (*e.g.* Raf-1) and tumor repression, and is an emerging target for rational chemotherapy of many cancers (Whitesell and Lindquist 2005). Novobiocin was found to bind to Hsp90 *in vitro* and led to markedly reduced cellular levels of oncogenic kinases *in vitro* and *in vivo* (mice) (Burlison *et al.* 2006; Marcu *et al.* 2000).

The biosynthetic gene clusters of all known aminocoumarin antibiotics have been identified, cloned and sequenced. For the clusters of novobiocin (Steffensky *et al.* 2000a), clorobiocin (Pojer *et al.* 2002) and coumermycin A_1 (Wang *et al.* 2000) the function of nearly all genes has been elucidated and the biosynthesis of these compounds are among the best-known pathways in secondary metabolite biosynthesis in Streptomycetes. The clusters of novobiocin, clorobiocin and coumermycin A_1 span over 23.4, 35.6 and 38.2 kb and code for 20, 29 and 31 putative proteins, respectively (Li and Heide 2006). The biosynthetic gene clusters of simocyclinone D8 and rubradirin have also been identified and sequenced (Galm *et al.* 2002; Kim *et al.* 2008; Recktenwald *et al.* 2002), but only few experimental data are available concerning their biosynthesis. The clusters of simocyclinone and rubradirin are larger, spanning over approximately 65 and 105 kb, respectively, and their precise borders have not yet been defined (Heide 2009a).

Comparison of the novobiocin gene cluster to the gene clusters of clorobiocin and coumermycin A_1 revealed a strikingly stringent accordance between the structure of the antibiotics and the organization of the biosynthetic genes.

The biosynthesis starts with the 3-amino-4,7-dihydroxy coumarin moiety, which is present in all aminocoumarins. Therefore, all clusters contain a group of three to four genes for its biosynthesis, *i.e. novHIJK*, *cloHIJK*, *couHIJK, rubC1C2C3* and *simHIJK.* NovH, CloH, CouH, SimH and RubC1 are non-ribosomal peptide synthetases (see 3.4.2) and comprise an adenylation domain (A) and a phosphopantetheinyl carrier domain (PCP) (Chen and Walsh 2001). After activation

and binding of L-tyrosine to the PCP domain, L-tyrosine is β-hydroxylated by the P450 enzyme NovI (or CloI, CouI, SimI and RubC2) (Chen and Walsh 2001). This is very similar to the biosynthesis of β-hydroxytyrosine in glycopeptides as vancomycin (see 3.5). The β-OH-tyrosine is further reduced by NovJK (or CloJK, CouJK, SimJK and RubC3) to β-keto-tyrosine (Pacholec *et al.* 2005b). These enzymes show similarities to 3-oxoacyl-[acylcarrierprotein]-reductases from polyketide and fatty acid biosynthesis. The final step of the biosynthesis of the aminocoumarin moiety has not yet been elucidated. Experiments showed that a hydroxyl group is introduced into position 2 of the aromatic ring during the biosynthesis and speculation arose that this occurs in the β-keto-tyrosyl-S-NovH moiety, followed by a nucleophilic attack of this hydroxyl group on the thioester. This would lead to cleavage of the thioester bond and cyclization of the aminocoumarin ring. The clorobiocin, coumermycin A_1 and simocyclinone gene clusters contain a small set of open readings frame, *cloY*, *couY* and *simY*, which show sequence similarity to the gene *mbtH* from the biosynthetic gene cluster of the siderophore mycobactin from *Mycobacterium tuberculosis* (see 3.6). Their gene products interact with CloH, CouH and SimH in the adenylation reaction of L-tyrosine and are center of the investigations in chapter 7.

The 3-dimethylallyl-4-hydroxybenzoyl moiety of novobiocin and clorobiocin is formed from 4-hydroxyphenylpyruvate by the gene products of *novFQR* and *cloFQR* (Heide 2009a). The prephenate dehydrogenases NovF and CloF provide 4-hydroxyphenylpyruvate that is prenylated by an aromatic prenyltransferase (see 3.3) CloQ (Pojer *et al.* 2003b) or NovQ (Ozaki *et al.* 2009). The bifunctional non-heme iron(II)-dependent oxygenase CloR and NovR yield the final product *via* two consecutive oxidative decarboxylation steps (Pojer *et al.* 2003a). The acyl moieties in coumermycin, rubradirin and simocyclinone are structurally completely different from the moiety of novobiocin and clorobiocin (Fig. 1). Their biosynthesis has not yet been experimentally examined.

The aminocoumarin moiety is linked to the respective acyl moiety by an amide synthetase, *e.g.* NovL, CloL, CouL, SimL, Orf4 or RubC1, which are described in 3.4.1. After linkage, the aminocoumarin moieties of novobiocin and coumermycin A_1 are methylated by the methyl transferases NovO and CouO respectively (Li *et al.* 2002). Clorobiocin and simocyclinone D8 contain a chlorine atom and correspondingly the gene cluster of clorobiocin contains a gene *clo-hal* encoding a FAD-dependent halogenase (Eustáquio *et al.* 2003a). The simocyclinone cluster also

contains a similar halogenase termed *simD4* (Trefzer *et al.* 2002) but there is no experimental data available. Biochemical investigations proved that the modification of the aminocoumarin ring occurs after formation of the amide bond (Anderle *et al.* 2007; Pacholec *et al.* 2005c).

The deoxysugar moiety 5-C-methyl-L-rhamnose is synthesized by the products of a group of five genes (*novSTUVW)* and these genes have orthologues in the other antibiotic gene clusters. The glycosyl transferases NovM, CloM and CouM catalyze the transfer of the dTDP-activated deoxysugar to the 7-hydroxy group of the aminocoumarin moiety (Freel Meyers *et al.* 2003). After linkage to the aminocoumarin ring, the deoxysugar is methylated by the SAM-dependent methyltransferases NovP, CloP or CouPv (Freel Meyers *et al.* 2004).

As final step in the biosynthesis the deoxysugar is acylated on its 3-hydroxy group. In novobiocin NovN transfers a carbamoyl moiety (Freel Meyers *et al.* 2004), while in clorobiocin and coumermycin A_1 the corresponding acyl moiety is a pyrrol-2-carboxylic acid formed by the gene products *cloN*1-7 and *couN*1-7.

For the regulation of the biosynthesis, all three clusters contain also two positive regulators, *novE* (Eustáquio *et al.* 2003b) and *novG* (Dangel *et al.* 2008; Eustáquio *et al.* 2005) and their orthologs. The gene $gyrB^R$ encodes an aminocoumarin-resistant gyrase B subunit. Because clorobiocin and coumermycin A_1 are potent inhibitors of topoisomerase IV, their gene clusters contain an additional resistance gene $parY^R$ (Schmutz *et al.* 2003) that encodes an aminocoumarin resistant topoisomerase IV subunit.

The intensive knowledge of aminocoumarin biosynthesis allows direct optimization of production yield as well as modification of specific biosynthetic steps and the resulting antibiotic. Therefore, it is possible to create a large diversity of new aminocoumarin derivatives using metabolic engineering, combinatorial biosynthesis and mutasynthesis to overcome known restriction like reduced uptake or adverse effects (Alt *et al.* 2011; Heide 2009b).

3. Prenyltransferases

Prenyltransferases catalyze the transfer of isoprenoid moieties to acceptor molecules. The group of allylic prenyltransferases catalyzes a reaction which proceeds through formation of a carbocation during the cleavage of pyrophosphate group from C-1 of the allylic substrate DMAPP or GPP followed by a stereoselective

electrophilic alkylation of the carbon–carbon double bond in IPP by the allylic substrate. They are further divided into two classes, *cis* (or *Z*) and *trans* (or *E*), depending upon the stereochemistry of the resulting products. An archetypical allylic prenyltransferase is the well-examined FPP synthase (Poulter 2006) from primary metabolism that is involved in the synthesis of membrane sterols in all living organisms. Like all trans-prenyldiphosphate synthetases, FPP synthase contains two conserved (N/D)DxxD motifs for binding of the allylic and homoallylic substrate. The substrate binds in form of a Mg^{2+} complex, and the aspartate-rich motifs provides chelation of the Mg^{2+} ion. The three dimensional structure of trans-prenyltransferases consists exclusively of α-helices (Liang *et al.* 2002).

Aromatic prenyltransferases can catalyze the transfer of prenyl moieties like DMAPP, GPP or FPP to aromatic acceptor molecules. They catalyze an electrophilic substitution at the aromatic acceptor molecule forming a C-C bond between C-1 (for regular prenylation) or C-3 (for reverse prenylation) of the allylic isoprenoid moieties. Some aromatic prenyltransferases have been investigated, mostly enzymes of primary metabolism *e.g.* UbiA from ubichinone biosynthesis (Melzer and Heide 1994), MenA and UBIAD1 from menachinone biosynthesis (Nakagawa *et al.* 2010; Suvarna *et al.* 1998), Slr1736 from tocopherol biosynthesis (Schledz *et al.* 2001) and Pds2 from plastochinone biosynthesis (Collakova and DellaPenna 2001). Similar enzymes are involved in the secondary metabolism of plants in formation of prenylated flavonoids and isoflavanoids as well as shikonin (Yazaki *et al.* 2009). Very recently several membrane bound prenyltransferases for secondary metabolites have been reported in microorganisms, *e.g.* farnesyltransferase AuaA from *Stigmatella aurantiaca* that catalyzes the prenylation of 2-methyl-4-hydroxyquinoline in the biosynthesis of aurachins (Stec *et al.* 2011). These prenyltransferases are integral membrane proteins and their structure has not been determined to date. However, two models of this group have been published (Bräuer *et al.* 2008; Ohara *et al.* 2009). Like FPP synthetases, members of this group exhibit (N/D)DxxD motifs for substrate binding and are Mg^{2+} dependent.

A new group of soluble aromatic prenyltransferases has been identified in the last years. The first bacterial enzyme was CloQ from the biosynthesis of clorobiocin (Pojer *et al.* 2003b). CloQ and NovQ (Ozaki *et al.* 2009) catalyze the prenylation of 4-hydroxyphenylpyruvate with DMAPP as prenyl donor. In 2005 the structure of NphB from the biosynthetic gene cluster of the meroterpenoid (prenylated polyketide)

naphterpin was elucidated (Kuzuyama *et al.* 2005). It showed a new type of β/α barrel fold with antiparallel strands. The α-β-β-α architecture of this fold led to the name ABBA prenyltransferases for this group of enzymes (Tello *et al.* 2008). Many more genes with sequence similarity have been identified from bacteria and fungi and characterized biochemically. These enzymes do not show a (N/D)DxxD motif and are, with exception of NphB, Mg^{2+} independent. They include PpzP from *Streptomyces anulatus* that is responsible for the prenylation reaction in the biosynthesis of endophenazine A and is a focus of the investigation of the publication in chapter 5.

Another group of aromatic prenyltransferases has been identified and characterized mostly from fungi but also from bacteria (Edwards and Gerwick 2004; Schultz *et al.* 2010; Steffan *et al.* 2009). They mostly catalyze the prenylation of indole moieties, are soluble proteins, do not contain a (N/D)DxxD motif and are Mg^{2+} independent. This group has similar properties but share no sequence similarity with enzymes of the CloQ/NphB group. Surprisingly, the three dimensional structure of FgaPT2 from *Aspergillus fumigatus* revealed the same α-β-β-α architecture (Metzger *et al.* 2009).

The prenylation of aromatic moieties has led to a huge variety of secondary metabolites in nature. Aromatic prenyltransferases are important enzymes for exploring novel isoprenoid substitutions in aromatic compounds because the presence of the isoprenoid chain can lead to impressive changes in biological activity. This is mostly attributed to an increased affinity for biological membranes and to an improved interaction with proteins (Botta *et al.* 2005b). The structures of ABBA prenyltransferases show a central cavity for the substrates that allow promiscuous prenylation of different aromatic substrates. Site directed mutagenesis guided by mechanistic insights from structural biology may broaden their substrate range. Therefore, prenyltransferases represent attractive tools in the chemoenzymatic generation of bioactive compounds (Botta *et al.* 2005a; Koehl 2005; Macone *et al.* 2009).

4. Adenylate-forming enzymes

Adenylate-forming enzymes catalyze the activation of the otherwise unreactive carboxylic acid by the transfer of ATP. They are involved in a variety of metabolic pathways such as ribosomal and non-ribosomal peptide synthesis, fatty acid

oxidation or enzyme regulation (Schmelz and Naismith 2009). Based on sequence similarity, a superfamily was formed that was divided into four groups (Fulda *et al.* 1994). This classification has been extended recently to include the structurally distinct aminoacyl-tRNA synthetases and enzymes involved in NRPS-independent siderophores (NIS). Class I now comprises the three subclasses of NRPS adenylation domains (Ia), acyl- or aryl CoA-synthetases (Ib) and oxidoreductases (Ic). Class II contains all aminoacyl-tRNA synthetases while class III includes the NIS (Schmelz and Naismith 2009). All three subclasses activate their substrate *via* adenylation but only subclass Ib forms a covalent thioester bond.

The condensation between the weakly nucleophilic carboxylic acid and weakly electrophilic phosphate forms a transition state with a negatively pentavalent phosphorus atom and results in an adenylated carboxylate. As this intermediate is very reactive, the enzyme catalyzes a second step, *i.e.* the reaction with a nucleophile (either amine, alcohol or thiol). This leads to the desired product while AMP is released (Fig. 2). Two families of adenylate-forming enzymes that are important for secondary metabolite production are introduced here in detail.

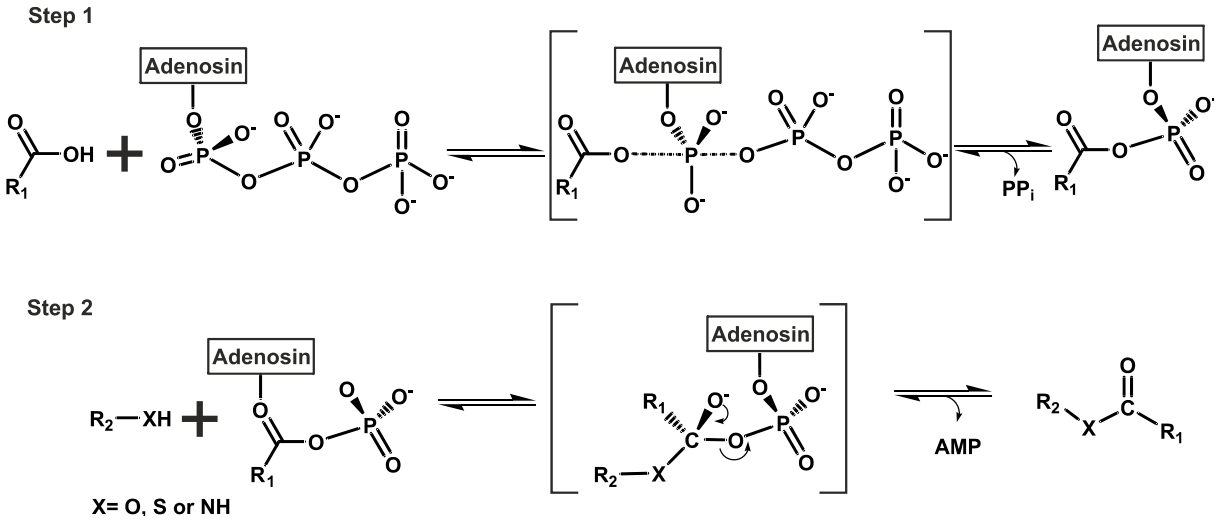


Figure 2 Adenylation reaction catalyzed by adenylate-forming enzymes. First step: adenylation of carboxylate. Second step: transfer of the nucleophile to an alcohol, amid or sulfhydryl moiety. Figure modified from (Schmelz and Naismith 2009)

4.1. Amide synthetases

Amide synthetases are similar to class I of adenylate-forming enzymes (Schmelz and Naismith 2009; Turgay *et al.* 1992), *i.e.* subclass Ia with A-domains and subclasses Ib and Ic with acyl- or aryl-CoA-synthetases and oxidoreductases. Amide synthetases share some sequence motifs typical for adenylate-forming

enzymes like 4-coumarate-CoA-ligase (Stuible *et al.* 2000) and A-domains from NRPS enzymes (Marahiel *et al.* 1997). But in contrast to non-ribosomal peptide synthetases, amide synthetases do not show a 5'-phosphopantetheinyl attachment site.

Amide synthetases catalyze one of the first steps in the biosynthesis of aminocoumarin antibiotics, the formation of the amide bond between aminocoumarin ring and acyl component *via* an amide bond. The enzymes in their respective biosynthetic gene clusters have been identified as NovL (Steffensky *et al.* 2000b), ClolL (Galm *et al.* 2004a), CouL (Schmutz *et al.* 2003), SimL (Luft *et al.* 2005; Pacholec *et al.* 2005a), Orf4 and RubC1 (Boll *et al.* 2011) (see chapter 6). The ATP-dependent reaction indicates an activation of the carboxy group of the acyl moiety *via* adenylation. While the amide synthetase activity could already be shown in 1975 in cell extracts of *S. niveus* (Kominek and Meyer 1975), the reaction mechanism was initially approved in 2000 by identification and characterization of NovL *via* an PP_i-ATP exchange assay (Steffensky *et al.* 2000b). After formation of an acyl-AMP-intermediate the acyl group is transferred to the amino group of the aminocoumarin ring (Fig. 2). Interestingly, coumermycin A_1 contains two amide bonds and the formation of both is catalyzed by CouL. Analysis of the two reaction products of CouL proved that they represent a monoamide and the diamide of 3-methylpyrrole-2,4-dicaboxylic acid and the aminocoumarin ring. Therefore, CouL is capable of catalyzing four enzymatic steps, *i.e.* two adenylation and two acyl transfer reactions, resulting in the formation of two amide bonds (Schmutz *et al.* 2003).

The different amide synthetases have varying substrate specificities and recently a high-throughput screening was established that allows fast creation and testing of amide synthetases with altered substrate specificity (Parajuli and Williams 2011). These distinctions in substrate specificity have been exploited and various acyl substrates have been offered to the amide synthetases. Because the acyl moiety of novobiocin and clorobiocin is of no importance for binding to the gyrase (Lafitte *et al.* 2002), a multiplicity of new aminocoumarin antibiotics have been generated by biochemical, genetic, and synthetic mutasynthesis in genetically engineered producer strains (Alt *et al.* 2011; Anderle *et al.* 2007; Galm *et al.* 2004a). Some of the aminocoumarin derivatives exhibited a higher inhibition of gyrase (Galm *et al.* 2004b) and derivatives with a catechol group showed improved uptake by bacteria (Alt *et al.*

2011). Therefore, amide synthetases represent an important tool in the chemoenzymatic generation of new antibiotics.

4.2 Non-ribosomal peptide synthetases

Non-ribosomal peptide synthetases (NRPS) are class I enzymes of adenylate-formation (Schmelz and Naismith 2009). They catalyze the formation of an amide bond between amino acids independent of the ribosome.

The first NRPS was identified in the biosynthesis of the cyclic decapeptide gramicidin S from *Bacillus brevis,* which activates the amino acids and tethers them one by one covalently to the NRPS (Gevers *et al.* 1968; Gevers *et al.* 1969; Kleinkauf *et al.* 1969).

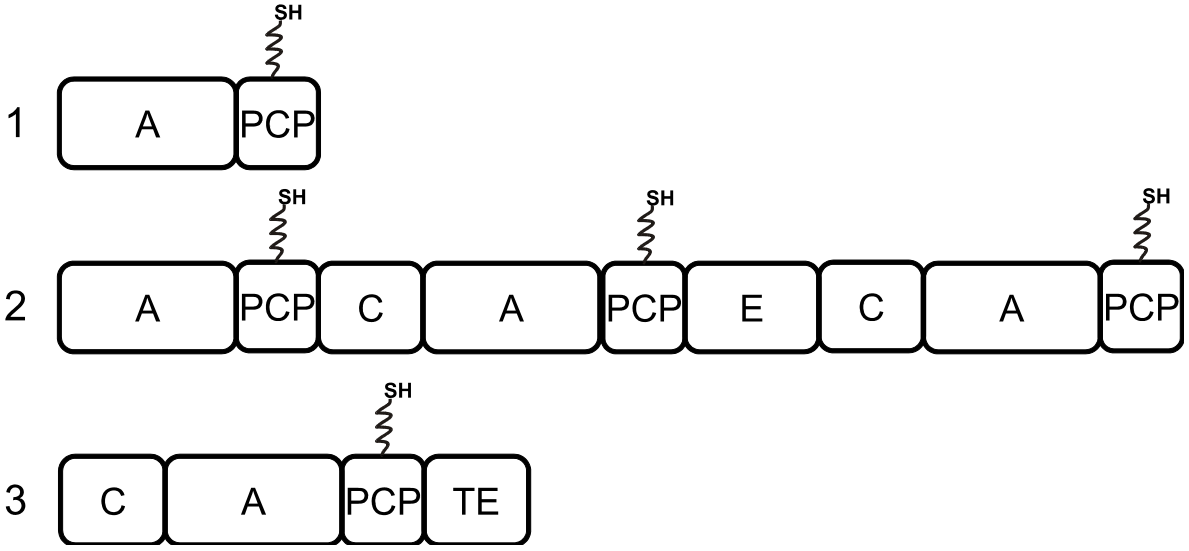


Figure 3 Schematic representations of NRPSs on the example of (1) NovH from novobiocin biosynthesis and (2) CepA and (3) CepC from vancomycin biosynthesis. (A) adenylation domain; (PCP) peptidyl carrier domain; (C) condensation domain; (E) epimerization domain; (TE) thioester domain.

Today many cyclic and linear peptides of medical and pharmaceutical importance are known to be biosynthesized by NRPS enzymes (Felnagle *et al.* 2008). Intensive studies revealed that these large enzymes often comprise different modules which again are divided in distinct domains (Fig. 3) (Marahiel and Essen 2009; Marahiel *et al.* 1997). Each domain assumes a certain function in the linking of amino acids. The adenylation domain (A) activates an amino acid (Turgay *et al.* 1992) and tethers it covalently to the phosphopantheteinyl cofactor of the peptidyl carrier domain (PCP) (Stachelhaus *et al.* 1996). The PCP domain transfers the amino acid to the condensation (C) domain where a condensation with a second amino acid occurs (Stachelhaus *et al.* 1998). In some NRPS an epimerization domain isomerizes

the amino acid from one stereochemistry to the other. During biosynthesis the amino acids remain covalently bound to the NRPS. For cleavage of the covalent thioester bond between PCP domain and substrate, some NRPS contain a thioesterase. For increased diversity the NRPS can contain additional modules for further modifications such as methylation or cyclization (Konz and Marahiel 1999). Furthermore, the NRPS can be part of a bigger hybrid enzyme that contains a polyketide synthetase (PKS). Bioinformatic studies of NRPS A-domains has led to the identification of a specificity-conferring code, which consists of approximately ten amino acid residues that line the substrate-binding pocket of the respective enzyme. This code allows the prediction of the specific amino acid that is adenylated by each A-domain (Challis *et al.* 2000; Stachelhaus *et al.* 1999).

By structure determination of single NRPS domains as well as X-ray and NMR studies of the interaction between different NRPS domains, especially the interaction between A and PCP domain (Yonus *et al.* 2008), new insights into NRP synthesis were gained. These findings allow the alteration of substrate specificity, domain swapping and NRPS re-engineering to broaden the range of secondary metabolites.

5. Vancomycin and glycopetide antibiotics

The class of glycopeptide antibiotics comprises a plethora of different compounds, all synthesized by Actinomycetes. Vancomycin is the most prominent representative of the class of glycopeptide antibiotics and was also the first member of this family discovered and isolated in the 1950s at Eli Lilly from *Streptomyces orientalis* (now *Amycolatopsis orientalis*) (Levine 2006). Shortly after its discovery the US Food and Drug Administration approved vancomycin for clinical use. Until today these antibiotics are of major importance for the treatment of infectious diseases particularly methicillin-resistant *Staphylococcus aureus* (MRSA). Glycopeptides are drugs of last resort because of their cumbersome intravenously application, side effects like ototoxicity, nephrotoxicity and red man syndrome (Levine 2006), and because semi-synthetic penicillins such as methicillin, nafcillin or cloxacillin exhibit better activity against non-MRSA staphylococci. Glycopeptides are not active against Gram-negative bacteria because their outer membrane is impermeable to large glycopeptide molecules (except some non-gonococcal species of *Neisseria* (Geraci and Wilson 1981)).

Glycopeptides target the cell wall biosynthesis of Gram-positive bacteria *via* interaction with the pentapetide on GlucNAc-MurNAc which is a precursor of peptidoglycan. The large hydrophilic molecule is able to bind to the terminal D-alanyl-D-alanine moieties of the peptidoglycan precursors by forming five hydrogen bonds with the amide bonds and C-terminus of the target (Nicolaou *et al.* 1999). This results in inhibition of the transpeptidation steps of peptidoglycan synthesis and thereby weakens the peptidoglycan layers (Kahne *et al.* 2005). Additionally, telavancin and oritavancin are able to disrupt bacterial membrane integrity and thereby increase membrane permeability (Zhanel *et al.* 2010). In the 1990s vancomycin-resistant enterococci and staphylococci spread in clinics (Nicolaou *et al.* 1999). These resistant enterococci either use peptidoglycan precursors terminated in D-Ala-D-Lac (VanA/VanB resistance) or less effectively D-Ala-D-Ser (VanC resistance) peptides for cell wall biosynthesis (Bugg *et al.* 1991; Pootoolal *et al.* 2002). This results in a ~1000-fold lower affinity for glycopeptides (Kahne *et al.* 2005). However, an aglycon derivative was created recently that is able to bind D-Ala-D-Ala as well as D-Ala-D-Lac motifs with good affinity (Xie *et al.* 2011).

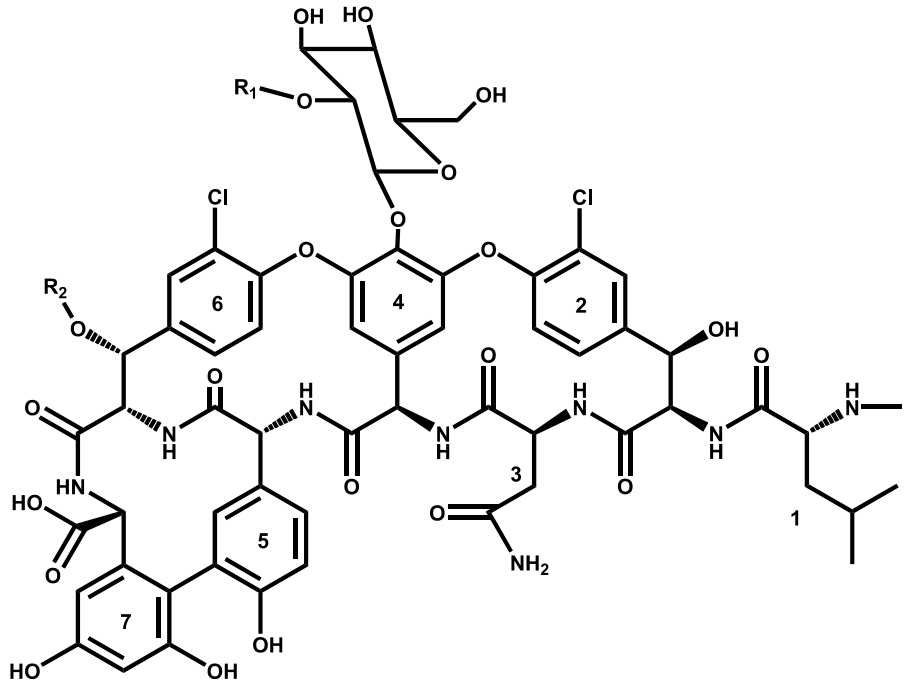


Figure 4 Vancomycin and related glycopeptides consisting of a crosslinked heptapetide backbone. Vancomycin: R_1= vancosamine; R_2= H; Balhimycin: R_1= H ; R_2= oxovancosamine ; Chloroeremomycin: R_1 and R_2= epivancosamine.

The biosynthesis of glycopeptides such as vancomycin (van Wageningen *et al.* 1998), balhimycin (Pelzer *et al.* 1999) and teicoplanin (Li *et al.* 2004; Sosio *et al.* 2004) were the aim of intensive studies and most biosynthetic steps of these molecules have been elucidated. All glycopeptides are comprised of seven amino

acids forming a heptapetide backbone (Fig. 4). They are assembled *via* non-ribosomal peptide synthases (NRPS). The biosynthesis itself can be divided into three phases. First, the synthesis of non-proteinogenic amino acids as building blocks. This precedes the assembly of the building blocks by NRPS to a linear heptapetide backbone (Pelzer *et al.* 1999; Süssmuth *et al.* 1999) and interlinking of the aromatic side chains. This assembly starts with recognition and activation of the respective amino acid by the adenylation domain of the NRPS, followed by thioester formation and subsequent condensation with the growing peptide chain. From observations in balhimycin biosynthesis is was concluded that the chlorination takes place on NRPS-bound intermediates and not on the free heptapeptide (Wohlleben *et al.* 2009). After the linear heptapeptide backbone is synthesized, glycopeptides undergo further modifications. This third step involves oxidative cross-linking (Bischoff *et al.* 2001) of the heptapetide and modification by glycosyltransferases (Chen *et al.* 2000) and a methyltransferase (O'Brien *et al.* 2000). Some derivatives additionally utilize acyltransferases and sulfotransferases (Lamb *et al.* 2006).

The vancomycin-type glycopeptides consist of two proteinogenic (Leu and Asn) and five non-proteinogenic amino acids, namely two β-hydroxytyrosines, two 3-hydroxyphenylglycins and a 3,5-dihydroxyphenylglycine (Fig. 4). The β-hydroxytyrosine is derived from L-tyrosine which is adenylated by a NRPS and then bound to its PCP domain. Subsequently, a monooxygenase catalyzes the β-hydroxylation of the NRPS bound L-tyrosine. This is very similar to the first steps of aminocoumarin biosynthesis (Chen and Walsh 2001; Puk *et al.* 2004). Here however, the product is cleaved from the PCP domain after β-hydroxylation by a thioesterase to be assembled by a different NRPS in the heptapeptide backbone (Mulyani *et al.* 2010). The β-hydroxylation of L-tyrosine and the influence of MbtH-like proteins on this reaction is the center of the investigation in chapter 7. The various vancomycin-like glycopeptides differ only in their glycosylation, methylation and chlorination patterns (Stegmann *et al.* 2010). Teicoplanin-like glycopeptides also consist of seven aminoacids but differ in their residues as well as in their aromatic cross-linking and further modifications.

Since glycopeptides have a similar structure they are good scaffolds for genetic engineering, combinatorial biosynthesis and chemoenzymatic modification. The extensive available knowledge about their biosynthesis allows alterations of the post-translational modifications or substitution of amino acids in the backbone (Weist *et al.*

2002; Weist *et al.* 2004), and leads to new antibiotics and new secondary metabolites.

6. MbtH-like proteins

The gene *mbtH* was found in *Mycobacterium tuberculosis* in the biosynthetic gene cluster of the virulence-conferring siderophore mycobactin (Quadri *et al.* 1998). It encodes a small protein with 71 amino acids. Subsequent studies showed that *mbtH*-like genes are often found in gene clusters containing NRPS domains for biosynthesis of antibiotics or siderophores.

The *mbtH*-like genes are exclusively found in bacteria but neither in Archaea nor Eukaryota and until now there are more than 400 *mbtH* homologs annotated in GenBank. MbtH homologs are most prevalent within the Actinobacteria with 188 homologs, among them 65 in Streptomycetes (Baltz 2011). Some genomes contain several *mbtH*-like genes and the largest number so far were observed in *Streptomyces clavigulerus* and *Streptomyces griseus,* which both contain seven *mbtH* homologs.

In Actinomycetes, *mbtH*-like genes are generally located in gene clusters of secondary metabolites encoding the biosynthesis of non-ribosomal peptides including the antibiotics vancomycin, daptomycin, teicoplanin, capreomycin as well as the aminocoumarin antibiotics clorobiocin and coumermycin A_1. Interestingly, the aminocoumarin biosynthetic gene clusters contain only one NRPS gene with an adenylation domain (A) and a peptidyl carrier domain (PCP) that is involved in the activation and binding of L-tyrosine (Chen and Walsh 2001). The *mbtH*-like gene in the clusters of clorobiocin and coumermycin A_1 is directly adjacent to the NRPS gene similar to the *nikP1* of nikkomycin where the *mbtH*-like gene is fused to the A-PCP gene into one open reading frame (Chen *et al.* 2002). The role and function of MbtH-like proteins in the biosynthesis of aminocoumarin and glycopeptide antibiotics like vancomycin is elucidated in chapter 7.

Gene inactivation studies of the *mbtH*-like gene *dptG* from the biosynthetic gene cluster of daptomycin showed a 50 % reduction in yield (Nguyen *et al.* 2006). However, production was not completely abolished and it is unknown if this is due to cross-pathway complementation of DptG by another MbtH-like protein in *S. roseosporus*. The same problem appeared in another study about MbtH-like proteins in the biosynthesis of balhimycin by Stegmann and coworkers (2006). The *mbtH*-like

gene *orf1* in *A. balhimycina* was inactivated but no effect on the balhimycin production could be observed. However, the authors mentioned that the genome contains two other *mbtH*-like genes that could take over the function of *orf1* (Stegmann *et al.* 2006).

First proof that these genes are essential for secondary metabolite production was provided by two studies about gene inactivation and complementation in the biosynthesis of clorobiocin (Wolpert *et al.* 2007), coelichelin and calcium-dependent antibiotic (Lautru *et al.* 2007). They proved that the gene products of *cdaX* from the calcium dependent antibiotic cluster, *cchK* from the coelichelin cluster and *cloY* from the clorobiocin cluster could complement each other to some extent. A strain where all *mbtH* homologs have been deleted was unable to produce the respective compound.

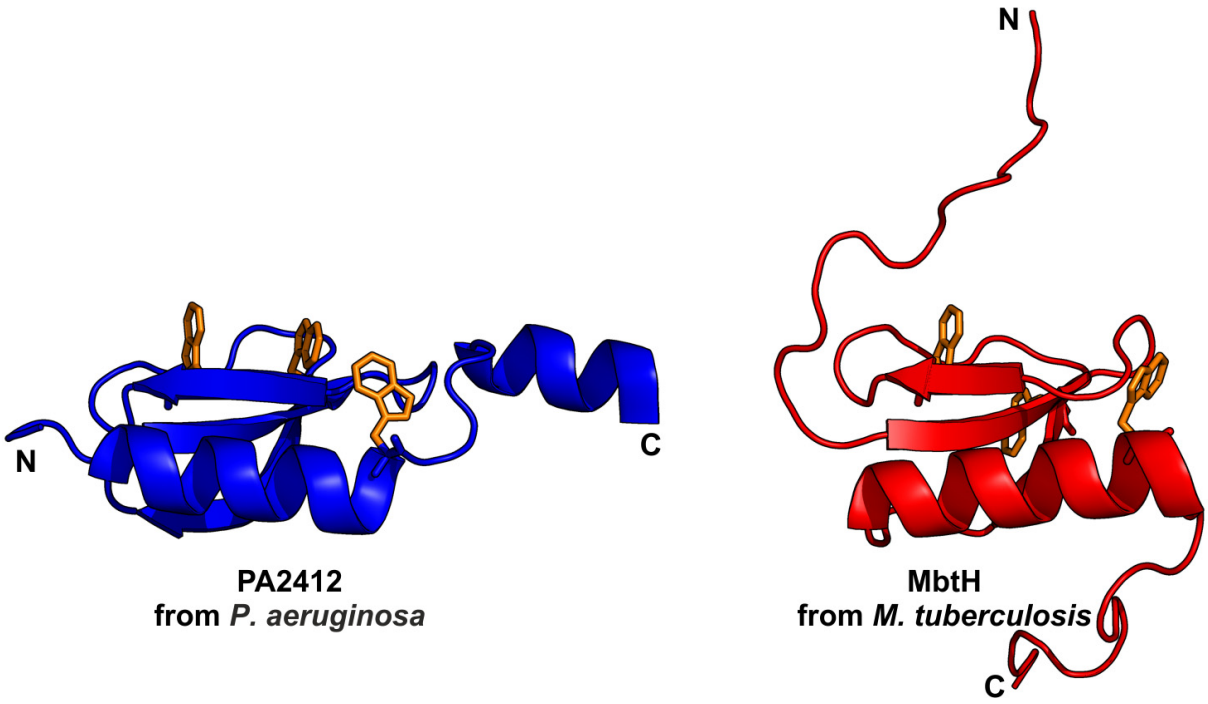


Figure 5 Structures of PA2412 (blue) and MbtH (red). The three conserved tryptophans are displayed in orange.

