J|A|C|S
COMMUNICATIONS

Published on Web 05/23/2006

Formylation Domain: An Essential Modifying Enzyme for the Nonribosomal Biosynthesis of Linear Gramicidin

Georg Schoenafinger, Nadine Schracke,† Uwe Linne, and Mohamed A. Marahiel*

Fachbereich Chemie/Biochemie, Philipps-Universität Marburg, Hans-Meerwein-Strasse, D-35032 Marburg, Germany

Received February 16, 2006; E-mail: marahiel@chemie.uni-marburg.de

A broad variety of pharmacologically important compounds is produced by nonribosomal peptide synthetases (NRPSs).[1] These enzymes consist of modules, each of which is responsible for the incorporation of a distinct building block into the peptide product during the often linear assembly line-like synthesis.[2] Modules can further be subdivided into catalytic core domains responsible for activation, condensation, covalent binding, and modification of each single building block.[3] Linear gramicidin is a 15 mer peptide antibiotic with several known isoforms (Supporting Information) produced by *Bacillus brevis* ATCC 8185 via nonribosomal peptide synthesis.[4−6] The bioactive form consists of two β-helical monomers which form an ion channel that can span the bacterial cell membrane due to its hydrophobicity, thus leading to cell death. Both peptide monomers are arranged in a linear head-to-head manner with the N-termini facing each other at the dimer interface.[5,7]

Characteristic for all linear gramicidins is the N-formylation of the N-terminal valine or isoleucine residues by a so far unknown mechanism.[8] Although N-formylation is a common mechanism for initiation of ribosomal protein biosynthesis in prokaryotes (for detailed review, see Laursen et al.[9]), it is exceptional in nonribosomal peptide synthesis.

Besides linear gramicidin, a second system, the anabaenopeptilides from *Anabaena* strain 90, is known to produce N-terminally formylated NRPS products.[10] Sequence analyses of the gene clusters responsible for the production of gramicidin and anabaenopeptilides (*lgr*A-E and *adp*A-F) revealed the presence of a putative N-terminal formylation domain in both cases: F (*adp*A) and F (*lgr*A).[4,10]

The aim of this study was to investigate the catalytic function of the putative formylation domain and the mode of action of N-formylation in linear gramicidin biosynthesis, testing its interaction with the natural neighboring domains and the timing of the initiation reaction of the NRPS assembly line.

The adenylation (A) domain of LgrA1 has been described earlier to activate valine (K_M = 0.84 mM), with minor side specificities toward leucine and isoleucine (K_M = 2.4 mM).[4] As formyl−valine was not accepted as a substrate by the A domain (Supporting Information), we hypothesized that valine bound to the 4′-ppant cofactor of the adjacent peptidyl carrier protein (PCP) is the substrate for the putative F domain.

We performed a formylation assay with the first module LgrA1 (F-A_1-PCP) as described in Supporting Information (Figure 1a−d), and subsequent analysis confirmed the formation of formyl−valine by both mass and retention time compared to an N_α-formyl−valine standard. Valine, however, was not detected—indicative of a 100% yield in this enzymatic modification reaction (Figure 1). When ATP was omitted, no enzyme-bound product could be detected, and also no formation of formyl−valine was observed from the solution in the absence of either F-A_1-PCP or Sfp (data not shown).

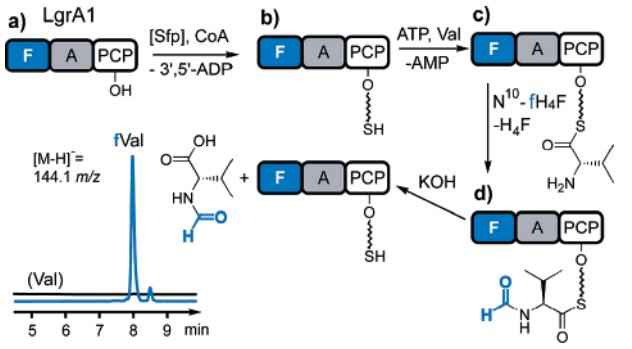

Figure 1. (a) The recombinant apo-enzyme F-A_1-PCP (LgrA1) is converted to its corresponding active holo-form by transfer of the 4′-ppant moiety of CoA, catalyzed by the 4′-ppant transferase Sfp. (b) Valine is ATP-dependently activated by the A domain and covalently linked to the thiol group of the holo-PCP. (c) Formylation of enzyme-bound valine is triggered by addition of the formyl donor N10-fH4F. (d) Formyl−valine is cleaved off by basic hydrolysis prior to detection.

In further experiments, F-A_1-PCP was analogously assayed with leucine and isoleucine, and both formyl−leucine and formyl−isoleucine could be detected with overall yields of a mere 5% (Supporting Information), and this indicates an F domain specificity toward Val, as is reflected by the natural product mixture.

