DOI: 10.1002/cbic.200800068

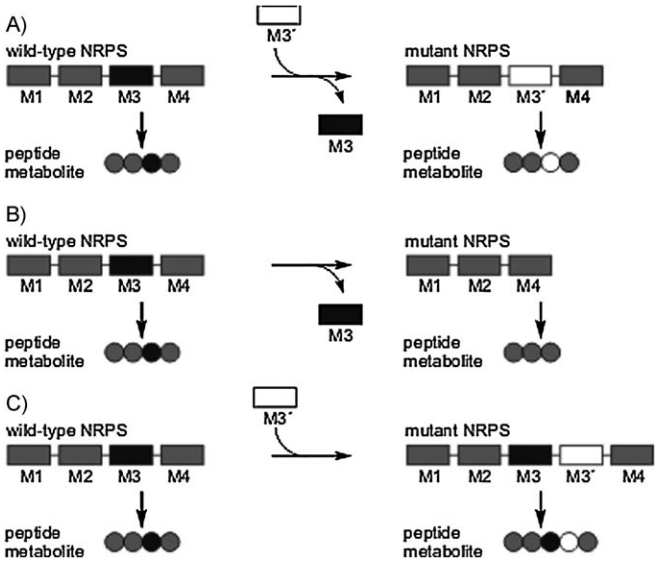
Module Extension of a Non-Ribosomal Peptide Synthetase of the Glycopeptide Antibiotic Balhimycin Produced by *Amycolatopsis balhimycina*

Diane Butz,[a] Timo Schmiederer,[a] Bianka Hadatsch,[b] Wolfgang Wohlleben,[b] Tilmann Weber,[b] and Roderich D. Süssmuth*[a]

Dedicated to Prof. Dr. Bernd Hamprecht

Non-ribosomal peptide synthetases (NRPSs) are among the fundamental multienzyme complexes in secondary metabolite biosynthesis.[1,2] A considerable number of antiinfectives and anticancer drugs, such as penicillins, vancomycin, and bleomycin, are synthesized by NRPSs of bacteria and fungi.[3] Non-ribosomal peptide synthesis is based on an arrangement of catalytic domains in modules, each of which is responsible for the activation and coupling of one amino acid. Adenylation (A) domains perform amino acid recognition, activation as aminoacyl-AMP, and transfer onto the phosphopantetheinyl residue of its C-terminal thiolation (T) domain. Subsequently, condensation (C) domains, connecting A-T didomains, catalyze the condensation of the T-domain-bound amino acids to the growing peptide chain. A thioesterase (TE) domain at the C terminus of the final module cleaves the peptide from the NRPS multienzyme complex. According to the colinearity rule, the sequence of the synthesized peptide is determined by the sequence of modules, particularly the amino acid specificity of the A domains, a concept that is basically valid for most peptides of non-ribosomal origin.[2] Other domains add further structural complexity to the peptide structures: for example, epimerization (E) domains for the synthesis of peptides containing D-amino acids. Recent investigations of complex non-ribosomal biosyntheses have increasingly revealed the involvement of tailoring enzymes as interaction partners of NRPSs during peptide assembly.[4,5]

A future biotechnological challenge based on molecular understanding of NRPSs lies in their engineering for the production of peptides containing nonproteinogenic amino acids or of peptides of increased structural complexity. Pioneering work in this field has been performed by Marahiel and co-workers with the NRPSs of surfactin, gramicidin, bacitracin, and tyrocidin in vitro and in vivo with *Bacillus* strains.[1,6] The in vivo approaches include site-directed mutagenesis of single amino acids in A domains, domain exchanges, and module exchanges on the genomic level, which accordingly led to altered NRPSs and thus to the synthesis of sequentially altered peptides of constant length (Scheme 1 A). NRPSs of actinomycetes have been altered at the example of CDA,[7] and in an extensive approach, Baltz's group has performed combinatorial biosynthesis with the NRPSs of the medically relevant lipopeptide antibiotic daptomycin.[8] In their work, various amino acid exchanges



Scheme 1. Schematic overview of three different types of NRPS module (M) manipulations, resulting in: A) an amino acid exchange, B) amino acid deletion, or C) amino acid insertion in the synthesized peptide metabolite.