Two structures of MbtH-like proteins have been determined: PA2412 from the pyoverdine cluster of *Pseudomonas aeruginosa via* crystallization (Drake *et al.* 2007) and the name-giving MbtH by using magnetic resonance and circular dichroism spectroscopy (Buchko *et al.* 2010). PA2412 is shaped like a thin arrowhead with three anti-parallel β-sheets followed by two α-helices, one nestled against the β-sheet and the second on the C-terminus forming the point of the arrow (Fig. 5) (Drake *et al.* 2007). The solution structure is similar to the aforementioned crystal structure

except for the C-terminus which is highly disordered in the second case (Fig. 5). The authors hypothesized that highly conserved but disordered regions as in the C-terminus are associated with binding to multiple partners. This might explain the promiscuity of MbtH-like proteins and the cross-talk between different MbtH-like proteins. For the second ordered helix observed in the PA2412 crystal structure it was speculated that this occurs due to crystal packing interactions or the binding of an unknown small ligand (Buchko *et al.* 2010). Sequence comparison of stand-alone MbtH-like proteins showed invariant regions, highly conserved regions and variable regions. The conserved residues, including the three highly conserved tryptophans (in MbtH W26, W36 and W56) all lie on one face of the protein. The localization of invariant residues on one face and variable on the other side suggests the interaction with two types of protein. Finally, both structures show no obvious active site of MbtH-like proteins.

Recent biochemical studies of MbtH-like proteins from Actinomycetes revealed their function. It was shown that MbtH-like proteins interact with adenylating enzymes which are part of NRPSs. The biosynthetic gene clusters of the antituberculosis antibiotics capreomycin from *Saccharothrix mutabilis* subsp*. capreolus* (Felnagle *et al.* 2007) and viomycin produced by *Streptomyces* sp. ATCC 11861 (Thomas *et al.* 2003) both contain an *mbtH*-like (*cmnN* and *vioN*) gene just upstream of a NRPS domain (*cmnO* and *vioO*). CmnO and VioO alone showed no adenylation activity with β-lysine or any other amino acid but the addition of CmnN or VioN resulted in the activation of β-lysine (Felnagle *et al.* 2010). Another study with the *mbtH*-like gene *pacJ* from the biosynthetic gene cluster for pacidamycin from *Streptomyces coeruleorubidus* (Zhang *et al.* 2010b) with the NRPS domains showed the same results. PacJ activated the NRPS protein PacL which was inactive alone (Zhang *et al.* 2010a). Both studies showed that the specific MbtH-like proteins could be substituted by homologs from different pathways, *e.g.* YbdZ from *E. coli* siderophore biosynthesis or KtzJ from kutzneride biosynthesis. To further explore the MbtH-NRPS interaction both groups performed site-directed mutagenesis with the three conserved tryptophan residues. The change of W56 to alanine (in reference to MbtH) resulted in protein that could neither be copurified nor stimulate the amino acid activation (Felnagle *et al.* 2010). The substitution of the corresponding residue W26 in MbtH led to a 50 % reduction of stimulated adenylation activity while the

substitution of W26 and W36 showed no stimulatory effect and the mutated protein was unable to bind to the NRPS enzyme (Zhang *et al.* 2010a).

The recent studies, including the one in chapter 7, showed that MbtH-like proteins are required for the adenylation of amino acids in non-ribosomal peptide synthesis. While it still remains unclear how they interact with NRPS, MbtH-like proteins are important for several applications. Since most important NRPS pathways include mbtH-like genes, they could serve as beacons for the identification of new bacterial strains to produce novel peptides (Baltz 2011). Furthermore, MbtH-like proteins play an important role in strain improvement, combinatorial biosynthesis and heterologous expression. Some NRPS domains are strictly dependent on MbtH-like proteins and are stimulated at different rates by various MbtH homologs. Therefore it is crucial for the success of such experiments to include the correct *mbtH*-like genes for the expression of NRPS gene clusters.

Chapter 4 • *References*

Alt, S., N. Burkard, A. Kulik, S. Grond and L. Heide (2011). "An artificial pathway to 3,4-dihydroxybenzoic acid allows generation of new aminocoumarin antibiotic recognized by catechol transporters of *E. coli*." Chemistry and Biology **18**(3): 304-313.

Anderle, C., S. Hennig, B. Kammerer, S. M. Li, L. Wessjohann, B. Gust and L. Heide (2007). "Improved mutasynthetic approaches for the production of modified aminocoumarin antibiotics." Chemistry & Biology **14**(8): 955-967.

Anderle, C., S. Hennig, B. Kammerer, S. M. Li, L. Wessjohann, B. Gust and L. Heide (2007). "Improved mutasynthetic approaches for the production of modified aminocoumarin antibiotics." Chemistry and Biology **14**(8): 955-967.

Baltz, R. H. (2011). "Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery." Journal of Industrial Microbiology and Biotechnology.

Bentley, S. D., R. Brosch, S. V. Gordon, D. A. Hopwood, S. T. Cole, C. M. Fraser, T. D. Read and K. E. Nelson (2004). Genomics of actinobacteria, the high G+C Gram-positive bacteria. Microbial Genomes. Totowa, NJ, Humana Press Inc.**:** 333-359.

Bérdy, J. (2005). "Bioactive microbial metabolites." Journal of Antibiotics **58**(1): 1-26.

Berger, J. and A. D. Batcho (1978). Coumarin-glycoside antibiotics. Antibiotics. Isolation, Separation and Purification. M. J. Weinstein and G. H. Wagman. Amsterdam, Oxford, New York, Elsevier Scientific Publishing Company. **15:** 101-158.

Bhuyan, B. K., S. P. Owen and A. Dietz (1965). "Rubradirin, a new antibiotic. I. Fermentation and biological properties." Antimicrobial Agents and Chemotherapy(Oct.): 91-96.

Bibb, M. J., P. R. Findlay and M. W. Johnson (1984). "The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences." Gene **30**(1-3): 157-166.

Bischoff, D., S. Pelzer, B. Bister, G. J. Nicholson, S. Stockert, M. Schirle, W. Wohlleben, G. Jung and R. D. Süssmuth (2001). "The biosynthesis of vancomycin-type glycopeptide antibiotics - the order of the cyclization steps." Angewandte Chemie International Edition **40**(24): 4688-4691.

Boll, B., S. Hennig, C. Xie, J. K. Sohng and L. Heide (2011). "Adenylate-forming enzymes of rubradirin biosynthesis: RubC1 is a bifunctional enzyme with aminocoumarin acyl ligase and tyrosine-activating domains." Chembiochem : a European journal of chemical biology **12**(7): 1105-1114.

Botta, B., G. D. Monache, P. Menendez and A. Boffi (2005a). "Novel prenyltransferase enzymes as a tool for flavonoid prenylation." Trends in Pharmacological Sciences **26**(12): 606-608.

Botta, B., A. Vitali, P. Menendez, D. Misiti and G. Delle Monache (2005b). "Prenylated flavonoids: pharmacology and biotechnology." Current Medicinal Chemistry **12**(6): 717-739.

Bräuer, L., W. Brandt, D. Schulze, S. Zakharova and L. Wessjohann (2008). "A structural model of the membrane-bound aromatic prenyltransferase UbiA from *E. coli*." Chembiochem : a European journal of chemical biology **9**(6): 982-992.

Buchko, G. W., C. Y. Kim, T. C. Terwilliger and P. J. Myler (2010). "Solution structure of Rv2377c-founding member of the MbtH-like protein family." Tuberculosis (Edinb) **90**(4): 245-251.

Bugg, T. D., G. D. Wright, S. Dutka-Malen, M. Arthur, P. Courvalin and C. T. Walsh (1991). "Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA." Biochemistry **30**(43): 10408-10415.

Burlison, J. A., L. Neckers, A. B. Smith, A. Maxwell and B. S. Blagg (2006). "Novobiocin: redesigning a DNA gyrase inhibitor for selective inhibition of hsp90." Journal of the American Chemical Society **128**(48): 15529-15536.

Challis, G. L., J. Ravel and C. A. Townsend (2000). "Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains." Chemistry and Biology **7**(3): 211-224.

Champoux, J. J. (2001). "DNA topoisomerases: structure, function, and mechanism." Annual Review of Biochemistry **70**: 369-413.

Chater, K. (1998). "Taking a genetic scalpel to the *Streptomyces* colony." Microbiology **144**: 1465-1478.

Chater, K. F. (2001). "Regulation of sporulation in *Streptomyces coelicolor* A3(2): a checkpoint multiplex?" Current Opinion in Microbiology **4**(6): 667-673.

Chater, K. F., S. Biro, K. J. Lee, T. Palmer and H. Schrempf (2010). "The complex extracellular biology of *Streptomyces*." FEMS Microbiology Reviews **34**(2): 171-198.

Chen, H., B. K. Hubbard, S. E. O'Connor and C. T. Walsh (2002). "Formation of β-hydroxy histidine in the biosynthesis of nikkomycin antibiotics." Chemistry and Biology **9**(1): 103-112.

Chen, H., M. G. Thomas, B. K. Hubbard, H. C. Losey, C. T. Walsh and M. D. Burkart (2000). "Deoxysugars in glycopeptide antibiotics: enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis." Proceedings of the National Academy of Sciences of the United States of America **97**(22): 11942-11947.

Chen, H. and C. T. Walsh (2001). "Coumarin formation in novobiocin biosynthesis: β-hydroxylation of the aminoacyl enzyme tyrosyl-*S*-NovH by a cytochrome P450 NovI." Chemistry and Biology **8**(4): 301-312.

Chen, Y., M. J. Smanski and B. Shen (2010). "Improvement of secondary metabolite production in *Streptomyces* by manipulating pathway regulation." Applied Microbiology and Biotechnology **86**(1): 19-25.

Collakova, E. and D. DellaPenna (2001). "Isolation and functional analysis of homogentisate phytyltransferase from *Synechocystis sp.* PCC 6803 and *Arabidopsis*." Plant Physiology **127**(3): 1113-1124.

Dangel, V., A. S. Eustáquio, B. Gust and L. Heide (2008). "*novE* and *novG* act as positive regulators of novobiocin biosynthesis." Archives of Microbiology **190**(5): 509-519.

Drake, E. J., J. Cao, J. Qu, M. B. Shah, R. M. Straubinger and A. M. Gulick (2007). "The 1.8 Å crystal structure of PA2412, an MbtH-like protein from the pyoverdine cluster of *Pseudomonas aeruginosa*." Journal of Biological Chemistry **282**(28): 20425-20434.

Duan, P. and G. You (2009). "Novobiocin is a potent inhibitor for human organic anion transporters." Drug Metablism and Disposition: The Biological Fate of Chemicals **37**(6): 1203-1210.

Edwards, D. J. and W. H. Gerwick (2004). "Lyngbyatoxin biosynthesis: sequence of biosynthetic gene cluster and identification of a novel aromatic prenyltransferase." Journal of the American Chemical Society **126**(37): 11432-11433.

Edwards, M. J., R. H. Flatman, L. A. Mitchenall, C. E. Stevenson, T. B. Le, T. A. Clarke, A. R. McKay, H. P. Fiedler, M. J. Buttner, D. M. Lawson and A. Maxwell (2009). "A crystal

structure of the bifunctional antibiotic simocyclinone D8, bound to DNA gyrase." Science **326**(5958): 1415-1418.

Eustáquio, A. S., B. Gust, T. Luft, S. M. Li, K. F. Chater and L. Heide (2003a). "Clorobiocin biosynthesis in *Streptomyces*. Identification of the halogenase and generation of structural analogs." Chemistry and Biology **10**(3): 279-288.

Eustáquio, A. S., S. M. Li and L. Heide (2005). "NovG, a DNA-binding protein acting as a positive regulator of novobiocin biosynthesis." Microbiology **151**(6): 1949-1961.

Eustáquio, A. S., T. Luft, Z. X. Wang, B. Gust, K. F. Chater, S. M. Li and L. Heide (2003b). "Novobiocin biosynthesis: inactivation of the putative regulatory gene *novE* and heterologous expression of genes involved in aminocoumarin ring formation." Archives of Microbiology **180**(1): 25-32.

Felnagle, E. A., J. J. Barkei, H. Park, A. M. Podevels, M. D. McMahon, D. W. Drott and M. G. Thomas (2010). "MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases." Biochemistry **49**(41): 8815-8817.

Felnagle, E. A., E. E. Jackson, Y. A. Chan, A. M. Podevels, A. D. Berti, M. D. McMahon and M. G. Thomas (2008). "Nonribosomal peptide synthetases involved in the production of medically relevant natural products." Mol Pharm **5**(2): 191-211.

Felnagle, E. A., M. R. Rondon, A. D. Berti, H. A. Crosby and M. G. Thomas (2007). "Identification of the biosynthetic gene cluster and an additional gene for resistance to the antituberculosis drug capreomycin." Applied and Environmental Microbiology **73**(13): 4162-4170.

Flatman, R. H., A. J. Howells, L. Heide, H. P. Fiedler and A. Maxwell (2005). "Simocyclinone D8, an inhibitor of DNA gyrase with a novel mode of action." Antimicrob Agents Chemother **49**(3): 1093-1100.

Freel Meyers, C. L., M. Oberthür, J. W. Anderson, D. Kahne and C. T. Walsh (2003). "Initial characterization of novobiocic acid noviosyl transferase activity of NovM in biosynthesis of the antibiotic novobiocin." Biochemistry **42**(14): 4179-4189.

Freel Meyers, C. L., M. Oberthür, H. Xu, L. Heide, D. Kahne and C. T. Walsh (2004). "Characterization of NovP and NovN: Completion of novobiocin biosynthesis by sequential tailoring of the noviosyl ring." Angewandte Chemie International Edition **43**(1): 67-70.

Fulda, M., E. Heinz and F. P. Wolter (1994). "The *fadD* gene of *Escherichia coli* K12 is located close to rnd at 39.6 min of the chromosomal map and is a new member of the AMP-binding protein family." Molecular and General Genetics **242**(3): 241-249.

Galm, U., M. A. Dessoy, J. Schmidt, L. A. Wessjohann and L. Heide (2004a). "*In vitro* and *in vivo* production of new aminocoumarins by a combined biochemical, genetic, and synthetic approach." Chemistry and Biology **11**(2): 173-183.

Galm, U., S. Heller, S. Shapiro, M. Page, S. M. Li and L. Heide (2004b). "Antimicrobial and DNA gyrase-inhibitory activities of novel clorobiocin derivatives produced by mutasynthesis." Antimicrobial Agents and Chemotherapy **48**(4): 1307-1312.

Galm, U., J. Schimana, H. P. Fiedler, J. Schmidt, S. M. Li and L. Heide (2002). "Cloning and analysis of the simocyclinone biosynthetic gene cluster of *Streptomyces antibioticus* Tü 6040." Archives of Microbiology **178**(2): 102-114.

Geraci, J. E. and W. R. Wilson (1981). "Vancomycin therapy for infective endocarditis." Reviews of Infectious Diseases **3 suppl**: S250-258.

Gevers, W., H. Kleinkauf and F. Lipmann (1968). "The activation of amino acids for biosynthesis of gramicidin S." Proceedings of the National Academy of Sciences of the United States of America **60**(1): 269-276.

Gevers, W., H. Kleinkauf and F. Lipmann (1969). "Peptidyl transfers in gramicidin S bisoynthesis from enzyme-bound thioester intermediates." Proceedings of the National Academy of Sciences of the United States of America **63**(4): 1335-1342.

Heide, L. (2009a). "The aminocoumarins: biosynthesis and biology." Natural Product Reports **26**: 1241–1250.

Heide, L. (2009b). "Genetic engineering of antibiotic biosynthesis for the generation of new aminocoumarins." Biotechnology Advances **27**(6): 1006-1014.

Hoeksema, H., J. L. Johnson and J. W. Hinman (1955). "Structural studies on streptonivicin, a new antibiotic." Journal of the American Chemical Society **77**: 6710-6711.

Holzenkämpfer, M., M. Walker, A. Zeeck, J. Schimana and H. P. Fiedler (2002a). "Simocyclinones, novel cytostatic angucyclinone antibiotics produced by *Streptomyces antibioticus* Tü 6040 - II. Structure elucidation and biosynthesis." Journal of Antibiotics **55**(3): 301-307.

Ikeda, M. and S. Nakagawa (2003). "The *Corynebacterium glutamicum* genome: features and impacts on biotechnological processes." Applied Microbiology and Biotechnology **62**(2-3): 99-109.

Kahne, D., C. Leimkuhler, W. Lu and C. Walsh (2005). "Glycopeptide and lipoglycopeptide antibiotics." Chemical Reviews **105**(2): 425-448.

Kim, C. G., J. Lamichhane, K. I. Song, V. D. Nguyen, D. H. Kim, T. S. Jeong, S. H. Kang, K. W. Kim, J. Maharjan, Y. S. Hong, J. S. Kang, J. C. Yoo, J. J. Lee, T. J. Oh, K. Liou and J. K. Sohng (2008). "Biosynthesis of rubradirin as an ansamycin antibiotic from *Streptomyces achromogenes* var. *rubradiris* NRRL3061." Arch. Microbiol. **189**(5): 463-473.

Kinashi, H. (2011). "Giant linear plasmids in *Streptomyces*: a treasure trove of antibiotic biosynthetic clusters." Journal of Antibiotics **64**(1): 19-25.

Kleinkauf, H., W. Gevers and F. Lipmann (1969). "Interrelation between activation and polymerization in gramicidin S biosynthesis." Proceedings of the National Academy of Sciences of the United States of America **62**(1): 226-233.

Koehl, P. (2005). "Relaxed specificity in aromatic prenyltransferases." Nature chemical biology **1**(2): 71-72.

Kominek, L. A. and H. F. Meyer (1975). "Novobiocic acid synthetase." Methods Enzymol. **43**: 502-508.

Konz, D. and M. A. Marahiel (1999). "How do peptide synthetases generate structural diversity?" Chemistry and Biology **6**(2): R39-R48.

Kuzuyama, T., J. P. Noel and S. B. Richard (2005). "Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products." Nature **435**(7044): 983-987.

Lafitte, D., V. Lamour, P. O. Tsvetkov, A. A. Makarov, M. Klich, P. Deprez, D. Moras, C. Briand and R. Gilli (2002). "DNA gyrase interaction with coumarin-based inhibitors: the role of the hydroxybenzoate isopentenyl moiety and the 5'-methyl group of the noviose." Biochemistry **41**(23): 7217-7223.

Lamb, S. S., T. Patel, K. P. Koteva and G. D. Wright (2006). "Biosynthesis of sulfated glycopeptide antibiotics by using the sulfotransferase StaL." Chemistry and Biology **13**(2): 171-181.

Lautru, S., D. Oves-Costales, J. L. Pernodet and G. L. Challis (2007). "MbtH-like protein-mediated cross-talk between non-ribosomal peptide antibiotic and siderophore biosynthetic pathways in *Streptomyces coelicolor* M145." Microbiology **153**(Pt 5): 1405-1412.

Levine, D. P. (2006). "Vancomycin: a history." Clinical Infectious Diseases **42 Suppl 1**: S5-12.

Lewis, R. J., O. M. P. Singh, C. V. Smith, T. Skarzynski, A. Maxwell, A. J. Wonacott and D. B. Wigley (1996). "The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography." EMBO Journal **15**(6): 1412-1420.

Li, S. M. and L. Heide (2006). "The biosynthetic gene clusters of aminocoumarin antibiotics." Planta Medica **72**(12): 1093-1099.

Li, S. M., L. Westrich, J. Schmidt, C. Kuhnt and L. Heide (2002). "Methyltransferase genes in *Streptomyces rishiriensis*: new coumermycin derivatives from gene-inactivation experiments." Microbiology **148**(10): 3317-3326.

Li, T. L., F. Huang, S. F. Haydock, T. Mironenko, P. F. Leadlay and J. B. Spencer (2004). "Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase." Chemistry and Biology **11**(1): 107-119.

Liang, P. H., T. P. Ko and A. H. Wang (2002). "Structure, mechanism and function of prenyltransferases." European Journal of Biochemistry **269**(14): 3339-3354.

Luft, T., S. M. Li, H. Scheible, B. Kammerer and L. Heide (2005). "Overexpression, purification and characterization of SimL, an amide synthetase involved in simocyclinone biosynthesis." Archives of Microbiology **183**(4): 277-285.

Macone, A., E. Lendaro, A. Comandini, I. Rovardi, R. M. Matarese, A. Carraturo and A. Bonamore (2009). "Chromane derivatives of small aromatic molecules: Chemoenzymatic synthesis and growth inhibitory activity on human tumor cell line LoVo WT." Bioorganic and Medicinal Chemistry **17**(16): 6003-6007.

Mancy, D., L. Ninet and J. PreudïHomme (1974). Antibiotic 18631 RP. U.S., Rhone-Poulenc S.A.

Marahiel, M. A. and L. O. Essen (2009). "Nonribosomal peptide synthetases mechanistic and structural aspects of essential domains." Methods in Enzymology **458**: 337-351.

Marahiel, M. A., T. Stachelhaus and H. D. Mootz (1997). "Modular peptide synthetases involved in nonribosomal peptide synthesis." Chemical Reviews **97**(7): 2651-2674.

Marcu, M. G., T. W. Schulte and L. Neckers (2000). "Novobiocin and related coumarins and depletion of heat shock protein 90-dependent signaling proteins." Journal of the National Cancer Institute **92**(3): 242-248.

Maxwell, A. and D. M. Lawson (2003). "The ATP-binding site of type II topoisomerases as a target for antibacterial drugs." Current Topics in Medicinal Chemistry **3**(3): 283-303.

Melzer, M. and L. Heide (1994). "Characterization of polyprenyldiphosphate: 4-hydroxybenzoate polyprenyltransferase from *Escherichia coli*." Biochimica et Biophysica Acta **1212**(1): 93-102.

Metzger, U., C. Schall, G. Zocher, I. Unsöld, E. Stec, S. M. Li, L. Heide and T. Stehle (2009). "The structure of dimethylallyl tryptophan synthase reveals a common architecture of aromatic prenyltransferases in fungi and bacteria." Proceedings of the National Academy of Sciences of the United States of America **106**(34): 14309-14314.

Mulyani, S., E. Egel, C. Kittel, S. Turkanovic, W. Wohlleben, R. D. Süssmuth and K. H. van Pée (2010). "The thioesterase Bhp is involved in the formation of beta-hydroxytyrosine during balhimycin biosynthesis in *Amycolatopsis balhimycina*." Chembiochem : a European journal of chemical biology **11**(2): 266-271.

Nakagawa, K., Y. Hirota, N. Sawada, N. Yuge, M. Watanabe, Y. Uchino, N. Okuda, Y. Shimomura, Y. Suhara and T. Okano (2010). "Identification of UBIAD1 as a novel human menaquinone-4 biosynthetic enzyme." Nature **468**(7320): 117-121.

Nguyen, K. T., D. Kau, J. Q. Gu, P. Brian, S. K. Wrigley, R. H. Baltz and V. Miao (2006). "A glutamic acid 3-methyltransferase encoded by an accessory gene locus important for

daptomycin biosynthesis in *Streptomyces roseosporus*." Molecular Microbiology **61**(5): 1294-1307.

Nicolaou, K. C., C. N. Boddy, S. Brase and N. Winssinger (1999). "Chemistry, Biology, and Medicine of the Glycopeptide Antibiotics." Angewandte Chemie International Edition **38**(15): 2096-2152.

O'Brien, D. P., P. N. Kirkpatrick, S. W. O'Brien, T. Staroske, T. I. Richardson, D. A. Evans, A. Hopkinson, J. B. Spencer and D. H. Williams (2000). "Expression and assay of an -methyltransferase involved in the biosynthesis of a vancomycin group antibiotic." Chemical Communications(1): 103-104.

Ohara, K., A. Muroya, N. Fukushima and K. Yazaki (2009). "Functional characterization of LePGT1, a membrane-bound prenyltransferase involved in the geranylation of *p*-hydroxybenzoic acid." Biochemical Journal **421**(2): 231-241.

Oppegard, L. M., B. L. Hamann, K. R. Streck, K. C. Ellis, H. P. Fiedler, A. B. Khodursky and H. Hiasa (2009). "In vivo and in vitro patterns of the activity of simocyclinone D8, an angucyclinone antibiotic from Streptomyces antibioticus." Antimicrobial Agents and Chemotherapy **53**(5): 2110-2119.

Ozaki, T., S. Mishima, M. Nishiyama and T. Kuzuyama (2009). "NovQ is a prenyltransferase capable of catalyzing the addition of a dimethylallyl group to both phenylpropanoids and flavonoids." Journal of Antibiotics **62**(7): 385-392.

Pacholec, M., C. L. Freel Meyers, M. Oberthur, D. Kahne and C. T. Walsh (2005a). "Characterization of the aminocoumarin ligase SimL from the simocyclinone pathway and tandem incubation with NovM,P,N from the novobiocin pathway." Biochemistry **44**(12): 4949-4956.

Pacholec, M., N. J. Hillson and C. T. Walsh (2005b). "NovJ/NovK catalyze benzylic oxidation of a ß-hydroxyl tyrosyl-*S*-pantetheinyl enzyme during aminocoumarin ring formation in novobiocin biosynthesis." Biochemistry **44**(38): 12819-12826.

Pacholec, M., J. Tao and C. T. Walsh (2005c). "CouO and NovO: C-Methyltransferases for tailoring the aminocoumarin scaffold in coumermycin and novobiocin antibiotic biosynthesis." Biochemistry **44**(45): 14969-14976.

Parajuli, N. and G. J. Williams (2011). "A high-throughput screen for directed evolution of aminocoumarin amide synthetases." Analytical Biochemistry.

Pelzer, S., R. Süßmuth, D. Heckmann, J. Recktenwald, P. Huber, G. Jung and W. Wohlleben (1999). "Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908." Antimicrobial Agents and Chemotherapy **43**(7): 1565-1573.

Pojer, F., R. Kahlich, B. Kammerer, S. M. Li and L. Heide (2003a). "CloR, a bifunctional non-heme iron oxygenase involved in clorobiocin biosynthesis." Journal of Biological Chemistry **278**(33): 30661-30668.

Pojer, F., S. M. Li and L. Heide (2002). "Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics." Microbiology **148**(Pt 12): 3901-3911.

Pojer, F., E. Wemakor, B. Kammerer, H. Chen, C. T. Walsh, S. M. Li and L. Heide (2003b). "CloQ, a prenyltransferase involved in clorobiocin biosynthesis." Proceedings of the National Academy of Sciences of the United States of America **100**(5): 2316-2321.

Pootoolal, J., J. Neu and G. D. Wright (2002). "Glycopeptide antibiotic resistance." Annual Review of Pharmacology and Toxicology **42**: 381-408.

Poulter, C. D. (2006). "Farnesyl diphosphate synthase. A paradigm for understanding structure and function relationships in E-polyprenyl diphosphate synthases." Phytochemistry Reviews **5**: 17-26.

Puk, O., D. Bischoff, C. Kittel, S. Pelzer, S. Weist, E. Stegmann, R. Süssmuth and W. Wohlleben (2004). "Biosynthesis of chloro-beta-hydroxytyrosine, a non-proteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics." Vaccine Research **186**: 6093-6100.

Quadri, L. E., J. Sello, T. A. Keating, P. H. Weinreb and C. T. Walsh (1998). "Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin." Chemistry and Biology **5**(11): 631-645.

Rappa, G., K. Shyam, A. Lorico, O. Fodstad and A. C. Sartorelli (2000). "Structure-activity studies of novobiocin analogs as modulators of the cytotoxicity of etoposide (VP-16)." Oncology Research **12**(3): 113-119.

Recktenwald, J., R. Shawky, O. Puk, F. Pfennig, U. Keller, W. Wohlleben and S. Pelzer (2002). "Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules." Microbiology **148**(4): 1105-1118.

Reusser, F. (1973). "Rubradirin, an inhibitor of ribosomal polypeptide biosynthesis." Biochemistry **12**(6): 1136-1142.

Reusser, F. (1979). "Inhibition of ribosomal and RNA polymerase functions by rubradirin and its aglycone." Journal of Antibiotics **32**(11): 1186-1192.

Reusser, F. B., B. Tarpley, W.G. Althaus, I. Zapotocky, B. (1988). Rubradirin derivatives for treatment of HIV infection. T. U. Company. **WO/1988/008707**.

Schimana, J., H. P. Fiedler, I. Groth, R. Süssmuth, W. Beil, M. Walker and A. Zeeck (2000). "Simocyclinones, novel cytostatic angucyclinone antibiotics produced by *Streptomyces antibioticus* Tü 6040. I. Taxonomy, fermentation, isolation and biological activities." Journal of Antibiotics **53**(8): 779-787.

Schledz, M., A. Seidler, P. Beyer and G. Neuhaus (2001). "A novel phytyltransferase from *Synechocystis* sp. PCC 6803 involved in tocopherol biosynthesis." FEBS Letters **499**(1-2): 15-20.

Schmelz, S. and J. H. Naismith (2009). "Adenylate-forming enzymes." Current Opinion in Structural Biology **19**(6): 666-671.

Schmutz, E., A. Mühlenweg, S. M. Li and L. Heide (2003). "Resistance genes of aminocoumarin producers: Two type II topoisomerase genes confer resistance against coumermycin A_1 and clorobiocin." Antimicrobial Agents and Chemotherapy **47**(3): 869-877.

Schmutz, E., M. Steffensky, J. Schmidt, A. Porzel, S. M. Li and L. Heide (2003). "An unusual amide synthetase (CouL) from the coumermycin A_1 biosynthetic gene cluster from *Streptomyces rishiriensis* DSM 40489." Eur. J. Biochem. **270**(22): 4413-4419.

Schultz, A. W., C. A. Lewis, M. R. Luzung, P. S. Baran and B. S. Moore (2010). "Functional characterization of the cyclomarin/cyclomarazine prenyltransferase CymD directs the biosynthesis of unnatural cyclic peptides." Journal of Natural Products **73**(3): 373-377.

Shiozawa, K., M. Oka, H. Soda, M. Yoshikawa, Y. Ikegami, J. Tsurutani, K. Nakatomi, Y. Nakamura, S. Doi, T. Kitazaki, Y. Mizuta, K. Murase, H. Yoshida, D. D. Ross and S. Kohno (2004). "Reversal of breast cancer resistance protein (BCRP/ABCG2)-mediated drug resistance by novobiocin, a coumermycin antibiotic." International Journal of Cancer **108**(1): 146-151.

Sohng, J. K., T. J. Oh, J. J. Lee and C. G. Kim (1997). "Identification of a gene cluster of biosynthetic genes of rubradirin substructures in *S. achromogenes var. rubradiris* NRRL3061." Molecules and Cells **7**(5): 674-681.

Sosio, M., H. Kloosterman, A. Bianchi, P. de Vreugd, L. Dijkhuizen and S. Donadio (2004). "Organization of the teicoplanin gene cluster in *Actinoplanes teichomyceticus* " Microbiology **150**(1): 95-102.

Stachelhaus, T., A. Huser and M. A. Marahiel (1996). "Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases." Chemistry and Biology **3**(11): 913-921.

Stachelhaus, T., H. D. Mootz, V. Bergendahl and M. A. Marahiel (1998). "Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain." Journal of Biological Chemistry **273**(35): 22773-22781.

Stachelhaus, T., H. D. Mootz and M. A. Marahiel (1999). "The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases." Chemistry and Biology **6**(8): 493-505.

Stec, E., D. Pistorius, R. Müller and S. M. Li (2011). "AuaA, a Membrane-Bound Farnesyltransferase from *Stigmatella aurantiaca*, Catalyzes the Prenylation of 2-Methyl-4-hydroxyquinoline in the Biosynthesis of Aurachins." Chembiochem : a European journal of chemical biology **12**(11): 1724-1730.

Steffan, N., A. Grundmann, W. B. Yin, A. Kremer and S.-M. Li (2009). "Indole prenyltransferases from fungi: A new enzyme group with high potential for the production of prenylated indole derivatives " Current Medicinal Chemistry **16**(2): 218-231.

Steffensky, M., S. M. Li and L. Heide (2000b). "Cloning, overexpression, and purification of novobiocic acid synthetase from *Streptomyces spheroides* NCIMB 11891." Journal of Biological Chemistry **275**(28): 21754-21760.

Steffensky, M., A. Mühlenweg, Z. X. Wang, S. M. Li and L. Heide (2000a). "Identification of the novobiocin biosynthetic gene cluster of *Streptomyces spheroides* NCIB 11891." Antimicrobial Agents and Chemotherapy **44**(5): 1214-1222.

Stegmann, E., H. J. Frasch and W. Wohlleben (2010). "Glycopeptide biosynthesis in the context of basic cellular functions." Current Opinion in Microbiology **13**(5): 595-602.

Stegmann, E., C. Rausch, S. Stockert, D. Burkert and W. Wohlleben (2006). "The small MbtH-like protein encoded by an internal gene of the balhimycin biosynthetic gene cluster is not required for glycopeptide production." FEMS Microbiology Letters **262**(1): 85-92.

Stuible, H., D. Büttner, J. Ehlting, K. Hahlbrock and E. Kombrink (2000). "Mutational analysis of 4-coumarate:CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes." FEBS Letters **467**(1): 117-122.

Su, Y., P. Hu, S. H. Lee and P. J. Sinko (2007). "Using novobiocin as a specific inhibitor of breast cancer resistant protein to assess the role of transporter in the absorption and disposition of topotecan." J Pharm Pharm Sci **10**(4): 519-536.

Süssmuth, R. D., S. Pelzer, G. Nicholson, T. Walk, W. Wohlleben and G. Jung (1999). "New Advances in the Biosynthesis of Glycopeptide Antibiotics of the Vancomycin Type from *Amycolatopsis mediterranei*." Angewandte Chemie International Edition **38**(13-14): 1976-1979.

Suvarna, K., D. Stevenson, R. Meganathan and M. E. Hudspeth (1998). "Menaquinone (vitamin K2) biosynthesis: localization and characterization of the *menA* gene from *Escherichia coli*." Journal of Bacteriology **180**(10): 2782-2787.

Tello, M., T. Kuzuyama, L. Heide, J. P. Noel and S. B. Richard (2008). "The ABBA family of aromatic prenyltransferases: broadening natural product diversity." Cellular and Molecular Life Sciences **65**(10): 1459-1463.

Thomas, M. G., Y. A. Chan and S. G. Ozanick (2003). "Deciphering tuberactinomycin biosynthesis: isolation, sequencing, and annotation of the viomycin biosynthetic gene cluster." Antimicrobial Agents and Chemotherapy **47**(9): 2823-2830.

Trefzer, A., S. Pelzer, J. Schimana, S. Stockert, C. Bihlmaier, H. P. Fiedler, K. Welzel, A. Vente and A. Bechthold (2002). "Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic." Antimicrobial Agents and Chemotherapy **46**(5): 1174-1182.

Tsai, F. T. F., O. M. Singh, T. Skarzynski, A. J. Wonacott, S. Weston, A. Tucker, R. A. Pauptit, A. L. Breeze, J. P. Poyser, R. O'Brien, J. E. Ladbury and D. B. Wigley (1997). "The high-resolution crystal structure of a 24-kDa gyrase B fragment from *E. coli* complexed with one of the most potent coumarin inhibitors, clorobiocin." Proteins **28**(1): 41-52.

Turgay, K., M. Krause and M. A. Marahiel (1992). "Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes." Molecular Microbiology **6**(4): 529-546.

Turgay, K., M. Krause and M. A. Marahiel (1992). "Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes." Mol Microbiol **6**(4): 529-546.

van Wageningen, A. M., P. N. Kirkpatrick, D. H. Williams, B. R. Harris, J. K. Kershaw, N. J. Lennard, M. Jones, S. J. Jones and P. J. Solenberg (1998). "Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic." Chemistry and Biology **5**(3): 155-162.

von Döhren, H. and U. Gräfe (2008). General Aspects of Secondary Metabolism. Biotechnology, Wiley-VCH Verlag GmbH**:** 1-55.

Wang, Z. X., S. M. Li and L. Heide (2000). "Identification of the coumermycin A_1 biosynthetic gene cluster of *Streptomyces rishiriensis* DSM 40489." Antimicrobial Agents and Chemotherapy **44**(11): 3040-3048.

