Four enzymes define the incorporation of coenzyme A in thienamycin biosynthesis

Michael F. Freeman*, Kristos A. Moshos†, Micah J. Bodner‡, Rongfeng Li§, and Craig A. Townsend*‡‡

Departments of *Biology and †Chemistry, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218

Edited by Christopher T. Walsh, Harvard Medical School, Boston, MA, and approved June 4, 2008 (received for review May 8, 2008)

The enzymatic activities of three proteins encoded by the thienamycin gene cluster of Streptomyces cattleya (ThnR, ThnH, and ThnT) have been shown to incrementally cleave CoA to afford the active side-chain component of the β-lactam antibiotic thienamycin. These results supersede proposals based on earlier radiochemical incorporation experiments. For 20 years it has been thought that cysteine was directly incorporated into the antibiotic. Specific, stepwise truncation of CoA to 4-phosphopantetheine, pantetheine, and finally cysteamine was observed with ThnR, ThnH, and ThnT, respectively, in a series of coupled enzymatic assays. Pantetheinylated carbapenem substrates were synthesized to address possible thienamycin biosynthetic intermediates and were shown to be effective substrates for the pantetheine-cleaving enzyme ThnT. Finally, a fourth gene, thnF, was shown to encode a protein capable of N-acetylation of a model compound containing cysteamine in the presence of acetyl-CoA, consistent with the production of the S. cattleya comatabolite, N-acetyltethiynamycin. Taken together, these four enzymes are proposed to siphon CoA from primary metabolism to create the side chains for the predominant S. cattleya carbapenem, thienamycin, and N-acetyltethiynamycin, in a process likely to be general for the broader class of these antibiotics.

β-lactam antibiotics | carbapenem | pyrophosphatase | acylase | phosphatase

Thienamycin has been the hallmark of the carbapenem class of β-lactam antibiotics since its discovery in the 1970s from the soil actinomycete Streptomyces cattleya (1). Recognized for their broad antibacterial spectra and resistance to most classes of β-lactamases, the carbapenems are of growing importance in the treatment of infectious diseases in humans. Carbapenems are distinct from β-lactams such as penicillins and cephalosporins where carbon replaces sulfur at the C-1 position in the bicyclic nucleus (2). To date, >50 known naturally occurring carbapenems and carbapenamers have been structurally characterized, yet most biosynthetic knowledge is limited to the simplest carbapenem, (SR)-carbapenem-3-carboxylic acid (2, 3) (Fig. 1A).

The gene cluster for (SR)-carbapenem-3-carboxylic acid was first discovered in Pseudobacterium carotovorum (formerly Erwinia carotovorum), but similar clusters have since been found in both Serratia marcescens and Photobacterium luminescens (4, 5). Nine ORFs comprise the cluster, but only three genes (carA, carB, and carC) are absolutely required for antibiotic production (6). Although crystal structures for each are known, and detailed mechanistic insights have been determined for two of the three enzymes involved in creating the simple carbapenem, very little is known about the biosyntheses of the more highly substituted carbapenemers such as thienamycin and the remainder of this antibiotic family (7, 8).

Early biochemical studies revealed that the primary metabolic sources of the thienamycin core paralleled those of the simple carbapenem, with glutamate giving rise to the pyrroline ring and acetate the β-lactam carbons (9). The C-6 hydroxethyl side chain was shown to originate from methionine through two successive methylation events, presumably occurring by adenosylmethionine-utilizing enzymes (9, 10). This result was further supported by the dependence on vitamin B12 and Co2+ for thienamycin production (2, 11).

Initial insights into the origin of the thienamycin C-2 side chain were based on incorporation studies that suggested cysteine was directly incorporated into the carbapenem. [35S]- and [U-14C]cysteine were shown to give >70% incorporation into thienamycin, whereas [35S]pantethine was very poorly incorporated into the antibiotic (9). These feeding studies were carried out in response to the discovery of the OA-6129 family of carbapenems, where an intact pantetheine side chain was observed in place of cysteamine (12). Interestingly, random mutagenesis of Streptomyces fulvoviridis A933 1719, a producer of the PS-5 series of carbapenems, resulted in a mutant that produced the OA-6129 series in lieu of PS-5 (Fig. 1C and D) (13). In addition, A933 acylase from S. fulvoviridis was purified from the parent strain and shown to hydrolyze the pantetheine side chain of the OA-6129 series to produce the cysteamine side chain found in thienamycin (14).

Despite intense efforts made through the mid 1980s, further progress has been thwarted on thienamycin biosynthesis because of the poor transformation systems for the host strain as well as the instability and low titers of the antibiotic (15). However, in 2003, Nunez and coworkers (16) isolated and sequenced the thienamycin gene cluster in S. cattleya, reopening the field to further research. Here, we report the first biochemical analysis of enzymes encoded by the thienamycin gene cluster. Four enzymes are now shown to be involved in processing the C-2 side chain in thienamycin biosynthesis. Three of these have been demonstrated to incrementally truncate CoA to cysteamine, with the fourth able to cap the amine of a cysteamine-containing model substrate with acetate. Taken together, these enzymes are proposed to be responsible for the predominant C-2 side chains of thienamycin and N-acetyltethiynamycin produced by S. cattleya (14).