were achieved by exchanges of one module (Scheme 1 A), of multiple modules, or even complete NRPSs. Following the logic of further manipulations of NRPS assembly lines leads to in-frame module deletion (Scheme 1 B), an experiment that has been carried out at the example of *Bacillus* surfactin synthetase. This module deletion resulted in the deletion of an amino acid and hence in a downsizing of the surfactin cycloheptapeptide to the corresponding cyclohexapeptide.[9] A third type of experiment, amino acid insertion by module extension (Scheme 1 C), complements the set of manipulations of NRPSs and may be regarded as a precursor to the arbitrary lining up of modules for the synthesis of any arbitrarily chosen amino acid sequence. To the best of our knowledge, this type of experiment has not previously been reported, although earlier bioinformatic studies on the nodularin and microcystin synthetases have given strong hints that module deletions and insertions are naturally occurring events.[10] Here we present results of work directed towards the achievement of this third type of

[a] *D. Butz, T. Schmiederer, Prof. Dr. R. D. Süssmuth*
Institut für Chemie, FG Organische Chemie, Technische Universität Berlin
Strasse des 17. Juni 124, 10623 Berlin (Germany)
Fax: (+49) 30-314-24783
E-mail: suessmuth@chem.tu-berlin.de

[b] *Dr. B. Hadatsch, Prof. Dr. W. Wohlleben, Dr. T. Weber*
Institut für Mikrobiologie/Biotechnologie, Universität Tübingen
Auf der Morgenstelle 28, 72076 Tübingen (Germany)

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author: experimental details.

NRPS manipulation. Furthermore, previous engineering attempts with NRPSs disregarded investigation into the relationship between peptide assembly and tailoring enzymes in post- or inter-NRPS biosynthesis steps.

A suitable system for exploring the NRPS manipulation discussed above is the biosynthesis gene cluster of the vancomycin-type glycopeptide antibiotic balhimycin (Figure 1). Characteristic structural features of glycopeptides are the nonproteinogenic aromatic amino acids that are conformationally fixed in their aromatic side chains.[11] The biosynthetic machinery for the assembly of balhimycin is coded and clustered in the genome of the producing strain *A. balhimycina*.[12,13] The seven modules that perform the non-ribosomal biosynthesis of the balhimycin heptapeptide backbone are organized into three NRPS genes: *bpsA/B/C*. BpsA and BpsB comprise three modules each, whereas BpsC consists of only one module (Figure 2A). Recent studies strongly suggest that the oxidative side chain cross-linkages (AB-, C-*O*-D-*O*-E-rings) of the heptapeptide backbone performed by P450-type monooxygenases OxyA/B/C occur on the NRPS thiolester-bound peptides.[4,5] The conformationally constrained peptide backbone carries further decorations—such as chlorination, *N*-terminal methylation, and glycosylation—performed by post- or inter-NRPS-acting enzymes. In conclusion, with regard to modification by various tailoring enzymes, the glycopeptide biosynthesis constitutes a complex example among non-ribosomal biosyntheses.

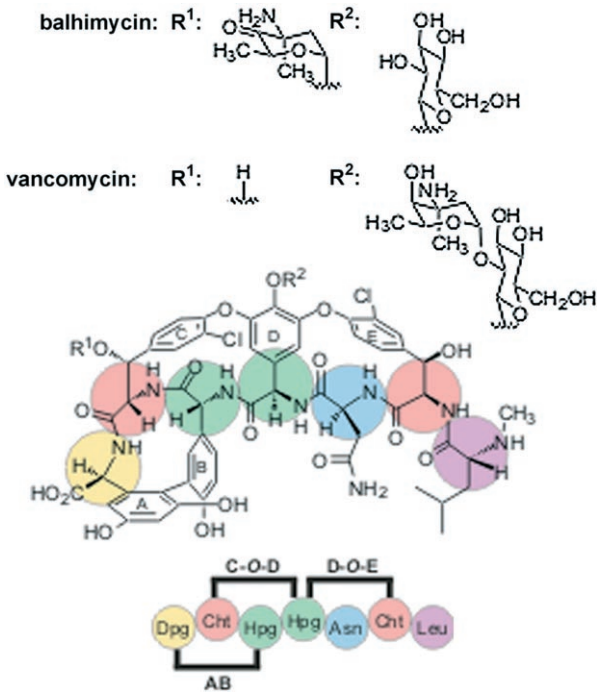


Figure 1. Chemical structures of the glycopeptide antibiotics balhimycin and vancomycin that share a common peptide aglycon but differ in their glycosylation patterns (color underlay symbolizes the different amino acids).

