ORGANIC LETTERS

2011
Vol. 13, No. 17
4518–4521

Enzymatic Timing and Tailoring of Macrolactamization in Syringolin Biosynthesis

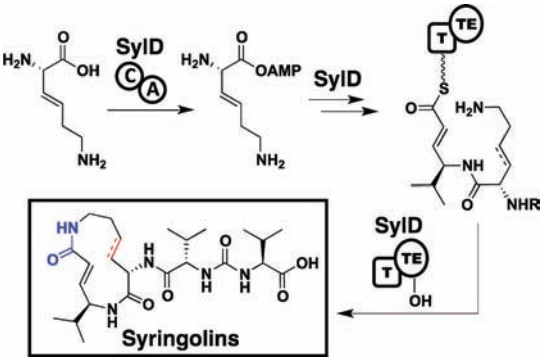
William M. Wuest,[†] Daniel Krahn,[‡] Markus Kaiser,[‡] and Christopher T. Walsh*,[†]

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, United States, and Zentrum für Medizinische Biotechnologie, Fakultät für Biologie, Universität Duisburg-Essen, 45117 Essen, Germany

christopher_walsh@hms.harvard.edu

Received June 21, 2011

ABSTRACT



The enzymatic activation of 3,4-dehydrolysine and subsequent formation of the 12-membered syringolin macrolactam were investigated. The timing of the desaturation was elucidated through the analysis of the initial adenylation domain of SylD. The SylD-TTE didomain was characterized and demonstrated to be the catalyst for formation of 12-membered macrocycles. When the SylD thioesterase domain was reacted with a family of acyclic CoA both natural and unnatural macrocycles were generated.

The syringolin family of natural products are 12-membered macrolactams produced by strains of *Pseudomonas syringae*.[1] The molecules serve as elicitors, which induce resistance to the fungal species *Pyricularia oryzae* in rice.[1] More recently, syringolin A has been identified as a virulence factor which irreversibly inhibits the 20S proteasome through a covalent mechanism.[2] Proteasome inhibitors such as the clinically used anticancer agent bortezomib represent a powerful class of chemotherapeutics.[3] The exciting biological activity and unique chemical architecture have consequently attracted the interest of the synthetic community and resulted in three total syntheses to date.[4] Supplementary studies have furthermore demonstrated that more potent and selective proteasome inhibitors can be generated by purposive derivatization of syringolins.[5] These findings have influenced our research efforts toward the exploitation of the biosynthetic gene cluster to create both natural and unnatural compounds in

[†] Harvard Medical School.
[‡] Universität Duisburg-Essen.

(1) Waspi, U.; Blanc, D.; Winkler, T.; Ruedi, P.; Dudler, R. *Mol. Plant-Microbe Interact.* **1998**, *11*, 727.

(2) Groll, M.; Schellenberg, B.; Bachmann, A. S.; Archer, C. R.; Huber, R.; Powell, T. K.; Lindow, S.; Kaiser, M.; Dudler, R. *Nature* **2008**, *452*, 755.

(3) (a) Groll, M.; Berkers, C. R.; Ploegh, H. L.; Ovaa, H. *Structure* **2006**, *14*, 451. (b) Orlowski, R. Z.; Kuhn, D. J. *Clin. Cancer Res.* **2008**, *14*, 1649.

(4) (a) Clerc, J.; Groll, M.; Illich, D. J.; Bachmann, A. S.; Huber, R.; Schellenberg, B.; Dudler, R.; Kaiser, M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6507. (b) Pirrung, M. C.; Biswas, G.; Ibarra-Rivera, T. R. *Org. Lett.* **2010**, *12*, 2402. (c) Dai, C.; Stephenson, C. R. J. *Org. Lett.* **2010**, *12*, 3453. (d) Clerc, J.; Schellenberg, B.; Groll, M.; Bachmann, A. S.; Huber, R.; Dudler, R.; Kaiser, M. *Eur. J. Org. Chem.* **2010**, 3991.