Results and Discussion

The thienamycin gene cluster consists of 22 ORFs, only two of which have apparent homologs in the nine-gene cluster for the simple carbapenem (SR)-carbapenem-3-carboxylic acid. ThnE and ThnM show 31.5% and 23.7% sequence identity to the first two biosynthetic proteins in the carbapenem pathway, CarB and CarA, respectively (www.ebi.ac.uk/emboss/align/). In contrast, a homolog of the last essential protein, CarC, is not apparent within the thienamycin cluster. The lack of a CarC homolog


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

To whom correspondence should be addressed at: Department of Chemistry, Johns Hopkins University, Remsen Hall, Room 252, 3400 North Charles Street, Baltimore, MD 21218. E-mail: ctownsend@jhu.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0804500105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA
E. coli tag, the codon-optimized achieved through the use of an N-terminal 6-histidine (6-His) synthetic double-stranded DNA linker of the optimized sequence. The expression host was Rosetta2(DE3), in which six rare codons were supplemented. The codons of thnH were suboptimal, with rare E. coli codons spread throughout the 675-bp gene. Therefore, the whole of thnH was codon-optimized and harmonized for E. coli.

ThnT (GenBank accession no. CAD18986) bears sequence similarity to the Nudix hydrolase family, showing 38% sequence similarity to a CoA pyrophosphatase from Deinococcus radiodurans (22). This family of enzymes hydrolyzes diphosphates of various nucleotide-containing molecules. Nudix hydrolases are characterized by a conserved Nudix box motif present directly upstream of the Nudix box (22). ThnR possesses slight variations of both motifs characteristic of a CoA pyrophosphatase and was thought to be the enzyme responsible for processing CoA or a CoA-containing precursor of thienamycin (Fig. S2).

ThnR expression was achieved in Rosetta2 either as an N-terminal 6-His tag fusion or as a C-terminal 6-His tag construct with the first 12-aa codon-optimized. Hydrolysis of CoA within the diphosphate moiety yielding 4-phosphopantetheine and adenosine 3′,5′-diphosphate was determined through CMQT derivatization. Importantly, ThnR was unable to yield pantetheine from CoA for subsequent ThnT cleavage, hinting that yet another enzyme in the cluster was responsible for connecting the activities of the two enzymes.

One ORF in the thienamycin cluster, thnH, encoded a protein with weak homology to the haloacid dehalogenase (HAD) superfamily of hydrolases, which was hypothesized to hydrolyze the phosphate from 4-phosphopantetheine (23). As with other genes in the cluster, thnH would not express in any of the E. coli hosts tested. The codons of thnH were suboptimal, with rare E. coli codons spread throughout the 675-bp gene. Therefore, the whole of thnH was codon-optimized and harmonized for E. coli.
expression through the annealing of overlapping 40-bp primers by using the polymerase cycling assembly (PCA) method (24). Success was achieved by separating the gene into two synthons of ~375 bp, followed by overlap extension of the two partial sequences to create the full-length, codon-optimized synthetic sequence. Expression was attained in E. coli BI21(DE3) pLysS, because the construct seemed to be toxic and inhibited growth of Rosetta2 before induction.

As hoped, ThnH (GenBank accession no. CAD18976) hydrolysis of 4-phosphopantetheine to yield pantetheine and phosphate was initially detected by using a phosphate-molybdate assay (data not shown) and further supported by HPLC and ESI mass spectrometric analysis (25). Thus, three enzymes are required from the thienamycin cluster to fully process CoA to cysteamine, the active side-chain component of thienamycin.

The ability for ThnR, ThnH, and ThnT to work in concert to completely and specifically process CoA to cysteamine was tested in a series of coupled enzyme assays. In addition to the three-enzyme coupled assay, reactions containing two of the three enzymes were incubated with CoA and MgCl₂ for 3 hours at 37°C. As shown in Fig. 2 A–E, all enzymes acted specifically to cleave only the molecules previously identified as substrates. The buildup of appropriate intermediates in CoA processing was present without the detection of subsequently hydrolyzed products. Interestingly, CoA prederivatized with CMQT was not a substrate for ThnR, suggesting a potential specificity toward CoA bearing only its free thiol. ThnH had a markedly reduced cleavage rate for prederivatized 4-phosphopantetheine, whereas ThnT easily cleaved all prederivatized pantetheine under the conditions tested (Fig. S3). From these results, it was still uncertain at which stage of CoA processing the side chain is attached to the carbapenam or carbapenem core during thienamycin biosynthesis. Based on the apparent structural homology to the S. fulvoviridis A933 1719 mutant 1501 OA-6129 series of carbapenems, it is reasonable to suggest attachment of the C-2 pantetheinyl side chain at C-2 was not seen under the conditions tested. ThnT seems to parallel the activity of hydrolyzing pantetheine-containing carbapenem substrates previously observed for the A933 acylase from S. fulvoviridis.