Considering the type of module for insertion into the balhimycin NRPS assembly line, we decided upon a D-hydroxyphenylglycine-coding (D-Hpg-coding) module, for several reasons. Firstly, the biosynthesis genes for the nonproteinogenic amino acid Hpg are an integral part of the balhimycin biosynthesis gene cluster, thus guaranteeing a sufficient supply of Hpg. Secondly, the substrate specificities of the domains have to be taken into account. Whereas the T, E, and TE domains have low to almost no substrate specificity, the A and C domains show moderate to high substrate specificities.[1] Consequently, the amino acid selected by its corresponding A domain must also match with the substrate specificity of the acceptor site of the upstream C domain. In contrast, the specificity of the donor site of a C domain for the upstream growing peptide chain is comparatively low.[14] Inserting a D-Hpg module between the two already existing D-Hpg modules 4 and 5 satisfactorily respects these regularities. Thirdly, Hpg is a central amino acid of the glycopeptide structure and is involved in reactions of intermediate and post-NRPS tailoring enzymes, so the insertion of a Hpg-coding module should probe the functionality of these enzymes.

The nature of the newly introduced module having been decided upon, the question of the origin and composition of the module had to be clarified. Finally, since NRPS manipulations are delicate with regard to loss of enzymatic function, it was of utmost importance to leave the sequential and conformational context of the NRPS as untouched as possible. We therefore decided to construct an Hpg module from balhimycin NRPS composed of the C-A didomain of Hpg-module 5 and the T-E didomain of Hpg-module 4 to yield module C5A5T4E4. In this case, module transitions to the neighboring modules 4 and 5 are kept natural, while the only nonnatural transition is located between the domains A5 and T4 of the artificial module C5A5T4E4 (Figure 2D).

The cloning strategy accomplished this plan through the design of a transformation plasmid pC5A5T4E4 containing the hybrid module C5A5T4E4 described above between flanking regions of the designated introduction site in the genome of *A. balhimycina*. With a double crossover strategy, C5A5T4E4 was integrated between modules 4 and 5 of *bpsB* to yield the strain *A. balhimycina* C5A5T4E4 (Figure 2B and C).

After verifying the correct integration of the C5A5T4E4 module in the *A. balhimycina* genome by means of PCR and Southern blot hybridization analyses (Figure 3A and Supporting Information) we demonstrated the module insertion on the proteinogenic level (Figure 3B and Supporting Information). Most importantly, enzymatic activity of the module-expanded NRPS had to be demonstrated on the metabolite level as final confirmation of a successful module expansion strategy. *A. balhimycina* C5A5T4E4 was therefore cultivated under standard conditions. Subsequently, the glycopeptide metabolite spectrum was analyzed by HPLC-ESI-MS and -MS/MS (Figures 3C and 4 and Supporting Information).

An octapeptide (**1**) and a heptapeptide (**2**) that both contained the new triple-Hpg motif were detected and their sequences were assigned from their MS/MS spectra. In addition, compounds **1** and **2** are both halogenated, but lack *N*-methylation of ^{1}Leu, glycosylation, and oxidative crosslinking in their side chains. Furthermore, the detection of metabolites **3** and **4**, with a mass difference of $\Delta m = 2$ amu, shows a certain sub-

