J|A|C|S COMMUNICATIONS

Published on Web 12/07/2009

SylC Catalyzes Ureido-Bond Formation During Biosynthesis of the Proteasome Inhibitor Syringolin A

Heidi J. Imker, Christopher T. Walsh,* and William M. Wuest

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115

Received October 28, 2009; E-mail: christopher_walsh@hms.harvard.edu

Nonribosomal peptide (NRPS) and polyketide synthetases (PKS) utilize numerous chemical reactions with precise control of regio- and stereochemistry on activated, tethered intermediates to construct diversely functionalized natural products with broad biological activity.[1] Typically these events occur through a controlled chain-elongation process where peptide and/or carbon−carbon bonds are constructed in an iterative manner. In NRPS the peptide scaffold is formed through successive assembly of amide bonds in a manner where chain polarity remains unidirectional. However, N-to-N terminal condensation via a ureido linkage leads to a reversal of chain polarity and introduces a new point of diversification to modulate biological activity.

The syringolin family of proteasome inhibitors are NRPS-PKS hybrid molecules with notable structural features including a 12-membered ring and *N*-terminal acylation via a ureido linkage.[2a] In syringolin A, Val_1 is *N*-acylated by an additional valine in a head-to-head condensation creating a ureido linkage and affording a negative terminus rather than the usual positive terminus. Other natural products containing a chain reversal by way of a ureido linkage include the anabaenopeptins,[2b] brunsvicamides,[2c] pacidomycins,[2d] mureidomycins,[2e] and napsamycins[2f] (Figure 1). Although the specific biological ramifications of chain reversal remain unknown at this time, *N*-acylation is known to dramatically influence the efficacy of syringolin A as a proteasome inhibitor.[3]

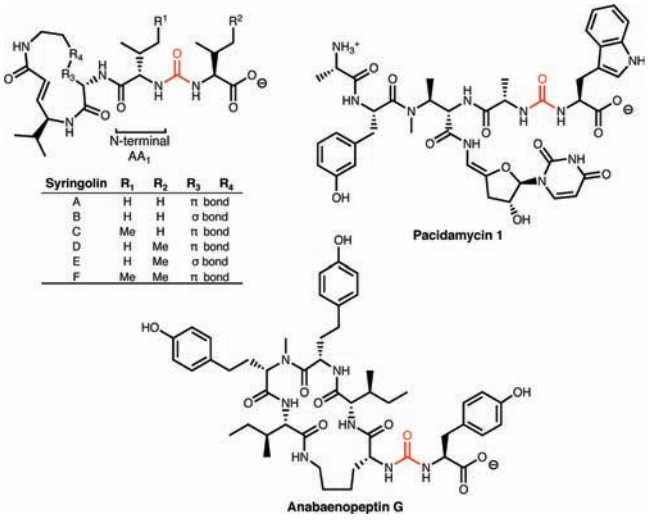


Figure 1. Ureido containing natural products.

Syringolin A was first isolated from *P. syringae* and characterized in 1998 by Dudler and co-workers as an elicitor of stress response in rice, aiding in resistance toward the phytotoxic fungi, *P. oryzae*.[2a] Recent studies have shown that the active site Thr_1 of the 20S eukaryotic proteosome interacts covalently and irreversibly with the α,β-unsaturated amide of the syringolin macrocylic core via a Michael addition mechanism.[4] The resulting inhibition is consistent with studies that show syringolin-induced cell death of neuroblastoma, ovarian,[5a] and leukemic[5b] cancer cells. The combined biomedical relevance and unusual ureido functionality led us to examine the formation of this linkage in syringolin biosynthesis.

Recent *in vivo* investigations by Dudler and co-workers into the biosynthetic origin of the ureido linkage of syringolin A revealed integration of either bicarbonate or carbon dioxide.[6] Feeding studies with [^{13}C]-bicarbonate followed by product characterization validated that incorporation was restricted to the carbonyl moiety. In this study we have focused on *in vitro* characterization of SylC, the NRPS enzyme presumed responsible for syringolin chain initiation, to evaluate its role in the generation of the ureido linkage subsequent to amino acid monomer activation.

