# Type II Thioesterase Restores Activity of a NRPS Module Stalled with an Aminoacyl-S-enzyme that Cannot Be Elongated

Ellen Yeh,[a] Rahul M. Kohli,[a] Steven D. Bruner,[b] and Christopher T. Walsh*[a]

Nonribosomal peptide synthetases (NRPSs) carry out the biosynthesis of numerous peptide natural products, including many with important clinical applications. The NRPS, organized into a series of modules, is an efficient, high-fidelity assembly line for the production of a particular peptide. Each module consists of domains, whose activities contribute to the accuracy of these assembly-line systems. The activation (A) domain uses ATP to selectively load an amino acid onto the module through formation of a thioester bond to the pantetheine arm of the thiolation (T) domain. Peptide-bond formation, catalyzed by the condensation (C) domain, is stringent for both side-chain identity and stereochemistry.[1–4] The C domain accepts an aminoacyl- or peptidylthioester from the preceding module for nucleophilic addition by the amine of the loaded amino acid; this generates the elongated peptide attached to the downstream module. The peptide product is synthesized one amino acid at a time until it reaches the final module. There, the fully synthesized chain is released by a type I thioesterase (TEI), the terminal domain of the NRPS assembly.

Despite the high fidelity of this process, an error in any step of the assembly-line synthesis severely impacts the efficiency of the system and creates a bottleneck that results in a build-up of unprocessed intermediates. For example, an error by the A domain, which can load amino acids other than that normally accepted by the C domain,[5–7] would prevent peptide-bond formation. The loaded module would be blocked until the incorrect amino acid was hydrolyzed (Figure 1). A type II thioesterase (TEII), whose gene is associated with the gene cluster of many NRPSs and related polyketide synthases (PKSs), improves the efficiency of product formation in these systems and has been proposed to edit modules through hydrolysis of acyl groups.[8–13] In the surfactin NRPS, TEII was shown to regenerate misacylated modules resulting from priming of the apo-module with acyl-CoA groups.[9] In this study we provide evidence to expand the editing function of TEIIs to include restoring the activity of modules stalled by loaded amino acids that cannot be processed.

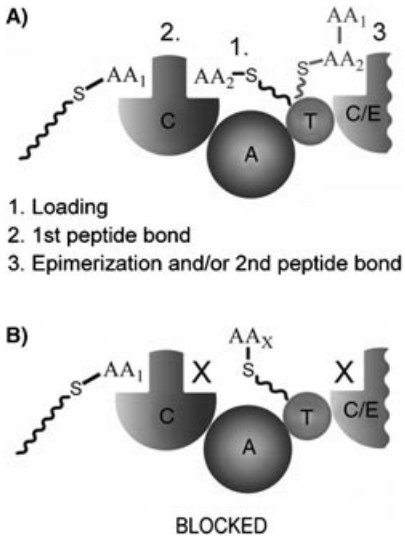


***Figure 1.*** *Model for the processing of correct versus incorrect amino acids by NRPS modules. A) Normal processing of an amino acid by a module, including initial loading of the amino acid followed by the later steps of peptide-bond formation and/or epimerization. B) Blocking of a module following loading of an incorrect amino acid ($AA_X$). "X" indicates the inability of C or E domains to recognize the loaded amino acid.*

*N*-acetylcysteamine (SNAC) thioesters have been used previously to assay NRPS domain activities.[2,13–15] Hydrolysis of SNAC substrates was used here to explore the specificity of the TEII from the tyrocidine biosynthetic operon, TycF.[16] TycF accepted a broad variety of aminoacyl-SNACs of different side-chain identity and stereochemistry with a 20-fold $k_{cat}/K_m$ range between the most- and least-active substrate (Table 1). A series of peptidyl-SNACs derived from the tyrocidine sequence was