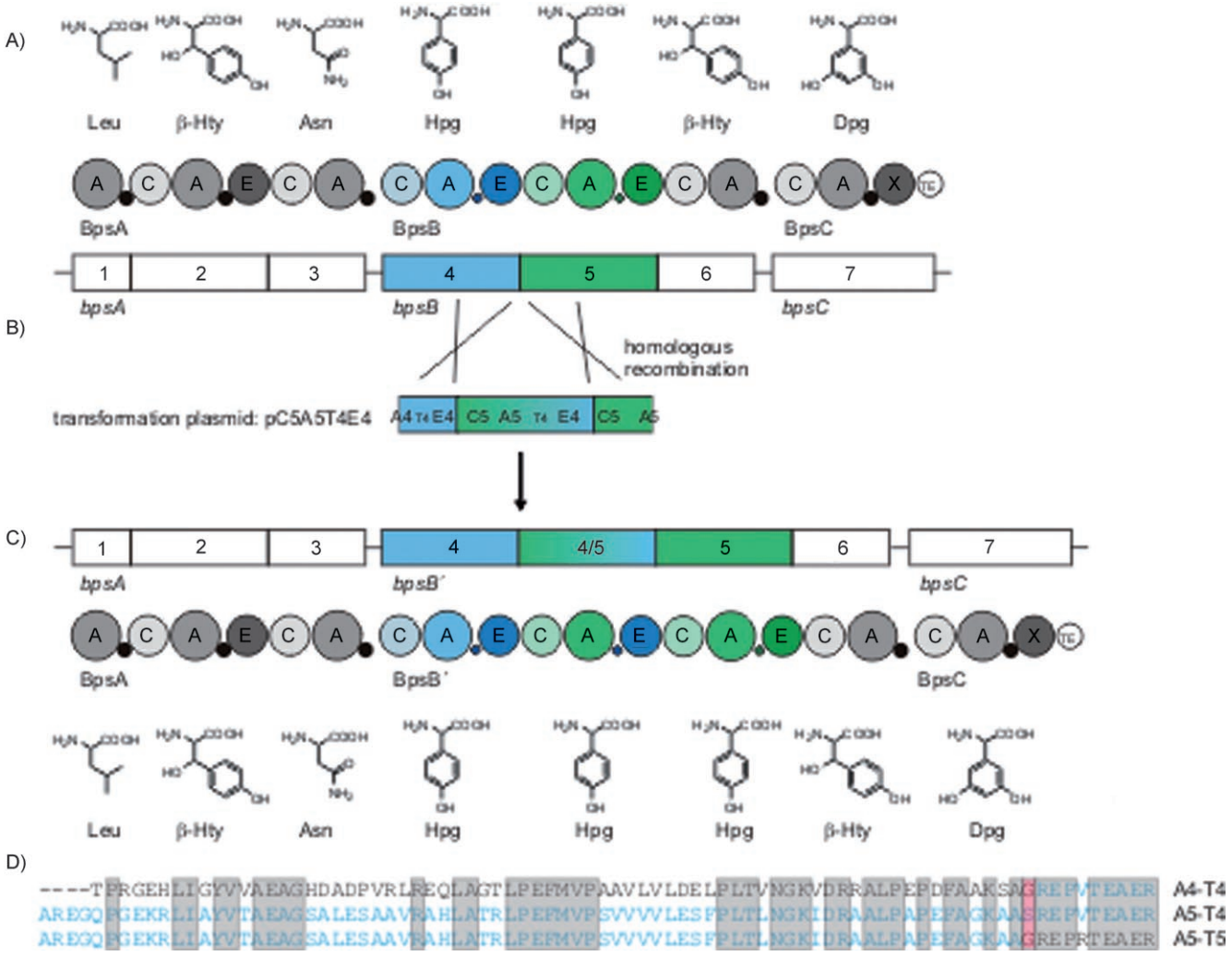


Figure 2. A) Non-ribosomal peptide synthetase enzymes and genes of balhimycin biosynthesis and the corresponding amino acid building blocks. B) Double crossover strategy for the insertion of one additional Hpg-coding module between modules 4 and 5 of *bpsB*. C) New arrangement of the modules in the *A. balhimycina* C5A5T4E4 mutant. D) Alignment of the wild-type linker regions A4-T4 and A5-T5 between the A and T domains of module 4 and the A and T domains of module 5 of BpsB with the artificial linker region A5-T4 constructed in the mutant strain *A. balhimycina* C5A5T4E4. The artificial linker region is based on fusion of the A5-T5 linker region and of the A4-T4 linker region.

strate tolerance of P450-type monooxygenases for this new peptide backbone. Interpretation of the MS/MS spectra of **3** and **4** clearly localizes the ring between amino acids ^{5}Hpg and $^{7}\beta$-hydroxytyrosine (β-Hty), which corresponds nicely to the first side chain crosslink (C-*O*-D) in balhimycin biosynthesis (Figure 1). In addition, we identified two compounds with molecular masses indicating bi- and tricyclic octapeptide structures, but their low concentrations impeded unambiguous assignment of the rings. In addition to compounds **1**–**4**, related derivatives were also detected (mass shift $\Delta m = 162$ amu). The MS/MS spectra of compounds **5** and **6**—derivatives of **1** and **2**—indicate unspecific glycosylation at β-^{2}Hty (Supporting Information). In addition to these hepta- and octapeptides we also identified the chlorinated, C-terminally truncated hexa- to dipeptides **7**–**11**, all of which lack side chain cyclizations, *N*-methylation, and glycosylation.