The full-length *syl*C gene was amplified from *P. syringae* B728a and cloned into an *E. coli* expression vector to generate a 147-kDa His_6-tag fusion. Overexpression and Ni-NTA purification provided soluble SylC, with the thiolation (T) domain in the apo-form as confirmed by subsequent phosphopantetheinylation with acetyl-coenzyme A (AcCoA) and the promiscuous phosphopantetheinyl transferase, Sfp.[7] The SylC adenylation (A) domain was first assayed using ATP-PP_i exchange, and both L-Val and L-Ile were preferentially reversibly adenylated.[8] Additionally, L-Thr and L-*allo*-Ile were also activated but at ∼40% the level of the natural substrates (Figure S2). The covalent loading of amino acid monomers onto the SylC thiolation (T) domain was next investigated. Calibration experiments utilizing Sfp and radiolabeled [1-^{14}C]-AcCoA demonstrated that ∼30% of SylC could be labeled during conversion from the apo-form to the holo-form by installation of the phosphopantetheinyl (Ppant) group on the T-domain (Figure 2, ○).[9] Intriguingly, when the HS-Ppant holo-form of SylC enzyme was formed by prior incubation with Sfp and unlabeled CoA and then subjected to amino acid loading via ATP and [1-^{14}C]-L-valine, approximately twice the level of radiolabeled protein was detected (Figure 2, □). These results indicated that 2 equiv of valine were incorporated into a single SylC enzyme and thus provided the first indication that SylC was in fact generating the ureido-Val-CO-Val-S-SylC as a covalently tethered thioester intermediate. [1-^{14}C]-L-Ile was also successfully loaded onto

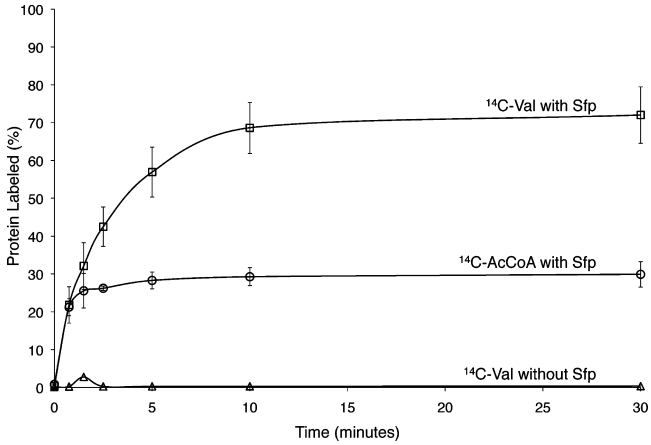
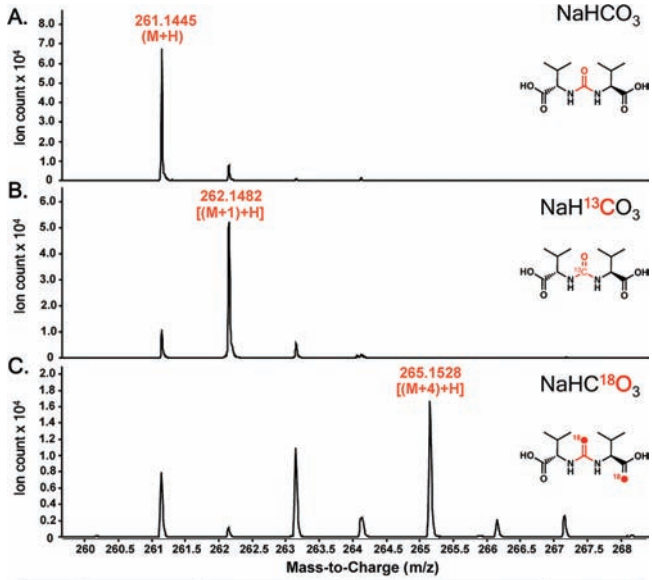


Figure 2. Comparison of extent of conversion from apo-SylC to holo-SylC by Sfp (○) to extent of [^{14}C]-L-Val loading onto the holo-SylC T-domain (□). SylC is isolated entirely in the apo-form as indicated by no [^{14}C]-L-Val loading if Sfp is omitted from the assay (△).