To link the truncation of CoA or, more specifically, pantetheine, to potential thienamycin biosynthetic precursors, 2-pantetheinyl carbapenams were selected to test the ability of ThnT to hydrolyze the pantetheinyl side chains after linkage to the β-lactam antibiotic core. Pantetheinylated carbapenams have been isolated in trace amounts from Streptomyces sp. OA-6129 (Fig. 3A) (26). Given that clear homologs of CarB and CarA, which are responsible for the biosynthesis of the carbapenam nucleus, correspond to ThnE and ThnM from the thienamycin gene cluster, we deduced that simple carbapenam formation was likely an early step. Although we now know the precursors of both the C-6 and C-2 side chains, the timing of their attachments, bicyclic ring inversion, and desaturation to the carbapenem are not known. To examine test substrates for the ThnT reaction that presented minimum steric demand but contained the probable carbapenam core, two potential substrates, 2-pantetheinyl carbapenams 1 and 2, were prepared (Fig. 3B). The synthesis proceeded from L-aspartic acid to 2-oxo-carbapenam 3 (R = p-methoxybenzyl) by established procedures (27, 28). Reduction of 3 gave a diastereomeric mixture of alcohols 4 (10:1 ratio), which underwent elimination in the presence of methanesulfonyl chloride and base to the carbapenem 5 (R = p-methoxybenzyl) (29). Heteroconjugate addition of pantetheine acetonide (30) to the carbapenem afforded a 2:1 mixture of (3R)-thioether products (31), which could be separated by HPLC and deprotected with trifluoroacetic acid and anisole (32) to give the (2R,3R,5R)- and (2S,3R,5R)-2-pantetheinyl carbapenams 1 and 2, respectively. Both the desired cysteamyl carbapenams and the β-lactam-hydrolyzed products were detected by dansyl chloride derivatization, followed by HPLC-ESI mass spectrometry (Fig. S4).

Fig. 2. The stepwise truncation of CoA. (A) Representation of the incremental truncation of CoA by ThnR, ThnH, and ThnT. ThnR cleaves CoA to 4-phosphopantetheine, whereas ThnH subsequently cleaves the phosphate to produce pantetheine, and ThnT hydrolyzes pantetheine to yield cysteamine (8–10). HPLC analyses of coupled enzyme assays to cleave CoA before CMQT derivatization: ThnH and ThnT (B), ThnR and ThnT, ThnR and ThnH (D), and ThnR, ThnH, and ThnT (E). The 4-phosphopantetheine is labeled as 4-P-Pant.; pantetheine is labeled as Pant.; cysteamine is labeled as Cyst. Peaks corresponding to enzymatic products are highlighted in red.

panetheinyl side chains at C-2 was not seen under the conditions tested. The acceptance of either thioether configuration does not allow discrimination between a pantetheinyl carbapenam or carbapenem as the true substrate of ThnT. ThnT seems to parallel the activity of hydrolyzing pantetheine-containing carbapenam substrates previously observed for the A933 acylase from S. fulvoviridis.

As hoped, ThnH (GenBank accession no. CAD18976) hydrolysis of 4-phosphopantetheine to yield pantetheine and phosphate was initially detected by using a phosphate-molybdate assay (data not shown) and further supported by HPLC and ESI mass spectrometric analysis (25). Thus, three enzymes are required from the thienamycin cluster to fully process CoA to cysteamine, the active side-chain component of thienamycin.

The ability for ThnR, ThnH, and ThnT to work in concert to completely and specifically process CoA to cysteamine was tested in a series of coupled enzyme assays. In addition to the three-enzyme coupled assay, reactions containing two of the three enzymes were incubated with CoA and MgCl₂ for 3 hours at 37°C. As shown in Fig. 2 A–E, all enzymes acted specifically to cleave only the molecules previously identified as substrates. The buildup of appropriate intermediates in CoA processing was present without the detection of subsequently hydrolyzed products. Interestingly, CoA prederivatized with CMQT was not a substrate for ThnR, suggesting a potential specificity toward CoA bearing only its free thiol. ThnH had a markedly reduced cleavage rate for prederivatized 4-phosphopantetheine, whereas ThnT easily cleaved all prederivatized pantetheine under the conditions tested (Fig. S3). From these results, it was still uncertain at which stage of CoA processing the side chain is attached to the carbapenam or carbapenem core during thienamycin biosynthesis. Based on the apparent structural homology to the S. fulvoviridis A933 1719 mutant 1501 OA-6129 series of carbapenems, it is reasonable to suggest attachment of the C-2 pantetheinyl side chain occurs before pantetheine cleavage.