Concomitantly with these peptides, balhimycin was detected by mass spectrometry. This finding can be explained by a reversion of *bpsB*′ to *bpsB* by homologoues recombination. However, we did not observe this reversion to wild-type BpsB in the SDS-PAGE at the protein level (Figure 3 B). It is conceivable that wild-type BpsB is more active than BpsB′ and therefore that even traces of BpsB, below the detection limit in silver-stained SDS-PAGE gels, might effect a production level of balhimycin in the range of the mutant metabolites. Module skipping[15] as an alternative explanation seems less likely. The determination of the absolute yields of the peptide metabolites was impeded by the lack of appropriate standard substances, and so we estimated the relative yields on the basis of ionization yields in HPLC-ESI-MS. Hence, the linear heptapeptide **2** is the main product and the linear octapeptide **1** and the monocyclic heptapeptide **4** are detected at five times and three times lower concentrations, respectively. All the other metabolites were found at concentrations $< 10\%$ of the main metabolite **2**. In addition, the productivity of the mutant was estimated as ~10% of the wild-type *A. balhimycina*.[16] The production

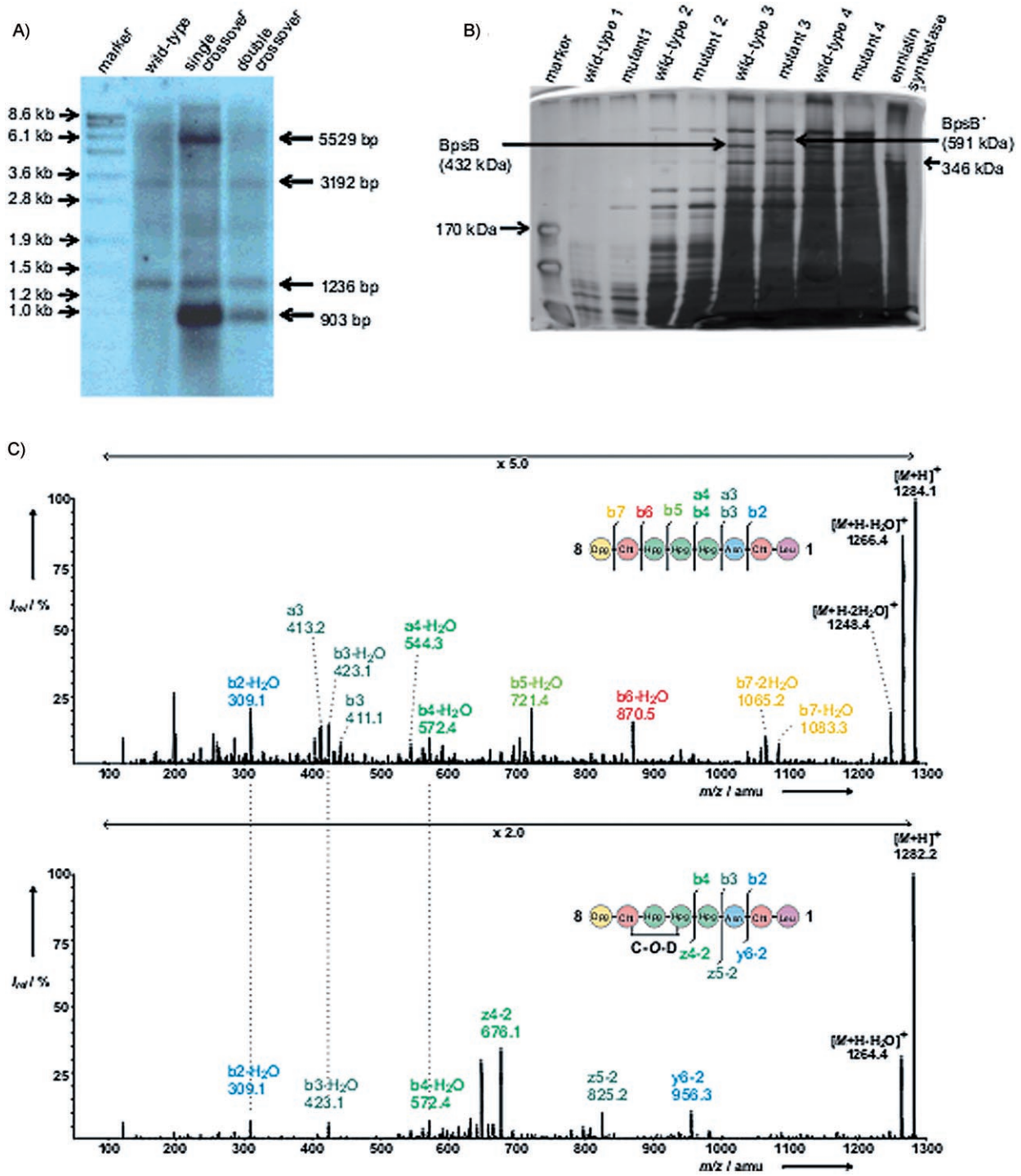