SylC at lower fractional stoichiometry, while the formation of [1-^{14}C]-Thr-S-SylC was not observed (data not shown).

To verify formation of the thioester bound Val-CO-Val moiety on the T-domain of SylC we turned to high-resolution mass spectrometry (HRMS). HRMS was utilized due to the accumulation of only picomolar quantities of the ureido-peptidyl-S-T intermediates as a result of the single turnover process. Quenching the assays with hydroxide, which hydrolyzed the T-domain thioester, allowed for the identification of the released ureido dipeptides by their corresponding masses.

In vitro assays with ATP, bicarbonate, L-Val and/or L-Ile, and SylC, followed by hydrolytic release of the peptidyl-S-T domain intermediates, gave all three of the ureido dipeptides (Val-CO-Val, Figure 3A; Val-CO-Ile, Ile-CO-Ile, Figure S5). Additionally, the enzyme accepted and processed L-*allo*-Ile, presumably an unnatural substrate, producing the symmetric ureido product (Figure S6). Next, [^{13}C]-bicarbonate was utilized as a chemical probe in SylC incubations to validate the carbon source of the carbonyl moiety of the ureido linkage. The appearance of an [M+1] mass peak by HRMS corroborates the feeding studies of Dudler and co-workers and confirms that SylC alone forms the ureido linkage from a bicarbonate source (Figure 3B).[6]



Panel	Bicarbonate	Relative Ion Abundance to (M+H)				Retention time (min)
		(M+H)	[(M+1)+H]	[(M+2)+H]	[(M+4)+H]	
A	$NaHCO_3$	1.0	0.11	0.015	--	5.73
B	$NaH^{13}CO_3$	1.0	4.4	0.45	--	5.57
C	$NaHC^{18}O_3$	1.0	0.15	1.6	2.1	5.69

Figure 3. HRMS of the (A) unlabeled valine ureido dipeptide, (B) [^{13}C]-labeled valine ureido dipeptide, and (C) [^{18}O]-labeled valine ureido dipeptide with normalization of ion abundance shown below.

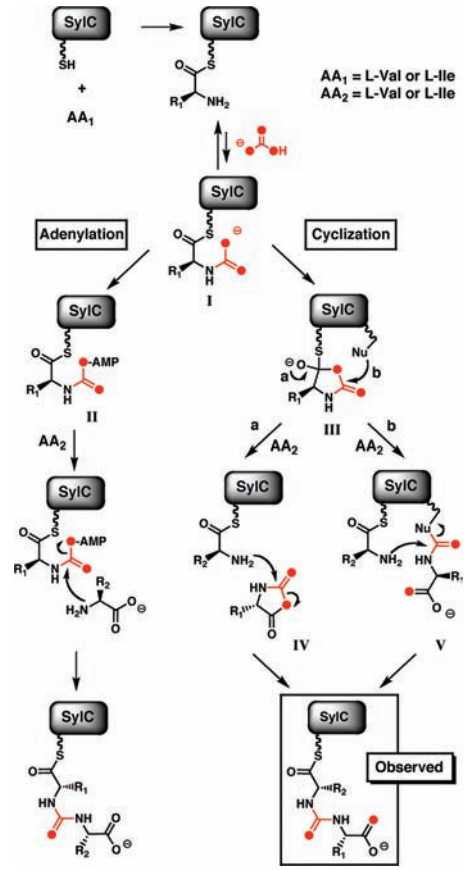


Figure 4. Proposed mechanisms of SylC-catalyzed formation of the ureido dipeptide showing two distinct pathways (adenylation and cyclization) which are distinguished by label incorporation when the reaction is carried out with [^{18}O]-bicarbonate.