(5) (a) Clerc, J.; Florea, B. I.; Kraus, M.; Groll, M.; Huber, R.; Bachmann, A. S.; Dudler, R.; Driessen, C.; Overkleeft, H. S.; Kaiser, M. *ChemBioChem* **2009**, *10*, 2638. (b) Clerc, J.; Li, N.; Krahn, D.; Groll, M.; Bachmann, A. S.; Florea, B. I.; Overkleeft, H. S.; Kaiser, M. *Chem. Commun.* **2011**, *47*, 385. (c) Archer, C. R.; Koomoa, D. L. T.; Mitsunaga, E. M.; Clerc, J.; Shimuzu, M.; Kaiser, M.; Schellenberg, B.; Dudler, R.; Bachmann, A. S. *Biochem. Pharmacol.* **2010**, *80*, 170.

10.1021/ol2016687
Published on Web 08/03/2011

higher yield and efficiency than are currently available by total synthesis. More specifically, we believe that exploration of the enzymatic machinery will improve cyclization of the acyclic precursor, a notably troublesome step in total syntheses of syringolins.[4d]

Dudler and co-workers initially identified the syringolin gene cluster in 1998.[6] The biosynthetic machinery is composed of three specific genes, *syl*B, *syl*C, and *syl*D, which are solely responsible for the construction of the natural products (Figure 1). SylB is a putative amino acid desaturase (DS), which we propose transforms lysine into 3,4-dehydrolysine. SylC is a nonribosomal peptide synthetase (NRPS) composed of a condensation (C), a C-terminal condensation domain (C*), an adenylation (A), and a thiolation (T) domain. Our group previously demonstrated that this enzyme is responsible for production of the ureido side chain and proceeds through a cyclized intermediate.[7] SylD is an NRPS-PKS (polyketide synthase) megasynthase[8] possessing two iterative CAT domains followed by a ketosynthase (KS), dehydratase (DH), acyltransferase (AT), ketoreductase (KR), thiolation (T), and thioesterase (TE) domains. Bioformatic analysis predicts that this protein is responsible for the production and cyclization of the macrolactam core structure.

Our continuing interest in both the biosynthesis and biological activity of the syringolins led us to investigate both the timing and roles of two enzymes of interest: SylB and the TE domain of SylD. SylB would be the first

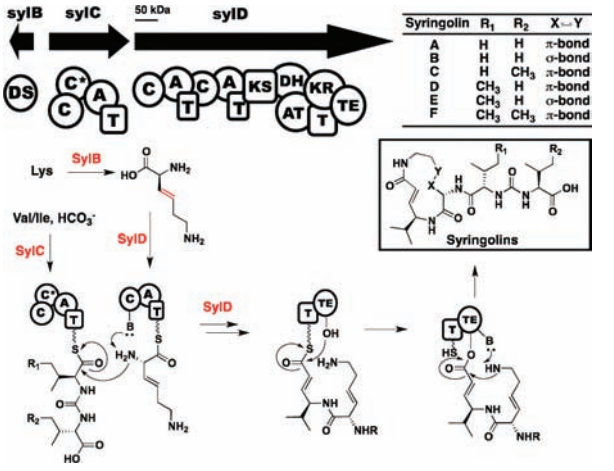


Figure 1. Syringolin gene cluster and biosynthetic pathway.

example of a desaturase that acts specifically on an amino acid, and the TE domain of SylD would be the first thioesterase responsible for the formation of a 12-membered macrolactam.[9] Additionally, the timing of desaturation, either before or after macrocyclization, could be determined.

Significant efforts were made to overexpress SylB heterologously *in vitro*; however all attempts were unsuccessful, perhaps because overexpression of an enzyme that converts lysine to a non-natural derivative, 3,4-dehydrolysine, would presumably be toxic to *E. coli*. In the absence of a soluble SylB, we refocused our interest on characterizing the A domain of SylD. The specificity of this enzyme would give strong evidence to the timing of lysine desaturation during biosynthesis: if the enzyme preferentially activates 3,4-dehydrolysine instead of lysine then presumably the desaturation would have to occur prior to activation and elongation by the SylD assembly line.