To link the truncation of CoA or, more specifically, pantetheine, to potential thienamycin biosynthetic precursors, 2-pantetheinyl carbapenams were selected to test the ability of ThnT to hydrolyze the pantetheinyl side chains after linkage to the β-lactam antibiotic core. Pantetheinylated carbapenams have been isolated in trace amounts from Streptomyces sp. OA-6129 (Fig. 3A) (26). Given that clear homologs of CarB and CarA, which are responsible for the biosynthesis of the carbapenam nucleus, correspond to ThnE and ThnM from the thienamycin gene cluster, we deduced that simple carbapenam formation was likely an early step. Although we now know the precursors of both the C-6 and C-2 side chains, the timing of their attachments, bicyclic ring inversion, and desaturation to the carbapenem are not known. To examine test substrates for the ThnT reaction that presented minimum steric demand but contained the probable carbapenam core, two potential substrates, 2-pantetheinyl carbapenams 1 and 2, were prepared (Fig. 3B). The synthesis proceeded from L-aspartic acid to 2-oxo-carbapenam 3 (R = p-methoxybenzyl) by established procedures (27, 28). Reduction of 3 gave a diastereomeric mixture of alcohols 4 (10:1 ratio), which underwent elimination in the presence of methanesulfonyl chloride and base to the carbapenem 5 (R = p-methoxybenzyl) (29). Heteroconjugate addition of pantetheine acetonide (30) to the carbapenem afforded a 2:1 mixture of (3R)-thioether products (31), which could be separated by HPLC and deprotected with trifluoroacetic acid and anisole (32) to give the (2R,3R,5R)- and (2S,3R,5R)-2-pantetheinyl carbapenams 1 and 2, respectively. Both the desired cysteamyl carbapenams and the β-lactam-hydrolyzed products were detected by dansyl chloride derivatization, followed by HPLC-ESI mass spectrometry (Fig. S4).
In addition to the hydrolyzing activity of the A933 acylase, the enzyme was also shown to acylate the amine of select carbapenems with a variety of acyl-CoAs (14). This acylase was presumed to create acetylated carbapenems such as epithienamycin A and C produced in *S. fulvoviridis* (Fig. 1C). These two enzymatic functions were reported as originating from a single enzyme despite the fact that a mutant strain of *S. fulvoviridis* producing the OA-6129 series of pantetheinylated carbapenems lost the hydrolyzing activity of the A933 acylase and yet retained acylating activity. This result suggests either the A933 acylase was a multidomainal protein or that another protein copurified with the A933 acylase.

By analogy to *S. fulvoviridis*, in *S. cattleya*, the final addition to the thienamycin side chain known to occur is acetylation to yield N-acetylthienamycin, a carbapenem coproduced with thienamycin in wild-type cultures (Fig. 1B). Because ThnT, the presumed A933 acylase homolog, was unable to acylate cysteamine-containing substrates with acetyl-CoA (data not shown), an enzyme capable of capping the reactive primary amine of thienamycin (or more likely a thienamycin precursor) with acetyl-CoA was sought within the cluster. ThnF (GenBank accession no. CAD18974) was found to have weak homology to the GCN5-related N-acetyltransferase (GNAT) superfamily, which utilizes acetyl-CoA to transfer acetyl groups to various primary amine-containing compounds. The GNAT superfamily contains four conserved sequence domains not readily identified in ThnF (Fig. S5). Despite the lack of apparent sequence conservation, *thnF* was cloned and expressed as an N-terminal 6-His fusion in Rosetta2 to test for possible acetyltransferase activity.

Initial efforts to detect acetyltransferase activity with cysteamine as the substrate were unsuccessful. Under the conditions tested, free thiol-containing potential substrates readily underwent transthioesterification with the acetyl group of acetyl-CoA. Cysteamine was especially problematic, however, because the molecule not only underwent transthioesterification, but also intramolecular rearrangement presumably occurred to produce N-acetyl cysteamine in the presence of acetyl-CoA alone. Therefore, model substrates cis- and trans-carboxyethylenecysteamine [cis- and trans-3-(2-aminoethylthio)acrylic acid] were synthesized and tested to eliminate the unwanted reactivity of the substrate itself and to better serve as carbapenem precursors (Fig. 4A).

ThnF was able to efficiently convert both the cis- and trans-carboxyethylenecysteamine to the corresponding acetylated products in the presence of acetyl-CoA (Fig. 4B-D). Acetylation of the *trans*-model substrate was not expected. This result can be partially explained through the observation that spontaneous isomerization of the *cis*- to the *trans*-carboxyethylenecysteamine occurs in acidic aqueous conditions and, therefore, could mask the specificity of ThnF for the two substrates. Nonetheless, a preference for coplanarity of the carboxylic acid with the cysteaminyl side chain seems to be governing ThnF activity because coupled reactions of ThnT and ThnF with the 2-pantetheinyl carbapenems did not yield any acetylated products (data not shown).