***Table 1.*** *TycF hydrolysis of SNAC substrates. A spectrophotometric assay used dithionitrobenzoic acid (DTNB) to monitor hydrolysis of SNAC substrates (0.1–5.0 mM).*[a]

| SNAC substrate | $k_{cat}/K_m$ [$mM^{-1}min^{-1}$] |
|---|---|
| Ac- | 0.13 |
| Ala- | 0.11 |
| Ac-D-Ala- | 0.06 |
| Ac-D-Leu- | 0.30 |
| D-Phe- | 0.41 |
| Phe- | 0.32 |
| Ac-Phe | 0.43 |
| Leu- | 0.24 |
| Ac-Leu- | 1.40 |
| Orn-Leu- | n.d.[b] |
| Ac-Orn-Leu- | –[c] |
| Val-Orn-Leu- | –[c] |
| Tyr-Val-Orn-Leu- | –[c] |
| Gln-Tyr-Val-Orn-Leu- | –[c] |

[a] Background hydrolysis, 20–50% of overall signal for aminoacyl-SNACs and $<2\%$ for all others, was subtracted. Nonsaturating kinetics were observed for all substrates tested. [b] Not determined due to high background hydrolysis; [c] Rates below detection limit of the assay.

[a] *E. Yeh, Dr. R. M. Kohli, Prof. C. T. Walsh*
*Department of Biological Chemistry and Molecular Pharmacology*
*Harvard Medical School*
*240 Longwood Avenue, Boston, MA 02115 (USA)*
*Fax: (+1) 617-432-0438*
*E-mail: christopher_walsh@hms.harvard.edu*

[b] *Prof. S. D. Bruner*
*Department of Chemistry, Boston College*
*Chestnut Hill, MA 02467 (USA)*

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

  DOI: 10.1002/cbic.200400077 

constructed, in which the C-terminal Leu was maintained but successively lengthened by one amino acid (Table 1, shaded). None of these peptidyl-SNACs were substrates for TycF. This result supports a previous report that TEII release of a tripeptide formed by a three-module system was very inefficient.[9] Therefore, editing by TEII could occur at the level of the aminoacyl intermediate after loading but prior to peptide-bond formation. Indeed, release of inefficiently processed peptidyl intermediates was found to be independent of TEII and instead might be catalyzed by the C domain.[17] Notably, acetyl-SNAC was accepted by TycF with a $k_{cat}/K_m$ of 0.13 $mM^{-1}min^{-1}$ compared to 0.06–1.4 $mM^{-1}min^{-1}$ for aminoacyl-SNACs. Based on these results, TEII hydrolysis of aminoacyl groups appears comparable to that of acetyl groups, the substrate demonstrated in the mispriming model.[9]

Hydrolysis of amino acids loaded on NRPS modules by TycF and SrfA-D, the TEII from surfactin biosynthetic operon, was demonstrated on the first module of gramicidin synthetase, PheATE. This module introduces D-Phe as the first amino acid in gramicidin and has high sequence identity to the first module of tyrocidine synthetase. In addition to A and T domains, an E domain acts to convert the loaded amino acid between L- and D-stereoisomers.[17,18] Following loading with optically pure L-Phe, D-Phe was generated on the module by the action of the E domain. Release of D-Phe from the module by TEII was monitored by its accumulation in solution (Figure 2). Similarly, rates for TEII cleavage of L-Phe from the module were also determined following loading with optically pure D-Phe. Both initial loading and epimerization of Phe were fast relative to TEII-catalyzed hydrolysis.[18] The $k_{cat}/K_m$ for D-Phe-S-enzyme