Watve, M. G., R. Tickoo, M. M. Jog and B. D. Bhole (2001). "How many antibiotics are produced by the genus *Streptomyces*?" Arch Microbiol **176**(5): 386-390.

Weist, S., B. Bister, O. Puk, D. Bischoff, S. Pelzer, G. J. Nicholson, W. Wohlleben, G. Jung and R. D. Süssmuth (2002). "Fluorobalhimycin-a new chapter in glycopeptide antibiotic research." Angewandte Chemie International Edition **41**(18): 3383-3385.

Weist, S., C. Kittel, D. Bischoff, B. Bister, V. Pfeifer, G. J. Nicholson, W. Wohlleben and R. D. Süssmuth (2004). "Mutasynthesis of glycopeptide antibiotics: variations of vancomycin's AB-ring amino acid 3,5-dihydroxyphenylglycine." Journal of the American Chemical Society **126**(19): 5942-5943.

Wherli, W. (1977). "Topics in current chemistry." Ansamycins. chemistry, biosynthesis and biological activity **72**(Berlin Heidelberg): 21-49.

Whitesell, L. and S. L. Lindquist (2005). "HSP90 and the chaperoning of cancer." Nature Reviews Cancer **5**(10): 761-772.

Wohlleben, W., E. Stegmann and R. D. Sussmuth (2009). "Chapter 18. Molecular genetic approaches to analyze glycopeptide biosynthesis." Methods in Enzymology **458**: 459-486.

Wolpert, M., B. Gust, B. Kammerer and L. Heide (2007). "Effects of deletions of *mbtH*-like genes on clorobiocin biosynthesis in *Streptomyces coelicolor*." Microbiology **153**: 1413-1423.

Xie, J., J. G. Pierce, R. C. James, A. Okano and D. L. Boger (2011). "A Redesigned Vancomycin Engineered for Dual d-Ala-d-Ala and d-Ala-d-Lac Binding Exhibits Potent Antimicrobial Activity Against Vancomycin-Resistant Bacteria." Journal of the American Chemical Society **133**(35): 13946-13949.

Yazaki, K., K. Sasaki and Y. Tsurumaru (2009). "Prenylation of aromatic compounds, a key diversification of plant secondary metabolites." Phytochemistry **70**(15-16): 1739-1745.

Yonus, H., P. Neumann, S. Zimmermann, J. J. May, M. A. Marahiel and M. T. Stubbs (2008). "Crystal structure of DltA. Implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains." Journal of Biological Chemistry **283**(47): 32484-32491.

Zhanel, G. G., D. Calic, F. Schweizer, S. Zelenitsky, H. Adam, P. R. Lagace-Wiens, E. Rubinstein, A. S. Gin, D. J. Hoban and J. A. Karlowsky (2010). "New lipoglycopeptides: a comparative review of dalbavancin, oritavancin and telavancin." Drugs **70**(7): 859-886.

Zhang, W., J. R. Heemstra, Jr., C. T. Walsh and H. J. Imker (2010a). "Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins." Biochemistry **49**(46): 9946-9947.

Zhang, W., B. Ostash and C. T. Walsh (2010b). "Identification of the biosynthetic gene cluster for the pacidamycin group of peptidyl nucleoside antibiotics." Proceedings of the National Academy of Sciences of the United States of America **107**(39): 16828-16833.

Chapter 5 • *Aromatic Prenylation in Phenazine Biosynthesis*

This manuscript has originally been published in The Journal of Biological Chemistry.

'Aromatic Prenylation in Phenazine Biosynthesis: Dihydrophenazine-1-carboxylate Dimethylallyltransferase from *Streptomyces anulatus*'

Orwah Saleh ‡, Bertolt Gust ‡, **Björn Boll ‡,** Hans-Peter Fiedler § and Lutz Heide ‡*

‡ Pharmazeutische Biologie, Pharmazeutisches Institut, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

§Mikrobiologisches Institut, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

* To whom correspondence should be addressed: E-mail: heide@uni-tuebingen.de
Phone: +49 (0)7071 29 72460
Fax: +49 (0)7071 29 5250

The Journal of Biological Chemistry Vol. 284, No. 21, pp. 14439-47, May 22, 2009

Received February 25, 2009, Published online, April 1, 2009,

DOI: 10.1074/jbc.M901312200

Preface − About the Manuscript

Many aromatic prenyltransferases have been isolated and characterized from Actinomycetes. So far none of them was capable to prenylate phenazines. The following publication describes the identification and characterization of the first aromatic prenyltransferase of prenylated phenazine 1-carboxylic acid biosynthesis. Sequence comparison with known aromatic prenyltransferases reveals striking similarity but obviously different substrate specifity. The enzyme was independent from the presence of divalent cations.

Author contributions:

- Orwah Saleh
  - decisively involved in initial hypothesis generation
  - decisively involved in planning, establishing and accomplishment of experimental setup
  - construction and screening of cosmid library, heterologous expression, gene inactivation, expression and purification of PpzP
  - generation of data (antimicrobial activity assays, HPLC analyses, LC-MS and Prenyltransferase assays)
  - decisively involved in all data discussions
  - writing of the manuscript
  - preparation of all figures
- Bertolt Gust
  - assisted in construction and screening of cosmid library, heterologous expression and gene inactivation
  - involved in data discussion and analyses
- **Björn Boll**
  - cloning, overexpression and purification of PpzP
  - enzyme assay to detect prenyltransferase activity of PpzP
- Hans-Peter Fiedler
  - LC-MS measurement
  - provided phenazine standards
- Lutz Heide
  - supervised the project
  - decisively involved in initial hypothesis generation
  - decisively involved in all data discussions and analyses
  - manuscript preparation

My personal part for this manuscript involved the initial cloning, expression and purification of PpzP. I performed activity assays which proved for the first time that PpzP is an active prenyltransferase and is capable of prenylating phenazines.

Abstract

The bacterium *Streptomyces anulatus* 9663, isolated from the intestine of different arthropods, produces prenylated derivatives of phenazine 1-carboxylic acid. From this organism, we have identified the prenyltransferase gene *ppzP*. *ppzP* resides in a gene cluster containing orthologs of all genes known to be involved in phenazine 1-carboxylic acid biosynthesis in *Pseudomonas* strains as well as genes for the six enzymes required to generate dimethylallyl diphosphate *via* the mevalonate pathway. This is the first complete gene cluster of a phenazine natural compound from streptomycetes. Heterologous expression of this cluster in *Streptomyces coelicolor* M512 resulted in the formation of prenylated derivatives of phenazine 1-carboxylic acid. After inactivation of *ppzP*, only nonprenylated phenazine 1-carboxylic acid was formed. Cloning, overexpression, and purification of PpzP resulted in a 37-kDa soluble protein, which was identified as a 5,10-dihydrophenazine 1-carboxylate dimethylallyltransferase, forming a C–C bond between C-1 of the isoprenoid substrate and C-9 of the aromatic substrate. In contrast to many other prenyltransferases, the reaction of PpzP is independent of the presence of magnesium or other divalent cations. The K_m value for dimethylallyl diphosphate was determined as 116 μM. For dihydro-PCA, half-maximal velocity was observed at 35 μM. K_{cat} was calculated as 0.435 s^{-1}. PpzP shows obvious sequence similarity to a recently discovered family of prenyltransferases with aromatic substrates, the ABBA prenyltransferases. The present finding extends the substrate range of this family, previously limited to phenolic compounds, to include also phenazine derivatives.

Introduction

The transfer of isoprenyl moieties to aromatic acceptor molecules gives rise to an astounding diversity of secondary metabolites in bacteria, fungi, and plants, including many compounds that are important in pharmacotherapy. However, surprisingly little biochemical and genetic data are available on the enzymes catalyzing the C-prenylation of aromatic substrates. Recently, a new family of aromatic prenyltransferases was discovered in streptomycetes (1), Gram-positive soil bacteria that are prolific producers of antibiotics and other biologically active compounds (2). The members of this enzyme family show a new type of protein fold with a unique α-β-β-α architecture (3) and were therefore termed ABBA prenyltransferases (1). Only 13 members of this family can be

Originally published in The *Journal of Biological Chemistry*. (2009), ***284***, *21*, *pp. 14439–47*

identified by sequence similarity searches in the data base at present, and only four of them have been investigated biochemically (3-6). Up to now, only phenolic compounds have been identified as aromatic substrates of ABBA prenyltransferases. We now report the discovery of a new member of the ABBA prenyltransferase family, catalyzing the transfer of a dimethylallyl moiety to C-9 of 5,10-dihydrophenazine 1-carboxylate (dihydro-PCA).2 Streptomyces strains produce many of prenylated phenazines as natural products. For the first time, the present paper reports the identification of a prenyltransferase involved in their biosynthesis.

Streptomyces anulatus 9663, isolated from the intestine of different arthropods, produces several prenylated phenazines, among them endophenazine A and B (Fig. 1A) (7). We wanted to investigate which type of prenyltransferase might catalyze the prenylation reaction in endophenazine biosynthesis. In streptomycetes and other microorganisms, genes involved in the biosynthesis of a secondary metabolite are nearly always clustered in a contiguous DNA region. Therefore, the prenyltransferase of endophenazine biosynthesis was expected to be localized in the vicinity of the genes for the biosynthesis of the phenazine core (*i.e.* of PCA).

In *Pseudomonas*, an operon of seven genes named *phzABCDEFG* is responsible for the biosynthesis of PCA (8). The enzyme PhzC catalyzes the condensation of phosphoenolpyruvate and erythrose-4-phosphate (*i.e.* the first step of the shikimate pathway), and further enzymes of this pathway lead to the intermediate chorismate. PhzD and PhzE catalyze the conversion of chorismate to 2-amino-2-deoxyisochorismate and the subsequent conversion to 2,3-dihydro-3-hydroxyanthranilic acid, respectively. These reactions are well established biochemically. Fewer data are available about the following steps (*i.e.* dimerization of 2,3-dihydro-3-hydroxyanthranilic acid, several oxidation reactions, and a decarboxylation, ultimately leading to PCA *via* several instable intermediates). From *Pseudomonas*, experimental data on the role of PhzF and PhzA/B have been published (8,9), whereas the role of PhzG is yet unclear. Surprisingly, the only gene cluster for phenazine biosynthesis described so far from streptomycetes (10) was found not to contain a *phzF* orthologue, raising the question of whether there may be differences in the biosynthesis of

Originally published in The *Journal of Biological Chemistry*. (2009), ***284***, 21, pp. 14439–47

phenazines between *Pseudomonas* and *Streptomyces*.

Screening of a genomic library of the endophenazine producer strain *S. anulatus* now allowed the identification of the first complete gene cluster of a prenylated phenazine, including the structural gene of dihydro-PCA dimethylallyltransferase.

Experimental Procedures

Bacterial Strains, Plasmids, and Culture Conditions

S. anulatus 9663 has been isolated previously from the intestine of different arthropods (7,11). It was grown in liquid YMG medium or on solid MS medium. For production of secondary metabolites, the medium described by Sedmera *et al.* (12) was used.

Escherichia coli XL1 Blue MRF, *E. coli* SURE (Stratagene, Heidelberg, Germany), *E. coli* BW 25113, and *E. coli* ET 12567 (pUB307) were used for cloning and were grown in liquid or on solid (1.5 % agar) Luria-Bertani or SOB medium at 37 °C. The REDIRECT technology kit for PCR targeting was obtained from Plant Bioscience Limited (Norwich, UK). For inactivation experiments, the *aac(3)IV/oriT* (apramycin resistance) cassette from pIJ773 (13) was used. Carbenicillin (50–100 μg ml^{-1}), apramycin (50 μg ml^{-1}), kanamycin (50 μg ml^{-1}), chloramphenicol (25 μg ml^{-1}), and nalidixic acid (20 μg ml^{-1}) were used for selection of recombinant strains.

Chemicals

Kanamycin and carbenicillin were purchased from Genaxxon BioSciences GmbH (Biberach, Germany); apramycin and nalidixic acid were from Sigma; chloramphenicol was from Merck; and phenazine 1-carboxylic acid was from InFormatik. Dimethyllallyl diphosphate was synthesized as described by Woodside *et al.* (14). Endophenazine A was isolated from cultures of *Streptomyces cinnamonensis* DSM 1024 as described by Gebhardt *et al. (7)*.

Genetic Procedures

Standard methods for DNA isolation and manipulation were performed as described by Kieser *et al.* (15) and Sambrook *et al.* (16). DNA fragments were isolated from agarose gels by using a PCR purification kit (Amersham Biosciences). Genomic DNA was isolated from Streptomyces strains by lysozyme treatment and phenol/chloroform extraction as described by Kieser *et al.* (15).

Construction and Screening of the Cosmid Library

Chromosomal DNA from *S. anulatus* was partially digested with *Sau*3AI, dephosphorylated and then ligated into the *Bam*HI sites of SuperCos 1 (Stratagene) according to the manufacturer's instructions. The ligation products were packaged with Gigapack III XL (Stratagene) and transduced into *E. coli* SURE. Colony hybridization was performed on Hybond-N membranes (Amersham Biosciences). *ephzA* from *S. cinnamonensis* (10) was used as hybridization probe for the first screening of the library. The digoxigenin-labeled *ephzA* was generated using the PCR digoxigenin labeling mix (Roche Applied Science) with the primers *ephzA*_for (5′-ATG AGC ACC CCC CTG ACC ACC-3′) and *ephzA*_rev (5′-TCA GGA GGG GAT CCA GTC CCG-3′).

A second screening was performed by PCR for the identification of the following genes: *phzD*, *phzF*, *hmgr* (3-hydroxy-3-methyl-glutaryl-CoA reductase), *hmgs* (3-hydroxy-3-methylglutaryl-CoA synthase), and *mdpd* (mevalonate diphosphate decarboxylase). The following primers were used: *phzD*_for (5′-CGC GCC GTC CTG (A/G)TN CA(C/T) GA(C/T) (A/C/T)T-3′) and *phzD*_rev (5′-CGG TGG TGG TCC CGG (G/C)(A/T)(A/G) AA(A/G) TCN (G/C)-3′); *phzF*_for (5′-CAT CCG GAT CTT GAC CCC NGT NAA (C/T)GA-3′) and *phzF*_rev (5′-GAG GGG CGC CCC AT(C/T) TCN CAN CC-3′); HMGR_for (5′-GGG CAT CGC CGC GAC CCT CGT GGA GGA GGG-3′) and HMGR_rev (5′-GCG ATG ACG GGG AGG CGC CGG GCG TTC TC-3′); HMGS_for (5′-GCC AAG TCC GCC GGN GTN TA(C/T) GT-3′) and HMGS_rev (5′-AGC CGG AAG GGG CCN GTN GT(C/T) TG-3′); MDPD_for (5′-GAC CCT GGA CGT CTT CCC NAC NAC NAC-3′) and MDPD_rev (5′-GCG TTC CGC TCG GC(A/G/T) AT(C/T) TCN-3′). PCRs were carried out with *Taq* polymerase.

Heterologous Expression of Cosmids ppzOS04 and ppzOS02

The plasmid pIJ787 (17) was first digested with *Dra*I and *Bsa*I, and the 4990-bp fragment containing the integrase cassette was used to replace the *bla* gene in the SuperCos 1 backbone in cosmids 11C7 and 18A9, using λ RED-mediated recombination (17,18), generating ppzOS02 and ppzOS04, respectively. Both cosmids were first transformed into the nonmethylating host *E. coli* ET12567, and the nonmethylated DNA was introduced into *Streptomyces coelicolor* M512 *via* triparental conjugation.

Inactivation of the Gene *ppzP*

An apramycin resistance cassette (*aac(3)IV*) was amplified from plasmid pUG019 (19) using the following primers: *ppzP*_F (5′-CGC CCC AAG GGT GTC TTG TCG ACC TGT GGA GGA AAA ATG TCT AGA ATT CCG GGG ATC CGT CGA CC-3′) and *ppzP*_R (5′-CAA GCC CTT GTC CTT CAC ATG CCG ACG GGT GAG GCG CTA ACT AGT TGT AGG CTG GAG CTG CTT C-3′). Underlined are restriction sites for *Xba*I and *Spe*I, used for later removal of the cassette. The resulting 1077-bp PCR product was used to replace the *ppzP* gene on cosmid ppzOS04 by λ RED-mediated recombination, resulting in cosmid ppzOS05. Deletion of the *aac(3)IV* cassette from ppzOS05 was carried out by digestion with *Xba*I and *Spe*I and religation, resulting in cosmid ppzOS09. The resulting construct was introduced into *S. coelicolor* M512 *via* triparental conjugation (15).

Production and Analysis of Secondary Metabolites

Exconjugants of all mutants as well as wild type *S. anulatus* were precultured for 48 h in liquid YMG medium (50 ml). 50 ml of production medium (20) was then inoculated with 2.5 ml of the precultures. The flasks were agitated on a rotary shaker at 30 °C and 200 rpm for 120 h. For cultivation of mutants, all liquid media contained kanamycin (50 μg ml^{-1}).

For isolation of endophenazine A, mycelia from 50-ml cultures were centrifuged at 3500 × g for 10 min. The supernatant was discarded, and the cells were extracted with methanol (10 ml) by vortexing. The extract was mixed with sodium acetate buffer (10 ml; 1 M, pH 4.0) and extracted with dichloromethane (5 ml). After separation of the organic phase, the solvent was evaporated, and the residue was redissolved in methanol (0.5 ml).

Extracts were analyzed with HPLC (Agilent 1100 series; Waldbronn, Germany) by using an Eclipse XDB-C18 column (4.6 × 150 mm, 5 μm; Agilent) at a flow rate of 1 ml min^{-1} with a linear gradient from 10 to 100 % of solvent B in 20 min (solvent A: water/phosphoric acid (999:1); solvent B, acetonitrile) and detection at 252 and 365 nm. Additionally, a UV spectrum from 200 to 400 nm was logged by a photodiode array detector. The absorbance at 365 nm was used for quantitative analysis, employing authentic reference samples of PCA and endophenazine A as external standards.

Analysis by LC-MS

The extracts were examined with LC-MS and LC-MS^2 analysis using a Nucleosil 100-C18 column (3 µm, 100 × 2 mm) coupled to an ESI mass spectrometer (LC/MSD Ultra Trap System XCT 6330; Agilent Technology). Analysis was carried out at a flow rate of 0.4 ml min^{-1} with a linear gradient from 10 to 100 % of solvent B in 15 min (solvent A: water/formic acid (999:1); solvent B: acetonitrile/formic acid (999.4:0.6)). Detection was carried out at 230, 260, 280, 360, and 435 nm. Electrospray ionization (positive and negative ionization) in Ultra Scan mode with capillary voltage of 3.5 kV and heated capillary temperature of 350 °C was used for LC-MS analysis. For LC-MS^2 and LC-MS^3, the analysis was carried out in positive ionization mode with capillary voltage of 3.5 kV at 350 °C. For LC-MS^2 identification of the enzymatic product endophenazine A, the mass 293 ± 0.5 Da was selected for fragmentation. In LC-MS^3, the mass 275 ± 1 Da was selected for fragmentation.

Overexpression and Purification of PpzP Protein

ppzP was amplified by using the primers *ppzP*_pHis_F (5′-ACC TGT GGA GAA TTC ATG TCA GAA TCC GCT GAG CT-3′) and *ppzP*_pHis_R (5′-CCG GAC GGG CTC GAG GCT ATC CGG CAT CGG CGG TCA-3′). The underlined letters represent *Eco*RI and *Xho*I restriction sites, respectively. The resulting PCR fragment was digested with *Eco*RI and *Xho*I and ligated into plasmid pHis8 (21) digested with the same restriction enzymes. The resulting plasmid, pHis8-OS01, was verified by restriction mapping and sequencing.

E. coli BL21(DE3)pLysS cells harboring plasmid pHis8-OS01 were cultivated in 2 liters of liquid TB medium containing 50 µg ml^{-1} kanamycin and grown at 37 °C to an A_{600} of 0.6. The temperature was lowered to 20 °C, and isopropyl 1-thio-β-d-galactopyranoside was added to a final concentration of 0.5 mM. The cells were cultured for a further 10 h at 20 °C and harvested. 30 ml of lysis buffer (50 mM Tris-HCl, pH 8.0, 1 M NaCl, 10% glycerol, 10 mM β-mercaptoethanol, 20 mM imidazole, 0.5 mg ml^{-1} lysozyme, 0.5 mM phenylmethylsulfonyl fluoride) were added to the pellet (40 g). After stirring at 4 °C for 30 min, cells were ruptured with a Branson sonifier by 4 °C. The lysate was centrifuged (55,000 × g, 45 min). Affinity chromatography with 4 ml of Ni^{2+}-nitrilotriacetic acid-agarose resin (Qiagen, Hilden, Germany) was carried out according to the manufacturer's instructions, using 2 × 2.5 ml of 250 mM

imidazole (in 50 mM Tris-HCl, pH 8.0, 1 M NaCl, 10% glycerol, 10 mM β-mercaptoethanol) for elution. Subsequently, a buffer exchange was carried out by PD10 columns (Amersham Biosciences), which were eluted with 50 mM Tris-HCl, pH 8.0, 1 M NaCl, 10% glycerol, and 2 mM 1,4-dithiothreitol. Approximately 30 mg of purified PpzP were obtained from 2 liters of cultures.

Assay for Prenyltransferase Activity

The reaction mixture (100 µl) contained 80 mM Na-TAPS (pH 7.5) (Sigma), 0.8 mM freshly prepared dihydro-PCA, 0.4 mM dimethylallyl diphosphate (DMAPP), 500 mM NaCl, and 5 µg ml^{-1} PpzP. For the preparation of dihydro-PCA, 90 µl of 100 mM freshly dissolved sodium dithionite (Merck) were mixed with 10 µl of 100 mM PCA dissolved in Tris-HCl, pH 8.0. 8 µl of this mixture was added to the incubation mixture. After incubation for 20 or 30 min at 30 °C, 15 µl of 100 mM sodium persulfate (Sigma) were added to oxidize dihydro-PCA and dihydroendophenazine A into PCA and endophenazine A, respectively. The mixture was immediately extracted with 100 µl of ethyl acetate/formic acid (975:25). After vortexing and centrifugation, 75 µl of the organic phase were evaporated. The residue was dissolved in 100 µl of methanol, and 35 µl thereof were analyzed by HPLC using an Eclipse XDB-C18 column (4.6 × 150 mm, 5 µm; Agilent) with the same mobile phase and gradient described above for secondary metabolite analysis. Detection was carried out at 252 and 365 nm. Quantitative analysis was carried out using the absorbance at 365 nm, as described under “Production and Analysis of Secondary Metabolites.”

Calculation of Kinetic Constants

K_m and K_{cat} were calculated using the GraphPad Prism software, version 5.01 for Windows (GraphPad Software Inc., La Jolla, CA).

Results

Cloning of a Biosynthetic Gene Cluster for Prenylated Phenazines

A genomic library of the phenazine producer strain *S. anulatus* was constructed in cosmid vector SuperCos 1 (Stratagene). 2000 independent cosmid clones were subjected to a colony blot screening with a labeled probe of the phenazine biosynthesis gene *ephzA* from *S. cinnamonensis* (10). From the genome size of streptomycetes (~8.5 megabases) and the average insert size of SuperCos 1 (~38 kb), it could be estimated that a

single genome locus would be represented, on average, in nine different cosmid clones. However, the screening revealed 26 positive clones, indicating that more than one genomic locus hybridized with *ephzA*. In order to identify those cosmids that contained the correct locus and that were likely to contain the entire gene cluster, these 26 cosmids were screened with degenerate primers for two further phenazine biosynthetic genes, *phzD* and *phzF*, and for three genes of the mevalonate pathway for isoprenoid biosynthesis (see "Experimental Procedures"). In another Streptomyces strain, it had been shown that the prenyl moiety of endophenazine A is formed *via* the mevalonate pathway (22), whereas in streptomycetes, the methyl erythritol phosphate pathway is used for the formation of isoprenoids of primary metabolism (23). Two cosmids, 18A9 and 11C7, were identified that gave PCR products with all five primer pairs. Sequencing of the PCR products confirmed that they represented *phzD*, *phzF*, and the three mevalonate pathway genes. Sequencing of cosmid 18A9 revealed a DNA region of 17.5 kb comprising 18 putative coding sequences (Fig. 1B), which together code for all enzymatic functions expected to be involved in endophenazine formation. The results of data base comparisons for these genes are listed in Table 1. Seven of the putative coding sequences, designated as *ppzA*, *ppzC*, *ppzD*, *ppzE*, *ppzF*, *ppzG*, and *ppzB*, showed obvious similarities to the seven genes that direct the biosynthetic pathway from phosphoenolpyruvate and erythrose-4-phosphate to PCA in *Pseudomonas* strains (8,9). In addition, the gene *ppzM* was identified, showing striking similarity to *phzM* from *P. aeroginosa*. *phzM* is proposed to be responsible for the N-methylation in the biosynthesis of pyocyanine (1-hydroxy-5-N-methylphenazine) (24). *ppzM* may therefore be responsible for the N-methylation reaction in the biosynthesis of endophenazine B (Fig. 1A).

Sequence analysis further revealed a group of six coding sequences with striking similarity to genes of the mevalonate pathway (23), leading from acetyl-CoA and acetoacetyl-CoA to isopentyl diphosphate and DMAPP (see Table 1). Orthologs of these six genes, arranged in the exact same order, are found in the biosynthetic gene cluster of furaquinocin (25), naphterpin (26), terepenticin (27), and BE 40644 (28). *S. anulatus*, therefore, is one of the rare *Streptomyces* species that possesses the genes of the mevalonate pathway, a

Originally published in The *Journal of Biological Chemistry*. (2009), ***284***, 21, pp. 14439–47

feature limited to ~1% of the strains of this genus (29).

TABLE 1: Genes in the endophenazine biosynthetic gene cluster of *S. anulatus* 9663

Gene	Amino acid	Proposed function	Ortholog identified by BLAST search	Identity/ Similarity
ppzP	299	Prenyltransferase	Fnq26, *S. cinnamonensis*	44/64
mk	345	Mevalonate kinase	MEVK, *Streptomyces* sp. KO-3988	67/79
mdpd	351	Diphosphomevalonate decarboxylase	MDPD, *Streptomyces* sp. KO-3988	73/80
pmk	371	Phosphomevalonate kinase	PMK, *Streptomyces* sp. KO-3988	68/75
ippi	363	Isopentenyldiphosphate δ-isomerase	Fni, *Streptomyces* sp. CL190	79/88
hmgr	353	3-Hydroxy-3-methylglutaryl coenzyme A reductase	HMGR, *Streptomyces* sp. CL190	86/92
hmgs	391	3-Hydroxy-3-methylglutaryl CoA synthase	HMGS, *Streptomyces* sp. CL190	78/87
ppzT	327	3-Oxoacyl-[acyl-carrier-protein] synthase	3-Oxoacyl-[acyl-carrier-protein] synthase, *Streptomyces* sp. KO-3988	79/86
ppzU	221	Flavodoxin	SGR_4078, *Streptomyces griseus*	59/75
ppzV	206	Hypothetical protein	Fnq22, *S. cinnamonensis*	62/77
ppzM	340	*N*-Methyltransferase	PhzM, *Pseudomonas aeruginosa*	44/58
ppzB	168	Phenazine biosynthesis	PhzB, *Pseudomonas fluorescens*	51/67
ppzG	233	FMN-dependent oxidase	PhzG, *P. fluorescens*	54/71
ppzF	279	*trans*-2,3-Dihydro 3-hydroxyanthranilate isomerase	PhzF, *P. fluorescens*	68/77
ppzE	646	2-Amino-2-desoxy-isochorismate synthase	PhzE, *P. fluorescens*	60/76
ppzD	207	2,3-Dihydro-3-hydroxy-anthranilate (DHHA) Synthase	PhzD, *P. fluorescens*	60/76
p*pzC*	392	3-Deoxy-d-arabino-heptulosonic acid 7-phosphate synthase	PhzC, *P. fluorescens*	59/73
ppzA	162	Phenazine biosynthesis	PhzA, *P. fluorescens*	78/87

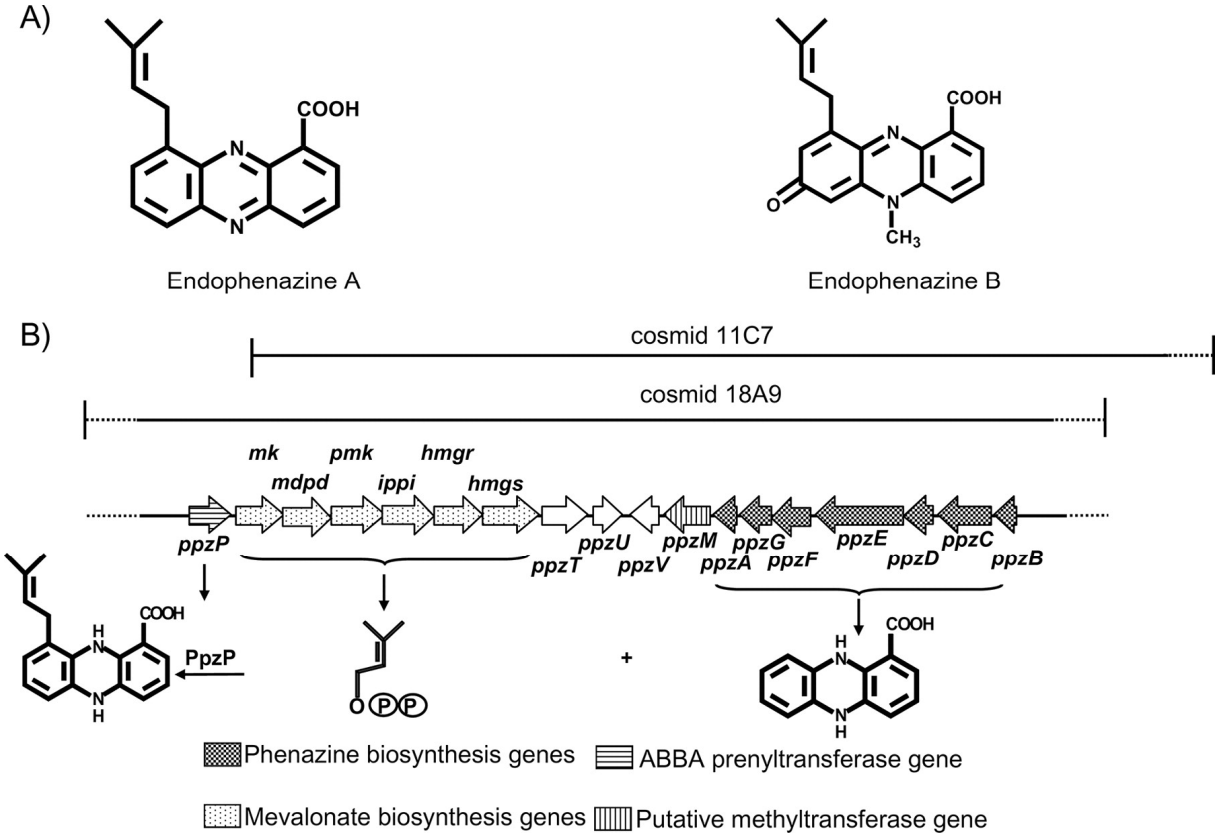


Figure 1: A, prenylated phenazines from *S. anulatus* 9663. B, biosynthetic gene cluster of endophenazine A.

Heterologous Expression of the Endophenazine Cluster in *S. coelicolor* M512

λ RED-mediated recombination in *E. coli* was used to replace the β-lactamase gene within the SuperCos 1 backbone of cosmid 18A9 with the cassette pIJ787 containing a tetracycline resistance gene and the integration function of the phage ΦC31 (13). The resulting cosmid ppzOS04 was introduced into the genome of *S. coelicolor* M512 using triparental conjugation (15).

As control, the same pIJ787 cassette was introduced in an empty SuperCos 1 vector, and the resulting construct, ppzOS30, was also integrated into the genome of *S. coelicolor* M512. The secondary metabolite production of the resulting integration mutants was analyzed by HPLC-UV and LC-MS, using the wild type strain *S. anulatus* as comparison. In accordance with previous results (7), *S. anulatus* was found to produce PCA and endophenazine A (Fig. 2A). The polar compound PCA was predominantly found in the culture medium, but some amount was also present in the mycelial extract. The more lipophilic, prenylated compounds endophenazine A and B were associated with the cells (*i.e.* were

mainly found in the pellet after centrifugation of the cultures). The heterologous expression strain *S. coelicolor*(ppzOS04) produced both phenazine 1-carboxylic acid (most of which was found in the medium) and endophenazine A (most of which was associated with the cells) (Fig. 2B). The presence of both compounds was also shown by LC-MS at *m/z* = 225 $[M + H]^+$ and *m/z* = 293 $[M + H]^+$ for PCA and endophenazine A, respectively. LC-MS^2 showed that both compounds had the same fragmentation pattern as the authentic reference samples (data not shown). In contrast, the control strain *S. coelicolor* M512(ppzOS30), which had integrated the empty vector, did not produce any phenazines (data not shown).

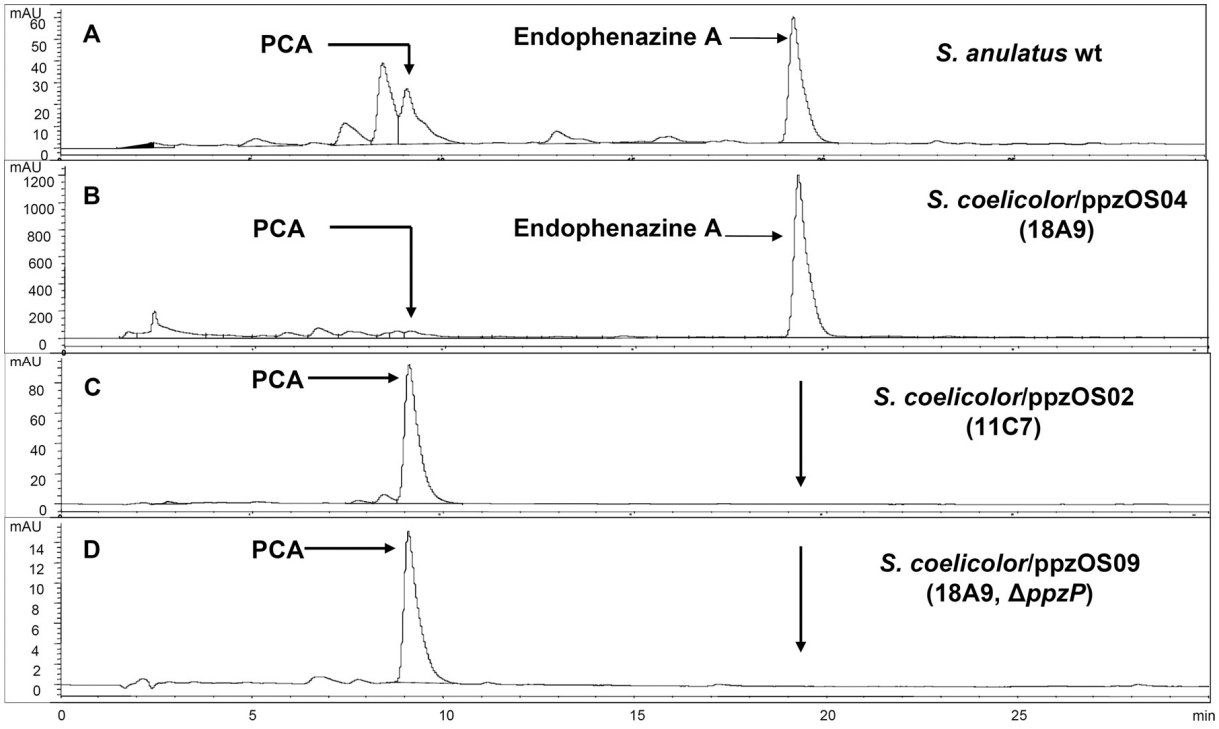


Figure 2: HPLC analysis of cell extracts of different *Streptomyces* strains (detection: UV, 365 nm). A, wild-type *S. anulatus*. B, *S. coelicolor* M512 containing the complete endophenazine cluster from cosmid 18A9. C, *S. coelicolor* M512 containing a partial endophenazine cluster (cosmid 11C7; see Fig. 1). D, *S. coelicolor* M512 containing the complete endophenazine cluster (cosmid 18A9) but with deletion of the putative prenyltransferase gene *ppzP*. PCA, phenazine 1-carboxylic acid. The polar compound PCA is predominantly found in the culture medium, but varying amounts are also present in the mycelial extracts, as shown here. The cultures also produced varying amounts of PCA-methyl ester, which is not separated from PCA in this HPLC system but could be identified in the LC-MS analysis (data not shown). The more lipophilic, prenylated compound endophenazin A is predominantly found in cell extracts, not in the culture medium.