Figure 3. Characterization of the *A. balhimycina* C5A5T4E4 mutant. A) Southern blot analysis of the genome. Both the wild type and the two crossover states show *Pst*I-digested DNA fragments of 3192 bp and 1236 bp. The crossover states are characterized by the additional 903 bp fragment, while the single crossover state additionally shows a 5529 bp band (see Supporting Information). B) Comparative SDS-PAGE gel from Sephadex fractions of wild-type and mutant proteins. The increase in molecular mass of BpsB′ versus BpsB is shown. The expected molecular masses are 432 kDa for BpsB and 591 kDa for BpsB′ (enniatin synthetase 346 kDa). The identification of the gel bands was done by mass fingerprinting (tryptic digest and ESI-MS/MS). C) The metabolite profile of *A. balhimycina* C5A5T4E4 mutant was analyzed by HPLC-ESI-MS and -MS/MS. The MS/MS spectra of the linear and the monocyclic octapeptide **1** and **3**, verifying the new peptide backbone, are shown representatively.

of balhimycin in the mutant strain was found in same range as the linear heptapeptide **2**.

In summary, we have shown that module extension of NRPS assembly lines is a possible means to generate new peptide backbones. Furthermore, tailoring enzymes are widely able to tolerate the elongated peptide backbone as a substrate. The epimerization of the additional Hpg (Supporting Information), together with β-hydroxylation and chlorination reactions,[17,18]