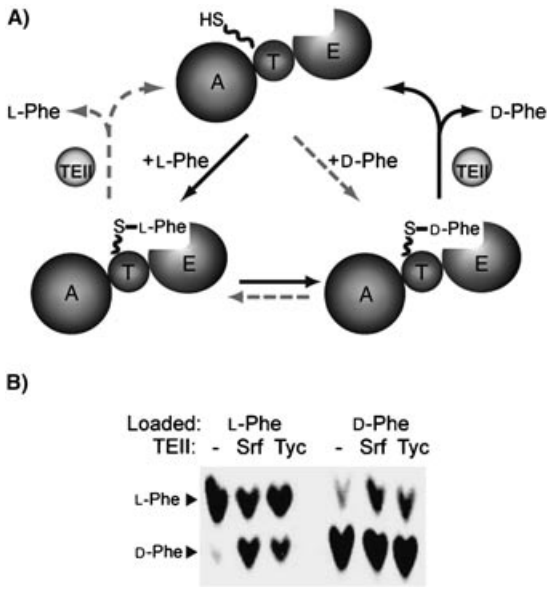


***Figure 2.*** *TEII acts on loaded NRPS modules. A) Schematic for the assay of TEII activity on loaded NRPS modules. Following loading of PheATE with optically pure Phe, release of the stereoisomer generated on the module is detected. Solid arrows, steps from L-Phe loading to TEII hydrolysis of D-Phe to determine rates for D-Phe-S-enzyme. Dashed arrows, steps from D-Phe loading to L-Phe-S-enzyme cleavage. B) Chiral silica TLC shows release of Phe stereoisomer opposite that loaded when TEII is present (24 h incubation).*

cleavage by TycF and SrfA-D were 150 and 180 $mM^{-1}min^{-1}$, respectively.[19] Values for the L-Phe-S-enzyme were 180 and 160 $mM^{-1}min^{-1}$. Reactions were not saturated at concentrations of PheATE up to 80 μM. Modest rates and high $K_m$ values were also observed for the PKS TEII from pikromycin, and thought to be consistent with a TEII largely dissociated from the PKS and able to scan carrier domains without interfering with normal processing of the PKS.[11] But unlike PKS TEIIs,[11] interactions with the T domaine might be important for NRPS TEII recognition, as there was a 500-fold decrease in rates observed for Phe-SNAC as compared to the Phe-S-enzyme. In addition to L- and D-Phe, TycF and SrfA-D also catalyzed cleavage of D-Trp and D-Tyr from PheATE following loading with the L-amino acid. Rates could not be obtained due to the slow initial loading of these noncognate amino acids.[6] As observed with SNAC substrates, TycF and SrfA-D recognized a broad range of amino acids loaded on a NRPS module.

Finally, we evaluated the ability of TEII to restore the activity of a stalled module, in which the loaded amino acid could not be processed by the NRPS dimodule PheATE/ProCAT. The action of PheATE and downstream module ProCAT generates a dipeptide intermediate that undergoes a self-catalyzed cleavage from the enzyme to form D-Phe-L-Pro-DKP (Figure 3 A, left). Previous studies of ProCAT showed that the C domain is D-amino acid specific for the upstream donor.[2,20] DKP formation was abolished when PheATE modules with inactivating mutations in the E domain were loaded with L-Phe and presented to the downstream ProCAT,[21] since these mutants were unable to generate the required D-Phe-S-enzyme (Figure 3 A, right). We used one of these mutants, H753A/N975A, to investigate the effect of TEII in this system. Following full loading with L-Phe, addition of the productive D stereoisomer (at equal concentration to L-Phe) did not lead to recovery of DKP formation as L-Phe continued to block the carrier-protein-loading site (Figure 3 B, box 1). However, addition of TycF or SrfA-D along with D-Phe restored product DKP formation by removing the stalled L-Phe and allowing loading and subsequent processing of D-Phe (boxes 2 and 3). Interestingly, the recovered activity varied with different E domain mutants; this suggests an effect of adjacent domains on TEII activity (Table 2). The specific function of the residues mutated in these mutants, whether in catalytic steps of E domain function or substrate recognition, is unknown. However, if in some mutants the loaded Phe were displaced from its binding pocket in the E domain, where it is believed to bind following loading,[3] it may be more available for TEII hydrolysis. The same would be true of incorrectly loaded substrates that are not recognized by later-acting domains (Figure 1 B). Lacking specificity for particular substrates, TEII recognition of unprocessed amino acids may be kinetically controlled by the increased half-lives of these unprocessed and unrecognized aminoacyl intermediates. A similar mechanism was proposed for PKS TEII recognition of aberrant acyl groups.[11] An attempt was made to compare kinetic parameters for cleavage of L-Phe from different E mutant modules by monitoring the loss of enzyme-bound radioactivity following loading with $^{14}$C-labelled L-Phe. Unfortunately, tight binding of the Phe-AMP intermediate resulted in fast reloading