Directly adjacent to the genes of the mevalonate pathway, a gene designated *ppzP* was identified (Fig. 1B). Its predicted gene product showed 44% identity to Fnq26, a prenyltransferase of the ABBA family that is involved in the biosynthesis of the meroterpenoid furanonaphtoquinone I (5).

LC-MS analysis showed that both *S. anulatus* and *S. coelicolor*(ppzOS04) also produced a small amount of endophenazine B as well as phenazine 1-carboxylic acid methyl ester (data not shown). Both compounds were identified by LC-MS^2 in comparison with authentic reference samples.

As mentioned above, the screening of the cosmid library had identified a second cosmid, 11C7. Sequencing of the termini of its insert and comparison with the sequence obtained from 18A9 showed that 11C7 lacked the gene *ppzP* (the putative prenyltransferase) and 744 nucleotides of the mevalonate kinase gene but contained all other genes of the putative endophenazine cluster (see Fig. 1B). Cosmid 11C7 was expressed heterologously in *S. coelicolor* M512 using the same method described above. The resulting strain *S. coelicolor*(ppzOS02) produced phenazine 1-carboxylic acid and its methyl ester but no prenylated phenazines (Fig. 2C).

The results of the heterologous expression experiment with 18A9 prove that this cosmid contains all of the genes required for endophenazine biosynthesis. The experiment with 11C7 indicates that *ppzP* may be responsible for the prenylation reaction in this pathway. However, 11C7 lacked not only *ppzP* but also the promoter of the putative operon of mevalonate biosynthesis genes (see Fig. 1B) as well as part of the mevalonate kinase gene. Therefore, the heterologous expression of 18A9 and 11C7 did not provide conclusive evidence for the function of *ppzP*.

Inactivation of the Prenyltransferase Gene *ppzP* and Heterologous Expression of the Resulting Construct in *S. coelicolor* M512

λ RED-mediated recombination was used to create an in-frame deletion of the gene *ppzP* within the integrative cosmid ppzOS04. For this purpose, the cassette pIJ773, containing the apramycin-resistant gene *acc(3)IV* and flanked with *Xba*I and *Spe*I restriction sites, was used to replace the coding sequence of the *ppzP* gene. Subsequently, the cassette was excised by *Xba*I and *Spe*I digestion and religation of the compatible overhangs, leaving only 6 bp as a “scar” sequence between the start and the stop codon of *ppzP*. The resulting integrative cosmid ppzOS09 was introduced into the genome of *S. coelicolor* M512 as described above. HPLC and LC-MS analysis of the cultures (Fig. 2D) showed phenazine 1-carboxylic acid and its methyl ester but no endophenazine A or

Originally published in The *Journal of Biological Chemistry*. (2009), ***284***, 21, pp. 14439–47

B. This strongly supported the hypothesis that *ppzP* is responsible for the prenylation reaction in endophenazine biosynthesis in the heterologous expression experiment. However, a polar effect of the deletion on the expression of the genes downstream of *ppzP* cannot be excluded with certainty.

Overexpression and Purification of PpzP

For the biochemical investigation of the putative prenyltransferase, the gene *ppzP* was cloned into a plasmid for expression as an N-terminally His-tagged protein (see "Experimental Procedures"). The resulting construct was introduced into *E. coli* BL21 (DE3)pLysS. Induction with isopropyl 1-thio-β-d-galactopyranoside resulted in expression of a protein of ~37 kDa, as determined by SDS-PAGE (Fig. S1), coinciding with the calculated molecular mass of 37.138 kDa. Ni^{2+} affinity chromatography resulted in a protein of ~90% purity. 30 mg of purified soluble PpzP were obtained from 2 liters of culture. This protein was used for the biochemical investigations described below.

Identification of PpzP as 5,10-Dihydrophenazine 1-Carboxylate 9-Dimethylallyltransferase

Phenazine 1-carboxylic acid has been suggested as the product of the core pathway of phenazine biosynthesis from which the differently substituted or dimerized phenazines are derived by various tailoring reactions (8). In different *Pseudomonas* strains, PCA is converted by the monooxygenase PhzO to 2-hydroxy-PCA, by the amide synthethase PhzH to phenazine 1-carboxamide, and by the consecutive action of PhzM (methyltransferase) and PhzS (oxidoreductase) to pyocyanine (1-hydroxy-5-N-methyl-phenazine).

However, when we incubated PCA with PpzP and DMAPP under various conditions, we did not observe product formation in HPLC analysis. From chemical reasoning, the reduced form of PCA (*i.e.* dihydro-PCA) (Fig. 3A), appeared to be an attractive candidate for the prenylation substrate, since it is much more reactive than PCA for an electrophilic attack at C-9. Dihydro-PCA is presumed to be the direct biosynthetic precursor of PCA (8). We therefore generated dihydro-PCA by reduction of PCA with sodium dithionite (Fig. 3A). When dihydro-PCA was incubated with PpzP and DMAPP, the time-dependent formation of an enzymatic product was

readily observed by HPLC (Fig. 3B). Due to the very rapid oxidation of dihydro-PCA and its derivatives, they could not be quantified in the reduced form. After incubation, the reaction mixture was therefore oxidized with sodium persulfate ($Na_2S_2O_8$), and substrate and product were analyzed in the oxidized form. The enzymatic product was thereafter identified as endophenazine A by LC-MS analysis showing the same retention time, UV spectrum, molecular ion (*m/z* = 293 in positive ionization), and fragmentation pattern in ESI-MS^2 and ESI-MS^3 as an authentic reference sample of endophenazine A.

Biochemical Properties of PpzP

In the enzymatic assay described under "Experimental Procedures," the formation of endophenazine A showed a linear dependence on the amount of PpzP (up to 1 μg) and on the reaction time (up to 45 min). The reaction was strictly dependent on the presence of active PpzP, DMAPP, and dihydro-PCA. Maximal product formation was observed at pH 7.5, with half-maximal values at pH 9.5 and pH 5.0. In sharp contrast to the trans-prenyldiphosphate synthases like FPP synthase (30) and to the aromatic prenyltransferase NphB of naphterpin biosynthesis (3), the catalytic activity of PpzP was independent of the presence of Mg^{2+} or any other divalent metal ions. The addition of EDTA (10 mM) even increased reaction velocity 1.5-fold. Similarly, the addition of 500 mM NaCl increased product formation 3-fold, whereas the addition of 50 mM $MgCl_2$ and 100 mM $CaCl_2$ increased product formation ~1.5-fold. Therefore, 500 mM NaCl was routinely included in all assays (see "Experimental Procedures"). Although 10 mM Zn^{2+} had no effect on the reaction, the addition of 10 mM $FeSO_4$ completely abolished the formation of endophenazine A.

PpzP was found to be specific for both DMAPP and dihydro-PCA. When geranyldiphosphate was used instead of DMAPP, no product formation was observed. Likewise, no prenylated products could be observed when PpzP was incubated with other phenazine substrates, such as phenazine, phenazine 1-carboxylic acid methyl ester, N-methyl-phenazine (as methyl sulfate salt), or pyocyanine (1-hydroxy-5-N-methyl-phenazine). These phenazine substrates were reduced with dithionite to their dihydro analogs in the same way as described for PCA. Since N-methyl-phenazine and pyocyanine are compounds with quaternary nitrogens, these reaction mixtures were analyzed

Originally published in The *Journal of Biological Chemistry*. (2009), ***284***, 21, pp. 14439–47

directly without prior extraction with ethyl acetate.

We also tested the aromatic substrates of previously examined ABBA prenyltransferases (*i.e.* 4-hydroxyphenylpyruvate, flaviolin (2,5,7-trihydroxynaphthoquinone), 1,3-dihydroxynaphthalene, and 1,6-dihydroxynaphthalene). Of these, only flaviolin was prenylated by PpzP in the presence of DMAPP. LC-MS confirmed that the product was a monoprenylated flaviolin derivative. However, the reaction velocity was only 0.5 % of that obtained with dihydro-PCA. A reaction mechanism of the C-prenylation of flaviolin under catalysis of Fnq26 was suggested by Haagen et al. (5). A similar mechanism may be expected for the prenylation of flaviolin by PpzP. However, the amount of the prenylated product was too low for a structural identification by NMR spectroscopy.

Using a constant concentration of dihydro-PCA (0.8 mM) and varying concentrations of DMAPP, a typical hyperbolic curve of product formation over substrate concentration was obtained (Fig. 3C), indicating that the reaction followed Michaelis-Menten kinetics. Nonlinear regression analysis resulted in a K_m value of 116 ± 9 μM and a K_{cat} of 0.435 ± 0.006 s^{-1}. Using different concentrations of dihydro-PCA in the presence of 0.4 mM DMAPP, a sigmoidal dependence of product formation on substrate concentration was observed (Fig. 3D). The half-maximal reaction velocity was obtained at ~35 μM dihydro-PCA.

Using this value (35 μM) as an estimate for the Km value for dihydro-PCA, the catalytic efficiency (K_{cat}/K_m) of PpzP was calculated as 12,400 M^{-1} s^{-1}. This is significantly higher than the value of 7.7 M^{-1} s^{-1} reported for the conversion of the (artificial) substrate 1,3-dihydroxynaphthalene by NphB (6) but is comparable with the value of 5280 M-1 s-1 calculated for the prenylation of the (genuine) substrate 4-HPP by CloQ (4). We therefore suggested that dihydro-PCA is the genuine substrate of PpzP.

Investigation of Fnq28 from *S. cinnamonensis* DSM 1042 for Dihydro-PCA Dimethylallyltransferase Activity

Our group recently reported a gene cluster of prenylated naphthoquinone and prenylated phenazine biosynthesis from *S. cinnamonensis* DSM 1042 (10). This strain produces, besides the prenylated naphthoquinone derivative furanonaphthoquinone I, the same prenylated phenazines as *S. anulatus* (10,12,20). The gene cluster from

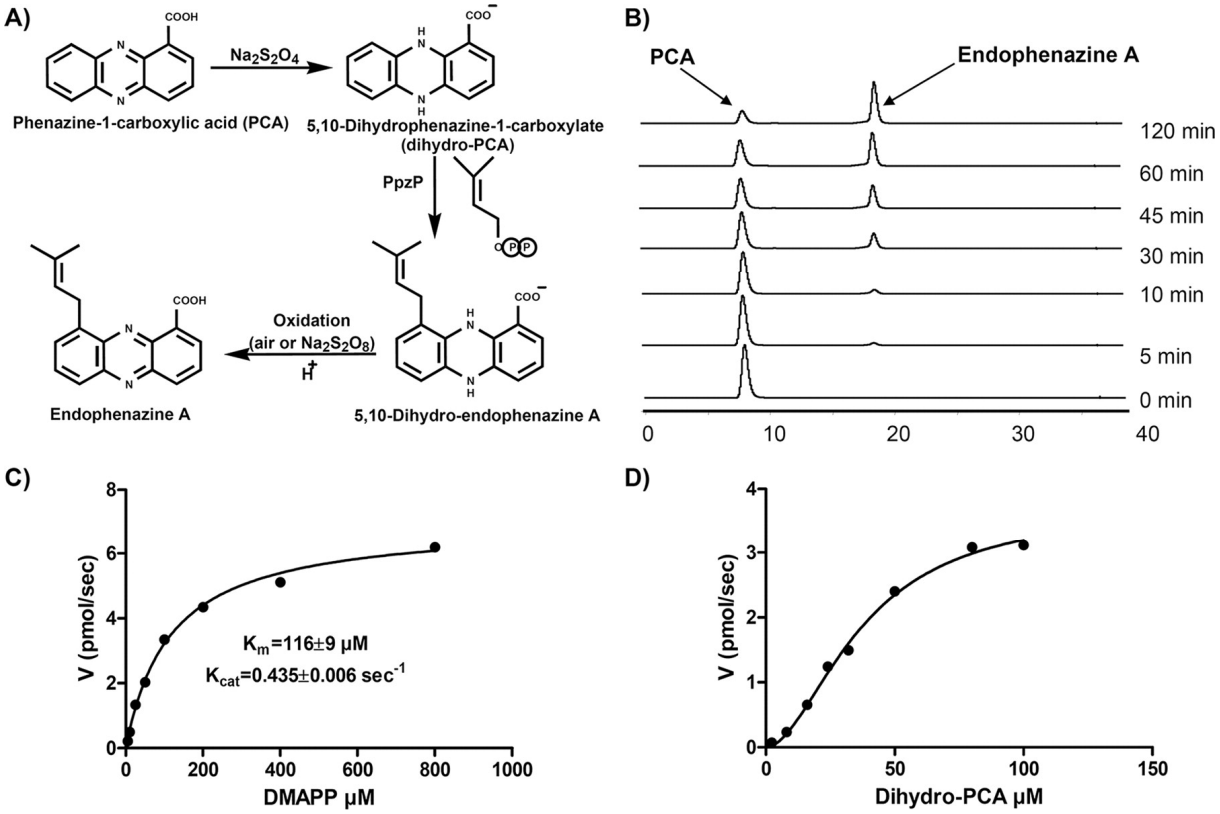


Figure 3: A, reaction scheme of the assay for dihydrophenazine 1-carboxylate dimethylallyltransferase activity. B, HPLC analysis of the dihydrophenazine 1-carboxylate prenyltransferase assay. C and D, product formation at different concentrations of DMAPP and dihydrophenazine 1-carboxylic acid. In C, dihydrophenazine 1-carboxylic acid was kept constant at 0.8 mM.In D, DMAPP was kept constant at 0.4 mM.

S. cinnamonensis contains genes assigned to furanonaphthoquinone I biosynthesis as well as a contiguous group of six genes (*i.e. ephzBCDEGA*) assigned to phenazine biosynthesis. By gene inactivation (10) and biochemical investigation (5), the prenyltransferase Fnq26 had been identified as the aromatic prenyltransferase of furanonaphthoquinone I biosynthesis. A second gene with similarity to ABBA prenyltransferases, designated as fnq28, was found immediately adjacent to the genes of the phenazine biosynthesis. Fnq28 shows 36% sequence identity with PpzP on the amino acid level. We tested whether Fnq28 also prenylates dihydro-PCA. For this purpose, Fnq28 was expressed as an N-terminally His-tagged protein, using the same method as employed for PpzP. Ni^{2+} affinity chromatography readily yielded a soluble protein of apparent homogeneity (data not shown). However, no prenylation of dihydro-PCA was observed with this protein, in clear contrast to the results with PpzP.

Fnq26 has been shown previously to transfer a geranyl moiety to flaviolin (2,5,7-trihydroxynaphthoquinone) or to

the artificial substrate 1,3-dihydroxynaphthalene. We now tested whether Fnq26 could also use dihydro-PCA and either DMAPP or geranyldiphosphate as substrates. However, again no prenylation products were detected. These observations are consistent with the results from inactivation experiments, which had shown that endophenazine A is still formed in *S. cinnamonensis* after inactivation of both fnq26 and fnq28 (10). Therefore, neither Fnq26 nor Fnq28 catalyzes the prenylation reaction of endophenazine biosynthesis in *S. cinnamonensis*. In this organism, the enzyme that functionally corresponds to PpzP of *S. anulatus* is yet to be identified.

Discussion

In this study, we have identified, for the first time, a prenyltransferase involved in the biosynthesis of prenylated phenazines. Sequence similarities and biochemical properties suggest that PpzP belongs to the recently discovered family of ABBA prenyltransferases (1). The present functional characterization of PpzP as 5,10-dihydrophenazine 1-carboxylate 9-dimethylallyltransferase now extends the substrate range of this family, previously limited to phenolic compounds, to include also phenazine derivatives.

At present, 13 genes with sequence similarity to ABBA prenyltransferases can be identified in the data base. A phylogenetic analysis of these genes (1) separates them into two clades. One of them comprises the 4-hydroxyphenylpyruvate 3-dimethylallyltransferases CloQ and NovQ from *Streptomyces* strains as well as four genes of unknown function from fungal genomes. The other clade comprises genes involved in the biosynthesis of prenylated naphthoquinones in different streptomycetes. A ClustalX analysis (data not shown) places PpzP into this second clade. Its closest ortholog is Fnq26, the prenyltransferase of furanonaphthoquinone I biosynthesis (5).

The x-ray structural analysis of the ABBA prenyltransferase NphB (formerly designated as Orf2) had revealed a novel protein fold, characterized by a β/α-barrel with 10 antiparallel β-strands. A structural model of PpzP (Fig. S2), generated using NphB as template, suggests a close similarity of the three-dimensional structure of the two proteins. The active center of NphB is localized within the spacious central cavity of the barrel, which has been

suggested to explain the promiscuity of this enzyme for different aromatic substrates (3).

NphB requires the presence of Mg^{2+} for catalytic activity. X-ray crystallography has shown that the Mg^{2+} ion is coordinated by a single aspartate residue, by four water molecules, and by one of the oxygen atoms of the α-phosphate group of the isoprenoid diphosphate (3). In contrast, PpzP does not require Mg^{2+} ions for its activity, and the same has been reported for the ABBA prenyltransferases CloQ (4), SCO7190 (3) and Fnq26 (5). Modeling studies suggested that positively charged residues (Lys54 in CloQ, Arg51 in SCO7190, and Arg50 in Fnq26) may functionally substitute for Mg^{2+} in the binding of the α-phosphate group of the isoprenyl diphosphate (1,3). An alignment of PpzP with the enzymes named above (Fig. S3) shows that also PpzP contains an arginine residue (Arg49) in the respective position of the first β strand. This may explain the Mg^{2+} independence of the PpzP reaction. Like the other ABBA prenyltransferases, PpzP does not contain the DDXXD motif typical for trans-prenyldiphosphate synthases (1).

In the biosynthesis of phenazines in different microorganisms, a common pathway leads to PCA, and this compound was believed to represent the branching point from which the pathways to differently substituted phenazines diverge (8). Our study now suggests that, at least for the prenylation of phenazines, the branching point is dihydro-PCA rather than PCA. This compound has been suggested (8) to be the product of the last enzyme-catalyzed step of the biosynthesis of the phenazine core, whereas its oxidation to PCA may occur non-enzymatically. Therefore, dihydro-PCA is likely to be available as substrate for PpzP in vivo.

Our study reports the first complete gene cluster of phenazine biosynthesis from streptomycetes, after a previously identified cluster turned out to be incomplete (10). Natural phenazines have important biological activities, including antibacterial, antitumor, antioxidant, and testosterone 5α-reductase-inhibiting activity (31). The present discovery of the endophenazine gene cluster from *S. anulatus* and the functional identification of PpzP as dihydro-PCA prenyltransferase may pave the way to the generation of new prenylated phenazines with improved biological activities by metabolic engineering and chemo-enzymatic synthesis.

Originally published in The *Journal of Biological Chemistry*. (2009), ***284***, 21, pp. 14439–47

Acknowledgments

We thank Kerstin Seeger for samples of Fnq26 and Fnq28, Ute Metzger and Elisa Haug-Schifferdecker for the synthesis of DMAPP, Andreas Kulik (Faculty of Biology, Tübingen University, Germany) for the LC-MS analysis, Christian Peifer (Medicinal Chemistry, Tübingen University, Germany) for the chemical synthesis of phenazine 1-carboxylic acid methyl ester, Joseph P. Noel (Salk Institute, La Jolla, CA) for plasmid pHis8, and Wulf Blankenfeldt (Max Planck Institute of Molecular Physiology, Düsseldorf, Germany) for helpful discussions.

Footnotes

The abbreviations used are: PCA, phenazine 1-carboxylic acid; LC, liquid chromatography; MS, mass spectrometry; ESI, electrospray ionization; DMAPP, dimethylallyl-diphosphate; TAPS, 3-{[2-hydroxy-1,1-bis(hydroxylmethyl)ethyl] amino}-1-propanesulfonic acid; HPLC, high pressure liquid chromatography; MS2, second level tandem mass spectrometry; MS3, third level tandem mass spectrometry.

References

1. Tello, M., Kuzuyama, T., Heide, L., Noel, J. P., and Richard, S. B. (2008) *Cell. Mol. Life Sci.* **65**, 1459-1463
2. Hopwood, D. A. (2007) *Streptomyces in Nature and Medicine: The Antibiotic Makers*, Oxford University Press, New York, USA
3. Kuzuyama, T., Noel, J. P., and Richard, S. B. (2005) *Nature* **435**, 983-987
4. Pojer, F., Wemakor, E., Kammerer, B., Chen, H., Walsh, C. T., Li, S. M., and Heide, L. (2003) *Proc. Natl. Acad. Sci. U.S.A.* **100**, 2316-2321
5. Haagen, Y., Unsöld, I., Westrich, L., Gust, B., Richard, S. B., Noel, J. P., and Heide, L. (2007) *FEBS Lett.* **581**, 2889-2893
6. Kumano, T., Richard, S. B., Noel, J. P., Nishiyama, M., and Kuzuyama, T. (2008) *Bioorg. Med. Chem.* **16**, 8117-8126
7. Gebhardt, K., Schimana, J., Krastel, P., Dettner, K., Rheinheimer, J., Zeeck, A., and Fiedler, H. P. (2002) *J. Antibiot. (Tokyo)* **55**, 794-800
8. Mavrodi, D. V., Thomashow, L. S., and Blankenfeldt, W. (2008) Biosynthesis and regulation of phenazine compounds in *Pseudomonas* spp. in *Pseudomonas* (Rehm, B. H. A. ed.), Wiley-VCH, Weinheim, Germany. pp 331-351
9. Ahuja, E. G., Janning, P., Mentel, M., Graebsch, A., Breinbauer, R., Hiller, W., Costisella, B., Thomashow, L. S., Mavrodi, D. V., and Blankenfeldt, W. (2008) *J. Am. Chem. Soc.* **130**, 17053-17061
10. Haagen, Y., Glück, K., Fay, K., Kammerer, B., Gust, B., and Heide, L. (2006) *Chembiochem* **7**, 2016-2027
11. Krastel, P., Zeeck, A., Gebhardt, K., Fiedler, H. P., and Rheinheimer, J. (2002) *J. Antibiot. (Tokyo)* **55**, 801-806
12. Sedmera, P., Pospíšil, S., and Novák, J. (1991) *J. Nat. Prod.* **54**, 870-872
13. Gust, B., Chandra, G., Jakimowicz, D., Yuqing, T., Bruton, C. J., and Chater, K. F. (2004) *Adv. Appl. Microbiol.* **54**, 107-128
14. Woodside, A. B., Huang, Z., and Poulter, C. D. (1993) *Organic Synthesis* **66**, 211-211

15. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000) *Practical Streptomyces genetics*, John Innes Foundation, Norwich, UK
16. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York
17. Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003) *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1541-1546
18. Datsenko, K. A., and Wanner, B. L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640-6645
19. Eustáquio, A. S., Gust, B., Galm, U., Li, S. M., Chater, K. F., and Heide, L. (2005) *Appl. Environ. Microbiol.* **71**, 2452-2459
20. Tax, J., Sedmera, P., Vokoun, J., Urban, J., Karnetová , J., Stajner, K., Vaněk, Z., and Krumphanzl, V. (1983) *Collect. Czech. Chem. Commun.* **48**, 527-532
21. Jez, J. M., Ferrer, J. L., Bowman, M. E., Dixon, R. A., and Noel, J. P. (2000) *Biochemistry* **39**, 890-902
22. Bringmann, G., Haagen, Y., Gulder, T. A., Gulder, T., and Heide, L. (2007) *J. Org. Chem.* **72**, 4198-4204
23. Kuzuyama, T., and Seto, H. (2003) *Nat. Prod. Rep.* **20**, 171-183
24. Mavrodi, D. V., Bonsall, R. F., Delaney, S. M., Soule, M. J., Phillips, G., and Thomashow, L. S. (2001) *J. Bacteriol.* **183**, 6454-6465
25. Kawasaki, T., Hayashi, Y., Kuzuyama, T., Furihata, K., Itoh, N., Seto, H., and Dairi, T. (2006) *J. Bacteriol.* **188**, 1236-1244
26. Takagi, M., Kuzuyama, T., Takahashi, S., and Seto, H. (2000) *J. Bacteriol.* **182**, 4153-4157
27. Isshiki, K., Tamamura, T., Sawa, T., Naganawa, H., Takeuchi, T., and Umezawa, H. (1986) *J. Antibiot. (Tokyo)* **39**, 1634-1635
28. Dairi, T. (2005) *J. Antibiot. (Tokyo)* **58**, 227-243
29. Sigmund, J. M., Clark, D. C., Rainey, F. A., and Anderson, A. S. (2003) *Microb. Ecol.* **46**, 106-112
30. Poulter, C. D. (2006) *Phytochem. Rev.* **5**, 17-26
31. Laursen, J. B., and Nielsen, J. (2004) *Chem. Rev.* **104**, 1663-1685

Supplemental Data

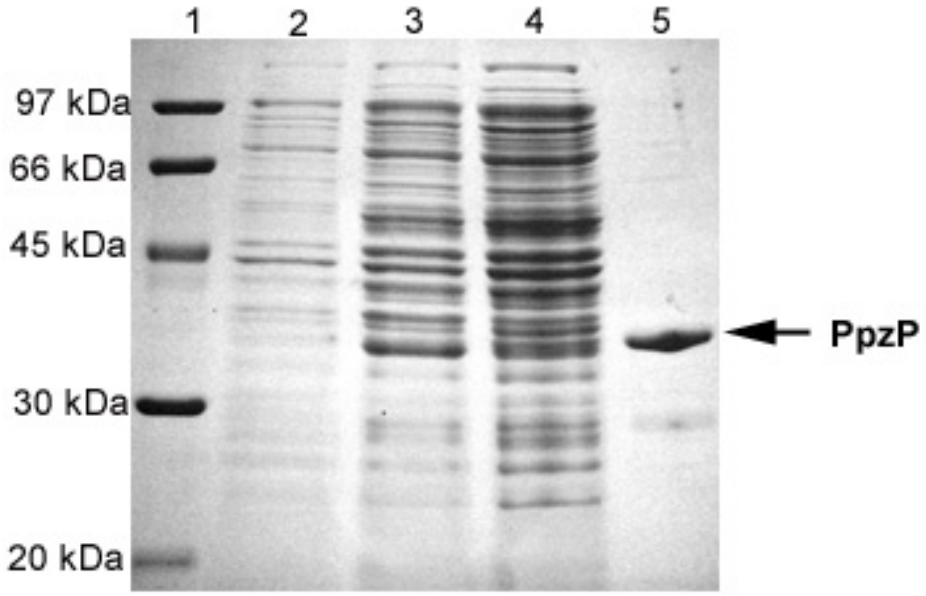


Fig. Fig. S1: Expression and purification of PpzP. Lane 1, molecular weight standards; lane 2, total protein before IPTG induction; lane 3, total protein after IPTG induction; lane 4, soluble protein after IPTG induction; lane 5, eluate from Ni-NTA-agarose. The 12% polyacrylamide gel was stained with Coomassie Brilliant Blue R-250.

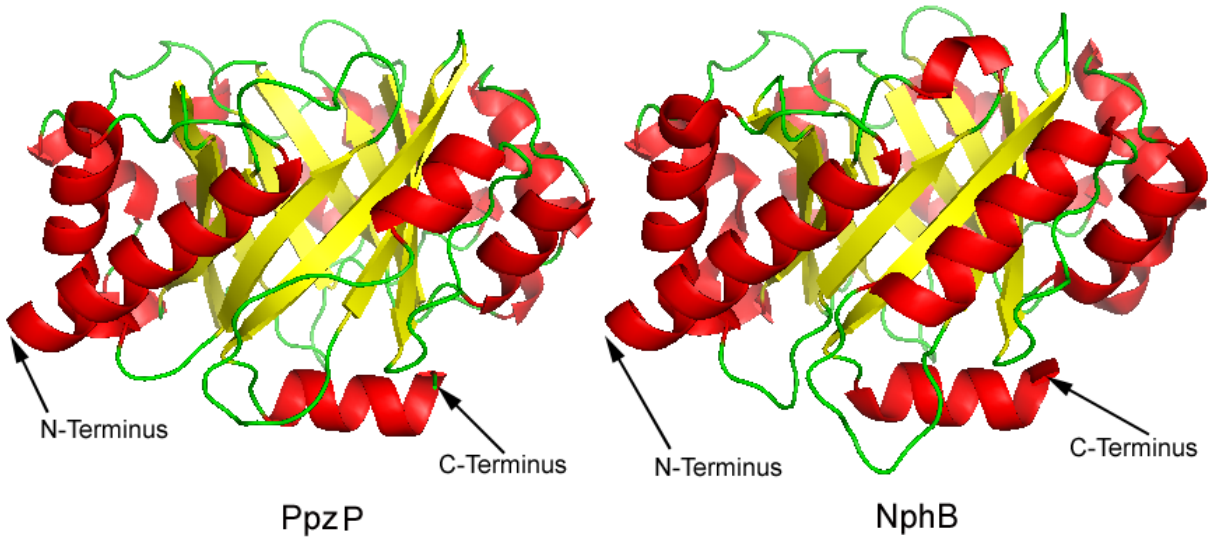


Fig. S2: Structural model of the dihydro-phenazine-1-carboxylate prenyltransferase PpzP, and experimentally determined structure of the prenyltransferase NphB. The unique fold of the ABBA prenyltransferases is characterized by a central barrel consisting of ten antiparallel β-sheets. The PpzP model was generated with the SWISS-MODEL program, using the PDB file 1zdyA as template.

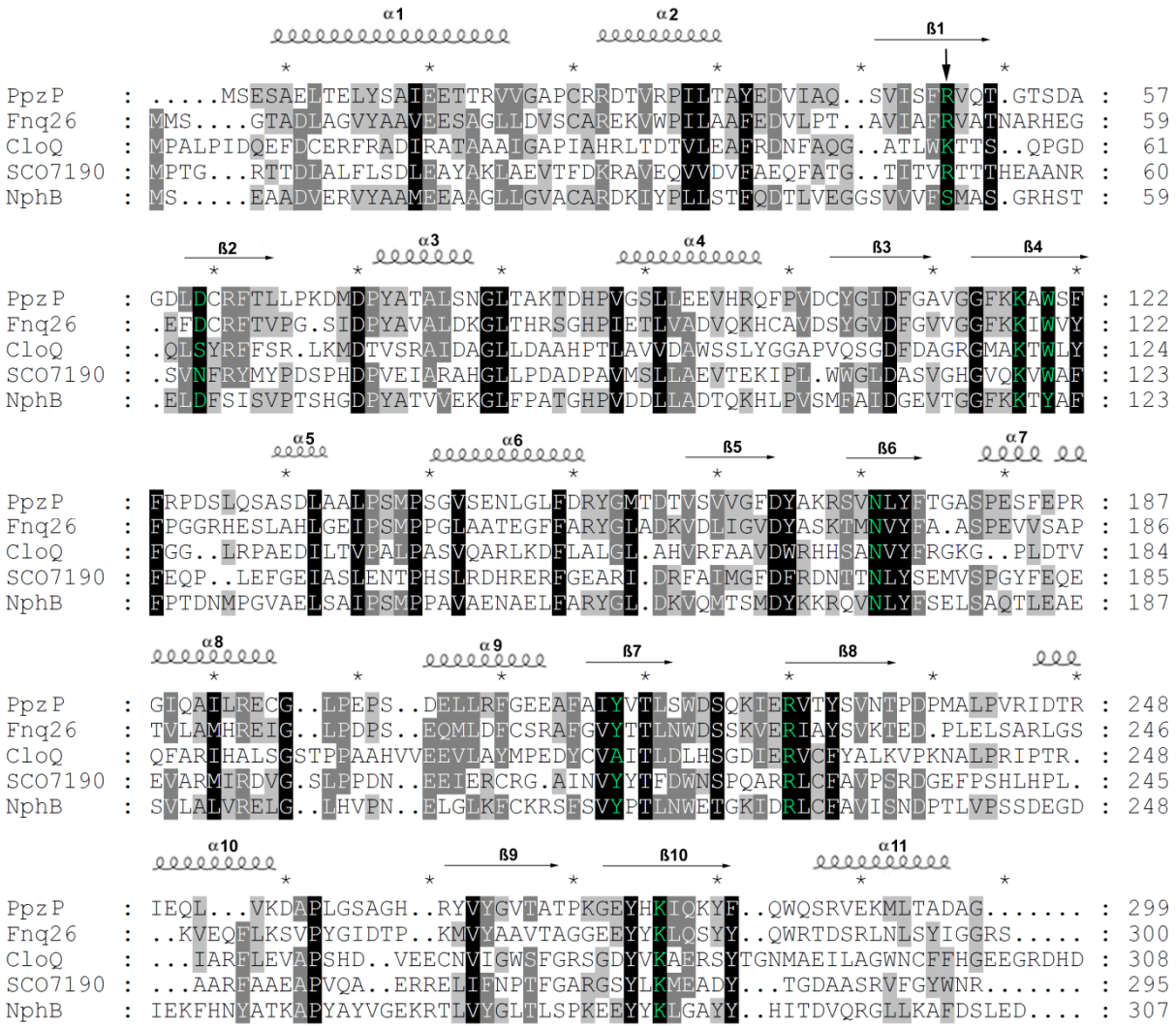


Fig. S3: Structure based multiple sequence alignment of the ABBA prenyltransferases PpzP, Fnq26, CloQ, SCO7190 and NphB. The structure of NphB and the structure models of CloQ and SCO7190 are described by Kuzuyama *et al*. 2005. The structure of PpzP and Fnq26 were generated with the SWISS-MODEL program using the PDP file 1zyA (NphB) as template. Arg49, suggested to coordinate the pyrophosphate group of the isoprenoid substrate, is indicated by an arrow.

Chapter 6 • *Adenylate-Forming Enzymes of Rubradirin Biosynthesis*

This manuscript has originally been published in ChemBioChem.

'Adenylate-Forming Enzymes of Rubradirin Biosynthesis: RubC1 Is a Bifunctional Enzyme with Aminocoumarin Acyl Ligase and Tyrosine-Activating Domains'

Björn Boll[a], Susanne Hennig[a], Chunsong Xie[b], Jae K. Sohng[c], Lutz Heide*[a]

[a] Pharmazeutisches Institut, Universität Tübingen, 72076 Tübingen (Germany),
[b] Center for Biomedicine and Health, Hangzhou Normal University, Hangzhou 310012, Zhejiang Province (China)
[c] Department for Pharmaceutical Engineering, Sun Moon University, 100, Kalsan-ri, Tangjeong-myeon, Asansi, Chungnam 336-708 (Republic of Korea)

* To whom correspondence should be addressed: E-mail: heide@uni-tuebingen.de
Phone: +49 (0)7071 29 72460
Fax: +49 (0)7071 29 5250

Keywords: adenylation; aminocoumarin; antibiotics; biosynthesis; rubradirin

ChemBioChem; Volume 12, Issue 7, pages 1105–1114, May 2, 2011

Received December 22, 2010; first published online March 29, 2011

DOI: 10.1002/cbic.201000778

Preface − About the Manuscript

Adenylate-forming enzymes are important in the biosynthesis of aminocoumarin antibiotics. They are involved in the formation of the aminocoumarin ring in form of NRPS and later in the generation of an amide bond between the aminocoumarin ring and an acyl moiety. Especially for the generation of new derivatives, the substrate specificity of these amide synthetases is interesting. The following publication describes the identification and characterization of the adenylate forming enzymes from the rubradirin biosynthetic gene cluster involved in amide bond formation. Interestingly, the cluster contains three putative amide synthetases from which two showed activity and the third could be converted by site directed mutagenesis. One of the active ones represents a fusion protein which also comprises a NRPS module. This is the first example of a biosynthetic pathway to an aminocoumarin-containing natural product in which such a fusion of these catalytic activities are found, possible capable of channeling between the active sites without release into the medium.