Sequence comparisons of the N-terminal part of the F domain with several formyl−methionyl transferases (FMTs) suggested the presence of a binding motif typical for N^{10}-formyltetrahydrofolate (fH_4F).[4] Nevertheless, the sterically similar but more stable and thus commercially available cofactor N^5-fH_4F was also accepted by the F domain and was, therefore, used in analogous formyltransferase assays. Both cofactors led to the formation of formylated products, with an 18-fold higher formyl transfer ratio for the N^{10}-fH_4F cofactor, which was determined by time-dependent formylation assays.

To investigate the influence of formyl−valine on the initiation of linear gramicidin biosynthesis, we performed elongation assays using the truncated dimodular system of LgrA F-A_1-PCP-C-A_2-PCP since the complete synthetase LgrA (with the C-terminal epimerization domain) was insoluble. As the second A domain activates glycine only,[4] we performed the assay with valine and glycine as substrate amino acids (Figure 2a−e). Subsequent HPLC-MS analysis revealed the formation of the expected dipeptide formyl−valyl−glycine (Figure 2f) only in the presence of the formyl donor. Omission of the formyl donor led to a complete loss of dipeptide formation (Figure 2f and g).

Similar results were obtained with either leucine or isoleucine as substrate amino acids for F-A_1-PCP, even though we detected a dramatic drop in formylation activity, which gives rise to the assumption that the F domain exhibits a higher specificity toward the valyl-4′-ppant. This is also supported by the fact that formyl−valine is found predominantly at position 1 in the gramicidins.[8] The fact that free valine or valyl-AMP is not a substrate for the F

† Present address: Biozentrum der Universität Basel, Klingelbergstrasse 50-70, CH-4056 Basel, Switzerland.

10.1021/ja0611240 CCC: $33.50

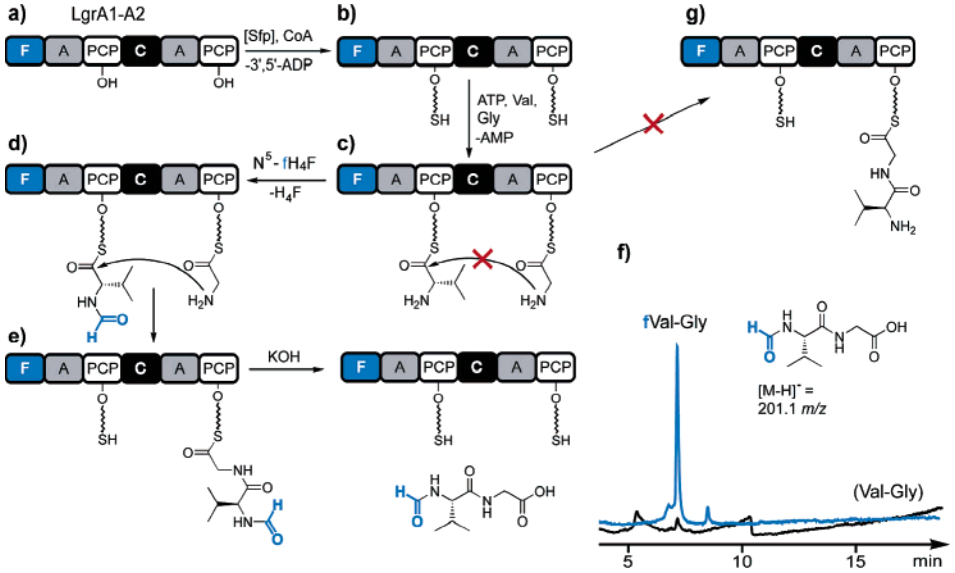

Figure 2. Formation of the dipeptide formyl−Val−Gly catalyzed by F-A_1-PCP-C-A_2-PCP. (a) Conversion of apo-enzymes to their corresponding holo-forms. (b) Valine and glycine are activated by the A domains and transferred onto the prosthetic group of the PCP domains. (c) Formylation of enzyme-bound valine by the F domain using N^5-fH_4F. (d) Only when the formylation of valine had occurred in the previous step was the elongation reaction catalyzed by the condensation (C) domain, generating the enzyme-bound dipeptide fVal−Gly. (e) Product is cleaved off by hydrolysis prior to detection. (f) HPLC-MS analysis of the assay reveals the formylated dipeptide fVal−Gly (blue), whereas no condensation product was found (black) when the cofactor was omitted (c, g).

domain points to the need and function of the 20 Å 4′-ppant moiety onto which the substrate is covalently bound as a thioester. The F domain is quite small (24 kDa) compared to FMTs due to the fact that the whole C-terminal subdomain of FMTs is missing. This subdomain directs the tRNA in such a way that the 3′-OH terminus with the methionine attached is positioned in the active site pocket located in the N-terminal subdomain.[12,13] As the F domain only accepts the enzyme-bound substrates, we propose that the 4′-ppant moiety takes over the part of positioning the substrate. This mode of action is consistent with the general idea of the positioning role of the 4′-ppant moiety. The presence of the formyl donor was again strictly required for the formation of the dipeptides formyl−isoleucyl−glycine and formyl−leucyl−glycine, respectively.