![Fig. 3. Natural and synthetic pantetheinyl carbapenems. (A) Carbapenems naturally produced by *Streptomyces* sp. OA-6129. (B) Scheme for the synthesis of 2-pantetheinyl carbapenems to mimic possible biosynthetic intermediates susceptible to ThnT.](image-url)

![Fig. 4. ThnF model reactions and HPLC analysis. (A) ThnF model reactions with cis- and trans-carboxyethylenecysteamine based on the two predominant carbapenems produced by *S. cattleya*. (B–D) HPLC analyses of ThnF reactions with the model substrates and acetyl-CoA. (B) Reaction control without enzyme. (C) ThnF reaction. (D) HPLC trace of synthetic cis- and trans-carboxyethylenecysteamine product standards.](image-url)
Table S1. Genes DNA (NRRL 8057) by using PCR primers (Sigma–Genosys) listed in cattleya

Malonyl-CoA is labeled as Mal-CoA. a C-6 thienamycin side chain of hydrogen, methyl, ethyl, or hydroxyethyl.
ent substrate preferences of enzymes ThnR, ThnH, ThnT, and ThnF. R denotes

Materials and Methods

Cloning. Genes thnR, thnH, thnT, and thnF were amplified from genomic S. cattleya DNA (NRRL 8057) by using PCR primers (Sigma–Genosys) listed in Table S1. Genes thnR, thnT, and thnF were cloned into pET28b(+) (Novagen) to create N-terminal 6-His tag fusions, whereas thnH was ligated into pET29a(+) to create a C-terminal 6-His tag construct. Positive clones were transformed into E. coli Rosetta2(DE3)-competent cells (Novagen).

thnR Codon Optimization. The first 15 codons of thnT were optimized for expression in E. coli by using a DraIII cut site within the thnT sequence. Primers thnT-codon-F and thnT-codon-R were annealed to create a DNA linker with sticky ends for Ndel and DraIII cut sites. The DNA linker and appropriately cut pET29b plasmid were then ligated at approximately a 5:1 molar ratio and subsequently transformed into DH5α-competent cells for sequence analysis.

thnH Codon Optimization. The entire thnH sequence was synthetically codon-optimized by using the PCA method (24, 35). The web site http://software.kosan.com/GeMS was used to codon optimize thnH as well as split the gene into 40-bp strands with 20-bp overlaps. Gene thnH was prepared in two synths of ~375 bp to maximize success of the PCA reaction. The 50-μl PCA mixtures contained 20 mM Tris–HCl (pH 8.8), 10 mM KCl, 6 mM (NH₄)₂SO₄, 0.1% (vol/vol) Triton X-100, 0.1 mg/ml bovine albumin, 0.2 mM each dNTP, 1.25 units Turbo DNA polymerase (Stratagene), 2.0 ng/μl concentrations of each primer, and 5% (vol/vol) DMSO and was run at 96°C for 9 min (hot start after 2 min), with 25 cycles at 96°C for 30 s, gradient annealings of 45–54°C for 30 s, and 72°C for 1 min. Subsequent PCRs were run with the outermost primers for an additional 35 cycles with 2 μl from each PCA reaction to further amplify each synthon. To create full-length thnH, 1 μl of each synthon PCR was amplified in another PCR with the only outside primers of the gene.

Protein Expression and Purification. ThnR, ThnT, and ThnF were expressed in E. coli Rosetta2(DE3) cells with 1 mM IPTG induction at 18°C for 16 h. ThnH was expressed in E. coli B121(DE3) pLysS cells. Nickel-NTA resin (Qiagen) was used to purify each protein according to the manufacturer’s specifications.

Coupled Enzyme and ThnF Activity Assays. All 500-μl reactions were performed in 50 mM Tris–HCl (pH 7.5) and incubated for 3 h at 37°C before derivatization. ThnR (10 μM), 2.0 μM ThnH, and 2.0 μM ThnT were incubated in various combinations with 2.5 mM CoA and 10.0 mM MgCl₂. ThnF activity assays contained 10 μM enzyme, 2.0 mM concentrations each of cis- and trans-carboxybethylyeneacysteamine, 2.5 mM acetyl-CoA, and 10.0 mM MgCl₂ in 100 mM Tris (pH 7.5).

Thio Derivatization and HPLC Analysis. Thiol-containing compounds were derivatized with CMQT, which was synthesized as described (20). Briefly, a typical derivatization reaction contained 200 μl of 2.5 mM thiol-containing compound and 100 μl of 25 mM CMQT in 100 mM Tris (pH 7.5). The reaction was mixed and acidified after 5–10 min with 40 μl of 72% (wt/vol) trichloroacetic acid. Standards (10 mM) of CoA, 4-phosphopantetheine, pantetheine, and cysteamine were derivatized and independently injected onto an Agilent 1100 Series HPLC in 20-μl aliquots onto a Phenomenex Luna 5-μm C18 100 Å (250 10.0 mm) column. A flow rate of 1.5 ml/min was used with a mobile phase consisting of acetonitrile (solvent A) and 10 mM NH₄HCO₃ (solvent B) adjusted with NaOH. A method of 95–85% solvent B from 0 to 20 min, 30% solvent B at 40 min, and 95% solvent B from 45 to 50 min was used for separation of all analytes. Monitoring at 355 nm, the retention times of each derivatized standard were: CoA-CMQT, 8.2 min; 4-phosphopantetheine-CMQT, 40.10 min; pantetheine-CMQT (C21H30N3O4S), 499.98 min; and cysteamine-CMQT at 36.9 min. Masses for the derivatized samples were: CoA-CMQT (C31H44N8O16P₃S), 909.18, synthetic standard (m/z 909.13), observed (m/z 909.04); 4-phosphopantetheine-CMQT (C₁₇H₂₃N₃O₁₀P₃S), theoretical (m/z 500.16), synthetic standard (m/z 499.98); pantetheine-CMQT (C₁₇H₂₃N₅O₅S), theoretical (m/z 420.20), synthetic standard (m/z 420.23); cysteamine-CMQT (C₁₇H₂₃N₅S), theoretical (m/z 219.10), synthetic standard (m/z 218.46), observed (m/z 218.11).