  

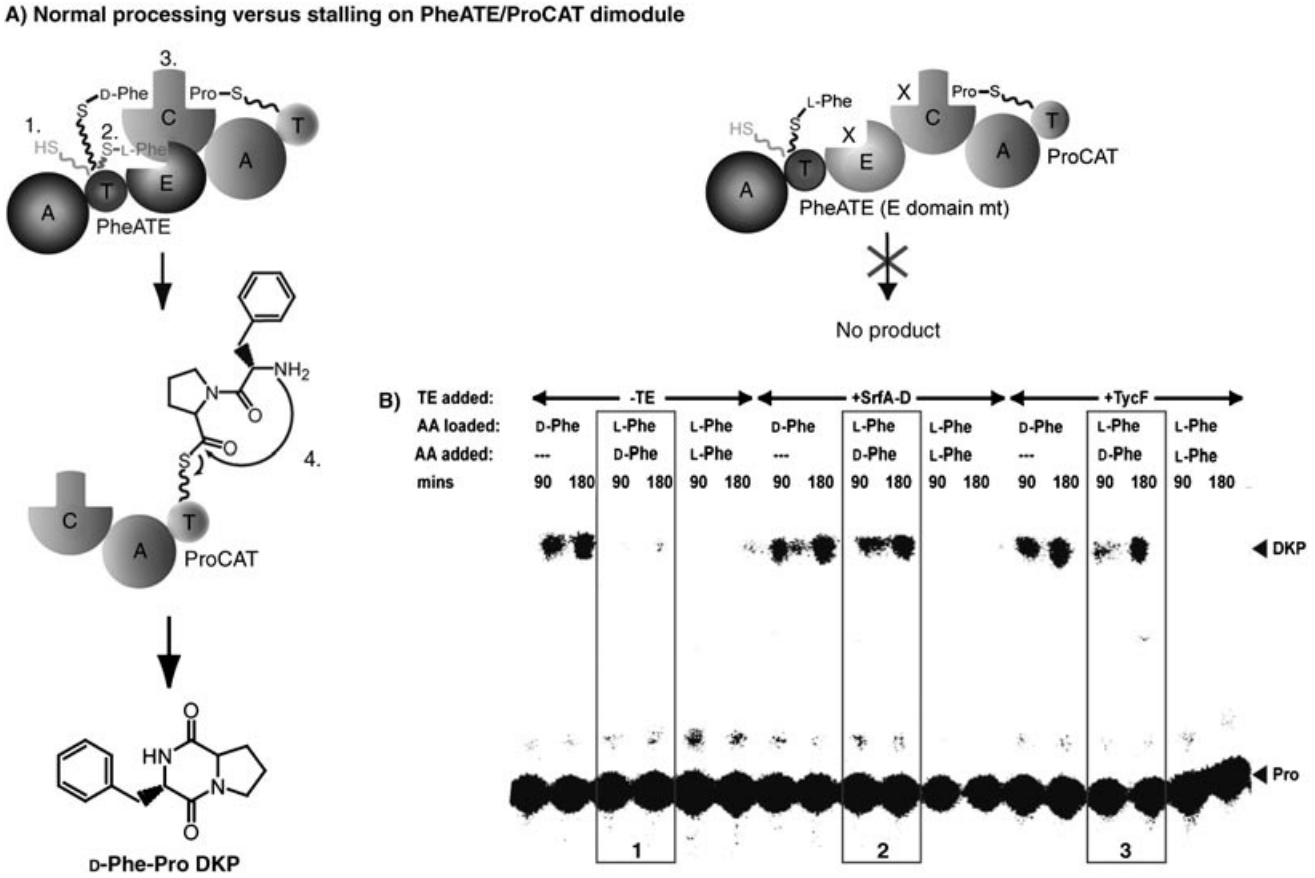


***Figure 3.*** *Restoration of product formation on stalled modules. A) D-Phe-Pro DKP formation by PheATE/ProCAT (left panel), including 1) loading L-Phe, 2) epimerization to D-Phe, 3) peptide bond formation with L-Pro, and 4) cyclization and release from module. When PheATE containing an E-domain mutation is used, the module is stalled after the first loading step (right panel; compare with Figure 1B). B) Preloaded H753A/N975A mutant (1.5 μM) and $^{14}$C-Pro-loaded ProCAT (1.5 μM) were mixed to initiate DKP product formation, which was detected by TLC. Boxes highlight conditions of stalling for which addition of TEII (1.5 μM) restored product formation.*

***Table 2.*** *Domain effect on recovery of DKP formation. Yields (t = 180 min) for stalled PheATE mutants ± TEII are shown. The baseline DKP formation for each mutant was taken as the yield obtained when only D-Phe was loaded and present in solution (italics), shown for comparison.*

| | Product yield of D-Phe-Pro DKP [μM] | | | |
|---|---|---|---|---|
| E domain mutation(s) | *D-Phe −TEII* | L-Phe loaded, D- and L-Phe in solution −TEII | +SrfA-D | +TycF |
| H753A | *12.6* | 0.0 | 2.5 | 1.8 |
| E892A | *5.3* | 0.0 | 0.0 | 0.0 |
| Y976A | *11.4* | 1.9 | 5.0 | 6.0 |
| H753A/N975A | *9.2* | 1.6 | 6.8 | 9.0 |

of the module such that TEII hydrolysis activity could not be detected (data not shown).

In summary, we have described several features of NRPS TEII activity consistent with a role for TEII in amino acid editing, including broad specificity for aminoacyl groups, cleavage of amino acids from NRPS modules at modest rates, and an ability to restore the activity of modules stalled with unprocessed aminoacyl intermediates. Moreover, the results describe