Accordingly, the SylD-CA didomain was amplified from *P. syringae* B728a and cloned into an *E. coli* expression vector to generate the 113 kDa C-term-His_6-tag fusion. Overexpression and Ni-NTA purification yielded 24.2 mg/L of soluble protein. The A domain was screened against L-lysine, L-3,4-dehydrolysine (dhl), D-lysine, D,L-4,5-dhl, ornithine, aspartate, and nonpolar derivatives by an ATP-PP_i radioactivity assay to identify the preferred substrate.[10] As expected L-3,4-dhl was preferentially reversibly adenylated with complete activation of the amino acid after 45 min (Figure S1). In contrast, L-lysine was reversibly adenylated at a much slower rate and reached a maximum of 70% activation. The other amino acids were activated at levels below 25% of the preferred substrate. These results correspond with what is observed in the syringolin family; in the absence of 3,4-dehydrolysine the A domain activates lysine effecting incorporation into the mature structure forming syringolin B and E (Figure 1). However in the presence of 3,4-dehydrolysine the compound is selectively activated and incorporated forming the most abundant derivative, syringolin A.

Next, our attention shifted to the SylD TE domain, which should be responsible for forming the 12-membered macrolactam. The SylD TTE-didomain construct was selected, as the availability of an adjacent T domain would permit installation of the acyl-pantetheinyl moieties of acyl CoAs for the investigation of the adjacent thioesterase domain. The enzyme was prepared in an analogous fashion to that of SylD-CA by heterologous overxpression in *E. coli* and purified as the 43 kDa C-term-His_6-tag fusion protein to provide 27.3 mg/L of soluble enzyme. The T domain was found to be in the apo-form as confirmed by subsequent phosphopantetheinylation with Bodipy-coenzyme A and the promiscuous phosphopantetheinyl transferase, Sfp (Figure S2).[11]

(6) Amrein, H.; Makart, S.; Granado, J.; Shakya, R.; Schneider-Pokorny, J.; Dudler, R. *Mol. Plant-Microbe Interact.* **2004**, *17*, 90.

(7) (a) Imker, H. J.; Walsh, C. T.; Wuest, W. M. *J. Am. Chem. Soc.* **2009**, *131*, 18623. (b) Ramel, C.; Tobler, M.; Meyer, M.; Bigler, L.; Ebert, M.-O.; Schellenberg, B.; Dudler, R. *BMC Biochemistry* **2009**, *10*, 62.

(8) Cane, D. E.; Walsh, C. T.; Khosla, C. *Science* **1998**, *282*, 63.

(9) A previous report demonstrated the enzymatic production of a 12-membered macrolactone ring through mutagenesis of the Erythromycin A thioesterase domain: Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1995**, *117*, 9105.

(10) Nielsen, J. B.; Hsu, M.-J.; Byrne, K. M.; Kaplan, L. *Biochemistry* **1991**, *30*, 5789.

(11) (a) Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M.; Marahiel, M. A.; Reid, R.; Walsh, C. T. *Chem. Biol.* **1996**, *3*, 923. (b) Reuter, K.; Mofid, M. R.; Marahiel, M. A.; Ficner, R. *EMBO J.* **1999**, *18*, 6823.

Prior to loading studies with the acyl CoAs noted below we undertook calibration of the percentage of the T domain in the apo-form by utilizing a radioactive acetylcoenzyme A (AcCoA) loading assay.[12] However, all initial experiments failed to provide any labeled proteins. This result could arise from two possibilities: either the T domain was inactive or both the T and TE domains were active (Figure 2A). By preincubating the T-TE didomain protein with phenylmethanesulfonylfluoride (PMSF), a known serine protease inhibitor,[13] one can eliminate TE activity and decipher between the two alternatives. Greater than 40% of the protein was radioactively labeled by AcCoA when pretreated with PMSF (Figure 2B) indicating both the T and TE domains of SylD were active *in vitro* and would permit their investigation with the synthetic CoA substrates.