Our experimental data presented here are strong evidence that the F domain is essential for the initiation of linear gramicidin biosynthesis. The fact that the corresponding A domain does not accept formyl−valine emphasizes the importance of the F domain within the assembly line synthesis performed by the NRPS. In addition, in vitro elongation assays in the presence of the first two modules F-A_1-PCP-C-A_2-PCP indicated likewise a substrate specificity of the C domain as no unformylated dipeptide product could be detected in the appropriate controls. We therefore believe that the C domain has a high specificity for formyl-valyl-4′-ppant as donor substrate, whereas the analogously presented valyl-4′-ppant is rejected, therefore inhibiting elongation. This is in accordance with the general view of a tight selection mechanism of C domains.[14] This in vitro result is furthermore in accordance with the fact that only formylated gramicidins are found in vivo.[8] In contrast, when isoleucine and leucine were used as substrates in the elongation assays, we detected only minor traces of formylated dipeptides and no unformylated ones. This apparent low turnover of isoleucine and leucine reflects the low occurrence of isoleucine in the gramicidins in vivo with respect to the incorporation of valine;[8] leucine at position 1 has not been described at all.

In summary, we present here the first biochemical study of a novel type of modifying domains in NRPS, the F domain. We have also demonstrated its crucial role for the initiation of the biosynthesis of linear gramicidin. Both the small size and the substrate tolerance of this domain make it an ideal candidate for future investigations involving the construction of hybrid NRPSs aiming at the production of linear N-terminally protected peptides in vitro and in vivo.

Acknowledgment. We thank Holger Schuhmann for the construction of pBAD202-*fatcat*, Dr. Martin Hahn and Antje Schäfer for excellent technical assistance. Furthermore, we thank Dr. Bärbel Buchenau and Prof. Dr. R. Thauer for the generous gift of pET24b+-*folD*. N.S. and G.S. wish to acknowledge their fellowships from the Studienstiftung des deutschen Volkes. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

Supporting Information Available: NRPS cluster responsible for the biosynthesis of linear gramicidin, chemoenzymatic synthesis of N^{10}-fH_4F, radioactive ATP/PP_i-exchange assays, enzyme preparation, assay and detection methods, and LCMS data on formylation assays with F-A_1-PCP. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Walsh, C. T. *Science* **2004**, *303*, 1805−10.
(2) Mootz, H. D.; Schwarzer, D.; Marahiel, M. A. *Chembiochem* **2002**, *3*, 490−504.
(3) Finking, R.; Marahiel, M. A. *Annu. Rev. Microbiol.* **2004**, *58*, 453−88.
(4) Kessler, N.; Schuhmann, H.; Morneweg, S.; Linne, U.; Marahiel, M. A. *J. Biol. Chem.* **2004**, *279*, 7413−19.
(5) Weinstein, S.; Wallace, B. A.; Morrow, J. S.; Veatch, W. R. *J. Mol. Biol.* **1980**, *143*, 1−19.
(6) Hotchkiss, R. D.; Dubois, R. J. *J. Biol. Chem.* **1940**, *132*, 791−2.
(7) Wallace, B. A. *BioEssays* **2000**, *22*, 227−34.
(8) Sarges, R.; Witkop, B. *J. Am. Chem. Soc.* **1965**, *87*, 2011−20.
(9) Laursen, B. S.; Sorensen, H. P.; Mortensen, K. K.; Sperling-Petersen, H. U. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 101−23.
(10) Rouhiainen, L.; Paulin, L.; Suomalainen, S.; Hyytiainen, H.; Buikema, W.; Haselkorn, R.; Sivonen, K. *Mol. Microbiol.* **2000**, *37*, 156−67.
(11) Reuter, K.; Mofid, M. R.; Marahiel, M. A.; Ficner, R. *EMBO J.* **1999**, *18*, 6823−31.
(12) Crepin, T.; Schmitt, E.; Blanquet, S.; Mechulam, Y. *Biochemistry* **2002**, *41*, 13003−11.
(13) Crepin, T.; Schmitt, E.; Blanquet, S.; Mechulam, Y. *Biochemistry* **2004**, *43*, 2635−44.
(14) Clugston, S. L.; Sieber, S. A.; Marahiel, M. A.; Walsh, C. T. *Biochemistry* **2003**, *42*, 12095−104.

JA0611240