Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 17 (2007) 3034–3037

The thioesterase domain from the pimaricin and erythromycin biosynthetic pathways can catalyze hydrolysis of simple thioester substrates

Krishna K. Sharma and Christopher N. Boddy*

Department of Chemistry, Syracuse University, Syracuse, NY 13244-4100, USA

Received 20 January 2007; revised 16 March 2007; accepted 19 March 2007
Available online 23 March 2007

Abstract—The recombinant polyketide synthase thioesterase domains from the pimaricin and 6-deoxyerythronolide B biosynthetic pathways catalyze hydrolysis of a number of simple *N*-acetylcysteamine thioester derivatives. This study demonstrates that thioesterases are not highly substrate selective in formation of the acyl-enzyme intermediate, in contrast to non-ribosomal peptide synthase thioesterase domains that show very high specificity for substrate loading. This observation has important implications for the engineering of biosynthetic pathways to produce polyketide products.

Macrocyclic polyketides include numerous important pharmaceutical agents such as the anticancer agent epothilone and the antibiotic rifamycin. These natural products are produced by modular polyketide biosynthesis.[1] Substantial effort has been undertaken to understand this biosynthetic machinery with the hope that the enzymatic pathways can be engineered to produce new, non-natural polyketides for applications in drug discovery.[2–4] A major challenge in accomplishing this goal is the intrinsic substrate specificity of the thioesterase catalytic domain. In this study, we probe the substrate specificity of the thioesterase domains from the pimaricin[5] (**3**) and 6-deoxyerythronolide (**1**) biosynthetic pathways.

Thioesterases are the final catalytic domains of modular polyketide synthases. These enzymes cyclize the completed linear polyketide and cleave the polyketide, which has been covalently linked to the enzyme, from the biosynthetic machinery. This allows turnover of the enzyme system. Substrates that are not processed by the thioesterase domain remain attached to the enzyme, inhibiting product formation.

Linear acyl chains are macrocyclized by thioesterase domains in a two-step process.[6] The first step is acylation of the thioesterase at an active-site serine residue to generate an acyl-enzyme intermediate. The second step is a nucleophilic attack by an intramolecular nucleophile resulting in cyclization and hydrolysis of the acyl-enzyme intermediate. In this study, we have focused on understanding the specificity of substrate loading, which is the first step of the mechanism.

The loading of a limited number of substrates onto two thioesterase domains has been investigated.[7–9] However, the substrates examined did not entirely represent the possible diversity observed in polyketide biosynthesis. The two thioesterase domains studied, from the 6-deoxyerythronolide and pikromycin biosynthetic pathways, produce very similar 14-membered macrolactones (**1** and **2**, Fig. 1). Investigating thioesterases that produce structurally diverse macrocycles will give better insight into thioesterase substrate specificity and help identify residues in the substrate-binding pocket responsible for specificity. We chose to examine the pimaricin (**3**) thioesterase domain since its native substrate differed in ring size, substitution pattern, and acyl-enzyme intermediate electronics from **1** and **2**.

The goals of this study included identifying similarities in substrate specificity between multiple thioesterase domains, correlating substrate specificity with substitution and stereochemistry of the native substrate, and deter-

Keywords: Antibiotics; Biosynthesis; Polyketides; Thioesterase; Hydrolysis; Acyl-enzyme intermediate.
* Corresponding author. Tel.: +1 315 443 5803; fax: +1 315 443 4070; e-mail: cnboddy@syr.edu

0960-894X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2007.03.060

mining if there is a correlation between enzyme activity and substrate electronics. Our results show that the loading step catalyzed by polyketide thioesterase domains is not inherently highly substrate selective, but that thioesterase domains do preferentially recognize more native-like substrates. Additionally, we found that electronics of the thioester substrate play a significant role in enzyme activity.

A library of five substrates was selected to probe the selectivity of the thioesterase domains. The substrates chosen, **6**, **7**, **10**, **12**, and **14** (Fig. 2), had no accessible intramolecular nucleophiles and therefore could not undergo thioesterase-catalyzed cyclization. The thioesterase domains therefore could only catalyze hydrolysis of the activated esters as shown in Figure 2. The substrates differed in the oxidation state and stereochemistry at the α and β positions. The chemical diversity at these centers is representative of the diversity generated during polyketide biosynthesis (β-keto, β-hydroxy, α,β-unsaturated and saturated thioesters), and represents the diversity seen among native polyketide substrates. All compounds were synthesized using standard techniques and were purified to homogeneity, as measured by ^{1}H NMR spectroscopy.[10–15]