Author contributions:

- Björn Boll
  - decisively involved in initial hypothesis generation
  - decisively involved in planning, establishing and accomplishment of experimental setup
  - generation of data (cloning of genes and expression and purification of protein, HPLC analyses, LC-MS, sequence analysis) leading to figures 1-5
  - decisively involved in all data discussions
  - writing of the manuscript
  - preparation of all figures
- Susanne Hennig
  - optimization of amide synthetase assays
- Chunsong Xie
  - synthesis of 8-unmethylated aminocoumarin ring
- Jae K. Sohng
  - provided of cosmids with rubradirin biosynthetic gene cluster and original producer strain *S. achromogenes* var. *rubradiris* NRRL 3061
- Lutz Heide
  - supervised the project
  - decisively involved in initial hypothesis generation
  - decisively involved in all data discussions and analyses
  - manuscript preparation

My personal part for this manuscript involved performance of all initial experiments showing activity of RubC1 and Orf4. From these findings I and Prof. L. Heide designed further experiments to investigate the substrate specificity as well as to examine the kinetic constants. I planned and investigated the influence of mutations on the activity of RubF6. For this manuscript I wrote major text parts, generated figures and was in charge of the final version.

Abstract

The biosynthesis of aminocoumarin antibiotics requires two acyladenylate-forming enzymes: one for the activation of L-tyrosine as a precursor of the aminocoumarin moiety and another for the linkage of an acyl moiety to the aminocoumarin moiety. Unexpectedly, the biosynthetic gene cluster of the aminocoumarin antibiotic rubradirin was found to contain three genes for putative acyladenylate-forming enzymes of aminocoumarin biosynthesis and conjugation. We expressed, purified, and investigated these three proteins. Orf4 (55 kDa) was shown to be an active aminocoumarin acyl ligase. RubF6 (56 kDa) was inactive, but could be converted into an active enzyme by site-directed mutagenesis. RubC1 (138 kDa) was shown to be a unique bifunctional enzyme, comprising an aminocoumarin acyl ligase, and tyrosine-adenylation and peptidyl-carrier domains. This natural hybrid enzyme is unique among known proteins. A hypothesis is proposed as to how such an enzyme could offer a particularly effective machinery for aminocoumarin antibiotic biosynthesis.

Introduction

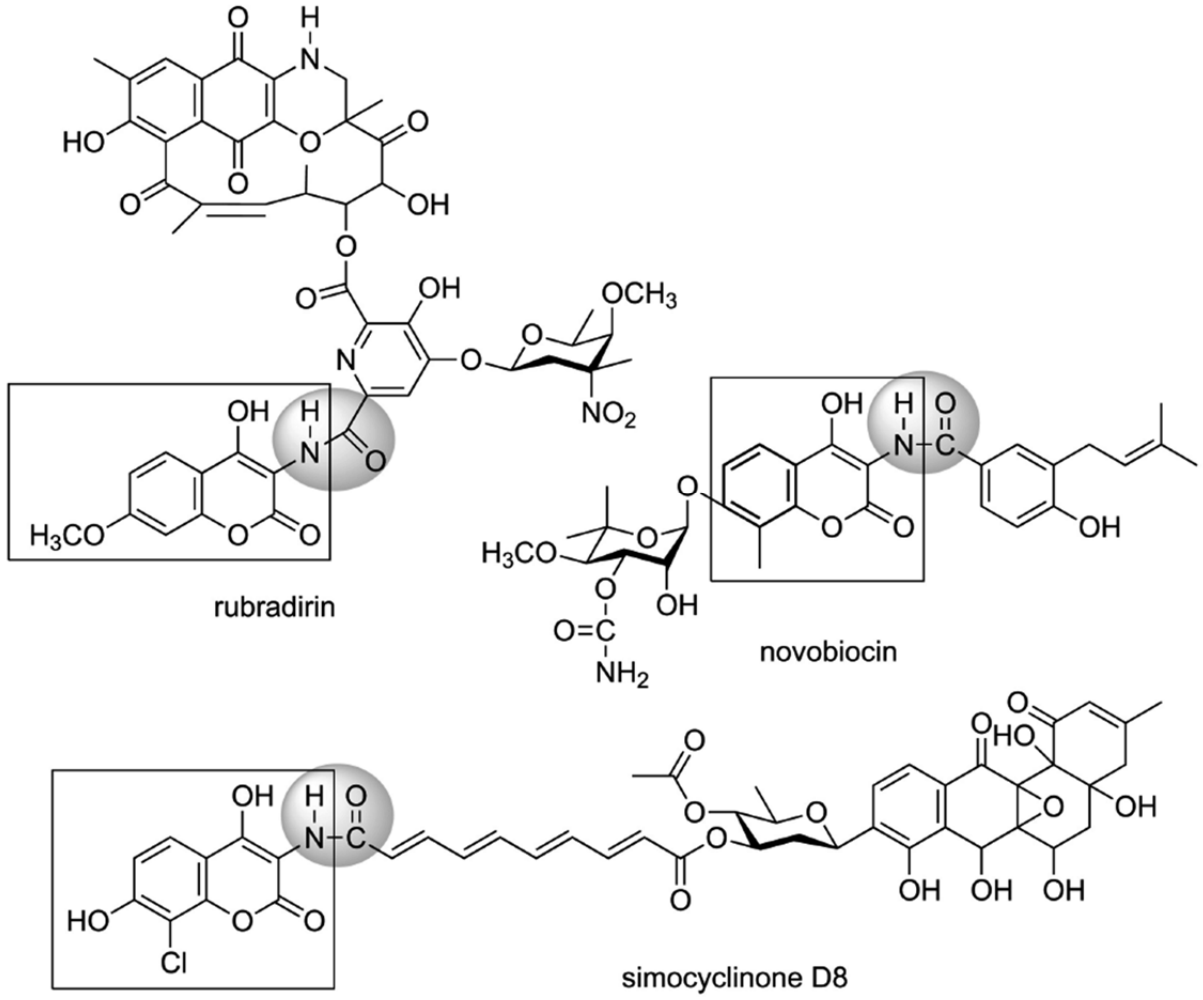
The activation of acyl moieties by acyladenylate formation plays a central role in the biosynthesis of many antibiotics, especially nonribosomal peptides and polyketides (1). In the biosynthesis of the aminocoumarin antibiotic novobiocin (Scheme 1), a potent inhibitor of bacterial gyrase (2), two acyladenylate-forming enzymes are involved: NovH (600 aa), which activates tyrosine as a precursor of the aminocoumarin moiety (3), and NovL (527 aa), which links the aminocoumarin moiety to an acyl moiety by an amide bond (4). The biosynthetic gene clusters of the aminocoumarin antibiotics clorobiocin (5), coumermycin A_1 (6) and simocyclinone (7,8) each contain an orthologue of *novH* and an orthologue of *novL*. The different substrate specificities of the aminocoumarin acyl ligases from these clusters - CloL, CouL, and SimL - have been utilized for the mutasynthetic and chemoenzymatic generation of new aminocoumarin antibiotics (9-12).

In 2008, the biosynthetic gene cluster of the aminocoumarin antibiotic rubradirin was published (13). Rubradirin is an ansamycin antibiotic isolated from *Streptomyces achromogenes* var. *rubradiris* NRRL3061. It exhibits activity against many Gram-positive bacteria, including *Staphylococcus aureus* strains with multiple antibiotic resistance (14). Rubradirin inhibits translation at the

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

bacterial ribosomes (15), and the rubradirin aglycone has been shown to inhibit bacterial RNA polymerase by a different mechanism from those of ansamycins such as rifamycins, streptovaricins, and tolypomycins (16). Further biological effects of the rubradirin aglycone have also been described (17,18).

Rubradirin is comprised of four distinct structural moieties (Scheme 1): the aminocoumarin moiety, the ansamacrolide moiety (termed rubransarol), the 3,4-dihydroxydipicolinate (DHDP) moiety, and the deoxysugar D-rubranitrose (2,3,6-trideoxy-3-C-4-O-dimethyl-3-C-nitro-D-xylo-hexose). The DHDP moiety is linked to the aminocoumarin moiety through an amide bond, the formation of which is expected to be catalyzed by an aminocoumarin acyl ligase.



Scheme 1. Structures of the aminocoumarin antibiotics rubradirin, novobiocin, and simocyclinone D8. The aminocoumarin moieties are highlighted by boxes, and the amide bonds linking them to the different acyl moieties are accentuated in gray.

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

Unexpectedly, the biosynthetic gene cluster of rubradirin was found to contain not one but three putative genes for aminocoumarin acyl ligases with sequence similarity to NovL (13), that is, *orf4*, *rubF6*, and *rubC1*. The predicted gene product of rubC1 showed a unique chimeric structure, with similarity to NovL at its N terminus and similarity to NovH at its C terminus. This suggested that it might catalyze two different steps in aminocoumarin antibiotic biosynthesis, that is, activation of tyrosine as well as aminocoumarin acyl ligation. Such a chimeric enzyme has not been previously described. In this study, we expressed and purified the three putative adenylate-forming enzymes of the rubradirin biosynthetic gene cluster, that is, Orf4, RubF6, and RubC1, and investigated them biochemically.

Results

Sequence analysis of Orf4, RubF6 and RubC1

Figure 1 shows an alignment of the predicted proteins Orf4 (521 aa), RubF6 (529 aa), and RubC1 (1317 aa) with NovL and NovH of novobiocin biosynthesis. Orf4, RubF6, and the N-terminal 514 aa of RubC1 show sequence similarity to NovL (527 aa). At the amino-acid level, their respective identities to NovL are 34, 32, and 35 %. BLAST searches revealed that these proteins possess similarity to acyl-CoA synthetases. Correspondingly, they show conserved motifs of adenylate-forming enzymes (Figure 1). Orf4 is extremely similar to RubF6 (88 % identity), but less so to the N-terminal part of RubC1 (49 % identity).

RubC1 is much larger than Orf4 and RubF6. While its N-terminal part shows similarity to NovL, its C-terminal 600 aa show 46 % identity to the tyrosine-activating enzyme NovH (600 aa; Figure 1). Using the NRPS-PKS software package (19) adenylation and peptidyl carrier protein (PCP) domains were identified in this portion of RubC1. The PCP domain contains a predicted attachment site for a 4′-phosphopantetheinyl cofactor, marked in Figure 1.

The specificity of adenylation domains for the activation of certain amino acids can often be predicted from the “specificity-conferring code”, which consists of approximately ten amino acid residues that line the substrate binding pocket of the respective enzyme (20). From this code, the NRPS-PKS software package (19) predicts that the adenylation domain in the C-terminal

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

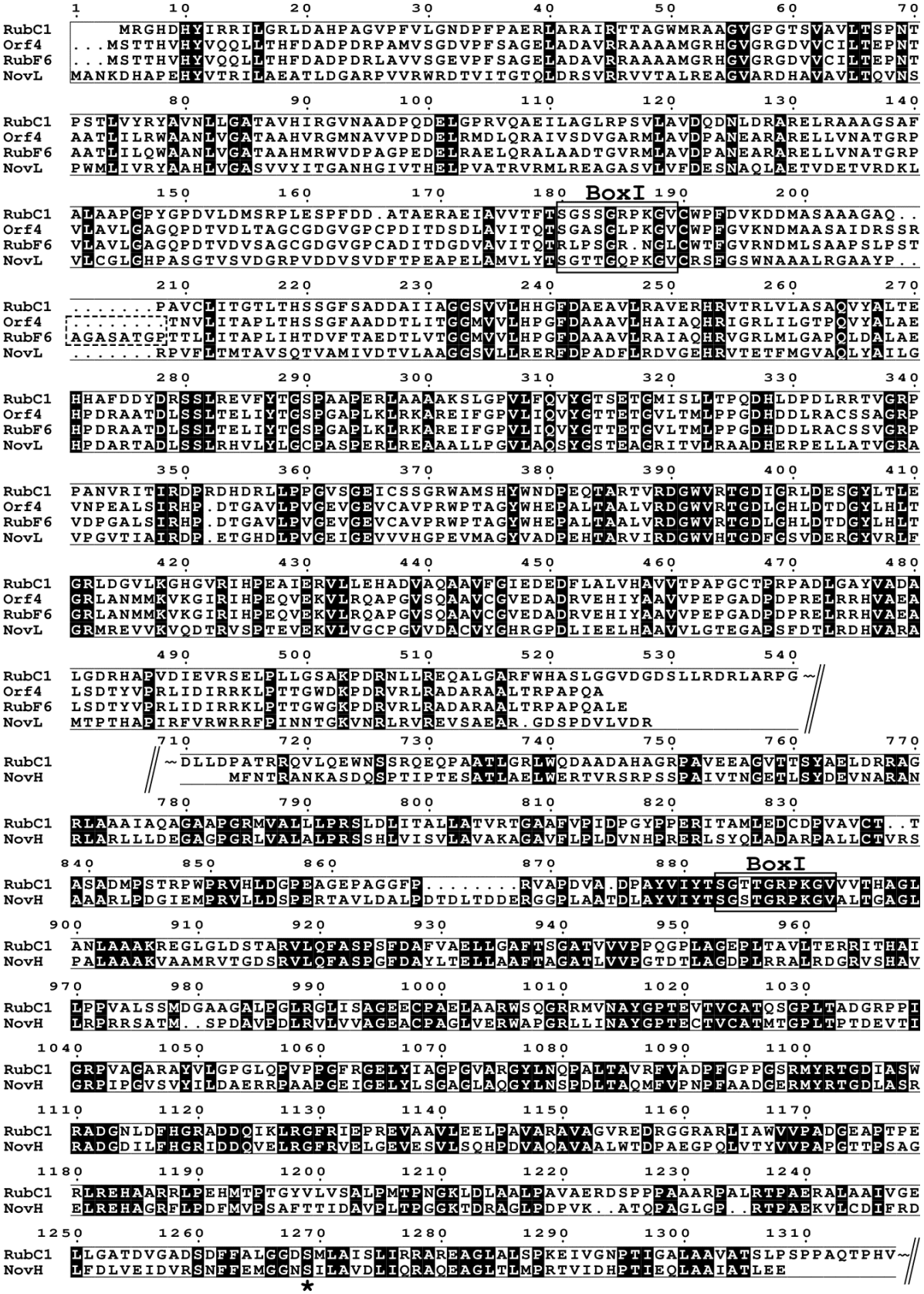


Figure 1. Alignment of RubC1, Orf4, and RubF6 with the aminocoumarin acyl ligase NovL and the L-tyrosine-activating enzyme NovH. Ser1269, representing the 4′-phosphopantetheinyl attachment site of the PCP domain, is marked with an asterisk. The BoxI motifs of adenylate-forming enzymes are highlighted with a frame. The (inactive) RubF6 shows poor conservation of this motif. The insertion of eight amino acids found in RubF6 is highlighted with a dotted frame.

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

part of RubC1 is specific for the activation of L-tyrosine; this is in accordance with the expected role of this part of RubC1 in the biosynthesis of the aminocoumarin moiety of rubradirin.

The central part of the predicted RubC1 protein, approximately from amino acid 515 to 710 (Figure 1), shows low sequence similarity to other NRPS proteins in BLAST searches, with E values >0.05. No conserved domains are detected by BLAST searches in this part of RubC1.

Expression and purification of Orf4, RubF6, and RubC1

Orf4, RubF6, and RubC1 were expressed in *E. coli* in the form of N-terminally His-tagged fusion proteins and purified by Ni^{2+}-affinity chromatography (Figure 2). RubC1 was coexpressed with the nonspecific 4′-phosphopantetheinyl transferase Sfp from *Bacillus subtilis* (21) to ensure phosphopantetheinylation of its PCP domain. The three proteins were obtained in yields of 26, 9, and 9 mg per liter of culture, respectively. The observed molecular weights in an SDS-PAGE analysis corresponded to the calculated values for the His-tagged proteins, that is, 55, 58, and 140 kDa, respectively (Figure 2).

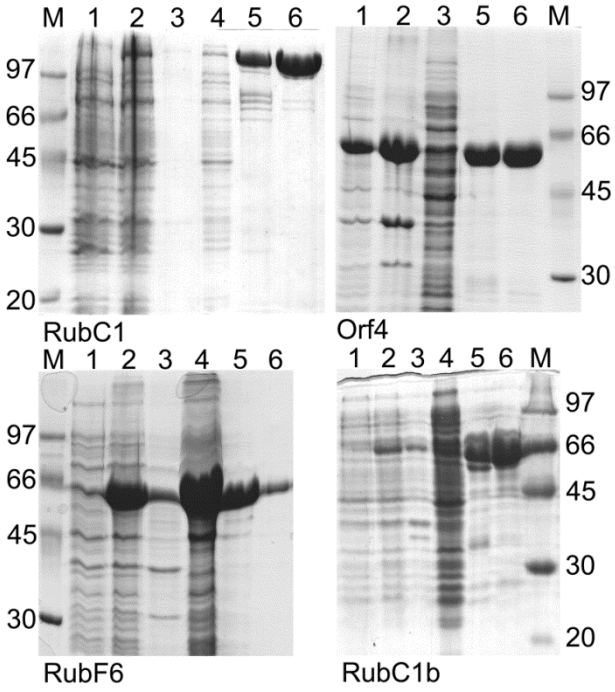


Figure 2. Expression and purification of Orf4, RubF6, RubC1, and the C-terminal domain of RubC1 (termed RubC1b). Lane 1: total protein before IPTG induction, lane 2: total protein after IPTG induction, lane 3: insoluble protein after IPTG induction, lane 4: soluble protein after IPTG induction, lane 5: protein after Ni^{2+}-affinity chromatography, lane 6: protein after size-exclusion gel chromatography, lane M: molecular mass standards. The calculated masses of the His-tagged proteins are 55.3 (Orf4), 57.9 (RubF6), 140.5 (RubC1), and 66.8 kDa (RubC1b). The 12 % polyacrylamide gel was stained with CBB R-250.

Orf4 is a functional aminocoumarin acyl ligase

Orf4 in the biosynthetic gene cluster of rubradirin shows clear sequence similarity to NovL (Figure 1). We first investigated the enzymatic activity of purified Orf4 with the same substrates as used previously in assays of NovL from novobiocin biosynthesis (4,22), namely 3-amino-4,7-dihydroxy-8-methyl-coumarin (hereafter referred to as

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

methyl-aminocoumarin) as the amino substrate and 3-dimethylallyl-4-hydroxybenzoic acid (3-DMA-4HBA) as the acyl substrate. In the presence of ATP and Mg^{2+}, the enzyme-dependent formation of novobiocic acid, that is, the amide formed from these two substrates was readily detected by HPLC. The identity of the product was confirmed by comparing its HPLC and mass spectra to those of an authentic reference compound.

3-DMA-4HBA is not an intermediate of rubradirin biosynthesis, but it has been found to be accepted as a non-natural substrate by other aminocoumarin acyl ligases, such as SimL of simocyclinone biosynthesis (23) and CouL of coumermycin A1 biosynthesis (24). The natural acyl substrate of the aminocoumarin acyl ligase reaction in rubradirin is unknown. If the formation of the amide bond precedes the formation of the glycoside and the ester bonds in the rubradirin molecule (Scheme 1), the expected substrate would be 3,4-dihydroxypyridine-2,6-dicarboxylic acid (3,4-dihydroxypipecolic acid). A commercially available analogue of this compound is 4-hydroxypyridine-2,6-dicarboxylic acid (chelidamic acid in its enol form). However, Orf4 did not catalyze any detectable product formation with chelidamic acid or with several other pyridine-2,6-dicarboxylic acid derivatives (Table S1 in the Supporting Information). We tested Orf4 with 24 different acyl substrates. Product formation was detected by HPLC and LC-MS with cinnamic acid, ferulic acid, and benzoic acid, as well as with 3-DMA-4HBA, as mentioned above. In addition, HPLC analysis showed product formation with coumaric acid and 4-hydroxybenzoic acid, but no unequivocal mass spectrometric data were obtained for the products formed from these compounds. The highest-yielding product formation was observed with cinnamic acid, as previously observed for the aminocoumarin acyl ligase SimL of simocyclinone biosynthesis (23). Therefore, cinnamic acid was used as the acyl substrate in further investigations of Orf4.

Optimal product formation was measured at pH 7.5 in Tris-HCl buffer, with half maximal activities at pH 8.5 and 6.8. The optimal ATP concentration was 5 mM. At constant methyl-aminocoumarin concentration (1.5 mM) with various concentrations of cinnamic acid, the K_m value was determined to be 2760 μM (Figure 3 A). This is considerably higher than the K_m values observed for the natural substrates of NovL, CouL and SimL: 19, 26 and 20.4

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

μM, respectively (4,23,24). However, this was expected as cinnamic acid is not the natural substrate of the ligase reaction in rubradirin biosynthesis. V_{max} was calculated as 92.3 pkat per mg protein, similar to the value of 90 pkat per mg protein measured for SimL with cinnamic acid (23). By keeping the cinnamic acid concentration constant at 6 mM and varying the concentration of the aminocoumarin substrate, the K_m value for methyl-aminocoumarin was determined to be 358 μM (Figure 3 A).

Methyl-aminocoumarin is a structural moiety of novobiocin (22); however, the aminocoumarin moiety of rubradirin is unsubstituted in position 8. We therefore synthesized the corresponding 8-unsubstituted aminocoumarin according to a published procedure (25). The K_m value of Orf4 for this substrate was found to be 73 μM, five times lower than for the non-natural methyl-aminocoumarin. Moreover, V_{max} was higher for the 8-unsubstituted aminocoumarin than for the methyl-aminocoumarin. In comparison, the K_m values of NovL, CouL, and SimL for methyl-aminocoumarin have been previously determined to be 131, 44, and 20.5 μM, respectively (4,23,24).

RubF6 is inactive but can be mutated to a functional aminocoumarin acyl ligase

RubF6 is currently annotated in GenBank as a “putative aminocoumarin ligase”. This 529 aa protein shares 88 % identity with Orf4 (Figure 1). However, we found the purified RubF6 to be completely inactive in assays for aminocoumarin acyl ligase activity, using all 24 acyl substrates listed in Table S1. This is in clear contrast to the results obtained with Orf4.

Closer inspection of the amino acid sequence showed two important differences between the inactive enzyme RubF6 and the active enzymes Orf4 and NovL. First, RubF6 has an insertion of eight amino acids in its sequence (AGASATGP, see Figure 1). Second, the sequence of RubF6 shows poor conservation of the BoxI motif, which is conserved in other adenylate-forming enzymes (Figure 1). Modeling of the structure of RubF6 with *Salmonella enterica* acetyl-CoA synthetase (PDB ID: 2P2M) (26) as a template showed that the mutations in the BoxI motif in RubF6 affect the loop that, in the *S. enterica* enzyme, is involved in ATP binding.

One of the sequence differences in this loop is a lysine residue that has been proven to be involved in the

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

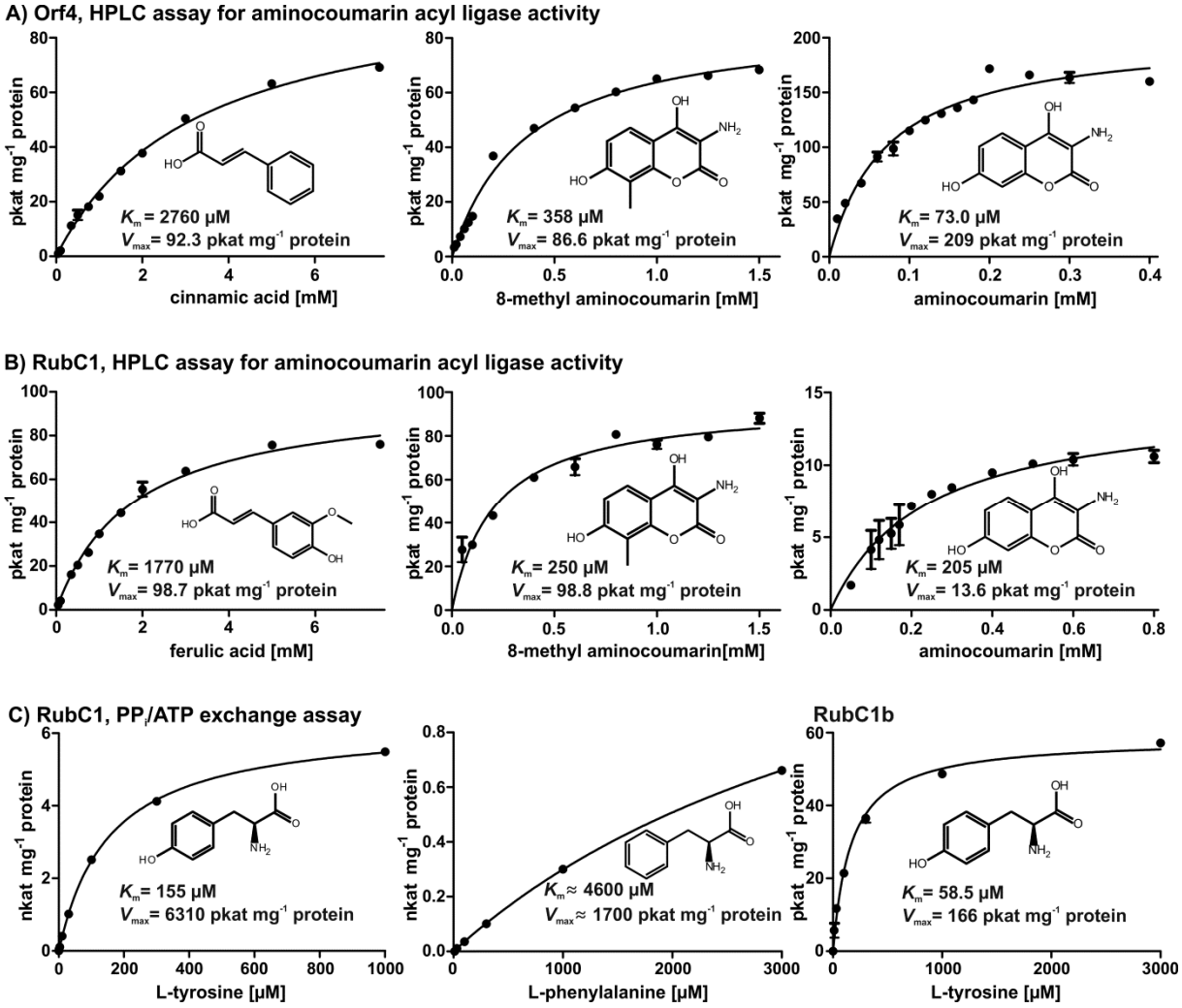


Figure 3. Determination of K_m and V_{max} values of A) Orf4 and B) RubC1 for acyl substrates and for two different aminocoumarin substrates in aminocoumarin acyl ligase assays. See text for further explanations. C) [^{32}P]PP$_i$-exchange assay of RubC1 and RubC1b with L-tyrosine and L-phenylalanine.

catalytic activity of adenylate-forming enzymes (27) but is missing from this loop in RubF6.

We investigated whether the activity of RubF6 could be restored by modification of the sequence using site-directed mutagenesis. Deletion of the eight-residue insertion in RubF6 did not result in detectable activity, neither did replacement of Asn184 in the BoxI motif with the amino acids Pro-Lys, found in the corresponding position of Orf4 and NovL. However, an active aminocoumarin acyl ligase was obtained when the mutation of Asn184 to Pro-Lys was combined with a mutation of amino acids 178–180 from Arg-Leu-Pro to Ser-Gly-Ser (the sequence found in SimL and RubC1). The resulting RubF6 mutant clearly showed activity in an assay with methyl-aminocoumarin and either cinnamic acid or 3-dimethylallyl-4-

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

hydroxybenzoic acid. However, the catalytic activity was only 4.6 pkat per mg protein with cinnamic acid as the best-accepted substrate. This is only 5 % of the activity observed for Orf4.

RubC1 is a bifunctional enzyme with aminocoumarin acyl ligase activity and tyrosine-activating activity

RubC1 is currently annotated in GenBank as a "putative acyl-coenzyme A ligase". Sequence analysis had shown that its C-terminal part is similar to the tyrosine-activating enzyme NovH, while its N-terminal part is similar to the aminocoumarin acyl ligase NovL (Figure 1). In the investigation described above, we showed that orf4 coded for a functional aminocoumarin acyl ligase. From the current hypothesis on rubradirin biosynthesis (13), only a single aminocoumarin acyl ligase is expected to be involved in this pathway. We therefore investigated whether the N-terminal part of RubC1 was a functional ligase or not.

When purified RubC1 was incubated with methyl-aminocoumarin and 3-DMA-4HBA, the formation of novobiocic acid, was readily detected by HPLC. The reaction was dependent on the presence of ATP, Mg^{2+}, and active enzyme. The identity of the product was confirmed by LC-MS by comparison with an authentic reference compound; the product was identical to those formed by NovL and Orf4 in the same assays.

Of the other 23 acyl substrates shown in Table S1, RubC1 catalyzed product formation only with cinnamic acid, 4-coumaric acid, and ferulic acid. The products showed the expected molecular masses in LC-MS analysis, and were identical to those formed by Orf4. The highest product yield was observed with ferulic acid, and this substrate was used in the further characterization of the enzyme.

Optimal product formation was observed with an ATP concentration of 6 mM. Mg^{2+} (1 mM) strongly stimulated the activity. In Tris-HCl buffer, the optimal pH was 8.5, compared to 8.0 for NovL and 7.5 for Orf4. Half-maximal reaction velocity was observed at pH values of 7.4 and 9.5. Addition of NaCl reduced the activity, with nearly complete inhibition at concentrations of 500 mM and higher.

The apparent K_m value for the non-natural substrate ferulic acid was determined as 1.7 mM (Figure 3 B), similar to the value of 2.7 mM observed for Orf4 with its preferred acyl substrate, cinnamic acid. Concentrations of ferulic acid exceeding 7.5 mM resulted in inhibition, and therefore only the data from concentrations up to 7.5 mM were

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

included in the nonlinear regression analysis for K_m determination. The apparent K_m value for methyl-aminocoumarin was 358 µM, similar to the value of 250 µM observed for Orf4. V_{max} was calculated as 98.7 pkat per mg protein. 92.3 pkat per mg protein had been calculated for Orf4, but since the molecular weight of RubC1 is 2.5 times higher than that of Orf4, the turnover number of RubC1 is higher than that of Orf4. Similarly, as observed for Orf4, the K_m value of RubC1 for the 8-unsubstituted aminocoumarin (205 µM) was lower than that of the 8-methylated compound (Figure 3 B). However, in the case of RubC1, this difference was less pronounced than for Orf4, and V_{max} was lower than observed for methyl-aminocoumarin. L-Tyrosine and L-phenylalanine were tested as alternative amino substrates of RubC1, but no amide formation was observed in the HPLC analysis.

Sequence analysis of RubC1 had shown an adenylation and a PCP domain at the C terminus, similar to NovH, which activates tyrosine as a precursor of the aminocoumarin moiety of novobiocin (3). To confirm the enzymatic activity of this C-terminal part of RubC1, we tested its ability to catalyze the adenylation of tyrosine using the ATP-[^{32}P]PP$_i$ exchange assay. This is a well-established assay commonly used for the investigation of aminoacyl-adenylating enzymes (28,29). It measures the incorporation of radioactive pyrophosphate into ATP. We used the same reaction conditions as employed previously for the investigation of NovH (3).

Upon incubation of RubC1 with L-tyrosine, ATP, Mg^{2+}, and [^{32}P]PP$_i$, radioactive ATP was formed. The reaction depended on the presence of active enzyme and each of the named assay components. As expected, RubC1 activity was specific for the amino acid: activity with L-phenylalanine was six times lower than with L-tyrosine, and no activity was observed with L-tryptophan (Figure 4 A).

The K_m and V_{max} values of RubC1 for L-tyrosine were determined as 155 µM and 6.31 nkat per mg protein. For NovH, the corresponding values were reported as 1390 µM and 4.0 nkat per mg protein, measured under identical conditions. For L-phenylalanine, RubC1 showed K_m and V_{max} values of

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

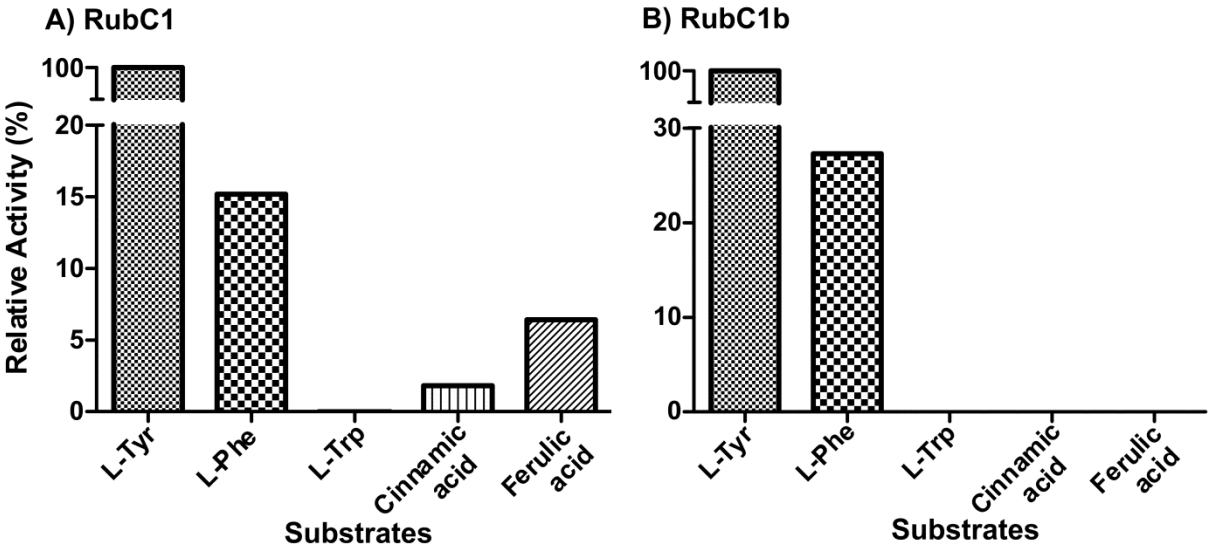


Figure 4. Relative adenylation activity of RubC1 and RubC1b towards different acids. Activity was assessed by ATP-[^{32}P]PP$_i$-exchange assays.

approximately 4600 μM and 1.7 nkat per mg protein, thus confirming that the affinity of this enzyme for tyrosine was much higher than for phenylalanine.

Auto-aminoacylation of RubC1 with [^{14}C]Tyr

The ability of holo-RubC1 to ligate the activated L-Tyr onto the thiol of the 4′-phosphopantetheinyl arm on the PCP domain was assayed by using L-[U-^{14}C]tyrosine as a substrate. As depicted in Figure 5, SDS-PAGE and autoradiographic analysis clearly showed ATP-dependent covalent loading of L-Tyr onto RubC1. This indicated that the PCP domain in RubC1, at the C-terminal side of the adenylation domain, is functional.

Tyrosine-adenylating activity, but not ferulic acid-adenylating activity is localized in the C-terminal part of RubC1

We subsequently tried to separately express the N-terminal part of RubC1, which is similar to the aminocoumarin acyl ligase NovL, and the C-terminal part of RubC1, which is similar to the L-tyrosine-activating enzyme NovH, as N-terminally His$_8$-tagged proteins in *E. coli*. The C-terminal part, comprising amino acids 689–1317 of RubC1 and hereafter termed RubC1b, was readily obtained as soluble protein in good yield (41 mg per L culture). In contrast, expression of the N-terminal part, comprising amino acids 1–521, resulted only in insoluble protein, despite repeated attempts under different culture conditions. We also tried

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

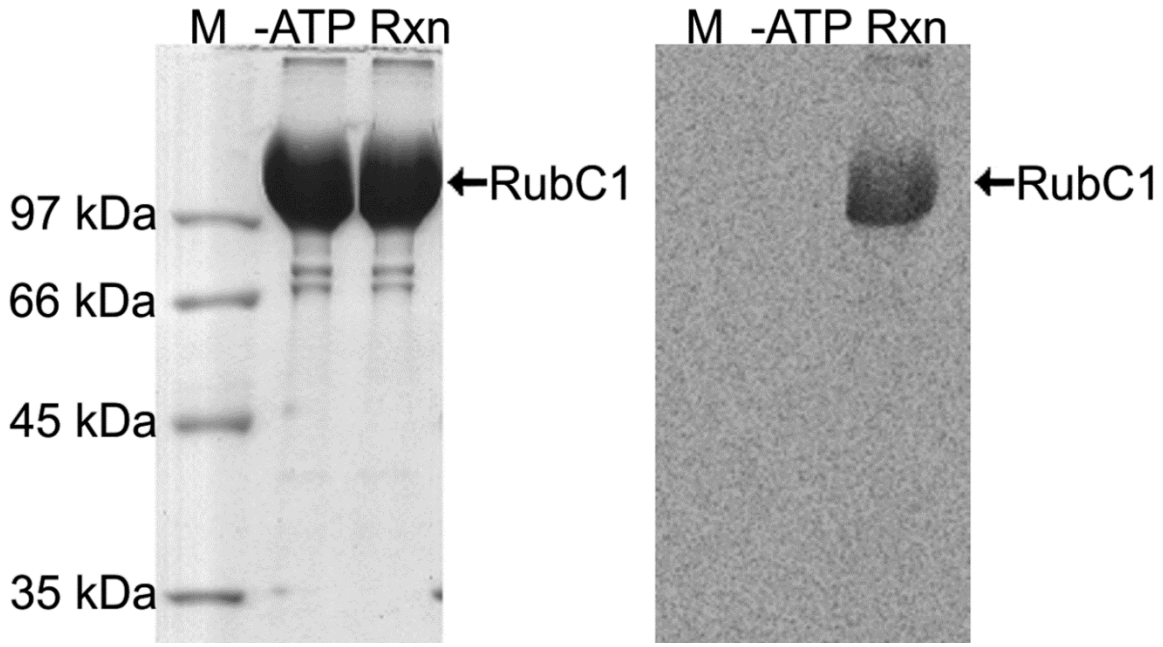


Figure 5. Demonstration of the ATP-dependent auto-aminoacylation of RubC1 with L-[U-^{14}C]tyrosine. A) CBB-stained 10 % SDS-PAGE gel of two reaction mixtures. Lane M: molecular weight markers, lane −ATP: control reaction without ATP; lane Rxn: complete reaction. B) Autoradiogram of the same gel.

to express a slightly larger N-terminal portion of RubC1, comprising aa 1-570, but again obtained only insoluble protein. Further biochemical experiments were therefore carried out only with the C-terminal fragment, RubC1b.