a model for amino acid editing by NRPS TEII in which TEII recognizes a variety of loaded amino acids at low affinity and may discriminate "correct" from "incorrect" amino acids based on the increased half-life of unprocessed intermediates displaced from the normal domain binding sites. A general, rather than specific, mode of recognition is also indicated in PKSs since TEIIs from heterologous PKS systems can be exchanged without loss of product yields.[12,22] In addition, given its preference for aminoacyl over peptidyl substrates, editing by TEII at the level of the aminoacyl intermediate would reflect the energetic cost of hydrolyzing longer peptides as well as the known stringency of the C domain for downstream acceptor site amino acid relative to the upstream donor.[1] When the NRPS calcium-dependent antibiotic (CDA) synthetase was engineered such that the A domain of the seventh module was mutated and unable to provide the proper amino acid substrate for the upstream C domain, the hexapeptide intermediate was released from the NRPS in significant yield, probably through C-domain action.[17] Clearly, release of these longer intermediates is undesirable for the energetic efficiency of NRPS systems and could be prevented by rigorous editing of the loaded amino acid prior to incorporation into the product. Together, the results point toward a model for TEII hydrolytic editing of misloaded NRPS modules, in addition to that previously described in editing of misprimed modules.[9] A comparable editing role in PKSs

  

has been reported.[8, 11, 13] Type II thioesterases may play an important role in removing nonproductive acyl groups that accumulate on NRPS and PKS modules, thereby conferring efficiency and accuracy upon nature's synthetic machinery.