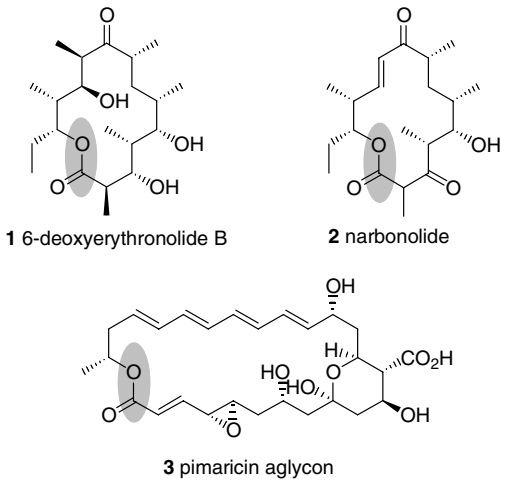

Figure 1. Structures of 6-deoxyerythronolide B, narbonolide (the precursor to pikromycin), and the pimaricin aglycon. The bonds generated by the thioesterase domains are highlighted.

The excised recombinant thioesterase domains from the pimaricin (pim TE) and 6-deoxyerythronolide B polyketide synthases (debs TE) were over-expressed in *Escherichia coli* and isolated in high purity by affinity chromatography.[16] Steady state kinetic parameters were determined for the thioesterase-catalyzed hydrolysis of substrates **6**, **7**, **10**, **12**, and **14** by quantifying the production of free thiol using Ellman's reagent.[7,8]

Because of the limited solubility of the substrates and the typically high Michaelis constants (K_M),[7,8] it was not possible to determine the turnover rate (k_{cat}) and the K_M independently. In most cases, the specificity constant (k_{cat}/K_M) was determined, as this could be easily measured at substrate concentrations below the dissociation constant where all substrates were sufficiently soluble in the reaction buffer. In the case of substrate **10** with the 6-deoxyerythronolide thioesterase, the turnover rate and dissociation constants were known.[7,8] Using our assay we determined $k_{cat} = 0.77 \pm 0.06\ min^{-1}$ and $K_M = 3.6 \pm 0.75$ mM and our experimental values agreed within the error with those previously reported ($k_{cat} = 0.68 \pm 0.16\ min^{-1}$ and $K_M = 9.4 \pm 4.5$ mM).[8] The specificity constants for the remaining substrates with both the 6-deoxyerythronolide and pimaricin thioesterase domains were then determined (Table 1).

Both the 6-deoxyerythronolide B and pimaricin thioesterases accelerated the hydrolysis reactions. The background hydrolytic rate for **7**, **10**, and **14** was $4.5 \pm 0.9 \times$

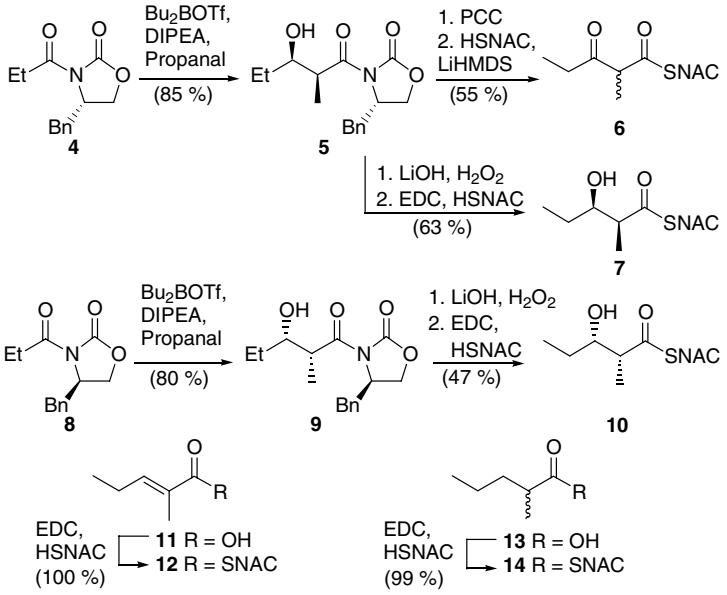

Figure 2. Synthesis of substrates used to probe the substrate specificity of the thioesterase domains.