These studies establish that SylC, a free-standing NRPS module with a *single* C-, A-, and T-domain, can activate L-Val or L-Ile *twice* and effect the chain reversal step to build the ureido linkage on the HS-Ppant arm of its T-domain. This constitutes novel chemistry for an NRPS module, and the lack of precedent led us to question the mechanistic details of ureido-linkage formation. We hypothesized that the SylC A-domain conventionally activates 1 equiv of amino acid as the aminoacyl-AMP, which is then captured by the thiol of the Ppant arm and loaded as the thioester onto the T-domain (Val-S-T SylC, Figure 4). It is likely that carboxylation of the amine of the tethered aminoacyl moiety by either bicarbonate or, more likely, carbon dioxide next forms a transient *N*-carboxy adduct, **I**. Examples of stabilized *N*-carboxy amines are known to participate in the catalytic mechanisms of RuBisCO[10a] and class D β-lactamases.[10b]

Formation of **I** would be followed by one of several possible transformations (Figure 4). The "adenylation" pathway involves *N*-carboxy-amino acid adenylation to form an activated mixed anhydride **II**, which would then react with a free amino acid

forming the resulting ureido dipeptide. Alternatively, in the "cyclization" pathway, the *N*-carboxy intermediate would cyclize intramolecularly, to generate a highly reactive covalently tethered thiohemiacetal intermediate, **III**. This transient thiohemiacetal would partition in one of at least two ways. In one scenario, formation of a "Leuch's anhydride" of type **IV**[11] would persist as a noncovalently bound species in the SylC active site during activation and loading of the second Val (Figure 4, pathway a). The nucleophilic amine of the second Val-S-T intermediate would capture **IV** regiospecifically and unravel it to the ureido-peptidyl-S-SylC.[12] Alternatively, the transient thiohemiacetal **III** would be intercepted first by a nucleophilic SylC residue (Nu) to form a covalent adduct **V** (Figure 4, pathway b).[13] This in turn would be captured by the nucleophilic amine of the second tethered Val yielding the ureido dipeptide. Presumably, the ureido-peptidyl-S-SylC serves as the upstream intermediate that is transferred to the next NRPS module during chain elongation in syringolin biosynthesis.

To distinguish between adenylation or cyclization mechanisms, we chose [^{18}O]-bicarbonate as a probe, since the adenylation pathway of Figure 4 would result in an [M+2] product while both arms of the cyclization pathway would furnish a ureido dipeptide with an enrichment of [M+4].[14] SylC incubations were performed in [^{18}O]-bicarbonate/[^{18}O]-water (80% total ^{18}O enrichment) after which the protein was precipitated with MeOH and the pellet washed three times to remove [^{18}O]-water and [^{18}O]-bicarbonate. The thioester-bound intermediates were released by chemical hydrolysis in unlabeled water and analyzed by HRMS. Inspection of these data shows clear enrichment of the [M+4] mass peak (Figure 3C). Notably, one of the two [^{18}O] atoms is incorporated into the amino acid carboxylate (Val), while the other is presumably localized in the ureido-group (Figure S9); the separation of the two oxygen atoms is fully consistent with intramolecular cyclization of the *N*-carboxyaminoacyl-S-T species (**I**). Capture of a Leuch's anhydride **IV** or the covalent adduct **V** could equally account for the incorporation and placement of two [^{18}O] atoms in ureido products.

In summary, we have completed the first *in vitro* characterization of enzymatic ureido-linkage formation. SylC, with a single C-, A-, and T-domain, iteratively activates two amino acid monomers and constructs the ureido linkage by incorporation of bicarbonate/CO_2 by cyclization of an initial *N*-carboxy-aminoacyl-S-Ppant enzyme intermediate.

Future studies will be directed at deducing evidence in favor of these or alternate mechanisms for formation of the ureido group and evaluating parallel systems for the peptide chain reversal in pacidamycins and anabaenopeptins. The mode of action of syringolin A is similar to that of the anticancer drug Velcade[15] and presents an opportunity to expand the proteasome inhibitor class of cancer therapeutics. We anticipate that exploration of the promiscuity inherent to SylC, as evidenced by formation of the unnatural L-*allo*-Ile-containing ureido-dipeptide, should allow for the production of new syringolin analogues.