Synthesis of Substrates. The 4-phosphopantetheine was synthesized essentially as described by Mandel (36). cis- and trans-carboxybethylyeneacysteamine, as well as cis- and trans-carboxybethylyene-N-acetylcysteamine [cis- and trans-3-(2-acetamidoethylthio)acrylic acid] were prepared according to published procedures (37). Details of the syntheses of (2R,3S,5R)- and (2S,3R,5R)-2-pantetheinyl carbanepens 1 and 2, respectively, can be found in SI Text and Fig. S6.

ThnF HPLC Assay. Alliquots (100 μl) were injected onto a Phenomenex Luna 5-μm C18 100 Å (250 10.0 mm) column. A flow rate of 1.5 ml/min was used with a mobile phase of acetonitrile (solvent A) and dH₂O with 0.1% (vol/vol)
ACKNOWLEDGMENTS. We thank Prof. Maurice J. Bessman and Dr. Wenlian Xu for helpful discussions and technical assistance with Nudix hydroxylases. MALDI-TOF mass spectrometry was performed by the Proteomics Core Facility at The Johns Hopkins University. This work was supported by National Institutes of Health Research Grant AI014937.


NMR (400 MHz, CDCl₃) δ 8.0 (dd, J = 13.6, 9.2, 8.4 Hz, 1H), 2.31 (ddd, J = 13.6, 5.2, 1.6 Hz, 1H), 2.62 (dd, J = 15.6, 1.6 Hz, 1H), 3.31 (dd, J = 15.6, 4.8 Hz, 1H), 3.79 (s, 3H), 4.11 (br m, 1H), 4.50 (d, J = 5.2 Hz, 1H), 4.93 (ddd, J = 4.4 Hz, 1H), 5.13 (ABq, J = 12.0 Hz, 2H), 6.87 (d, J = 8.8 Hz, 2H), 7.29 (d, J = 8.8 Hz, 2H); δ 13C NMR (400 MHz, CDCl₃) δ 40.22, 42.45, 52.20, 55.22, 65.36, 66.99, 79.29, 113.96, 127.31, 130.14, 159.73, 168.91, 176.42.

Addition of Panetheine Acetamide 9 to p-Methoxybenzyl Carbapenam 5. The procedure of Bateson was used (31). To a solution of p-methoxybenzyl carbapenam 5 (175 mg, 0.64 mmol) in DMF (1 ml) was added panetheine acetamide 9 (224 mg, 0.70 mmol) (30) and Et₃N (135 µl, 0.96 mmol). After 2 h, the solution was diluted with saturated NaHCO₃ (20 ml) and extracted three times with 5 ml of 0.1 N acetic acid. The combined organic layers were dried with brine and sodium sulfate and then concentrated. The resulting yellow oil (408 mg) was a 2:1:1 mixture of triethers 6-8 determined by correlation of the 1H NMR chemical shift of the C₃ proton with the literature ([2S,3R]-6 4.37; [2R,3R]-7 4.71; [2R,3S]-8 4.09). The diastereomers were separated by HPLC [Phenomenex Luna 5µ silica (2) 100A 250 × 10 mm 5µ, 3% methanol in ethyl acetate mobile phase, observed at 265 nm, retention time (min): pantetheine acetamide 9, 13.4; 6, 15.2; 7, 17.7, 18.25].

(2S,3R,5R) p-Methoxybenzyl-2-(2-(3,2,5,5-Tetramethyl-1,3-Dioxane-6-Carboxamido) Propanamido)Ethyl(hio)Carbapenam (5): 1H NMR (400 MHz, CDCl₃) δ 0.97 (s, 3H), 1.04 (s, 3H), 1.42 (s, 3H), 1.46 (s, 3H), 1.62 (ddd, J = 14.4, 6.8 Hz, 1H), 2.43 (t, J = 6.4 Hz, 2H), 2.60 (sym m, 2H), 2.74 (ddd, J = 14.4, 7.6, 1.6 Hz, 1H), 2.82 (dd, J = 16.0, 2.0 Hz, 1H), 3.26 (d, J = 11.6 Hz, 1H), 3.30 (d, J = 16.0, 5.2 Hz, 1H), 3.35 (m, 2H), 3.67 (d, J = 11.6 Hz, 1H), 3.72 (ddd, J = 7.6, 6.8, 5.2 Hz, 1H), 3.81 (s, 3H), 3.89 (sym m, 1H), 4.07 (s, 1H), 4.37 (d, J = 5.2 Hz, 1H), 5.12 (ABq, J = 12.0 Hz, 2H), 6.23 (br t, 1H), 6.89 (d, J = 8.8 Hz, 2H), 7.03 (br t, 1H), 7.30 (d, J = 8.8 Hz, 2H); δ 13C NMR (400 MHz, CDCl₃) δ 16.88, 18.88, 22.11, 29.45, 32.22, 39.24, 34.77, 35.89, 38.32, 44.37, 51.96, 52.47, 55.28, 66.29, 67.33, 71.41, 71.73, 99.03, 114.03, 127.16, 130.25, 155.50, 160.50, 169.77, 170.03, 171.16, 175.46.
was taken up in water (10 ml) and 1 equivalent of KHCO₃ was added. The solution was washed two times with 3 ml of ethyl acetate and lyophilized to give the desired product as a white solid.