Table 1. Specificity constants for the 6-deoxyerythronolide B (debs) and pimaricin (pim) thioesterase domains with substrates **6**, **7**, **10**, **12**, and **14**

Substrate	debs TE k_{cat}/K_M (M^{-1} s^{-1})	pim TE k_{cat}/K_M (M^{-1} s^{-1})
6	4.2 ± 0.4	nd
7	0.47 ± 0.06	0.04 ± 0.01
10	3.6 ± 1.4	0.02 ± 0.04
12	nd	1.3 ± 0.1
14	0.17 ± 0.01	0.24 ± 0.02

10^{-6} min^{-1}. This hydrolytic rate is comparable to rates documented for other *N*-acetylcysteamine thioesters.[17] Compounds **6** and **12** underwent faster background hydrolysis with rates of $k = 2.4 \pm 0.4 \times 10^{-5}$ min^{-1} and $k = 1.2 \pm 0.2 \times 10^{-5}$ min^{-1}, respectively. This is due to the conjugation between the thioesters and the α,β π system, which lowers LUMO energy, increasing hydrolysis rates. Assuming Michaelis constants in the range of 5–50 mM, which is generally seen for polyketide synthase thioesterases, a rate enhancement of 10^4–10^5 orders of magnitude is seen for both enzymes. This is comparable to the hydrolytic rate enhancement seen for epothilone thioesterase domain.[17]

Our kinetic studies on the pimaricin and 6-deoxyerythronolide B thioesterase domains provided insight into the substrate specificity for the selected substrates. The hydrolysis of substrates **6**, **7**, **10**, and **14** by the 6-deoxyerythronolide B thioesterase and substrates **7**, **10**, **12**, and **14** by the pimaricin thioesterase domain shows that the enzymes display limited substrate specificity. That is, regardless of the functional groups and stereochemistry present on the substrates the thioesterase can generate the requisite acyl-enzyme intermediate. The only substrates that were not hydrolyzed by the two thioesterase domains were compound **12** with the 6-deoxyerythronolide B thioesterase domain and compound **6** with the pimaricin thioesterase domain. Due to the higher rate of background hydrolysis of these two highly activated thioester substrates, it is difficult to detect levels of thioesterase-catalyzed hydrolysis that were slower than the background hydrolysis rate. It is therefore possible that **12** and **6** are also being hydrolyzed by the 6-deoxyerythronolide B and pimaricin thioesterases, respectively, but hydrolysis was below the detection limit of our assay. Development of an assay capable of discerning hydrolytic rates below background hydrolysis is necessary to confirm this hypothesis.

While it appears from the above data that most substrates can be loaded onto polyketide synthase thioesterase domains, there is a hierarchy of preferred substrates. In examining the specificity constants from 6-deoxyerythronolide B thioesterase-catalyzed hydrolysis, substrates **6** and **10** are hydrolyzed with the greatest efficiency. Substrate **10** has the same substitution pattern and stereochemistry as the native substrate for the 6-deoxyerythronolide B thioesterase domain. We therefore postulate that the thioesterase prefers the native substitution pattern and stereochemistry at the α and β carbons of the substrate. Substrate **6** however has a slightly greater specificity constant than **10**. This value may be partially influenced by the 3-keto group in substrate **6**, which makes the thioester highly activated and very prone to hydrolysis. Thus, the rapid hydrolysis of **6** does not represent preferred substrate recognition by the enzyme but also reflects an increased lability of the thioester bond.

Our hypothesis that the thioesterase prefers substrates that have native-like substitution and stereochemistry at the α and β carbons is supported by the data collected for the pimaricin thioesterase domain. Substrate **12**, which best represents the α,β unsaturated thioester found in the native pimaricin substrate, is hydrolyzed with the highest specificity constant in this study.

We propose that it may be a common feature of polyketide thioesterase domains that they are able to load most thioesters to generate acyl-enzyme intermediates. This observation is in contrast to non-ribosomal synthetase (NRPS) thioesterase domains, which exhibit high levels of substrate selectivity for the loading step of thioesterase-catalyzed chemistry.[18–21] NRPS thioesterase domains load linear peptides to generate peptidyl-enzyme intermediates similar to polyketide synthase thioesterase domains. Once the peptidyl-enzyme intermediate is formed, the NRPS thioesterase domains can catalyze hydrolysis and macrocyclization of a wide assortment of synthetic peptides.[22]

The formation of the peptidyl-enzyme intermediate is highly dependent upon the substitution of the thioester activated amino acid of the linear peptide chain. Replacement of the C-terminal amino acid residue in the substrate for the surfactin thioesterase leads to complete loss of hydrolytic and macrocyclization activity.[19] Similarly, substitution of the C-terminal and penultimate amino acid residues of the substrates for the tyrocidine[20] and fengycin[18] thioesterases leads to substantial degradation of hydrolysis and macrocyclization rates. These results suggest that recognition of C-terminal functional groups is important in NRPS peptidyl-enzyme intermediate formation.