The [^{32}P]PP$_i$-exchange assay readily confirmed the L-tyrosine-activating activity of RubC1b, with a similar substrate preference for L-tyrosine in comparison to L-phenylalanine and L-tryptophan as observed for RubC1 holoenzyme (Figure 4 B). The K_m value for L-tyrosine was determined as 59 µM (Figure 3 C). However, the activity of RubC1b was lower than that of the holoenzyme, by approximately two orders of magnitude (Figure 3). From these results we conclude that, in the absence of the N terminus, the structure of the protein is not optimal.

Ferulic acid is adenylated by the RubC1 holoenzyme with approximately 7 % of the reaction velocity observed for tyrosine. In contrast, no adenylation of ferulic acid was observed with RubC1b (detection limit: 0.1 % of the velocity obtained with tyrosine). This indicated that the ferulic acid adenylation activity was located in the N-terminal rather than the C-terminal part of RubC1, in agreement with the hypothesis that the aminocoumarin feruloyl ligase reaction

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

described above is catalyzed by the N-terminal part of the molecule.

The fact that RubC1 shows higher activity for the adenylation of L-tyrosine than of ferulic acid (Figure 4 A) was expected, as L-tyrosine is a natural substrate of this enzyme, while ferulic acid is not.

Discussion

All naturally occurring aminocoumarin antibiotics contain an aminocoumarin moiety that is linked to an acyl moiety by an amide bond. The biosynthesis of these antibiotics requires at least two acyladenylate-forming enzymes. First, an enzyme for the adenylation of L-tyrosine and its subsequent attachment to the 4′-phosphopantetheinyl cofactor of a peptidyl carrier protein (PCP). The prototype of such a tyrosyl-activating enzyme is NovH (3), consisting of an adenylation and a PCP domain. Second, the biosynthesis requires an enzyme for the adenylation of an acyl moiety and its subsequent transfer to the amino group of the aminocoumarin moiety. The prototype of such an aminocoumarin acyl ligase is NovL (4). NovL is similar in sequence to the acyl-CoA ligases (30). In contrast to nonribosomal peptide synthases, NovL contains neither a PCP nor a condensation domain, and it does not form a covalent bond with its substrate (4).

The biosynthetic gene cluster of rubradirin (13) contains two genes with similarity to *novL* - *orf4* and *rubF6* - and additionally a unique hybrid gene, *rubC1*, which represents a fusion of orthologues of *novL* and *novH* with a 600 bp linker region. In this study, we have shown that Orf4 is an active aminocoumarin acyl ligase with biochemical properties similar to those of SimL of simocyclinone biosynthesis (23,31). However, RubF6, which shares 88 % identity with Orf4 on the amino acid level, was inactive. Site-directed mutagenesis of the active site allowed us to convert RubF6 into an active aminocoumarin acyl ligase. The high sequence identity of *rubF6* to *orf4* suggests that these genes might have resulted from a gene-duplication event. One of the copies might then have mutated to an inactive form.

The natural acyl substrate of the aminocoumarin acyl ligase reaction in rubradirin biosynthesis is unknown. It is possible that the entire acyl moiety - consisting of rubransorol, the deoxynitrosugar, and dihydroxy-pipecolinic acid - is assembled before its transfer to the aminocoumarin moiety, similar to the biosynthetic reaction sequence of simocyclinone (23).

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

However, this complex acyl moiety of rubradirin has not been isolated or synthesized, and therefore could not be tested as substrate in this study.

The substrate specificity of RubC1 in the aminocoumarin acyl ligase assay was similar to that of Orf4 (Table S1). Our data indicate that both Orf4 and the N-terminal domain of RubC1 may be able to act as aminocoumarin acyl ligases in rubradirin biosynthesis.

RubC1 is a bifunctional enzyme. In addition to its aminocoumarin acyl ligase activity, RubC1 catalyzes the adenylation of tyrosine and its covalent attachment to a PCP domain. Both sequence analysis and biochemical investigation suggested that the latter functions are localized to the C-terminal part of RubC1, which is therefore expected to have identical function to that of NovH in novobiocin biosynthesis (Scheme 1) (2,3).

BLAST searches revealed many homologues to either the N- or the C-terminal part of RubC1, but not a single homologue with significant similarity over the full-length protein. The hybrid enzyme RubC1 is different from any NRPS or NRPS-PKS hybrid described previously (32,33).

The biosynthetic gene cluster of simocyclinone contains a gene for the aminocoumarin acyl ligase SimL and, 28 bp downstream thereof, a gene for the tyrosine-activating enzyme SimH (Figure 6). In contrast to NovH (600 aa), SimH (997 aa) shows an N-terminal extension of approximately 400 aa. The function of this extension, if any, is unknown. If the genes *simL* and *simH* were fused to a single coding sequence, the resulting gene would code for a protein with a very similar domain structure to RubC1. However, the 400 aa N-terminal extension of SimL does not show sequence similarity to the 200 aa linker region in RubC1, thus excluding a direct evolutionary relationship of *rubC1* to *simL* or *simH*.

The biochemical functions that RubC1 can fulfill in the biosynthesis of rubradirin are depicted in Scheme 2, in analogy to the established function of the five enzymes NovHIJKL of novobiocin biosynthesis, which are transcribed as a single operon (Figure 6) (34). The genes *rubC1*, *rubC2*, and *rubC3* in the rubradirin biosynthetic gene cluster resemble this operon. *rubC2* and *rubC3* are orthologues of *novI* and *novJ*, respectively. RubC3 alone might carry out the same reaction as NovJ/NovK, as all three enzymes belong to the FabG family of 3-ketoacyl-(acyl-carrier protein) reductases (35). RubC1 might adenylate

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

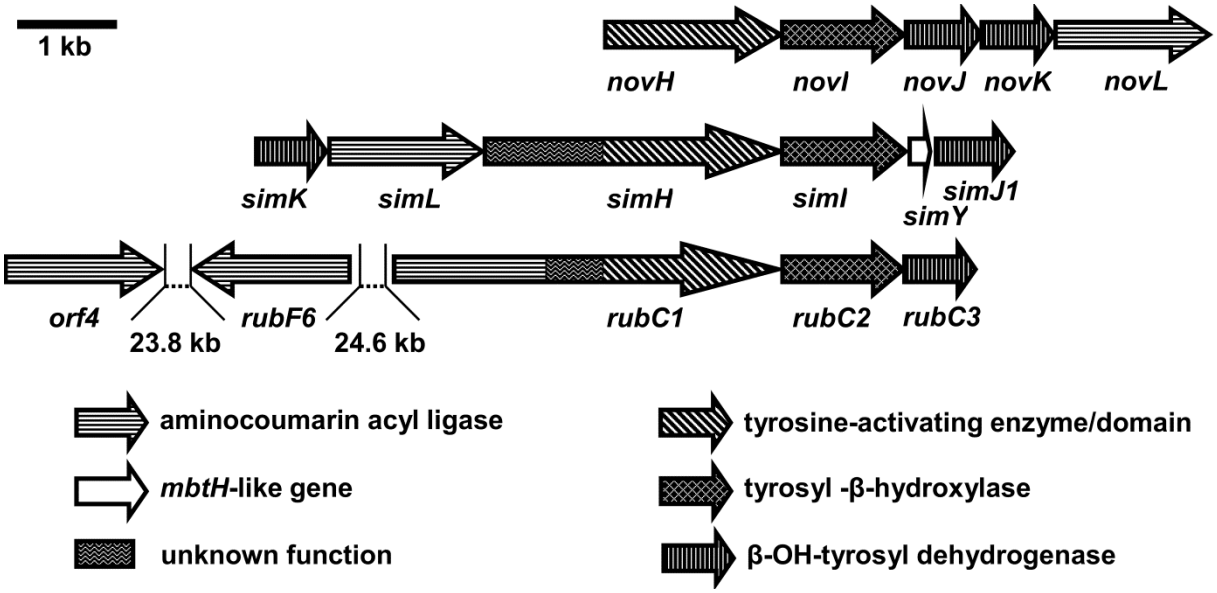


Figure 6. Organization of genes for adenylate-forming enzymes of aminocoumarin biosynthesis in the biosynthetic gene clusters of novobiocin (*nov*), simocyclinone (*sim*), and rubradirin (*rub*). The rubradirin gene cluster spans 106 kb, and the genes orf4, rubF6, and rubC1 are found at different loci within this cluster.

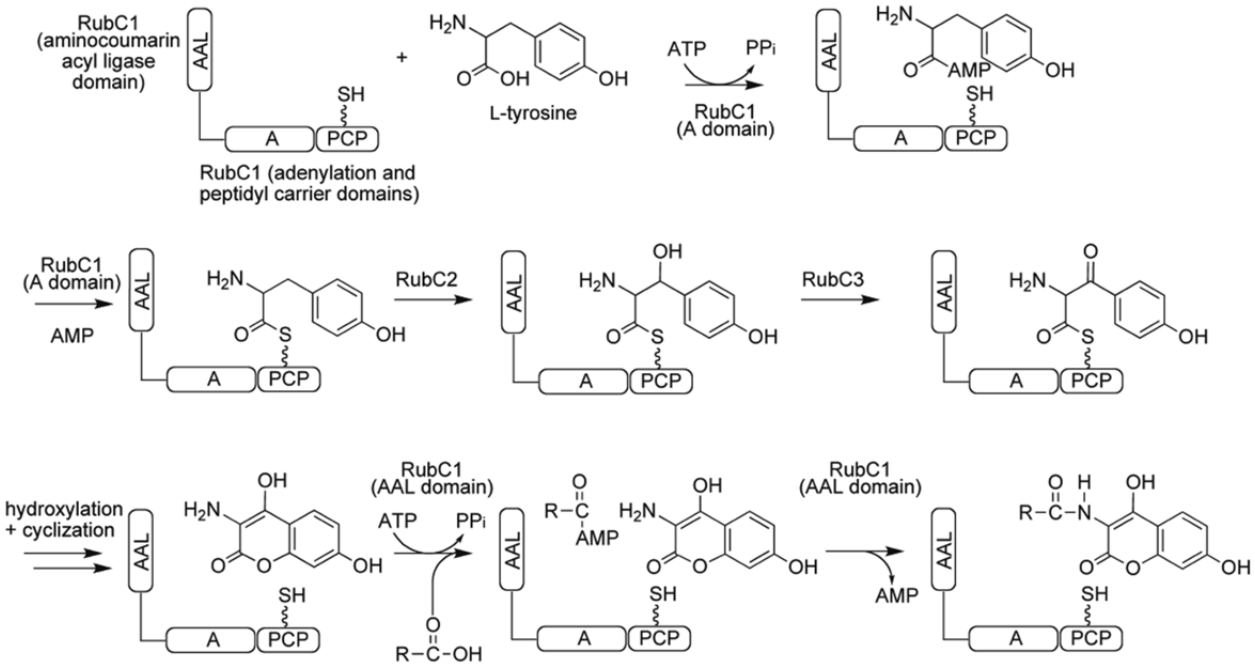
tyrosine and link it to the 4′-phosphopantetheinyl cofactor of its PCP domain. Subsequent β-hydroxylation of the tyrosyl moiety by RubC2 and oxidation by RubC3 would be carried out on the covalent tyrosyl-S-RubC1 complex (Scheme 2). Upon cyclization, the aminocoumarin moiety would be released from its covalent bond to RubC1 resulting in the substrate for the aminocoumarin acyl ligase (36). However, rather than being released into the cytoplasm, this moiety could be directly channeled to the N-terminal domain of RubC1, where its amino group would be acylated under catalysis of the aminocoumarin acyl ligase domain of RubC1 (Scheme 2). By this channeling mechanism, the unique bifunctional enzyme RubC1 could offer an efficient machinery for aminocoumarin antibiotic biosynthesis.

Experimental Section

Chemicals and radiochemicals

Tetrasodium [^{32}P]pyrophosphate (3.38 TBq mmol^{-1}) was obtained from Perkin–Elmer, L-[U-^{14}C]tyrosine (14.65 GBq mmol^{-1}) from Moravek Biochemicals and Radiochemicals. 3-Amino-4,7-dihydroxy-8-methyl-coumarin and novobiocic acid were kindly provided by Pharmacia & Upjohn, Inc. (Kalamazoo, MI). 3-Dimethylallyl-4-hydroxybenzoic acid was obtained by

Originally published in *ChemBioChem.* (2011), **12**, 1105-14



Scheme 2. Hypothetical role of RubC1 in the biosynthesis and the acylation of the aminocoumarin moiety of rubradirin. aaL=aminocoumarin acyl ligase domain; A=tyrosine adenylation domain; PCP=peptidyl carrier domain.

hydrolysis of novobiocin, as described previously (22). Acyl substrates for Orf4 and RubC1 were purchased from Sigma–Aldrich, Fluka, Lancaster and Merck.

Cloning of the genes *orf4*, *rubF6*, and *rubC1*

The genes *orf4*, *rubF6*, and *rubC1* of the rubradirin cluster were amplified from a cosmid library of *S. achromogenes* var. *rubradiris* NRRL 306113 by polymerase chain reaction. Primers *orf4*_f_*Nde*I (5′-GGG AAT TC**C ATA TG**A GTA CCA CGC ATG TGC ATT ATG-3′) *orf4*_r_*Xho*I (5′-GAG ATG GTT **CTC GAG** TCA GGC CTG GGG GGC G-3′), *rubC1*_f_*Nco*I (5′-CAT G**CC ATG G**CA TGC GGG GGC ATG ATC ACT ACA-3′) and *rubC1*_r_*Eco*RI (5′-CCG **GAA TTC** TCA TCG GCG GCC TCC ACG-3′) were used to amplify *orf4* and *rubC1*. Primers *orf4*_r_*Xho*I and *rubF6*_r_*Xho*I (5′-GAT CCA TCC **CTC GAG** CTA GGC CTG GGG GGC G-3′) were used to amplify *rubF6*. *rubC1b* was amplified with primers *rubC1b*_f_*Nco*I (5′-CAT G**CC ATG G**GC GCC TGC TCG GGG CAT TTC G-3′) and *rubC1*_r_*Eco*RI (see above). The introduced *Nde*I, *Xho*I, *Nco*I and *Eco*RI restriction sites of each primer are highlighted in bold. The amplified products were purified by gel electrophoresis, digested with the corresponding restriction enzymes, and ligated into pET28a for expression of

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

Orf4 and RubF6 as N-terminally His_6-tagged constructs, and into pHis8 (37) for expression of RubC1 and RubC1b N-terminally His_8-tagged constructs. The resulting plasmids pBB02 and pBB04 were transformed into *E. coli* Rosetta2 (DE3), while pBB05 and pBB11 were transformed into *E. coli* BL21(DE3) with pSU20_*sfp*, a plasmid containing the gene for the Sfp phosphopantetheinyl transferase from *Bacillus subtilis* (21) for protein overexpression of holo-RubC1 and holo-RubC1b.

Site-directed mutagenesis of RubF6 was carried out by PCR amplification of the template pBB04 using the QuickChange Site Directed Mutagenesis Kit (Stratagene). Reactions were performed according to the manufacturer's instructions with primers *rubF6*_N184PK_f (5′-AGG CTG CCG AGC GGA CGG cca AAa GGC CTC TGC TGG ACC TTC-3′), *rubF6*_RL178SG_f (5′-GGT GAT CAC CCA GAC CAG tgg GCC GAG CGG ACG GAA CGG-3′) and their reverse complements. The base changes are indicated by small letters. The PCR program consisted of an initial denaturation at 94 °C for 2 min followed by 18 cycles of 94 °C for 10 s, 55 °C for 30 s, and 68 °C for 9 min. For the deletion of the insertion sequence of *rubF6*, the Phusion Site-Directed Mutagenesis Kit (New England Biolabs) was used according to the manufacturer's instructions with the primers *rubF6*_Δ207–215_f (5′P-CGT GGA GGG CAG GGA CGG-3′) and *rubF6*_Δ207–215_f (5′P-ACG ACC CTC CTG ATC ACC GC-3′). The template DNA was digested with 10 units of *Dpn*I for 1 h at 37 °C before transformation. The correct DNA sequences of the entire genes were confirmed by sequencing.

Purification of His-tagged proteins

An overnight culture in Luria-Bertani medium (35 mL, 50 μg mL^{-1} kanamycin, 25 μg mL^{-1} chloramphenicol) of cells harboring the respective expression plasmid were used to inoculate terrific broth (38) (1 L, 50 μg mL^{-1} kanamycin, 25 μg mL^{-1} chloramphenicol). The cells were grown at 37 °C to an OD_{600} of 0.6, cooled to 20 °C, induced with isopropyl-β-D-thiogalactopyranoside (IPTG; 0.4 mM), and allowed to grow for an additional 14 h at 20 °C. The cells from each culture were harvested by centrifugation (15 min at 4800 x g) and resuspended in buffer A (25 mL per 10 g cells, 50 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol, 10 % glycerol). 1 % Tween 20 and lysozyme (0.5 mg mL^{-1}) were added, resuspended cells were

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

broken by a Branson sonifier, and the cell debris was removed by centrifugation (45 min at 35 000 x g). The supernatant was applied to a nickel–nitrilotriacetic acid–agarose resin column (GE Healthcare) according to the manufacturer's instructions by using a linear gradient of 0–60 % imidazole (250 mM, in 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 % glycerol, 10 mM β-mercaptoethanol) over 60 min for elution. Fractions containing the protein were pooled and further purified on a HiLoad 26/60 Superdex 200 column (Amersham Pharmacia Biotech) that had been equilibrated with Tris-HCl (20 mM, pH 8.0), NaCl (150 mM), and dithiothreitol (2 mM), concentrated by using an Amicon Ultra 10 000 MWCO centrifugal filter (Millipore), flash frozen in liquid nitrogen, and stored at −80 °C. Concentrations of the purified RubC1 and RubC1b were measured spectrophotometrically at 280 nm by using the calculated extinction coefficients of 0.820, 0.884, 0.735, and 0.712 g L^{-1} for Orf4, RubF6, RubC1, and RubC1b, respectively. His-tagged constructs were used without further modifications.

ATP-[^{32}P]PP_i-exchange assays with RubC1 and RubC1b

ATP-[^{32}P]PP_i-exchange assays (100 µL) contained Tris-HCl (75 mM, pH 8.0), $MgCl_2$ (5 mM), tris(2-carboxyethyl)phosphine hydrochloride (TCEP; 5 mM), ATP (2 mM), amino acid substrate (1.5 mM), RubC1 (0.5 µM), and [^{32}P]pyrophosphate (1 mM, Perkin–Elmer). The reactions were initiated by the addition of RubC1 or RubC1b, allowed to proceed for 5 min at 30 °C, and then quenched with a suspension of activated charcoal (500 µL, 1.6 %, w/v) in quenching buffer (4.5 %, w/v tetrasodium pyrophosphate and 3.5 % perchloric acid in water). The charcoal was pelleted by centrifugation, washed with quenching buffer, resuspended in water (0.5 mL), and added to scintillation liquid (9 mL). The radioactivity was quantified in a scintillation counter. Data reported are means of two independent reactions. RubC1 (4 µM) was used for incubations with L-phenylalanine as the substrate, and the reactions were allowed to continue for 10 min at 30 °C. Nonlinear regression was performed with Graph Pad Prism 5.0 (GraphPad Software Inc., La Jolla, USA).

Amide synthetase assays

Assay mixtures (100 µL) contained Tris-HCl (95 mM, pH 7.5 for Orf4 and pH 8.5 for RubC1), 3-amino-4,7-dihydroxy-8-methyl-coumarin (0.5 mM in DMSO), the

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

respective carboxylic acid substrate (6 mM), ATP (6 mM), and $MgCl_2$ (1 mM). Reactions were initiated with Orf4 or RubC1 and carried out for 30 min at 30 °C. The reactions were terminated by addition of an equal volume of methanol at 4 °C. After incubation at -20 °C for 20 min, the mixtures were centrifuged to remove precipitated protein (5 min, 15 700 g). The supernatant was analyzed by RP-HPLC using a Multosphere RP18–5 column (250×4 mm, 5 µm; C+S Chromatographie Service, Düren, Germany), with a linear gradient of 50-100 % solvent B (99 % CH_3OH, 1 % HCOOH) in solvent A (99 % H_2O, 1 % HCOOH) over 28 min, flow rate 1 mL min^{-1}. UV detection was carried out at 330 nm. The identity of the products was confirmed by LC-MS.

Auto-aminoacylation of RubC1

The incubation mixture (100 µL) included Tris-HCl (95 mM, pH 8.0), MgCl2 (5 mM), L-[U-^{14}C]tyrosine (0.1 mM, Moravek Biochemicals and Radiochemicals), TCEP (5 mM), ATP (2 mM), and RubC1 (50 µM). The reaction was allowed to proceed for 30 min at 30 °C, and 10 µL of the mixture were used for electrophoresis on 10 % SDS-PAGE. A control reaction was carried out in parallel in which ATP was omitted. For visualization, the gel was stained with Coomassie brilliant blue solution. The dried gel was exposed to a photostimulable phosphor plate for 72 h before development with a phosphor imager.

Acknowledgements

The authors are grateful to Bertolt Gust, Tübingen University, for valuable advice. We also thank Alex Tao for help with the synthesis of unmethylated aminocoumarin. This study was supported by a grant of the Deutsche Forschungsgemeinschaft (SFB766; to L.H.)

References

1. Marahiel, M. A., and Essen, L. O. (2009) *Methods Enzymol.* **458**, 337-351
2. Heide, L. (2009) *Nat. Prod. Rep.* **26**, 1241–1250
3. Chen, H., and Walsh, C. T. (2001) *Chem. Biol.* **8**, 301-312
4. Steffensky, M., Li, S. M., and Heide, L. (2000) *J. Biol. Chem.* **275**, 21754-21760
5. Pojer, F., Li, S. M., and Heide, L. (2002) *Microbiology* **148**, 3901-3911
6. Wang, Z. X., Li, S. M., and Heide, L. (2000) *Antimicrob. Agents Chemother.* **44**, 3040-3048
7. Galm, U., Schimana, J., Fiedler, H. P., Schmidt, J., Li, S. M., and Heide, L. (2002) *Arch. Microbiol.* **178**, 102-114
8. Trefzer, A., Pelzer, S., Schimana, J., Stockert, S., Bihlmaier, C., Fiedler, H. P., Welzel, K., Vente, A., and Bechthold, A. (2002) *Antimicrob. Agents Chemother.* **46**, 1174-1182
9. Galm, U., Dessoy, M. A., Schmidt, J., Wessjohann, L. A., and Heide, L. (2004) *Chem. Biol.* **11**, 173-183
10. Anderle, C., Hennig, S., Kammerer, B., Li, S. M., Wessjohann, L., Gust, B., and

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

Heide, L. (2007) *Chem. Biol.* **14**, 955-967
11. Heide, L. (2009) *Biotechnol Adv* **27**, 1006-1014
12. Freel Meyers, C. L., Oberthür, M., Heide, L., Kahne, D., and Walsh, C. T. (2004) *Biochemistry* **43**, 15022-15036
13. Kim, C. G., Lamichhane, J., Song, K. I., Nguyen, V. D., Kim, D. H., Jeong, T. S., Kang, S. H., Kim, K. W., Maharjan, J., Hong, Y. S., Kang, J. S., Yoo, J. C., Lee, J. J., Oh, T. J., Liou, K., and Sohng, J. K. (2008) *Arch. Microbiol.* **189**, 463-473
14. Bhuyan, B. K., Owen, S. P., and Dietz, A. (1965) *Antimicrob. Agents Chemother.*, 91-96
15. Reusser, F. (1973) *Biochemistry* **12**, 1136-1142
16. Wherli, W. (1977) *Ansamycins. chemistry, biosynthesis and biological activity* **72**, 21-49
17. Reusser, F. (1979) *J. Antibiot. (Tokyo)* **32**, 1186-1192
18. Reusser, F. B., B. Tarpley, W.G. Althaus, I. Zapotocky, B. (1988) Rubradirin derivatives for treatment of HIV infection. (Company, T. U. ed.
19. Ansari, M. Z., Yadav, G., Gokhale, R. S., and Mohanty, D. (2004) *Nucleic Acids Res.* **32**, 405-413
20. Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) *Chem. Biol.* **6**, 493-505
21. Quadri, L. E. N., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) *Biochemistry* **37**, 1585-1595
22. Kominek, L. A., and Meyer, H. F. (1975) *Methods Enzymol.* **43**, 502-508
23. Luft, T., Li, S. M., Scheible, H., Kammerer, B., and Heide, L. (2005) *Arch. Microbiol.* **183**, 277-285
24. Schmutz, E., Steffensky, M., Schmidt, J., Porzel, A., Li, S. M., and Heide, L. (2003) *Eur. J. Biochem.* **270**, 4413-4419
25. Tao, J., Hu, S., Pacholec, M., and Walsh, C. T. (2003) *Org. Lett.* **5**, 3233-3236
26. Reger, A. S., Carney, J. M., and Gulick, A. M. (2007) *Biochemistry* **46**, 6536-6546
27. Stuible, H., Büttner, D., Ehlting, J., Hahlbrock, K., and Kombrink, E. (2000) *FEBS Lett.* **467**, 117-122
28. Otten, L. G., Schaffer, M. L., Villiers, B. R., Stachelhaus, T., and Hollfelder, F. (2007) *Biotechnol. J.* **2**, 232-240
29. Lee, S. G., and Lipmann, F. (1975) *Methods Enzymol.* **43**, 585-602
30. Schmelz, S., and Naismith, J. H. (2009) *Curr. Opin. Struct. Biol.* **19**, 666-671
31. Pacholec, M., Freel Meyers, C. L., Oberthur, M., Kahne, D., and Walsh, C. T. (2005) *Biochemistry* **44**, 4949-4956
32. Koglin, A., and Walsh, C. T. (2009) *Nat. Prod. Rep.* **26**, 987-1000
33. Strieker, M., Tanović, A., and Marahiel, M. A. (2010) *Curr. Opin. Struct. Biol.* **20**, 234-240
34. Dangel, V., Härle, J., Goerke, C., Wolz, C., Gust, B., Pernodet, J. L., and Heide, L. (2009) *Microbiology* **155**, 4025-4035
35. Pacholec, M., Hillson, N. J., and Walsh, C. T. (2005) *Biochemistry* **44**, 12819-12826
36. Pacholec, M., Tao, J., and Walsh, C. T. (2005) *Biochemistry* **44**, 14969-14976
37. Jez, J. M., Ferrer, J. L., Bowman, M. E., Dixon, R. A., and Noel, J. P. (2000) *Biochemistry* **39**, 890-902
38. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York

Supplemental Data

Table S1. Aminocoumarin acyl ligase activities of Orf4, RubF6 and RubC1 with different acyl substrates (pkat mg^{-1} protein)

Substrates	RubC1	Orf4	RubF6
cinnamic acid	10.4	92.3	< 0.2
ferulic acid	98.7	9.9	< 0.2
p-coumaric acid	2.9	9.3	< 0.2
caffeic acid	< 0.2	< 0.2	< 0.2
3-dimethylallyl-4-hydroxybenzoic acid	6.3	15.8	< 0.2
benzoic acid	< 0.2	0.5	< 0.2
4-hydroxybenzoic acid	< 0.2	0.9	< 0.2
vanillic acid	< 0.2	< 0.2	< 0.2
2,3-dihydroxybenzoic acid	< 0.2	< 0.2	< 0.2
3,4-dihydroxybenzoic acid	< 0.2	< 0.2	< 0.2
3,5-dimethyl-4-hydroxybenzoic acid	< 0.2	< 0.2	< 0.2
3-bromo-4-hydroxybenzoic acid	< 0.2	< 0.2	< 0.2
3,5-dibromo-4-hydroxybenzoic acid	< 0.2	< 0.2	< 0.2
4-aminobenzoic acid	< 0.2	< 0.2	< 0.2
4-amino-3-hydroxybenzoic acid	< 0.2	< 0.2	< 0.2
4-hydroxy-3-propylbenzoic acid	< 0.2	< 0.2	< 0.2
(3,4-dihydroxyphenyl)acetic acid	< 0.2	< 0.2	< 0.2
(3,4-dihydroxyphenyl)hydroxyacetic acid	< 0.2	< 0.2	< 0.2
3-(3,4-dihydroxyphenyl)propanoic acid	< 0.2	< 0.2	< 0.2
3,5-dimethylpyrrolidine-2,4-dicarboxylic acid	< 0.2	< 0.2	< 0.2
chelidamic acid	< 0.2	< 0.2	< 0.2
pyridine-2,6-dicarboxylic acid	< 0.2	< 0.2	< 0.2
2,6-dimethyl-3,5-pyridinedicarboxylic acid	< 0.2	< 0.2	< 0.2
6-methylpyridine-2,3-dicarboxylic acid	< 0.2	< 0.2	< 0.2

Originally published in *ChemBioChem.* (2011), **12**, 1105-14

Chapter 7 • *The Role of MbtH-like Proteins in the Adenylation of Tyrosine*

This manuscript has originally been published in The Journal of Biological Chemistry.

'The Role of MbtH-like Proteins in the Adenylation of Tyrosine during Aminocoumarin and Vancomycin Biosynthesis'

Björn Boll, Tatjana Taubitz and Lutz Heide[]

Pharmazeutisches Institut, Universität Tübingen, 72076 Tübingen (Germany)

* To whom correspondence should be addressed: E-mail: heide@uni-tuebingen.de
Phone: +49 (0)7071 29 72460
Fax: +49 (0)7071 29 5250

Keywords: Antibiotics, Bacteria, Enzyme mechanisms, Natural product biosynthesis, Peptide biosynthesis, Adenylation, Aminocoumarin, Biosynthesis, MbtH, Vancomycin

The Journal of Biological Chemistry Vol. 286, Publication ahead of print

Received August 02, 2011; first published online September 02, 2011

Preface − About the Manuscript

MbtH-like proteins reside in many bacterial gene clusters of non-ribosomal formed secondary metabolites. The function of these small proteins remained a mystery for a long time. Recently, several biochemical studies shed light on their function. The following publication describes the role of MbtH-like proteins in aminocoumarin and vancomycin biosynthesis. MbtH-like proteins were found to bind to the L-tyrosine adenylating enzymes and thereby greatly stimulate their activity. The study contributes to understand the diversity of these proteins, their promiscuity in interacting with alternate adenylating domains, and the stoichiometry of the interaction. Additionally, a single point mutation near the active center resulted in MbtH-independent activity of a formerly MbtH-dependent adenylating enzyme.

Author contributions:

- Björn Boll
  - decisively involved in initial hypothesis generation
  - decisively involved in planning, establishing and accomplishment of experimental setup
  - generation of data (cloning, expression and purification of protein, generation of *ybdZ* knock-out strain, performance of all activity assays, size exclusion experiments and sequence analysis), leading to figures 3-6 and S1-S4
  - decisively involved in all data discussions
  - writing of the manuscript
  - preparation of all figures
- Tatjana Taubitz
  - assisted in cloning and purification of SimH, NovH and CloY
- Lutz Heide
  - supervised the project
  - decisively involved in initial hypothesis generation
  - decisively involved in all data discussions and analyses
  - manuscript preparation

My personal part for this manuscript involved the initial experiment that proofed the influence of MbtH-like proteins of L-tyrosine activating enzymes. From these results I and Prof Dr. L. Heide designed further experiments to investigate the specific effect and strengthen our hypothesis. The results reported in this manuscript that expands the understanding of the MbtH-like proteins like their physical interaction and the augmentation of catalytic activity of the interacting enzymes. Future work on amino acid adenylation will greatly benefit from knowledge of the important role of MbtH-like proteins in the enzymatic adenylation steps during combinatorial peptide synthesis using non-ribosomal peptide synthetase scaffolds.
For this manuscript I wrote major text parts, generated all figures and was in charge of the final version.

DOI: 10.1074/jbc.M111.288092

Summary

MbtH-like proteins consist of approximately 70 amino acids and are encoded in the biosynthetic gene clusters of non-ribosomally formed peptides and other secondary metabolites derived from amino acids. Recently, several MbtH-like proteins have been shown to be required for the adenylation of amino acid in non-ribosomal peptide synthesis. We now investigated the role of MbtH-like proteins in the biosynthesis of the aminocoumarin antibiotics novobiocin, clorobiocin and simocyclinone D8 and of the glycopeptide antibiotic vancomycin. The tyrosine-adenylating enzymes CloH, SimH and Pcza361.18, involved in the biosynthesis of clorobiocin, simocyclinone D8 and vancomycin, respectively, required the presence of MbtH-like proteins in a 1:1 molar ratio, forming heterotetrameric complexes. In contrast, NovH involved in novobiocin biosynthesis showed activity in the absence of MbtH-like proteins. Comparison of the active centers of CloH and NovH showed only one amino acid to be different, *i.e.* L383 *versus* M383. Mutation of this amino acid in CloH (L383M) indeed led to MbtH-independent adenylation activity. All investigated tyrosine-adenylating enzymes exhibited remarkable promiscuity for MbtH-like proteins from different pathways and organisms. YbdZ, the MbtH-like protein from the expression host *E. coli*, was found to bind to adenylating enzymes during expression and to influence their biochemical properties markedly. Therefore, the use of *ybdZ*-deficient expression hosts is important in biochemical studies of adenylating enzymes.

Introduction

The adenylation of amino acids is a key step in the biosynthesis of many antibiotics (*e.g.* vancomycin, daptomycin), immunosuppressants (*e.g.* cyclosporine A), siderophores (*e.g.* enterobactin, mycobactin) and other bioactive molecules (1). The activated amino acids can be assembled to peptides by non-ribosomal peptide synthases (NRPSs), leading *e.g.* to the backbone of vancomycin, or can serve as precursors of non-peptidic antibiotics like novobiocin, clorobiocin and simocyclinone D8 (Fig. 1). Approximately half of the biosynthetic gene clusters for non-ribosomally formed peptides, as well as the gene clusters for clorobiocin and simocyclinone D8, contain so-called *mbtH*-like genes. These small genes are named after *mbtH* contained in the gene cluster for the siderophore mycobactin in *Mycobacterium tuberculosis* which codes for a 71 aa protein. The function of

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

mbtH-like genes has remained enigmatic for many years. First proof that these genes are essential for secondary metabolite production was provided by a gene inactivation and complementation study of our group in clorobiocin biosynthesis (2) and by a similar study in the biosynthesis of coelichelin and calcium-dependent antibiotic (3). *In vivo* investigations were complicated by the fact that many bacterial genomes contain several *mbtH*-like genes which can functionally replace each other. The importance of a specific *mbtH*-like gene for the biosynthesis of a secondary metabolite can only be assessed after all other *mbtH*-like genes in the genome have been inactivated (2,3). However, *in vivo* studies could not define the precise physiological function of the *mbtH*-like genes, *e.g.* in catalysis, regulation, transport or protein-protein interactions. The three-dimensional structures of two MbtH-like proteins have been experimentally determined (4,5), but again this did not allow to recognize their function.