(2S,3R,5R) Potassium-2-(2-(3-(2,4-Dihydroxy-3,3-Dimethylbutamido-Propanamido) Ethylthio)Carbapenam 1: \(^1\)H NMR (400 MHz, CDCl₃) \(\delta\) 0.91 (s, 3H), 0.94 (s, 3H), 1.86 (ddd, \(J = 13.6, 9.6\) Hz, 1H), 2.53 (t, \(J = 6.4\) Hz, 2H), 2.80 (ddd, \(J = 13.6, 6.8\) Hz, 1H), 2.90–3.05 (m, 3H), 3.40–3.58 (m, 7H), 3.77 (ddd, \(J = \sim7.2\) Hz, 1H), 4.01 (s, 1H), 4.15 (sym m, 1H), 4.19 (d, \(J = 6.0\) Hz, 1H).

(2R,3R,5R) Potassium-2-(2-(3-(2,4-Dihydroxy-3,3-Dimethylbutamido-Propanamido) Ethylthio)Carbapenam 2: \(^1\)H NMR (400 MHz, CDCl₃) \(\delta\) 0.91 (s, 3H), 0.95 (s, 3H), 2.29 (m, 1H), 2.53 (t, \(J = 6.4\) Hz, 2H), 2.88 (m, 2H), 2.98 (m, 1H) 3.38–3.62 (m, 7H), 3.98 (br m, 1H), 4.02 (s, 1H), 4.38 (br m, 1H), 4.61 (d, \(J = 5.6\) Hz, 1H).

Fig. 1. The autocatalytic cleavage site of ThnT. (A) Sequence alignment of probable intramolecular cleavage sites, indicated with solid arrows, of members from the DmpA/OAT superfamily. Corresponding amino acid numbers flank the appropriate sequences. Proteins aligned: DmpA (GenBank accession no. CAA66259) from Ochrobactrum anthropi; BapA-PS (GenBank accession no. BAE02664) from Pseudomonas sp. MCI3434; BapA (GenBank accession no. AAX93858) from Sphingosinicella xenopeptidilytica; NylC (GenBank accession no. BAA05088) from Flavobacterium sp. KI723T1; M.tur. (GenBank accession no. CAA98097) from Mycobacterium tuberculosis H37Rv; M.lep. (GenBank accession no. AAA50889) from Mycobacterium leprae; ThnT (GenBank accession no. CAD18988) from Streptomyces cattleya. (B) SDS/PAGE (15%) of ThnT autocatalytic cleavage time course at room temperature. Open arrow indicates full-length, unprocessed Nhis-ThnT (41,321.08 Da), and solid arrows indicate predicted cleavage products of 29,229.4 Da and 12,109.7 Da, respectively. Lane 1: Bio-Rad Precision Plus Protein Standards (molecular masses of pertinent standards labeled on the gel); lanes 2–10: corresponding cleavage time points denoted in hrs. Total protein (6.7 μg) loaded per lane.
**Fig. S2.** Nudix box and NuCoA motifs of CoA pyrophosphatases compared with ThnR sequence. X represents any amino acid, and U denotes a bulky hydrophobic amino acid, usually isoleucine, leucine, or valine. Differences in ThnR motifs are highlighted in red.

Nudix Box: \( GX_3EX_3REUXXEXGU \)
ThnR Nudix Box: \( GX_3DX_3RESXEXGU \)