Our data suggest that polyketide synthase thioesterases are less restrictive in substrate selection and loading. This low substrate specificity has important implications for the engineering of polyketide biosynthetic pathways to produce non-natural products. Based on this study it is highly likely that linear acyl chains produced by an engineered polyketide synthase pathway will be loaded onto the thioesterase domain forming the acyl-enzyme intermediate. These intermediates are known to undergo hydrolysis, especially if macrocyclization cannot take place.[23] This implies that linear free acids of the desired non-natural compounds can be produced from engineered systems without worrying about engineering the specificity of the thioesterase domains. Additionally it indicates that the brunt of the effort in thioesterase protein engineering should be aimed at understanding and modulating the substrate specificity and selectivity of the macrocyclization step.

Acknowledgment

This research was supported by Syracuse University. We thank C. Khosla (Stanford University) for the plasmid encoding the 6-deoxyerythronolide B thioesterase domain.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2007.03.060.

References and notes

1. Hill, A. M. *Nat. Prod. Rep.* **2006**, *23*, 256.
2. Floss, H. G. *J. Biotechnol.* **2006**, *124*, 242.
3. Weissman, K. J.; Leadlay, P. E. *Nat. Rev. Microbiol.* **2005**, *3*, 925.
4. Cane, D. E.; Walsh, C. T.; Khosla, C. *Science* **1998**, *282*, 63.
5. Aparicio, J. F.; Fouces, R.; Olivera, N.; Mendes, M. V.; Martin, J. F. *Chem. Biol.* **2000**, *7*, 895.
6. Kohli, R. M.; Walsh, C. T. *Chem. Commun.* **2003**, 297.
7. Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. *Chem. Biol.* **1999**, *6*, 117.
8. Lu, H.; Tsai, S.-C.; Khosla, C.; Cane, D. E. *Biochemistry* **2002**, *41*, 12590.
9. Weissman, K. J.; Smith, C. J.; Hanefeld, U.; Aggarwal, R.; Bycroft, M.; Staunton, J.; Leadlay, P. F. *Angew. Chem. Int. Ed* **1998**, *37*, 1437.
10. Chan, P. C.-M.; Chong, J. M.; Kousha, K. *Tetrahedron* **1994**, *50*, 2703.
11. Evans, D. A.; Bartroli, J.; Shih, T. L. *J. Am. Chem. Soc.* **1981**, *103*, 2127.
12. Evans, D. A.; Ng, H. P.; Clark, J. S.; Rieger, D. L. *Tetrahedron* **1992**, *48*, 2127.
13. Gilbert, I. H.; Ginty, M.; O'Neill, J. A.; Simpson, T. J.; Staunton, J.; Willis, C. L. *Bioorg. Med. Chem. Lett.* **1995**, *15*, 1587.
14. Oikawa, Y.; Sugano, K.; Yonemitsu, O. *J. Org. Chem.* **1978**, *43*, 2087.
15. Arthur, C.; Cox, R. J.; Crosby, J.; Rahman, M. M.; Simpson, T. J.; Soulas, F.; Spogli, R.; Szafranska, A. E.; Westcott, J.; Winfield, C. J. *Chem. Bio. Chem.* **2002**, 253.
16. The pim TE did not hydrolyze the pimaricin. pim TE (2.0 mM) was treated with pimaricin (2.0 mM) in 50 mM phosphate buffer at pH 8.0. LCMS analysis of the reaction showed no hydrolysis after 24 h at 37 °C.
17. Boddy, C. N.; Schneider, T. L.; Hotta, K.; Walsh, C. T.; Khosla, C. *J. Am. Chem. Soc.* **2003**, *125*, 3428.
18. Samel, S. A.; Bjorn, W.; Marahiel, M. A.; Essen, L.-O. *J. Mol. Biol.* **2006**, *359*, 876.
19. Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. *Biochemistry* **2002**, *41*, 13350.
20. Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T. *Nature* **2000**, *407*, 215.
21. Trauger, J. W.; Kohli, R. M.; Walsh, C. T. *Biochemistry* **2001**, *40*, 7092.
22. Kohli, R. M.; Walsh, C. T.; Burkart, M. D. *Nature* **2002**, *418*, 658.
23. Boddy, C. N.; Hotta, K.; Tse, M. L.; Watts, R. E.; Khosla, C. *J. Am. Chem. Soc.* **2004**, *126*, 7436.