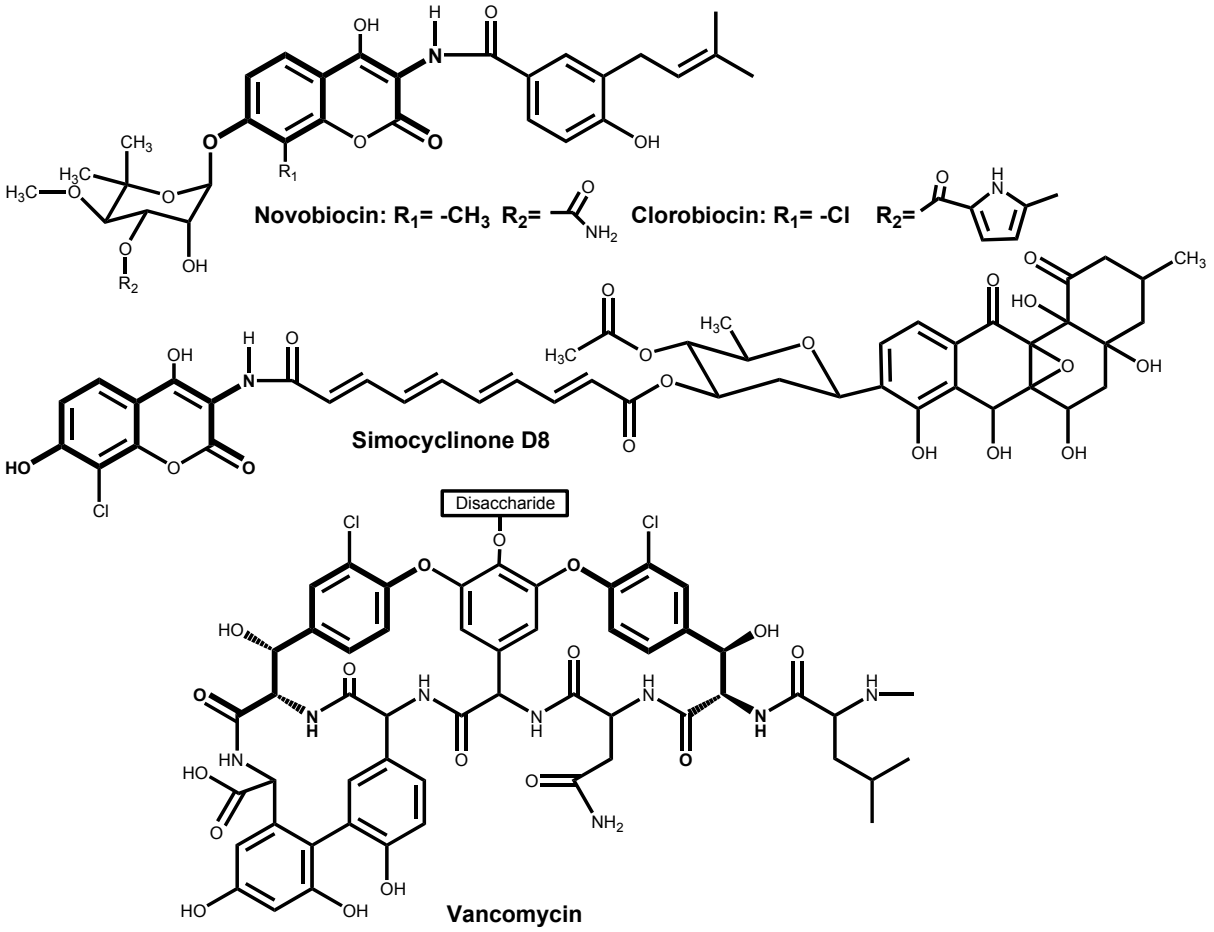


Figure 1: Structures of the aminocoumarin antibiotics novobiocin, clorobiocin and simocyclinone D8 and of the glycopeptide antibiotic vancomycin. Parts derived from L-tyrosine are drawn bold.

The first biochemical evidence for the function of MbtH-like proteins in non-ribosomal peptide biosynthesis has recently been provided in two rapid reports by Felnagle *et al.* (6) and Zhang *et al.* (7). Additional data were presented as part of a study on glidobactin biosynthesis (8). These reports showed that MbtH-like proteins interact with adenylating enzymes which are part of NRPSs. *In vitro*, the adenylating activity of these enzymes was strongly stimulated by addition of MbtH-like proteins. Out of the five adenylation domains which activate the different amino acids required for capreomycin biosynthesis, three were dependent on the presence of the MbtH-like protein CmnN, while the two others were not. The reason for this difference is unknown (6). The heterologous expression of adenylating enzymes in *E. coli* was found to be difficult or even impossible unless the respective MbtH-like protein was co-expressed simultaneously (6-8). From these data, Imker *et al.* (8) concluded that MbtH-like proteins act as activators, chaperones or both in NRPS assembly line. MbtH-like proteins form complexes with the adenylating enzymes, but the stoichiometry of these complexes has remained unclear. After purification of such complexes, the molar ratio of adenylating enzyme to MbtH-like protein was reported as 1:0.42 by Felnagle *et al.* (6), and as 1:1.7 by Imker *et al.* (8). If the adenylating enzyme CmnO and the MbtH-like protein CmnN were purified separately, a mixture of both in a 1:1 molar ratio showed only low activity. 10-fold higher activity was observed when the MbtH-like protein was added in a 16- to 32-fold molar excess (6). Therefore, the composition of the complex of MbtH-like proteins with adenylating enzymes is yet obscure.

As pointed out in a recent review on the occurrence and functions of MbtH proteins (9), the requirement of many adenylating enzymes for MbtH-like proteins implicates that the correct use of *mbtH*-like genes is a crucial factor for the success of combinatorial biosynthesis experiments. Our group is working extensively on the combinatorial biosynthesis of new aminocoumarin antibiotics (10,11). We were therefore interested to investigate the role of *mbtH*-like genes in the formation of these antibiotics. The biosynthesis of aminocoumarins involves the adenylation of L-tyrosine, followed by its attachment to a peptidyl carrier protein (PCP) domain and its *ß*-hydroxylation by a cytochrome P450 enzyme (Fig. 2A) (12,13). The same reaction sequence is part of the biosynthesis of vancomycin

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

and related glycopeptide antibiotics (14). In contrast to non-ribosomal peptide biosynthesis, however, the resulting *ß*-hydroxytyrosine (*ß*-OH-Tyr) is not transferred by a condensation domain to a peptide backbone. Rather, in the biosynthesis of vancomycin and the related balhimycin, *ß*-OH-Tyr is liberated from the PCP domain by a thioesterase and subsequently activated by another specific adenylation domain of the vancomycin or balhimycin NRPS (Fig. 2A) (15).

In aminocoumarin biosynthesis, *ß*-OH-Tyr is oxidized and cyclized to 4,7-dihydroxy-3-aminocoumarin, which is liberated from the PCP domain and

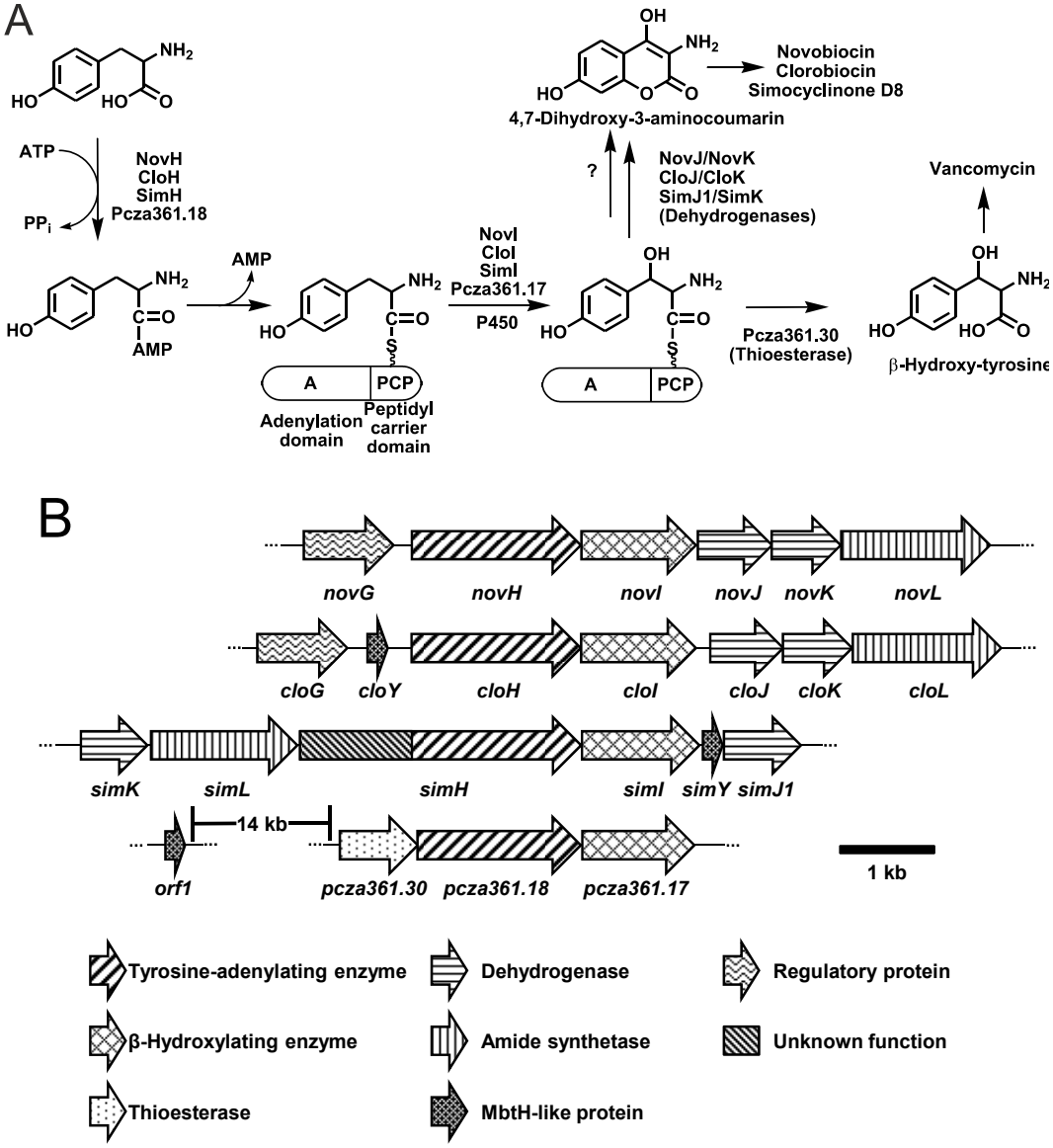


Figure 2: A) Adenylation and β-hydroxylation of tyrosine in the biosynthesis of aminocoumarin antibiotics and vancomycin. B) Genes for the formation of 4,7-dihydroxy-3-aminocoumarin and β-hydroxytyrosine in the biosynthetic gene clusters of novobiocin (*nov*), clorobiocin (*clo*), simocyclinone (*sim*), and vancomycin (*pcza*).

subsequently connected to an acyl moiety via an amide bond (Fig. 2A). In contrast to non-ribosomal peptide biosynthesis, formation of this amide bond does not involve the intermediary attachment of the acyl moiety to a PCP domain (12,16).

The biosynthetic gene clusters of vancomycin, clorobiocin and simocyclinone contain the *mbtH*-like genes *orf1van*, *cloY* and *simY*, respectively. In the vancomycin cluster, the gene *orf1van* is located 14 kb upstream from the genes coding for the enzymes for tyrosine adenylation (*i.e.* Pcza361.18), *ß*-hydroxylation and thioester cleavage (Fig. 2B). In the simocyclinone and clorobiocin clusters, the *mbtH*-like genes *simY* and *cloY* are located directly adjacent to the genes responsible for tyrosine adenylation and *ß*-hydroxylation, *i.e.* to *simH/simI* and *cloH/cloI*, albeit on different sides (Fig. 2B). Surprisingly, the novobiocin cluster does not contain an *mbtH*-like gene, despite otherwise exactly identical organization of the genes for aminocoumarin biosynthesis (Fig. 2B). *In vivo* studies have suggested that the formation of the aminocoumarin moiety of clorobiocin requires an intact *mbtH*-like gene, but the identical reaction sequence in novobiocin biosynthesis does not (2). Given the fact that the tyrosine-adenylating enzyme NovH has the same size (600 aa) as CloH, and both proteins share 83 % identity in their amino acid sequence (Fig. S1, Supplemental Data), this difference in the requirement for MbtH-like proteins is puzzling. We decided to express the tyrosine-adenylating enzymes NovH, CloH, SimH and Pzca361.18 as well as the cognate MbtH-like proteins, and to biochemically investigate these enzymes and the complexes formed by them.

Experimental Procedures

Chemicals and radiochemicals.

Tetrasodium [^{32}P]pyrophosphate (3.38 TBq mmol^{-1}) was obtained from Perkin Elmer. L-tyrosine was purchased from Merck.

Cloning of the genes *novH, cloH simH, cloY, simY, cdaX, pcza361.18* and *orf1van*

The genes *novH, cloH, simH, cloY and simY* were amplified from cosmids containing the respective clusters (17-19) by polymerase chain reaction. The gene *cdaX* was amplified from chromosomal DNA of *S. coelicolor* M512. Primers *novH*_f_*Nde*I (5'-GGG AAT TC**C ATA TG**T TCA ACA CAC GTG CGA AC-3') *novH*_r_*Xho*I (5'-GCC **CTC GAG** TCA CTC CTC CAG GGT CGC TA-3'), *cloH*_f_*Nde*I (5'-GGG AAT

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

TC**C ATA TG**T TAA ACA CGG GTC TGA ACA A-3'), *cloH*_r_*Xho*I (5'- GCC **CTC GAG** TCA CTC CCC GAG GGT CG -3') *simH*_f_*Nde*I (5'-GGG AAT TC**C ATA TG**G CCA TGC CAT CCG GCA-3'), *simH*_r_*Xho*I (5'-GCC **CTC GAG** TCA CTT CAC GGC CGT TGT GG -3'), *cloY*_f_*Nde*I (5'-GGG AAT TC**C ATA TG**G CGA CGA ACC CGT TCG A-3'), *cloY*_r_*Xho*I (5'-GCC **CTC GAG** CTA CTC GCC ACC CAT CGC-3'), *simY*_f_*Nde*I (5'-GGG AAT TC**C ATA TG**G CCA ACC CGT TTG ACG A-3'), *simY*_r_*Xho*I (5'-GCC **CTC GAG** TCA GCT GGG GTC CGT CG-3'), *cdaX*_f_*Nde*I (5'-GGG AAT TC**C ATA TG**A CCA ATC CGT TCG AAG ACG-3'), *cdaX*_r_*Xho*I (5'- GCC CTC GAG TCA GTT GCC GGT GCT CAT CG-3'), were used to amplify *novH, cloH, simH, cloY, simY* and *cdaX*. The introduced *Nde*I and *Xho*I restriction sites of each primer are highlighted in bold. The amplified products were purified by gel electrophoresis, digested with corresponding restriction enzymes and ligated into pET28a for expression as N-terminally His_6-tagged fusion proteins. The nucleotide sequences of *pcza361.18* and *orf1van* were optimized for expression in *E. coli* and synthesized commercially by Mr. Gene (Regensburg, Germany). The two genes were excised from their vectors with *Nde*I and *Xho*I and ligated into vector pET28a using the same restriction sites. For coexpression of CloH and CloY, both genes were ligated in the dual expression vector pETDUET1 (Novagene). The *cloH* gene was amplified with primers introducing a thrombin restriction site at the N-terminus: *cloH*_f_*BamH*I (5'-CG**G GAT CC**C CTG GTC CCG CGT GGT TCC TTA AAC ACG GGT CTG AAC AAG GC-3'); *cloH*_r_*Not*I (5'-A TAA GAA T**GC GGC CGC** TCA CTC CCC GAG GGT CG-3'). Restriction sites of each primer are highlighted in bold, the thrombin cleavage site is underlined. *cloY* was amplified with the same primers as before and ligated *via* the restriction sites *Nde*I and *Xho*I, resulting in an untagged protein. The correct DNA sequences of the entire genes were confirmed by sequencing. The resulting plasmids pBB28 (*simY*), pBB32 (*cdaX*), pBB34 (*orf1van*) and pBB37 (*cloY*) were transformed into *E. coli* Rosetta2 (DE3). pBB25 (*simH*), pBB26 (*cloH*), pBB35 (*cloH* and *cloY*), pBB43 (*pcza361.18*) and pBB44 (*novH*) were transformed into *E. coli* BL21(DE3) either with or without carrying an *ybdZ* deletion, and carrying pSU20_*sfp*, a plasmid containing the gene for the Sfp phosphopantetheinyl transferase from *Bacillus subtilis* for expression of the holo-enzymes (20).

Generation of *ΔybdZ E. coli* BL21(DE3) strain

The *E. coli* BL21(DE3) Δ*ybdZ* mutant was generated in BL21(DE3)/pIJ790 using Red/ET-mediated recombination (21). An apramycin resistance cassette [*acc(3)IV*] was amplified from plasmid pIJ773 (21). The primers used for PCR were as follows: *ybdZ*_f: (5'-*CCT CTG GCA ACC ACT TTT CCA TGA CAG GAG TTG AAT ATG* TGT AGG CTG GAG CTG CTT C-3') and *ybdZ*_r (5'-*TGC CGG GCT GTG CGG CGA CCA AAG GTA AAT GCT GGC TCA* ATT CCG GGG ATC CGT CGA CC-3'). Italic letters represent 39 nucleotides homologous to the regions up- and downstream of *ybdZ*, allowing Red/ET-mediated recombination. The amplicon was electroporated into a BL21(DE3)/pIJ790 strain after induction of the λ RED recombination system by arabinose. Strains containing the *ybdZ::acc(3)IV* mutation were selected on LB apramycin (100 µg ml^{-1}) at 37 °C leading to loss of the temperature-sensitive plasmid pIJ790. The genotype of the resulting mutants was confirmed by PCR with chromosomal DNA.

Site directed mutagenesis of CloH

Site directed mutagenesis of CloH was carried out by PCR amplification of the template *cloH*_pGEMT using the QuickChange Site Directed Mutagenesis Kit (Stratagene). Reactions were performed according to the manufacturer's instructions with primers of *cloH*_L383M_f (5'- CCC GAC TTG ACC GCG CAG aTG TTC GTG GCC AAC CCG T -3'), and the reverse complement. The base changes are indicated by small letters. The PCR program consisted of an initial denaturation at 94 °C for 2 min followed by 18 cycles of 94 °C for 10 s, 55 °C for 30 s, and 68 °C for 6 min. The template DNA was digested with 10 units of *Dpn*I for 1 h at 37 °C before transformation. The correct DNA sequence of the entire gene was confirmed by sequencing, and the DNA fragment was cloned in pET28a *via Nde*I and *Xho*I.

Purification of His-tagged Proteins

35 ml of an overnight culture in Luria-Bertani medium (50 µg ml^{-1} kanamycin, 25 µg ml^{-1} chloramphenicol) of cells harboring the respective expression plasmid were used to inoculate 1 liter of terrific broth (50 µg ml^{-1} kanamycin, 25 µg ml^{-1} chloramphenicol). The cells were grown at 37 °C to an OD_{600} of 0.6, cooled to 20 °C, induced with 0.4 mM of IPTG and allowed to grow for an additional 14 h at 20 °C. The cells from each culture were harvested by centrifugation (15 min at 4800×g) and

resuspended in 25 ml of buffer A (50 mM Tris–HCl pH 8.0, 0.5 M NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol, 10 % glycerol) per 10 g cells. 1% Tween 20 and 0.5 mg/ml lysozyme were added, resuspended cells were broken by a Branson sonifier and the cell debris was removed by centrifugation (45 min at 35,000×g). The supernatant was applied to a nickel-nitrilotriacetic acid-agarose resin column (GE Healthcare) according to the manufacturer's instructions, using a linear gradient of 0–60% 250 mM imidazole (in 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol, 10 mM β-mercaptoethanol) in 60 min for elution. Fractions containing the protein were pooled and further purified with a HiLoad 26/60 Superdex 200 column (Amersham Pharmacia Biotech) that had been equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM NaCl and 2 mM dithiothreitol, concentrated using an Amicon Ultra 10,000 MWCO centrifugal filter (Millipore) and stored at -80 °C. Concentrations of the purified proteins were measured spectrophotometrically at 280 nm using the calculated extinction coefficients. The N-terminal His-tags of the MbtH-like proteins were removed by incubation with 0.4 units thrombin (Sigma) per milligram of MbtH protein for 8 h at 4 °C before gel filtration. The tyrosine-adenylating enzymes were used without further modifications. The MbtH-like proteins were obtained in the following amounts per liter culture CloY 6.9mg, SimY 27 mg, CdaX 23.8 mg and Orf1van 30.2 mg. From the *ΔybdZ* expression hosts, the tyrosine-adenylating enzymes were obtained in the following amounts: NovH 38.6 mg, CloH 2.83 mg, SimH 9.4 mg, Pcza361.18 10.3 mg. From the *ybdZ*$^{+}$ expression hosts, the yields were: NovH 64 mg, CloH 11.3 mg, Pcza361.18 21.6 mg. The protein yield for the coexpression of CloH and CloY was 4.8 mg per liter culture using the *ΔybdZ* expression host. The mutant protein CloHL383M yielded 1.46 mg per 100 ml culture.

ATP-[^{32}P]PP$_i$ Exchange Assays

ATP-[^{32}P]PP$_i$ exchange assays (100 µl) contained 95 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 2 mM ATP, 1.5 mM L-tyrosine, 1 µM of the respective tyrosine-activating enzyme, 1.2 µM of the respective MbtH-like protein (unless other amounts are indicated) and 1 mM [^{32}P]pyrophosphate (Perkin Elmer). The reactions were initiated by the addition of the tyrosine-activating enzyme and allowed to proceed for 5 min at 30 °C, and then quenched with 500 µl of a

suspension of activated charcoal (1.6% w/v) in quenching buffer (4.5% (w/v) tetrasodium pyrophosphate and 3.5% perchloric acid in water). The charcoal was pelleted by centrifugation, washed with quenching buffer, resuspended in 0.5 ml of water and added to 9 ml scintillation liquid. Radioactivity was quantified in a scintillation counter. Data reported are means of two independent reactions. Activity was expressed in katal (kat) (22). For investigation of enzyme kinetics, nonlinear regression was performed with Graph Pad Prism 5.0 (GraphPad Software Inc., La Jolla, USA).

Analytical gel filtration

Analytical gel filtration was performed using a Superdex 200, 10/300 GL column (GE Healthcare) with a 24 ml bed resin and a buffer system of 20 mM Tris-HCl, pH 8.0, and 150 mM NaCl. A standard curve plotting the log of molecular weight standards versus the calculated K_{av} was generated using the following protein standards: ribonuclease A, chymotrypsinogen, ovalbulmin, albumin, aldolase, catalase and ferritin (GE Healthcare). This standard curve was used to calculate the observed molecular weight of SimH and SimH/SimY complexes.

Results

Expression and purification of tyrosine adenylating enzymes and MbtH-like proteins

The tyrosine-adenylating enzymes NovH, CloH, SimH and Pcza361.18 as well as the MbtH-like proteins CloY, SimY and Orf1van were expressed in *E. coli* in form of N-terminally His-tagged proteins and purified by Ni^{2+} affinity chromatography, followed by gel chromatography. The MbtH-like proteins were subjected to thrombin cleavage in order to remove the His-tag before gel chromatography (see Experimental Procedures).

The genome of *E. coli* contains an *mbtH*-like gene, *ybdZ*, in the gene cluster for the siderophore enterobactin. Felnagle *et al.* (6) have shown that purified YbdZ can activate adenylating enzymes of different NRPSs, albeit with low efficiency. In order to exclude the possibility that YbdZ would copurify with the expressed adenylating enzymes and interfere with subsequent investigations, we deleted the *ybdZ* gene from the *E. coli* expression host, utilizing the same Red/ET-mediated recombination strategy as described in a previous study of our group (11). The expression and purification of the adenylating enzymes was carried out both using the unmodified and using the *ΔybdZ*

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

expression host. Notably, the yields of the adenylating enzymes from the *ΔybdZ* expression strain were 2-4 times lower than from the unmodified strain (see Experimental Procedures). Using the *ΔybdZ* strain, most of the expressed CloH was insoluble, leading to a low yield of soluble protein (2.8 mg per liter culture). This is in accordance with previously reported difficulties in expressing adenylating enzymes without the corresponding MbtH-like proteins (6-8). In contrast, NovH was readily obtained in a yield of 38.6 mg per liter culture. Fig. S2 (Supplemental Data) shows an SDS-PAGE analysis of the adenylating enzymes purified from the *ΔybdZ* strain, and of the purified MbtH-like proteins.

We also purified the MbtH-like protein CdaX encoded in the gene cluster for the calcium-dependent antibiotic (CDA) of *Streptomyces coelicolor*. Our previous *in vivo* studies had shown that *cdaX* can functionally replace *cloY* in the biosynthesis of the aminocoumarin moiety of clorobiocin (2).

The gene products of the MbtH-like proteins Orf1van (AAL90876.1), CloY (AAN65223), SimY (AAG34186) and CdaX (CAB38589) comprise 70, 71, 69 and 71 amino acids, respectively. As shown in a phylogenetic analysis of MbtH-like proteins by Zhang *et al.* (7), CdaX and CloY are situated in close proximity in one branch of the phylogenetic tree, and Orf1van and SimY in another branch. Sequence identity among these proteins is around 60 %.

The tyrosine-adenylating enzymes Pcza361.18 (CAA11773) and CloH (AAN65224) comprise 580 and 600 amino acids, respectively, and show 43 % identity to each other. In contrast, the gene product of *simH* (AAL15600) comprises 997 amino acids. Like Pcza361.18 and CloH, SimH contains an adenylation and a PCP domain, and these domains show 56 % sequence identity to CloH. However, SimH carries additional 400 amino acids at is N-terminus. This domain shows moderate similarity to the condensation domains of NRPSs, but no function can be assigned to it in simocyclinone biosynthesis. As mentioned above, NovH (AAF67501) has exactly the same size as CloH, and these two proteins share 83 % sequence identity (Fig. S1, Supplemental Data).

The tyrosine-activating enzymes CloH, SimH and Pcza361.18 are dependent on MbtH-like proteins

The purified proteins were investigated for their tyrosine-adenylating activity using the well-established pyrophosphate exchange assay,

following the procedure described in a previous study on NovH (13). When assayed alone, CloH, SimH and Pcza361.18 showed a very low activity (Fig. 3). Upon addition of the respective MbtH-like protein, however, adenylating activity was readily detectable. A control with MbtH-like proteins alone showed no activity.

The results depicted in Fig. 3 were obtained with proteins expressed in the *ΔybdZ E. coli* strain. If CloH and Pcza361.18 were expressed in an *E. coli* strain with intact *ybdZ*, results were markedly different: in this case, tyrosine-adenylating activity of CloH and Pcza361.18 alone could clearly be detected, amounting to 20-30 % of the activity measured after addition of the respective MbtH-like protein. This suggests that YbdZ of *E. coli* had been copurified with the heterologously expressed adenylating enzymes. In contrast, SimH purified from strains with intact *ybdZ* gene did not show more activity than protein purified from a *ΔybdZ* strain.

The tyrosine-activating enzyme NovH is active in the absence of MbtH-like proteins

In clear contrast to CloH, SimH and Pcza361.18, NovH purified from the *ΔybdZ* strain showed activity in the absence of an MbtH-like protein (Fig. 3). This MbtH-independent activity is in agreement with our previous *in vivo* study which showed that biosynthesis of novobiocin, but not of clorobiocin, can be readily observed in strains which completely lack *mbtH*-like genes (2). However, activity of NovH was only moderate, and was stimulated markedly upon addition of CloY (Fig. 3). In view of the fact that the novobiocin cluster does not contain an *mbtH*-like gene, this result was unexpected.

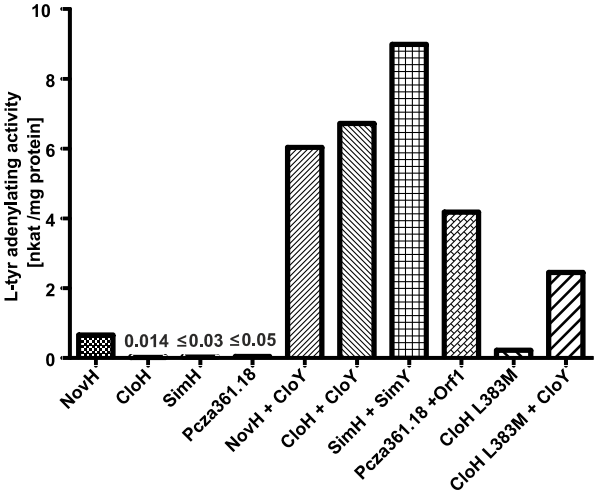


Figure 3: Activities of L-tyrosine-adenylating enzymes in the absence and presence of MbtH-like proteins. The MbtH-like proteins were added in a 1.2-fold molar excess over the tyrosine-adenylating enzymes. Data represents the mean value of two independent reactions. The value for CloH is the mean of six independent reactions.

Enzyme kinetics of the adenylation of L-tyrosine

We determined the K_m and k_{cat} values for the adenylation of tyrosine catalyzed by either NovH alone or by a mixture of NovH and CloY, using the

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

pyrophosphate exchange assay. The K_m value for L-tyrosine was nearly identical in both cases, but addition of CloY clearly increased the observed turnover number (Table 1). We noticed, however, that the observed K_m value (275 μM) was more than five times lower than the K_m value of 1390 μM determined for the same enzyme in an earlier study by Chen and Walsh (13). In that study, NovH had been expressed in an *E. coli* strain with intact *ybdZ*. This prompted us to repeat the investigation of NovH, this time using a protein expressed in an *ybdZ*$^+$ strain. Indeed, that protein showed a K_m value of 1278 μM for L-tyrosine, very similar to the value previously determined by Chen and Walsh (13). The turnover number of the enzyme from the *ybdZ*$^+$ strain (0.12 s^{-1}) was 1.6 times higher than that from the *ΔybdZ* strain. It appears therefore likely that YbdZ had been copurified with NovH, similar as we observed for CloH and Pcza361.18. These data show that MbtH-like proteins influence both the turnover number of the tyrosine-adenylating enzyme and their K_m value for the amino acid. Although the copurified YbdZ increased the turnover number observed for NovH, it actually decreased the catalytic efficiency (k_{cat} K_m^{-1}) due to the increased K_m (Table 1).

As mentioned above, CloH, SimH and Pcza361.18 showed only very low activity in the absence of MbtH. Therefore, no kinetic investigations could be performed with these proteins alone. However, when the cognate MbtH-like protein was added, kinetic data could be readily obtained. The K_m values for L-tyr ranged from 85 to 186 μM and the k_{cat} values from 0.50 to 1.91 s^{-1} (Table 1). For all investigated proteins, the dependency of the reaction velocity on L-tyrosine concentration is depicted in Fig. S3 (Supplemental Data).

Table 1: Kinetic parameters of tyrosine-adenylating enzymes in the presence of MbtH-like proteins. Unless indicated otherwise, proteins were expressed in an E. coli strain in which the mbtH-like gene ybdZ has been deleted. Tyrosine-adenylating enzymes and MbtH-like proteins were mixed in a molar ratio 1:1.2. The reaction velocities determined at different tyrosine concentrations and the statistical variation of the parameters are depicted in Fig. S3 (Supplemental Data).

	K_m for L-tyr [μM]	k_{cat} [s^{-1}]	k_{cat} K_m^{-1} [s^{-1} M^{-1}]
NovH alone	275	0.079	290
NovH alone, expressed in *ybdZ*$^+$ *E. coli*	1278	0.120	94
NovH + CloY	277	0.438	1580
CloH and CloY	186	0.497	2670
SimH and SimY	164	1.01	6180
Pcza361.18 and Orf1van	85	1.91	22200

Stoichiometry of the SimH/SimY and CloH/CloY complexes

Previous investigations (6-8) had suggested that adenylating enzymes and MbtH-like proteins form complexes, but the stoichiometry of these complexes had remained unclear (see Introduction). We now investigated the adenylating activity of SimH and CloH in the presence of different amounts of SimY and CloY, respectively. As depicted in Fig. 4, activity steadily increased with increasing amounts of MbtH-like protein until a molar ratio of 1:1 was reached. Addition of further amounts of MbtH-like protein did not lead to a further increase of activity, suggesting that the active complex contains both proteins in a molar ratio of 1:1.

We subsequently carried out an analytical gel chromatography of SimH alone and of a SimH/SimY mixture. Calibration of the column with reference proteins allowed to determine the apparent molecular weight of the eluted proteins. SimH (theoretical molecular mass 107.3 kDa) aggregated in aqueous solution, showing a dominant peak at approximately 650 kDa (Fig. 5A). In contrast, the mixture of SimH and SimY showed a dominant peak at 208 kDa. This is in reasonable agreement with the calculated molecular weight of a heterotetrameric $(SimH)_2(SimY)_2$ complex (229 kDa). We isolated this peak from the analytical column. SDS-PAGE analysis readily showed both SimH and SimY protein as components of the complex (Fig. 5B). Similar results were obtained for CloH and CloY (Fig. S4, Supplemental Data). However, CloH alone formed dimers in solution, and therefore in analytical gel chromatography the difference between CloH alone and the CloH/CloY complex was small.

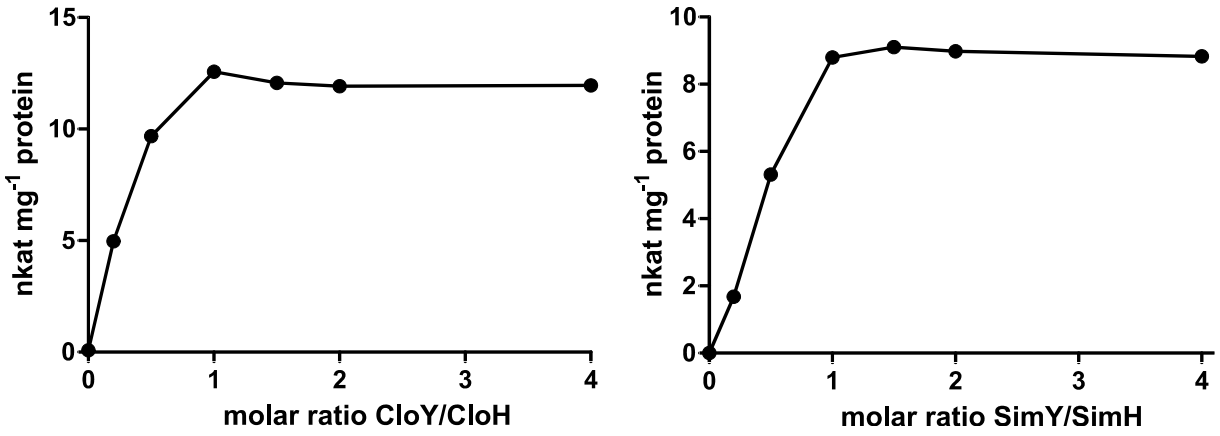


Figure 4: L-tyrosine-adenylating activity of CloH and SimH in the presence of different amounts of the MbtH-like proteins CloY and SimY.