NuCoA motif: \( LLTXR(SA)X_3RX_3GX_3FPGG \)
ThnR NuCoA motif: \( LVXR(SR)X_3RX_3DX_3FPGG \)
Fig. S3. Enzymatic reactions of substrates prederivatized with CMQT run for 3 h at 37°C. (A) ThnR with derivatized CoA. (B) ThnH with derivatized 4-phosphopantetheine. (C) ThnT with derivatized pantetheine. ThnR was unable to accept CMQT-CoA as a substrate, whereas ThnH minimally produced CMQT-pantetheine from CMQT-4-phosphopantetheine. Only ThnT was able to efficiently accept and cleave a thiol-derivatized substrate. The retention times of each derivatized standard were: CoA-CMQT, 8.2 min; 4-phosphopantetheine-CMQT, 11.7 min; pantetheine-CMQT, 18.9 min; and cysteamine-CMQT at 29.3 min.
Fig. S4. HPLC analyses of ThnT reactions with cis- and trans-pantetheinyl-carbapenams. Reactions (200 μl) containing 2 μM ThnT and 2.5 mM substrate in 100 mM Tris (pH 7.5) were run for 1–3 h at 37°C. Aliquots (100 μl) were derivatized with 100 mg/ml (wt/vol) dansyl-Cl and 100 μl 0.1 M NaHCO₃ and incubated for 1 h at ambient temperature. The resulting samples were then 0.2 μm filtered, and 30 μl was loaded onto a Phenomenex Luna 5 μm Phenyl-Hexyl 100 Å (250 × 10.0 mm) column. A flow rate of 1.0 ml/min. was used with solvents acetonitrile (solvent A) and dH₂O with 0.1% (vol/vol) TFA (solvent B). Monitoring at 254 nm, a method of 95% solvent B from 0–5 min, 50% solvent B at 25 min, 5% solvent B from 35–45 min, and 95% solvent B from 50–60 min was used for optimum separation of all analytes. The trans-cysteaminyl-carbapenam eluted from the column at 24.1 min with the corresponding β-lactam hydrolyzed product at 17.5 min, whereas the cis-cysteaminyl-carbapenam eluted at 23.7 min and its hydrolyzed product at 17.0 min. (A) ThnT reaction with the cis-pantetheinyl-carbapenam. (B) Cis-pantetheinyl-carbapenam without enzyme. (C) ThnT reaction with the trans-pantetheinyl-carbapenam. (D) Trans-pantetheinyl-carbapenam without enzyme. (E) ThnT reaction without substrate. Cysteamine-containing products are highlighted in red.
Fig. S6. Synthesis of 2-pantetheinyl carbapenams. Reagents: NaBH₄, THF/MeOH, −78°C, 80%, (4a:4b; 10:1 ratio) (a); Et₃N, MsCl, CH₂Cl₂, 80% (b); DMF, Et₃N, 9, 86%, (6:7:8; 2:1:1 ratio) (c); CH₂Cl₂, TFA, anisole, 0°C, 90% (d).
<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ThnTexp-codon-F</td>
<td>5'-CAAGACATATGGACCGGAGAGCTGGTTCGTCGCTC-3'</td>
</tr>
<tr>
<td>F-thnT-exp</td>
<td>5'-GATGATCATCCATATGGACGGCCCGAGGCCGG-3'</td>
</tr>
<tr>
<td>R-thnT-exp</td>
<td>5'-CGGAAATTTTCAGGACGGGCAAGCTGC-3'</td>
</tr>
<tr>
<td>thnT-codon-F</td>
<td>5'-TATGGACCGGAGAAGCTGGTTCGTCGTCGCTGTTGAAAGGTTCGTCGCTGACCCGG-3'</td>
</tr>
<tr>
<td>thnT-codon-R</td>
<td>5'-GTGAAACATCTCCGAGACAGCAGAGACAGACGATCTCCGACGTTGACC-3'</td>
</tr>
<tr>
<td>ThnR-F</td>
<td>5'-CAAGACATATGGCGCAGCGGGTGTCG-3'</td>
</tr>
<tr>
<td>ThnR-R</td>
<td>5'-CCAAGACATATGGAGCGACGGACGG-3'</td>
</tr>
<tr>
<td>ThnR-link-F</td>
<td>5'-CACTATCTCAGGAGGACGACAAGCGCGCCGGC-3'</td>
</tr>
<tr>
<td>ThnR-link-R</td>
<td>5'-CCAGGACATATGGACGGCGCAGGTCGACGCTCGGATGACCTGACGGG-3'</td>
</tr>
<tr>
<td>thnH-opt-1F</td>
<td>5'-GCCCTACGAGCACCAATTACGCATATGCGTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-2F</td>
<td>5'-GCCAGCTTCTACCTCTGCTGAGCATCTGCTGAGAGATGAGGTGGTTGATGATCCCTGCTG-3'</td>
</tr>
<tr>
<td>thnH-opt-3F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-4F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-5F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-6F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-7F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-8F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-9F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-10F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-11F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-12F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-13F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-14F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-15F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-16F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-17F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-opt-18F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-codon-F</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-codon-R</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnH-codon-chis-R</td>
<td>5'-GCCCTACGAGCACGCGCAGGCCTCTATCTACGCTGACCACGGGTACCCGGGAGGTGTTTGC-3'</td>
</tr>
<tr>
<td>thnF-F</td>
<td>5'-CAAGACATATGGTTTGCAGGACATCTGTCGCTACCT-3'</td>
</tr>
<tr>
<td>thnF-R</td>
<td>5'-CAAGACATATGGTTTGCAGGACATCTGTCGCTACCT-3'</td>
</tr>
<tr>
<td>ThnF-chis-R</td>
<td>5'-CAAGACATATGGTTTGCAGGACATCTGTCGCTACCT-3'</td>
</tr>
</tbody>
</table>