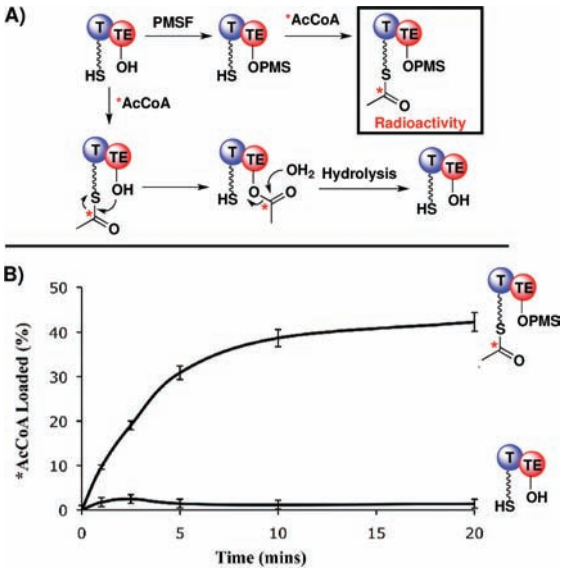
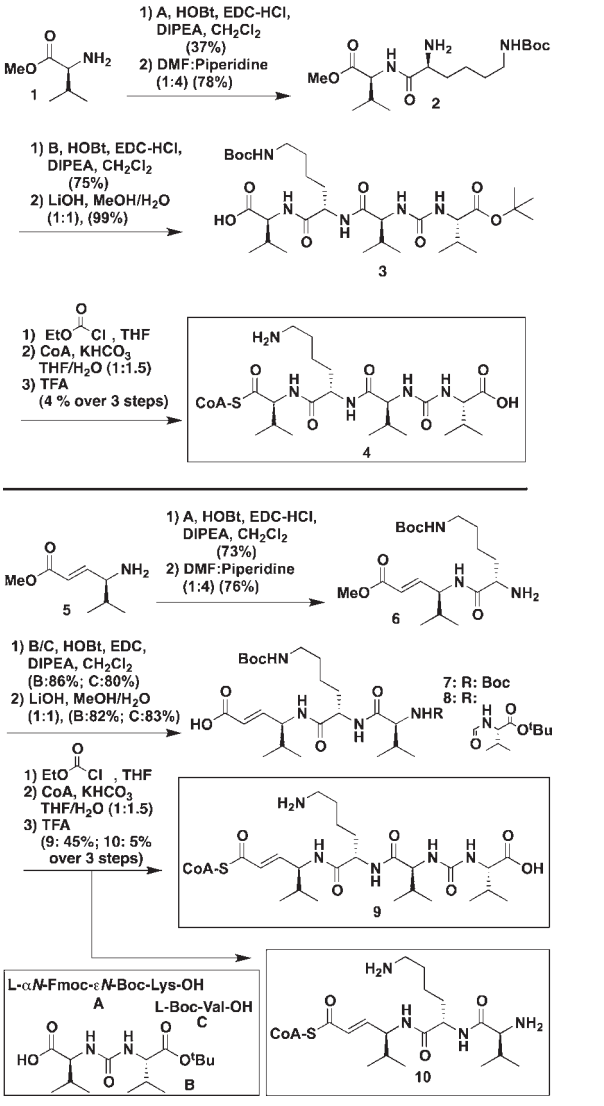


Figure 2. AcCoA loading of SylD-TTE.

To create a substrate for the TTE domain it was necessary to synthesize a small family of syringolin B acyclic precursors (Scheme 1). Due to the unusual polypeptide nature of the syringolins, possessing a polarity switch through the ureido linkage, we sought to investigate derivatives that differed at both the C- and N-terminus. The syntheses began with the coupling of amines **1** and **5** with bis-protected lysine (**A**), followed by cleavage of the Fmoc group. The corresponding amines were then appended with either the ureido side chain **B** or Boc-Valine (**C**) and saponified, thereby providing linear precursors **3**, **7**, and **8**. The carboxylic acids were then coupled with coenzyme A (CoA) and globally deprotected to yield the acyclic forms of syringolin B (**9**), the N-truncated variant (**10**), and the C-truncated derivative (**4**). The syntheses

(12) Linne, U.; Marahiel, M. A. *Methods Enzymol.* **2004**, *388*, 293.
(13) Caffrey, P.; Green, B.; Packman, L. C.; Rawlings, B. J.; Staunton, J.; Leadley, P. F. *Eur. J. Biochem.* **1991**, *195*, 823.

Scheme 1. Synthesis of Acylic CoA Precursors



were accomplished in six linear steps and provided milligram quantities of material for the enzymatic assays.