## Acknowledgements

*This work was supported in part by the National Institutes of Health Medical Scientist Training Program fellowship (E.Y. and R.M.K.), the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation (S.D.B.), and NIH Grant GM 20011 (C.T.W.). We also thank Lusong Luo for helpful discussion of kinetic assays and Fred Vaillancourt and Tanya Schneider for review of the manuscript.*

**Keywords:** biosynthesis · natural products · nonribosomal peptide synthetases · type II thioesterase

[1] P. J. Belshaw, C. T. Walsh, T. Stachelhaus, *Science* **1999**, *284*, 486.
[2] D. E. Ehmann, J. W. Trauger, T. Stachelhaus, C. T. Walsh, *Chem. Biol.* **2000**, *7*, 765.
[3] U. Linne, M. A. Marahiel, *Biochemistry* **2000**, *39*, 10439.
[4] S. L. Clugston, S. A. Sieber, M. A. Marahiel, C. T. Walsh, *Biochemistry* **2003**, *42*, 12095.
[5] T. A. Keating, Z. Suo, D. E. Ehmann, C. T. Walsh, *Biochemistry* **2000**, *39*, 2297.
[6] L. Luo, M. D. Burkart, T. Stachelhaus, C. T. Walsh, *J. Am. Chem. Soc.* **2001**, *123*, 11208.
[7] D. Schwarzer, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **2001**, *8*, 997.
[8] M. L. Heathcote, J. Staunton, P. F. Leadlay, *Chem. Biol.* **2001**, *8*, 207.
[9] D. Schwarzer, H. D. Mootz, U. Linne, M. A. Marahiel, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14083.
[10] A. Schneider, M. A. Marahiel, *Arch. Microbiol.* **1998**, *169*, 404.
[11] B. S. Kim, T. A. Cropp, B. J. Beck, D. H. Sherman, K. A. Reynolds, *J. Biol. Chem.* **2002**, *277*, 48028.
[12] M. Kotowska, K. Pawlik, A. R. Butler, E. Cundliffe, E. Takano, K. Kuczek, *Microbiology* **2002**, *148*, 1777.
[13] Z. Hu, B. A. Pfeifer, E. Chao, S. Murli, J. Kealey, J. R. Carney, G. Ashley, C. Khosla, C. R. Hutchinson, *Microbiology* **2003**, *149*, 2213.
[14] J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel, C. T. Walsh, *Nature* **2000**, *407*, 215.
[15] C. C. Tseng, S. D. Bruner, R. M. Kohli, M. A. Marahiel, C. T. Walsh, S. A. Sieber, *Biochemistry* **2002**, *41*, 13350.
[16] Details on all experimental procedures are provided as Supporting Information.
[17] G. C. Uguru, C. Milne, M. Borg, F. Flett, C. P. Smith, J. Micklefield, *J. Am. Chem. Soc.* **2004**, *126*, 5032.
[18] L. Luo, C. T. Walsh, *Biochemistry* **2001**, *40*, 5329.
[19] The E domain acts reversibly such that the module, when loaded with either L-Phe or D-Phe, is a mixture of 40% L- and 60% D-Phe-S-enzyme.[17] Since L-Phe could compete with D-Phe for TEII hydrolysis (and vice versa), the values reported by this assay underestimate TEII cleavage rates for the optically pure Phe-S-enzyme.
[20] T. Stachelhaus, H. D. Mootz, V. Bergendahl, M. A. Marahiel, *J. Biol. Chem.* **1998**, *273*, 22773.
[21] T. Stachelhaus, C. T. Walsh, *Biochemistry* **2000**, *39*, 5775.
[22] A. R. Butler, N. Bate, E. Cundliffe, *Chem. Biol.* **1999**, *6*, 287.
[23] T. Stachelhaus, M. A. Marahiel, *J. Biol. Chem.* **1995**, *270*, 6163.

Received: March 19, 2004

  