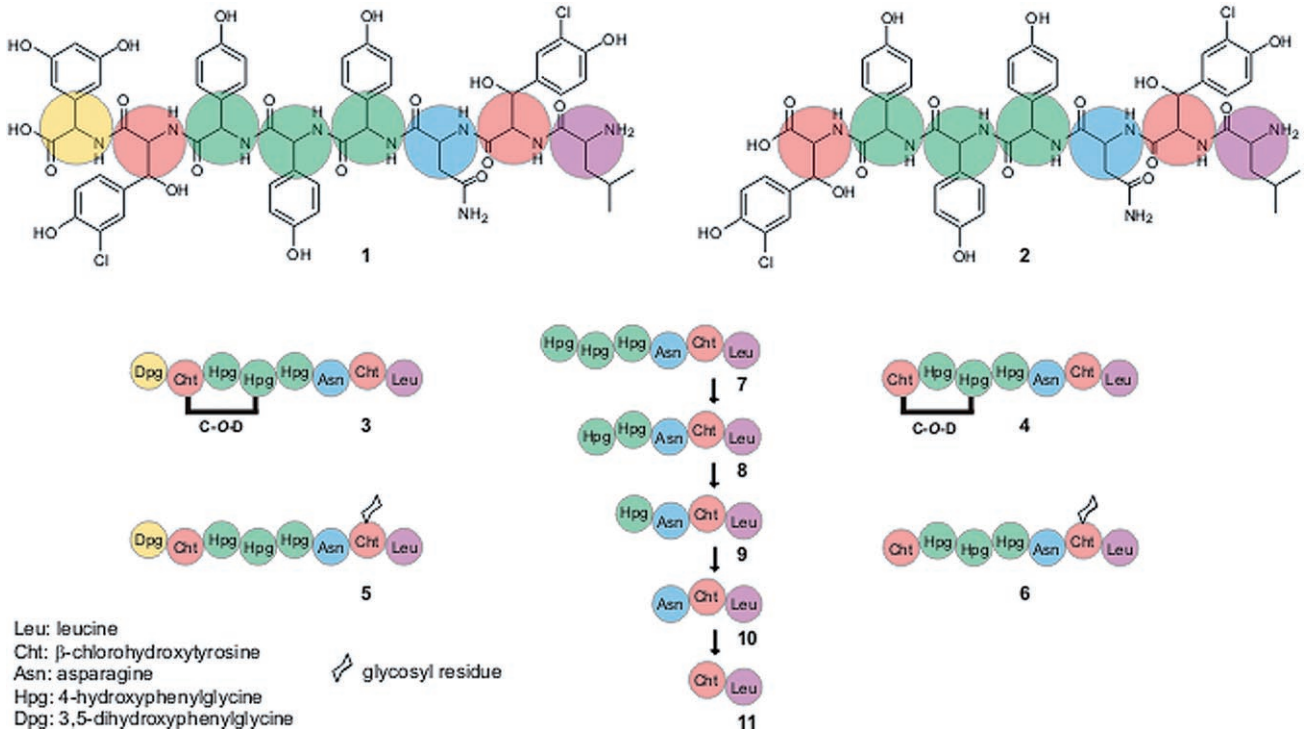


Figure 4. Overview of the peptide metabolite profile synthesized by *A. balhimycina* C5A5T4E4 (characterized by HPLC-ESI-MS/MS).

are constitutively found in all metabolites. These modifications occur before (β-hydroxylation) or during (epimerization and chlorination) NRPS assembly and are not affected by the module extension. In contrast, of three specific side-chain crosslinking oxygenases acting on NRPS-bound hexa- and heptapeptides[4,5] only OxyB is clearly functional and tolerates the new substrates generated by module extension. Because of high substrate specificity, the subsequent oxygenases OxyA and OxyC are practically incapable of substrate recognition.[5,11,19] Consequently, subsequent *N*-methylation and glycosylation as late- or post-NRPS biosynthetic steps are absent. This can be explained in terms of a significant structural difference between the linear and monocyclic compounds and the aglycon as the natural substrate. The observed glycosylation pattern of **5** and **6** is different from that in wild-type balhimycin and is therefore considered unspecific. Truncated peptide metabolites detected in minor amounts likely correspond to hydrolysis products of the NRPS assembly line as described previously.[5] As already suggested earlier, manipulation in NRPS assembly lines leads to a decrease in turnover, concomitantly promoting hydrolysis of biosynthetic intermediates.[7]

The feasibility of this module extension approach should encourage future attempts to line up modules for the directed synthesis of complex peptides of arbitrarily composed sequences. It is obvious that manipulations in NRPSs remain delicate, as subtle changes may lead to immediate loss of biosynthetic activity. However, the strategy of combining modules with highly similar sequential contexts might prove successful in this respect.

Acknowledgements

The authors acknowledge support from the Cluster of Excellence "Unifying Concepts in Catalysis" coordinated by the Technische Universität Berlin and funded by the Deutsche Forschungsgemeinschaft. This research was also supported by grants from the European Union (COMBIG-TOP, LSHG-CT-2003-503491, ActinoGEN, LSHM-CT-2004-005224) and by Schering AG (Germany). We thank Dr. Ullrich Keller for discussions on NRPS purification and Dr. Rainer Zocher for generously supplying enniatin synthetase. Furthermore we thank Dr. Andreas Reicke, Dr. Evi Stegmann, and Sigrid Stockert for the A. balhimycina *C5A5T4E4 mutant fermentation and Graeme Nicholson for the amino acid analysis and the ESI-FTICR-MS measurements.*

Keywords: combinatorial biosynthesis · genetic engineering · glycopeptides · module extension · non-ribosomal peptide synthetases