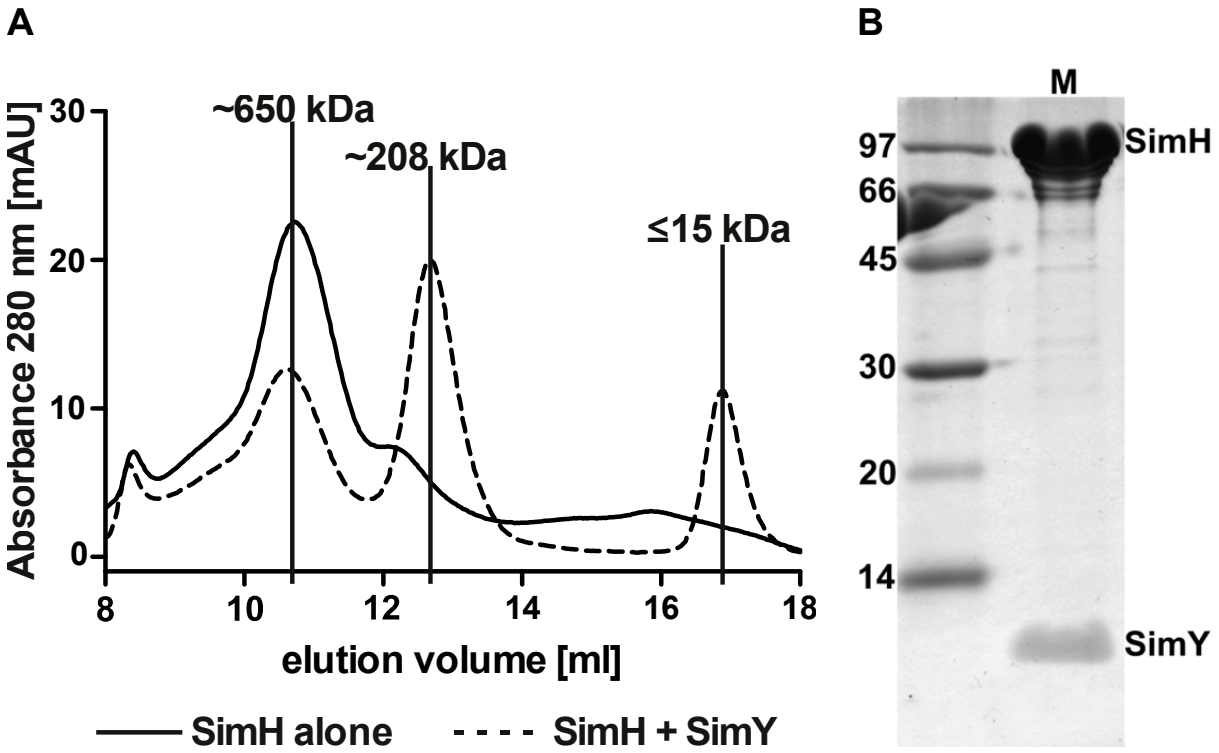


Figure 5: A) Molecular weight determination of the SimH/SimY complex by analytical gel chromatography. B) SDS-PAGE of the peak eluting at 208 kDa.

Coexpression of CloH and CloY

Felnagle *et al.* (6) had reported that the adenylating enzymes CmnO and VioO showed moderate activity upon addition of the MbtH-like proteins CmnN and VioN. However, much higher activity could be obtained when CmnO or VioO (as N-terminally his-tagged proteins) were simultaneously expressed with CmnN or VioN (as untagged proteins), followed by copurification of the proteins in form of the resulting complexes. From the supplemental data of the publication of Felnagle *et al.* (6), it can be estimated that the activities of the coexpressed CmnO/CmnN and VioO/VioN were 10-fold and 75-fold higher than the activities of complexes obtained by mixing of separately purified proteins. The reason for this observation is unknown.

We decided to coexpress N-terminally his-tagged CloH and untagged CloY using a pETDUET1 vector (Novagene). As expected, purification by Ni^{2+} affinity chromatography and by gel chromatography resulted in an active tyrosine-adenylating enzyme, indicating that CloH and CloY had formed a complex resulting in their copurification. However, in contrast to the results of Felnagle *et al.* (6), this complex was not more active, but 50 % less active than a mixture of separately purified CloH and CloY proteins. Addition of CloY to the complex increased the activity but again not to a value higher than observed for a

mixture of separately purified CloH and CloY. This indicates that some CloY may have been lost from the complex during purification, causing a loss of activity which could be restored by external addition of CloY.

Stimulation of tyrosine-adenylating enzymes by cognate and noncognate MbtH-like proteins

Each of the four tyrosine-adenylating enzymes NovH, CloH, SimH and Pcza361.18 was assayed with each of the MbtH-like proteins CloY, CdaX, SimY and Orf1van in a molar ratio of 1:1.2. As shown in Fig. 6, all enzymes were stimulated by all MbtH-like proteins. In each case, addition of CloY resulted in the highest activity. Therefore, SimH and Pcza361.18 did not show preference for their cognate MbtH-like proteins, *i.e.* SimY and Orf1van, respectively.

Generation of an L383M mutant of CloH

NovH showed tyrosine-activating activity in the absence of an MbtH-like protein. This is in contrast to CloH, which shows 83 % sequence identity to NovH. A sequence alignment (Fig. S1, Supplemental Data) shows no conspicuous differences between both proteins. No experimentally determined structure is available for NovH or CloH. We therefore modeled their structure after PheA (PDB ID: 1AMU). PheA is the phenylalanine-adenylating domain of the GrsA protein, which is part of the NRPS responsible for gramicidin S biosynthesis (23). PheA (556 aa) shows 37 and 39 % identity to the adenylation domains of NovH and CloH, respectively. A comparison of the structural models of NovH and CloH showed that all amino acids which are different between the two proteins are located distantly from the active center, with one single exception: in position 383, which is close to the active center, NovH contains a methionine and CloH a leucine residue. PheA contains a lysine residue in the corresponding position (K396). Modeling suggested that the size and nature of the residue in this position may influence the orientation of the side chains of a conserved tyrosine residue (Y397 in NovH/CloH, Y409 in PheA) and a neighboring glutamate residue (E316 in NovH/CloH and E327 in PheA). These residues assist in the binding of the Mg^{2+} ion in the active center. The positions of these residues were similar in the structure of PheA and the model of NovH, but different in CloH.

We carried out a site-directed mutagenesis of CloH, mutating the genuine L383 to M as found in NovH. The resulting mutant protein clearly

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

showed some adenylating activity in the absence of CloY (0.22 nKat mg^{-1}), amounting to 33 % of the activity of NovH. In the presence of CloY, the mutant protein showed an activity of 2.45 nKat mg^{-1} *i.e.* 36 % of the value determined for the genuine CloH in the presence of CloY. Therefore the L383M mutation was successful in generating an MbtH-independent activity, but it also reduced the optimal activity of the enzyme in the presence of CloY.

Discussion

The present study shows that the L-tyrosine-adenylating enzymes CloH and SimH of aminocoumarin antibiotic biosynthesis, and Pcza361.18 of vancomycin biosynthesis, require the presence of MbtH-like proteins for their catalytic activity. In the absence of MbtH-like proteins, their activity is lowered by 99.0-99.8 %. This is in accordance with our previous observations that inactivation of all *mbtH*-like genes in a clorobiocin producer strain lowered production of this aminocoumarin antibiotic by 99.3 % (2).

In contrast, the L-tyrosine-activating enzyme NovH of novobiocin biosynthesis showed significant activity also in the absence of any MbtH-like protein. Again, this is in accordance with the *in vivo* data which showed that even after inactivation of all *mbtH*-like genes, novobiocin was still produced, in approximately half of the amount formed in an *mbtH*$^+$ strain (2).

For optimal activity of the adenylating enzyme, the respective MbtH-like protein was required in a molar ratio of 1:1. The two proteins formed complexes with each other which coeluted during chromatographic purification. Analytical gel chromatography indicated that the complex contained two monomers of the adenylating enzyme. We therefore suggest that the adenylating enzymes form heterotetrameric complexes with the MbtH-like proteins, *i.e.* of the type $(SimH)_2(SimY)_2$.

We found that the tyrosine-adenylating enzymes SimH and CloH required the MbtH-like proteins SimY and CloY in a molar ratio of 1:1. In contrast, Felnagle *et al.* (6) found that the β-lysine adenylating enzymes CmnO and VioO require the MbtH-like proteins CmnN and VioN in 16- to 32-fold molar excess for optimal activity *in vitro*. In contrast to the enzymes investigated in our study, CmnO and VioO are *in vivo* part of large NRPS assembly lines, composed of several proteins. The absence of the other proteins of the NRPS may possibly affect their conformation, activity and stability *in*

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

vitro, influencing also their interaction with MbtH-like proteins. In contrast, NovH, CloH, SimH and Pcza361.18 most likely do not interact directly with NRPS assembly lines, but release their products into solution after enzymatic modification (Fig. 2A). We could obtain these tyrosine-adenylating enzymes with good activity by separate expression and subsequent mixing with the MbtH-like proteins *in vitro*. In contrast, Felnagle *et al.* (6) reported that separate expression of CmnO and VioO and subsequent mixing with the MbtH-like proteins CmnN and VioN gave low activity. Coexpression of CmnO with CmnN (or VioO with VioN) gave much higher activities, indicating some misfolding of the proteins during separate expression.

The strong requirement of adenylating enzymes for MbtH-like proteins raises the question whether amino acid residues of MbtH-like proteins are involved in catalysis in the active center of the adenylating enzyme. We modeled the structure of CloH and NovH after the experimentally determined structure of PheA, a phenylalanine adenylating domain from the biosynthetic gene cluster of gramicidin S (23). This gene cluster does not contain an *mbtH*-like gene, and purified PheA shows high activity without addition of an MbtH-like protein (24). The genome sequence of the gramicidin S producer strain is not available, but the closely related strain *Brevibacillus brevis* NBRC 100599 has been sequenced. It shows an NRPS gene cluster closely related to the gramicidin cluster, but no *mbtH*-like gene in the entire genome. PheA is therefore expected to be an MbtH-independent enzyme. Modeling of NovH and CloH showed that all amino acids expected to be in contact with the substrates ATP and L-tyrosine and with the cofactor Mg^{2+}, were identical in both proteins. It appears therefore unlikely that amino acid residues of CloY form direct contacts with the substrates in the active center of the CloH/CloY complex. Rather, binding of CloY may induce a conformational change in the structure of CloH which enhances activity. Some support for this hypothesis can be derived from our mutational experiment: exchange of Leu383 in CloH for Met (as found in NovH) clearly led to an activity of the enzyme in absence of MbtH-like proteins, suggesting that this amino acid exchange is one of the structural differences which are responsible for the MbtH-independent activity of NovH.

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

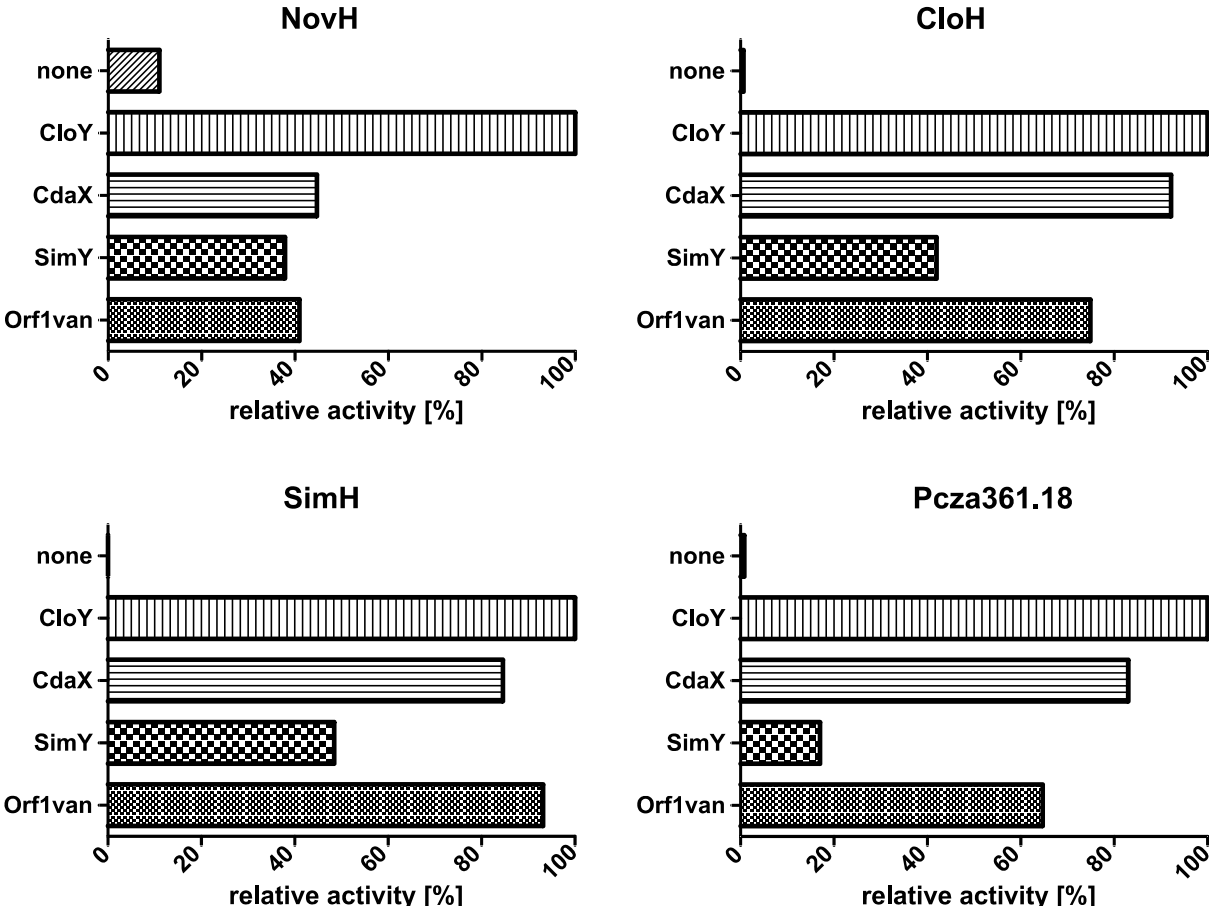


Figure 6: Activity of L-tyrosine adenylating enzymes in the presence of different MbtH-like proteins. Activity was determined using the PPi exchange assays.

M383 does not have direct contacts to the substrates but appears to influence the conformation of other residues in the active center. Conformational changes in adenylate-forming enzymes have been described previously, especially concerning a rotation of the PCP domain during thioester formation. The catalytic site is formed by the interface between the A and PCP domains, which are connected by a flexible hinge. It includes a highly conserved and functionally important loop similar to the P-loop (25) found in ATPases and GTPases. This loop wraps around the triphosphate of ATP. During the adenylate formation changes in the active center as well as a displacement of the P-loop of ~ 3.5 Å can be observed (26). The thioester formation includes a major structural change where the PCP domain rotates by ~ 140° burying the ATP-binding site (27). MbtH-like proteins presumably affect only the adenylate-forming part of the reaction, and the available data do not indicate their involvement in the rotation of the PCP domain.The present study, in accordance with previous *in vivo* (2,3) and *in vitro* studies (6-8), shows that adenylating enzymes have a

remarkable promiscuity for MbtH-like proteins from various pathways and organisms (Fig. 6). *E.g.*, the tyrosine-activating enzymes investigated in our study were efficiently activated by CdaX, the MbtH protein encoded in the biosynthetic gene cluster of the calcium-dependent antibiotic (CDA) from *Streptomyces coelicolor* A(3)2 (3). The peptide antibiotic CDA does not contain a tyrosyl residue or a residue derived from tyrosine (28). Therefore, the genuine function of CdaX is most likely related to the adenylation of another amino acid than tyrosine. For experiments in combinatorial biosynthesis involving genes for aminoacyl-adenylate forming enzymes, inclusion of an *mbtH*-like gene may be crucial, but there is flexibility regarding which *mbtH*-like gene is chosen.

The biosynthetic gene cluster of novobiocin does not contain an *mbtH*-like gene, and correspondingly NovH showed activity in the absence of an MbtH-like protein. Unexpectedly, however, the activity of NovH was still markedly stimulated by MbtH-like proteins such as CloY (Fig. 6). This finding prompted us to initiate a genome sequencing of the producer strain *Streptomyces spheroides* NCIMB 11891 (syn. *S. niveus*) and to search for other *mbtH* orthologs which may assist in novobiocin biosynthesis in this strain. Completion of the sequence is still in progress, but available data confirm that the novobiocin gene cluster and its immediate vicinity do not contain an *mbtH*-like gene. However, the genome contains at least two *mbtH* orthologs situated distantly from the novobiocin cluster. The stimulation of NovH by MbtH-like proteins shows that, even when a given gene cluster does not contain an *mbtH*-like gene, a stimulation of the biosynthesis by *mbtH*-like genes cannot be excluded. This may need to be considered in the design of combinatorial biosynthesis experiments.

More than 400 *mbtH*-like genes are currently found in the database. Drake *et al.* (5) solved the crystal structure of an MbtH-like protein from *Pseudomonas aeruginosa*. The protein displays a new protein fold and is shaped like a thin arrow head, with the point of the arrow formed by the C-terminal α-helix. Sequence comparison of 155 MbtH-like proteins showed that the conserved residues, including the three highly conserved tryptophans, all lie on one face of the protein, and the authors suggested that this face may interact with conserved components of NRPSs. Buchko *et al.* (4) determined by NMR the solution structure of another MbtH-like protein from *Mycobacterium*

Originally published in The *Journal of Biological Chemistry. (2011), 286,* Publication ahead of print

tuberculosis. The solution structure was similar to the aforementioned crystal structure except for the C-terminus which was highly disordered in solution, despite high sequence conservation of this region in the family of MbtH-like proteins. The authors pointed out that conserved but disordered regions of proteins are associated with binding to multiple partners, and suggested that binding via the disordered C-terminal region may explain the promiscuity of MbtH-like proteins for interaction with biosynthetic enzymes from different pathways.

We modeled the structure of our MbtH-like proteins after the published structures and tried to dock them to the tyrosine-adenylating enzymes using the Hex Protein Docking Server (http://hexserver.loria.fr). However, this resulted in several possible solutions. Therefore, the structure of these complexes remains speculative until a crystal structure can be determined.

Many biochemical studies have been published on amino acid-adenylating enzymes in secondary metabolism, especially in non-ribosomal peptide synthesis (29-31). In nearly all cases these enzymes have been expressed in *E. coli* strains containing an intact *ybdZ* gene. Our data suggest that YbdZ can copurify with, adenylating enzymes and that this complex formation can considerably affect the biochemical properties of the purified enzymes. We therefore suggest to use only *ΔybdZ* expression strains for investigations of amino acid-adenylating enzymes in future, and to carefully re-evaluate previous data for possible interference by YbdZ.

References

1. Strieker, M., Tanović, A., and Marahiel, M. A. (2010) *Curr. Opin. Struct. Biol.* **20**, 234-240
2. Wolpert, M., Gust, B., Kammerer, B., and Heide, L. (2007) *Microbiology* **153**, 1413-1423
3. Lautru, S., Oves-Costales, D., Pernodet, J. L., and Challis, G. L. (2007) *Microbiology* **153**, 1405-1412
4. Buchko, G. W., Kim, C. Y., Terwilliger, T. C., and Myler, P. J. (2010) *Tuberculosis (Edinb)* **90**, 245-251
5. Drake, E. J., Cao, J., Qu, J., Shah, M. B., Straubinger, R. M., and Gulick, A. M. (2007) *J. Biol. Chem.* **282**, 20425-20434
6. Felnagle, E. A., Barkei, J. J., Park, H., Podevels, A. M., McMahon, M. D., Drott, D. W., and Thomas, M. G. (2010) *Biochemistry* **49**, 8815-8817
7. Zhang, W., Heemstra, J. R., Jr., Walsh, C. T., and Imker, H. J. (2010a) *Biochemistry* **49**, 9946-9947
8. Imker, H. J., Krahn, D., Clerc, J., Kaiser, M., and Walsh, C. T. (2010) *Chem. Biol.* **17**, 1077-1083
9. Baltz, R. H. (2011) *J. Ind. Microbiol. Biotechnol.*, in press (DOI 10.1007/s10295-10011-11022-10298)
10. Heide, L. (2009c) *Biotechnol. Adv.* **27**, 1006-1014
11. Alt, S., Burkard, N., Kulik, A., Grond, S., and Heide, L. (2011) *Chem. Biol.* **18**, 304-313
12. Heide, L. (2009) *Nat. Prod. Rep.* **26**, 1241–1250
13. Chen, H., and Walsh, C. T. (2001) *Chem. Biol.* **8**, 301-312

14. Recktenwald, J., Shawky, R., Puk, O., Pfennig, F., Keller, U., Wohlleben, W., and Pelzer, S. (2002) *Microbiology* **148**, 1105-1118
15. Mulyani, S., Egel, E., Kittel, C., Turkanovic, S., Wohlleben, W., Süssmuth, R. D., and van Pée, K. H. (2010) *Chembiochem* **11**, 266-271
16. Steffensky, M., Li, S. M., and Heide, L. (2000) *J. Biol. Chem.* **275**, 21754-21760
17. Steffensky, M., Mühlenweg, A., Wang, Z. X., Li, S. M., and Heide, L. (2000a) *Antimicrob. Agents Chemother.* **44**, 1214-1222
18. Pojer, F., Li, S. M., and Heide, L. (2002) *Microbiology* **148**, 3901-3911
19. Galm, U., Schimana, J., Fiedler, H. P., Schmidt, J., Li, S. M., and Heide, L. (2002) *Arch. Microbiol.* **178**, 102-114
20. Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) *Biochemistry* **37**, 1585-1595
21. Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003) *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1541-1546
22. Dybkaer, R. (2002) *Clin. Chem.* **48**, 586-590
23. Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) *EMBO J.* **16**, 4174-4183
24. Stachelhaus, T., and Marahiel, M. A. (1995) *J. Biol. Chem.* **270**, 6163-6169
25. Yonus, H., Neumann, P., Zimmermann, S., May, J. J., Marahiel, M. A., and Stubbs, M. T. (2008) *J. Biol. Chem.* **283**, 32484-32491
26. Kochan, G., Pilka, E. S., von Delft, F., Oppermann, U., and Yue, W. W. (2009) *J. Mol. Biol.* **388**, 997-1008
27. Wu, R., Reger, A. S., Lu, X., Gulick, A. M., and Dunaway-Mariano, D. (2009) *Biochemistry* **48**, 4115-4125
28. Hojati, Z., Milne, C., Harvey, B., Gordon, L., Borg, M., Flett, F., Wilkinson, B., Sidebottom, P. J., Rudd, B. A., Hayes, M. A., Smith, C. P., and Micklefield, J. (2002) *Chem. Biol.* **9**, 1175-1187
29. Schmelz, S., and Naismith, J. H. (2009) *Curr. Opin. Struct. Biol.* **19**, 666-671
30. Koglin, A., and Walsh, C. T. (2009) *Nat. Prod. Rep.* **26**, 987-1000
31. Lautru, S., and Challis, G. L. (2004) *Microbiology* **150**, 1629-1636

Supplemental Data

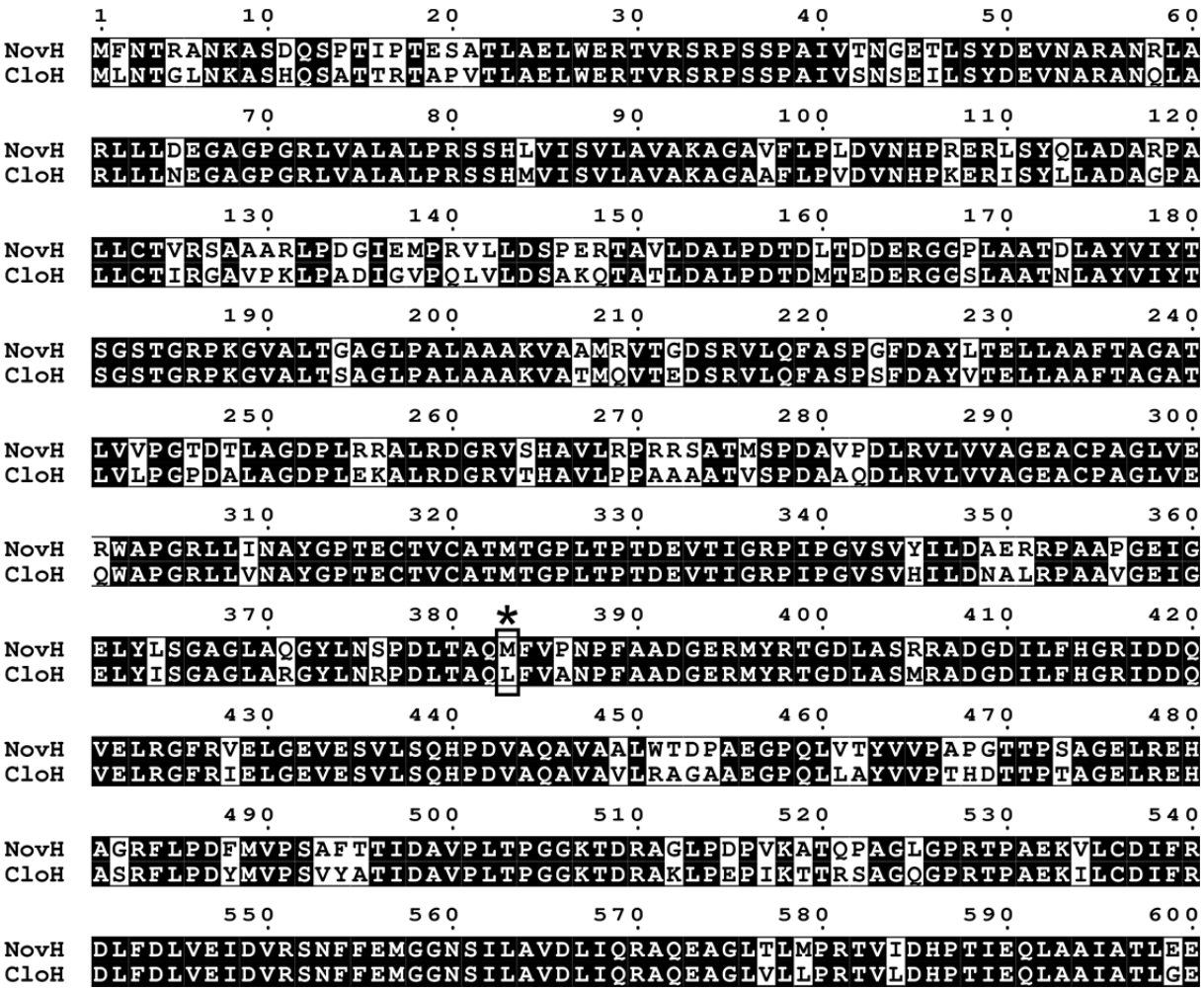


Fig. S1. Alignment of the amino acid sequences of the adenylate-forming enzymes NovH and CloH. The site of the mutation of CloH L383M is marked with an asterisk.

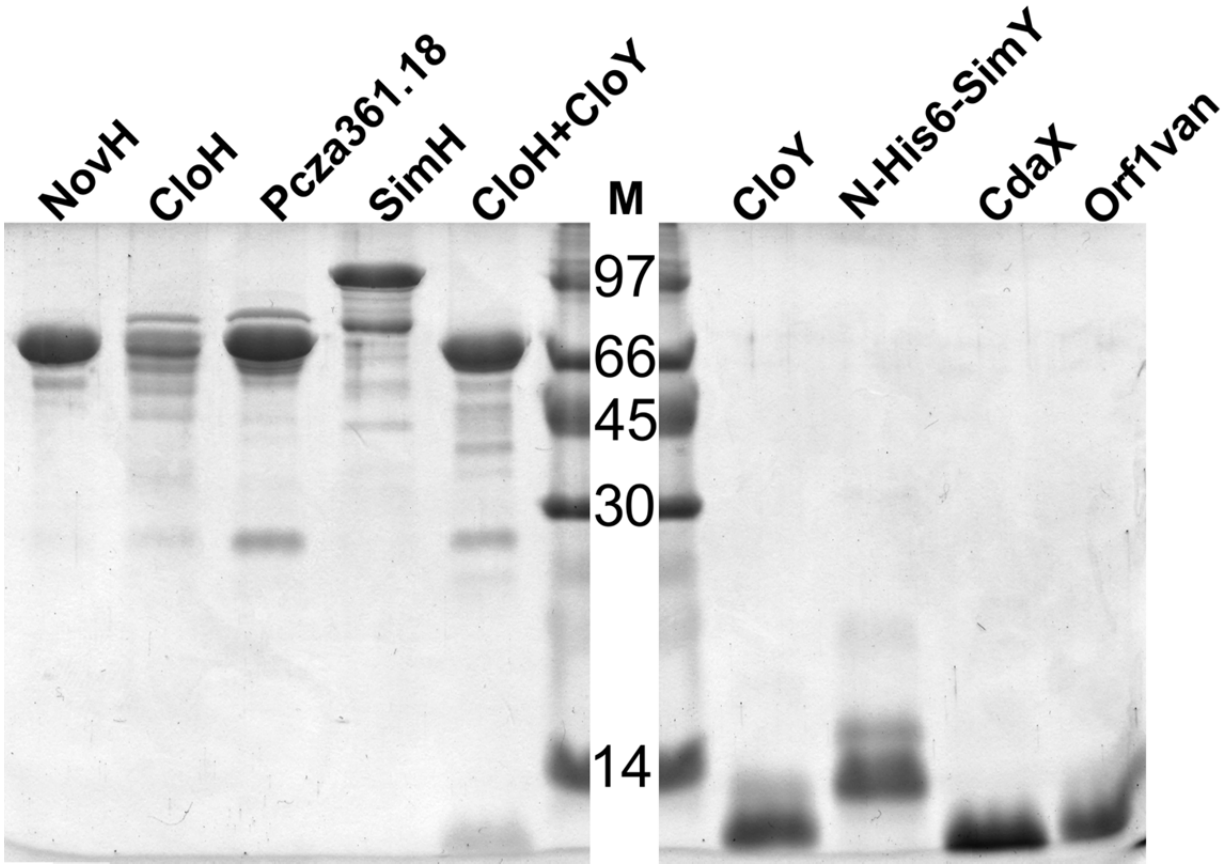


Fig. S2. SDS-PAGE of the purified tyrosine-adenylating enzymes CloH, SimH, NovH and Pcza361.18, and of the MbtH-like proteins CloY, SimY, CdaX and Orf1van. The calculated masses of the his-tagged tyrosine-adenylating enzymes are 65.7 for NovH, 65.2 for CloH, 63.2 for Pcza361.18 and 107.3 kDa for SimH. The calculated masses for the MbtH-like proteins are 8.3 kDa for CloY, 9.4 kDa for N-His_6-SimY, 8.2 kDa for CdaX, 8.1 for Orf1van,. The polyacrylamide gel was stained with Coomassie Brilliant Blue R-250.

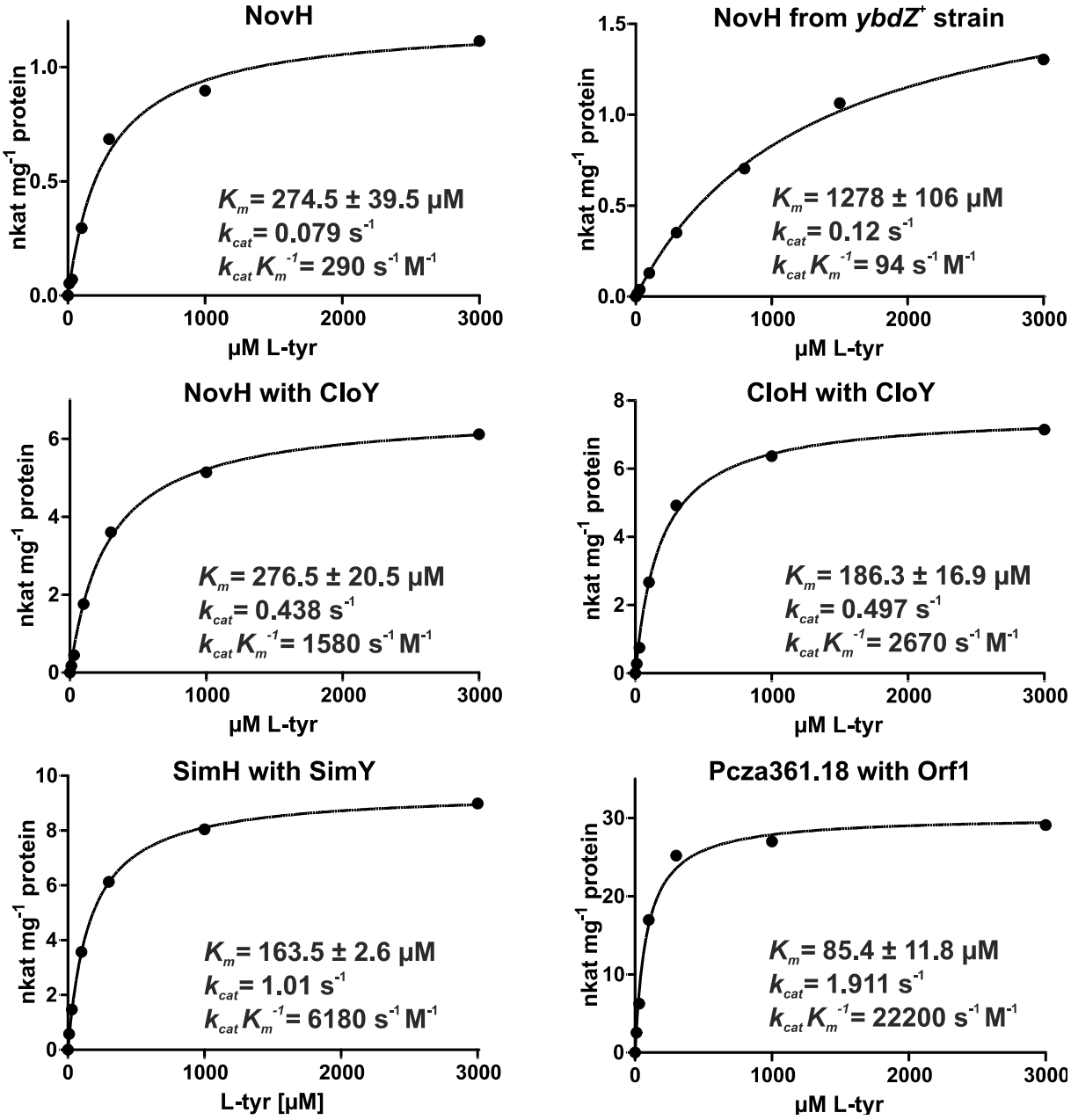


Fig. S3. Michaelis-Menten kinetics of the activation of L-tyrosine by CloH, SimH, NovH and Pcza361.18 in the absence and presence of MbtH-like proteins.

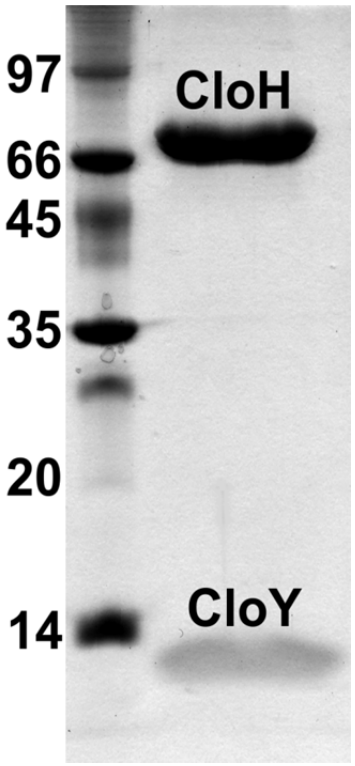


Fig. S4. SDS-PAGE of the CloH/CloY complex after analytical gel filtration.

Abbrevations

°C	degree Celsius
µ	micro
Å	Ångström (10^{-10} m)
aa	amino acids
aac(3)IV	apramycin resistance gene
Amp	ampicillin
ATP	adenosine triphosphate
bp	base pair
Cm	chloramphenicol
CoA	coenzyme A
Da	Dalton
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside 5´-triphosphate
DTT	1,4-dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylendiamine tetra-acetic acid
h	hour
HCl	hydrochloric acid
HCOOH	formic acid
His8	octahistidine
HPLC	high performance liquid chromatography
IPTG	isopropyl-*β*-thiogalactoside
k	kilo
KAc	potassium acetate
Kan	kanamycin
kb	kilo base pairs
k_{cat}	turnover rate
K_m	Michaelis-Menten constant
L	liter
lacZ	gene portion for α-complementation of *β*-galactosidase
LC	liquid chromatography
M	molar
m	milli
min	minute
MS	mass spectrometry
MW	molecular weight
m/z	mass-to-charge ratio
n	nano
NADP	nicotine amide adenine dinucleotide phosphate
NaOH	sodium hydroxide
OD_{600}	optical density at 600 nm

ORF	open reading frame
p	pico
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
RNA	ribonucleic acid
RNase	ribonuclease
RP	reverse phase
rpm	rounds per minute
s	second
S.	*Streptomyces*
SDS	sodium dodecyl sulfate
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
U	unit
V_{max}	maximal reaction velocity
×*g*	ground acceleration
X-gal	5-bromo-4-chloro-3-indolyl-*β*-D-galactopyranoside