With both the synthetic and enzymatic materials in hand the *in vitro* macrocyclization reaction was next investigated. High-resolution mass spectrometry (HRMS) was implemented to monitor the enzymatic reaction on the scale of microgram quantities of cyclized compound. The assays were run in parallel with a control reaction which inactivated the TE domain with PMSF to ensure that the cyclization was enzymatically catalyzed. The wild type acyclic syringolin B precursor was successfully loaded onto the T domain as the S-pantetheinyl thioester by use of Sfp phosphopantetheinyl transferase. Subsequent enzyme-mediated transfer to the TE domain, intramolecular cyclization, and release provided the natural product which was identified by HRMS (Scheme 2). Additionally, an N-truncated variant that lacks the ureido function and instead possesses an amine terminus was also cyclized demonstrating that the charge of that side chain does not play a role in

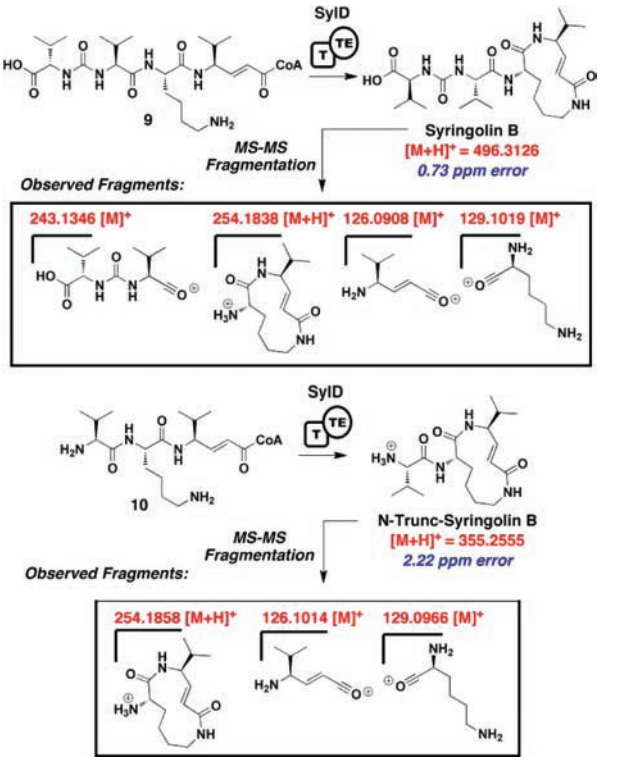
the cyclization event. Conversely when the C-truncated derivate was used as the substrate, no observable product was formed, presumably due to the excessive ring strain created by the contraction of the ring size from a 12-membered to a 10-membered ring. In addition, no cyclized product was observed in any of the control assays ensuring that the cyclization was indeed catalytic and required TE domain integrity.

MS-MS fragmentation was utilized to further confirm the identity of the cyclic products. The peptide backbone of the syringolin family of compounds produces a clean fragmentation pattern under optimized conditions. Fragment ions were identified with very low error (< 10 ppm) of the resulting ureido and macrocyclic fragments for syringolin B. Furthermore, we were able to identify the macrocyclic fragments as well with both the malonyl and lysine subunits (Scheme 2). In addition, the syringolin B fragment lacking valine and valine-CO were also identified (Supporting Information). Analogously, when the same procedure was implemented toward the N-truncated derivative **10** the same fragments, except for the ureido moiety, were detected, in accordance with the structural assignment. These results confirm that the SylD TE domain is solely responsible for the cyclization of the macrolactam ring of the syringolin family of natural products, making it the first *in vitro* example of a naturally occurring enzyme that forms a 12-membered macrocyclic lactam.

In summary, this work provides evidence for the timing and functionalization of the syringolin family of natural products. We have shown that 3,4-dehydrolysine is the preferred substrate of the initial A domain of SylD, strongly suggesting that the desaturation occurs before incorporation. Additionally, we have confirmed the activity of the TE domain of SylD *in vitro* and demonstrated that the enzyme is responsible for formation of the 12-membered macrolactam ring. Future work will investigate the use of the TE domain as a surrogate for the analogous, low yielding synthetic transformation allowing for the expedient synthesis of syringolin analogs.

Acknowledgment. This work is supported by NIH Grants GM49338 (C.T.W.) and GM091779 (W.M.W.).

Scheme 2. MS-MS Fragmentation of Full and N-Truncated Syringolin B



The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement No. 258413 (M.K.). The authors would like to thank Prof. S. Lindow at UC-Berkeley for supplying the B728a strain of *P. syringae*.

Supporting Information Available. Experimental details, characterization of all new compounds, and supporting figures. This material is available free of charge via the Internet at http://pubs.acs.org.