[1] S. A. Sieber, M. A. Marahiel, *Chem. Rev.* **2005**, *105*, 715–738.
[2] H. D. Mootz, D. Schwarzer, M. A. Marahiel, *ChemBioChem* **2002**, *3*, 490–504.
[3] a) F. von Nussbaum, M. Brands, B. Hinzen, S. Weigand, D. Häbich, *Angew. Chem.* **2006**, *118*, 5194–5254; *Angew. Chem. Int. Ed.* **2006**, *45*, 5072–5129; b) F. Wolter, S. Schoof, R. D. Süssmuth in *Topics in Current Chemistry: Glycopeptide Chemistry* (Ed.: V. Wittmann), Springer, Berlin, **2007**, 143–185.
[4] K. Zerbe, K. Woithe, D. B. Li, F. Vitali, L. Bigler, J. A. Robinson, *Angew. Chem.* **2004**, *116*, 6877–6881; *Angew. Chem. Int. Ed.* **2004**, *43*, 6709–6713.
[5] D. Bischoff, B. Bister, M. Bertazzo, V. Pfeifer, E. Stegmann, G. J. Nicholson, S. Keller, S. Pelzer, W. Wohlleben, R. D. Süssmuth, *ChemBioChem* **2005**, *6*, 267–272.

[6] U. Keller, F. Schauwecker, *Comb. Chem. High Throughput Screening* **2003**, *6*, 527–540.
[7] G. C. Uguru, C. Milne, M. Borg, F. Flett, C. P. Smith, J. Micklefield, *J. Am. Chem. Soc.* **2004**, *126*, 5032–5033.
[8] K. T. Nguyen, D. Ritz, J. Q. Gu, D. Alexander, M. Chu, V. Miao, P. Brian, R. H. Baltz, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17462–17467.
[9] H. D. Mootz, N. Kessler, U. Linne, K. Eppelmann, D. Schwarzer, M. A. Marahiel, *J. Am. Chem. Soc.* **2002**, *124*, 10980–10981.
[10] M. C. Moffitt, B. A. Neilan, *Appl. Environ. Microbiol.* **2004**, *70*, 6353–6362.
[11] D. Bischoff, S. Pelzer, B. Bister, G. J. Nicholson, S. Stockert, M. Schirle, W. Wohlleben, G. Jung, R. D. Süssmuth, *Angew. Chem.* **2001**, *113*, 4824–4827; *Angew. Chem. Int. Ed.* **2001**, *40*, 4688–4691.
[12] S. Pelzer, R. Süssmuth, D. Heckmann, J. Recktenwald, P. Huber, G. Jung, W. Wohlleben, *Antimicrob. Agents Chemother.* **1999**, *43*, 1565–1573.
[13] J. Recktenwald, R. Shawky, O. Puk, F. Pfennig, U. Keller, W. Wohlleben, S. Pelzer, *Microbiology* **2002**, *148*, 1105–1118.
[14] P. J. Belshaw, C. T. Walsh, T. Stachelhaus, *Science* **1999**, *284*, 486–489.
[15] S. C. Wenzel, B. Kunze, G. Höfle, B. Silakowski, M. Scharfe, H. Blöcker, R. Müller, *ChemBioChem* **2005**, *6*, 375–385.
[16] S. R. Nadkarni, M. V. Patel, S. Chatterjee, E. K. S. Vijayakumar, K. R. Desikan, J. Blumbach, B. N. Ganguli, *J. Antibiot.* **1994**, *47*, 334–341.
[17] O. Puk, P. Huber, D. Bischoff, J. Recktenwald, G. Jung, R. D. Süssmuth, K. H. van Pee, W. Wohlleben, S. Pelzer, *Chem. Biol.* **2002**, *9*, 225–235.
[18] O. Puk, D. Bischoff, C. Kittel, S. Pelzer, S. Weist, E. Stegmann, R. D. Süssmuth, W. Wohlleben, *J. Bacteriol.* **2004**, *186*, 6093–6100.
[19] R. D. Süssmuth, S. Pelzer, G. Nicholson, T. Walk, W. Wohlleben, G. Jung, *Angew. Chem.* **1999**, *111*, 2096–2099; *Angew. Chem. Int. Ed.* **1999**, *38*, 1976–1979.

Received: January 30, 2008
Published online on April 9, 2008