Acknowledgment. This work was supported by NIH Grant GM49338 (C.T.W.). The authors would like to thank Prof. S. Lindow at UC-Berkeley for supplying the B728a strain of *P. syringae* and Prof. R. Dudler at the University of Zurich for discussion of results prior to publication.

Supporting Information Available: Supplemental figures, experimental procedures, and HRMS characterization. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Cane, D. E.; Walsh, C. T.; Khosla, C. *Science* **1998**, *282*, 63.
(2) (a) Waspi, U.; Blanc, D.; Winkler, T.; Ruedi, P.; Dudler, R. *Mol. Plant-Microbe Interact.* **1998**, *11*, 727. Amrein, H.; Makart, S.; Granado, J.; Shakya, R.; Schneider-Pokorny, J.; Dudler, R. *Mol. Plant-Microbe Interact.* **2004**, *17*, 90. (b) Harada, K.-I.; Mayumi, T.; Shimada, T.; Suzuki, M.; Kondo, F.; Watanabe, M. F. *Tetrahedron Lett.* **1993**, *34*, 6091. (c) Walther, T.; Renner, S.; Waldmann, H.; Arndt, H.-D. *ChemBioChem* **2009**, *10*, 1153. (d) Chem, R. H.; Buko, A. M.; Whittern, D. N.; McAlpine, J. B. *J. Antibiot.* **1989**, *42*, 512. (e) Isono, F.; Sakaida, Y.; Takahashi, S.; Kinoshita, T.; Nakamura, T.; Inukai, M. *J. Antibiot.* **1993**, *46*, 1203. (f) Chatterjee, S.; Nadkarni, S. R.; Vijayakumar, E. K.; Patel, M. V.; Ganguli, B. N.; Fehlhaber, H. W.; Vertesy, L. *J. Antibiot.* **1994**, *47*, 595.
(3) 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*, 6705.
(4) 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.
(5) (a) Coleman, C. S.; Rocetes, J. P.; Park, D. J.; Wallick, C. J.; Warn-Cramer, B. J.; Michel, K.; Dudler, R.; Bachmann, A. S. *Cell Prolif.* **2006**, *39*, 599. (b) 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.
(6) Ramel, C.; Tobler, M.; Meyer, M.; Bigler, L.; Ebert, M.-O.; Schellenberg, B.; Dudler, R. *BMC Biochemistry* **2009**, *10*, 62.
(7) 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. Reuter, K.; Mofid, M. R.; Marahiel, M. A.; Ficner, R. *EMBO J.* **1999**, *18*, 6823.
(8) Nielsen, J. B.; Hsu, M.-J.; Byrne, K. M.; Kaplan, L. *Biochemistry* **1991**, *30*, 5789.
(9) Linne, U.; Marahiel, M. A. *Methods Enzymol.* **2004**, *388*, 293.
(10) (a) Lorimer, G. H.; Miziorko, H. M. *Biochemistry* **1980**, *19*, 5321. (b) Li, J.; Cross, J. B.; Vreven, T.; Meroueh, S. O.; Mobashery, S.; Schlegel, H. B. *Proteins* **2005**, *61*, 246.
(11) Hirschmann, R.; Strachan, R. G.; Schwam, H.; Schoenewaldt, E. F.; Joshua, H.; Barkemeyer, B.; Veber, D. F.; Paleveda, W. J.; Jacob, T. A.; Beesley, T. E.; Denkewalter, R. G. *J. Org. Chem.* **1967**, *32*, 3415.
(12) Leuch anhydrides have been utilized as peptide coupling reagents; however no dipeptide or resulting diketopiperazines were observed as side products from SylC incubations by HRMS.
(13) Wong, B. J.; Gerlt, J. A. *J. Am. Chem. Soc.* **2003**, *125*, 12076.
(14) ^{18}O-Enrichment represented by the filled oxygens.
(15) Groll, M.; Berkers, C. R.; Ploegh, H. L.; Ovaa, H. *Structure* **2006**, *14*, 451.

JA909170U