Substrate Recognition by the Peptidyl-(S)-2-mercaptoglycine Synthase TglHI during 3-Thiaglutamate Biosynthesis

Martin I. McLaughlin, Yue Yu, and Wilfred A. van der Donk*

Cite This: ACS Chem. Biol. 2022, 17, 930−940

Introduction

The genome of the plant pathogen Pseudomonas syringae pv. maculicola str. ES4326 contains a gene cluster (tgl) encoding the biosynthetic machinery for the production of 3-thiaglutamate, a founding member of the recently discovered family of natural products known as pearlins (Figure 1).

Although the physiological role of 3-thiaglutamate is still unknown, its structural similarity to glutamate suggests that it may function as an antimetabolite to interfere with glutamate signaling in plants; other pearlins are reported to have antibacterial activity (ammosamide C) and immunosuppressive properties (lymphostin). Pearlins biosynthesis has commonalities with that of ribosomally translated and posttranslationally modified peptides (RiPPs), although the pearlins themselves are not ribosomally synthesized. During pearlin and RiPP biosynthesis, a precursor peptide undergoes a series of posttranslational modifications in which the biosynthetic enzymes use a RiPP recognition element (RRE) for substrate recognition (Figure 1a). Pearlins biosynthesis is defined by a unique biosynthetic step in which a precursor peptide of ribosomal origin is nonribosomally elongated in an adenosine 5′-triphosphate (ATP)- and aminoacyl-tRNA-dependent manner by an enzyme known as a peptide aminoacyl-tRNA ligase (PEARL; Figure 1b).

PEARLs are related to the glutamyl transfer domains of LanB dehydratases, which catalyze the condensation of glutamate (from Glu-tRNA\(^{\text{Glu}}\)) to the side chains of Ser and Thr residues in the precursor peptide during class I lanthipeptide biosynthesis. In the pearlin pathways characterized to date, subsequent modifications of the PEARL-appended amino acid(s) are followed by proteolysis of the nonribosomal peptide bond, yielding the original precursor peptide and a mature natural product containing no ribosomally incorporated material. The regenerated precursor may then serve as a scaffold for another round of biosynthesis (Figure 1b), a strategy reminiscent of other secondary metabolite assembly systems where intermediates are tethered to carrier proteins as thioesters (polyketide synthases, nonribosomal peptide synthetases, and the recently characterized closthioamide biosynthetic machinery) or amides (noncanonical amino acid biosynthetic systems in actinomycetes). The use of a catalytic scaffold may also save energy compared to the stoichiometric precursor peptide synthesis used for traditional RiPPs.
During 3-thiaglutamate biosynthesis, TglB, the PEARL encoded in the \textit{tgl} gene cluster, catalyzes peptide bond formation between L-cysteine (as cysteinyl-tRNACys) and the C-terminus of the 50-residue precursor peptide TglA. The scaffolded peptidyl-cysteine residue then undergoes a net four-electron oxidation consisting of excision of the C\textsubscript{\beta} methylene group, oxidation of C\textsubscript{\beta} to formate, and joining of the thiol group directly to the \alpha position to form an (S)-2-mercaptoglycine residue with retention of stereochemical configuration at C\textsubscript{\alpha}. This complex set of reactions is catalyzed by the proteins TglH and TglI, which are insoluble separately but form a soluble TglHI complex upon coexpression of His\textsubscript{6}-TglH with TglI in \textit{Escherichia coli}. TglH, the putative catalytic subunit, is a member of the DUF692 protein family (Pfam: PF05114), which also includes the nonheme iron oxidase MbnB, the catalytic subunit of the MbnBC complex that generates the oxazolone and thioamide moieties of methanebactin. The use of nonheme iron oxygenases and other radical-utilizing enzymes to form and break C–S bonds has been observed during the biosynthesis of a variety of secondary metabolites, such as sactipeptides and other cyclic thioether-containing RiPPs, ergothioneine, and quinohemoprotein amine dehydrogenase. The reactions catalyzed by TglHI and MbnBC are unique, however, and their mechanisms are incompletely understood. TglI bears no sequence homology to MbnC or other proteins of known function, but residues 171–269 are predicted by a sequence-based profile hidden Markov model to constitute an RRE. RRE domains are often encoded in RiPP gene clusters, where they generally mediate recognition of precursor peptides by their posttranslational modification enzymes. TglB and the peptidase TglG in the \textit{tgl} gene cluster are also predicted to contain RREs, both of which are unrelated by sequence to the TglI RRE. TglI has been shown to bind a 20-residue C-terminal fragment of TglA as well as full-length TglA variants with single substitutions in several of the C-terminal eleven residues. Cysteinylination was observed on fragments as small as 12 residues. In contrast, preliminary activity experiments with TglHI indicated partial activity toward a 41mer TglACys C-terminal fragment lacking residues 1–10 and no activity toward further N-terminally truncated 31mer and 21mer fragments. Considered together, these results suggest that the RREs of TglB and TglI might have evolved to recognize entirely different regions of the TglA precursor peptide—a natural “hybrid” of two different RiPP posttranslational modification systems similar to other examples that are found naturally and to the artificial RiPP hybrids constructed recently in several laboratories. However, from a practical perspective, a 41mer or 51mer TglACys substrate would complicate detailed structural and mechanistic investigations of the highly unusual TglHI reaction. We therefore aimed to investigate TglHI-TglA recognition using...
TglACys analogs and determine the minimal region of the TglACys peptide required for TglHI activity.

RESULTS AND DISCUSSION

TglHI was expressed and purified as previously reported. Iron quantification using Ferene S determined that different preparations of protein contained 0.7–0.9 equiv of Fe as isolated, which increased to 2.3–2.7 equiv after reconstitution with excess Fe and ascorbate in an anaerobic chamber. These results are consistent with the previous report of 2.5 equiv of iron after reconstitution. Both preparations catalyzed the same transformation and as-isolated TglHI was used for all in vitro assays as it provided the most consistent iron content.

We initially set out to use binding assays to identify the approximate region(s) of TglA important for TglHI recognition. However, TglA fluorescently labeled at its N-terminus did not show appreciable binding to TglHI by fluorescence polarization experiments, presumably because the label is relatively far removed from the binding site on TglA. Rather than preparing and testing a series of fluorescently labeled truncated peptides, we decided to use activity assays to determine the important regions on the substrate for catalysis. Wild-type TglA and a set of variants with overlapping (Ala) substitutions collectively spanning residues 6 through 45 (sequences in Table S1) were purchased as synthetic peptides. Each synthetic TglA peptide was incubated in vitro in one pot with TglB, P. syringae tRNA\textsuperscript{Cys}, cysteinyl-tRNA synthetase (CysRS), and cysteine to attach the C-terminal Cys as well as with TglHI and O\textsubscript{2} (in ambient air). After overnight reaction at room temperature, the products were analyzed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (Table S2). Consistent with previous results with truncated peptides,\textsuperscript{7} TglB added Cys to all TglA variants except TglA\textsuperscript{[38–45]}. TglHI catalyzed the oxidation of the C-terminal cysteine added to the wild-type TglA and (Ala)\textsubscript{s} variants TglA\textsuperscript{[6–13]A} through TglA\textsuperscript{[26–33]A} but not TglA\textsuperscript{[30–37]} or TglA\textsuperscript{[34–41]} (Figure 2a). Because TglB did not cysteinylate TglA\textsuperscript{[38–45]}, TglHI activity could not be assessed for this peptide; instead, TglACys\textsuperscript{[38–45]} was synthesized on a small scale from a double-stranded DNA (dsDNA) template by in vitro transcription-coupled translation (IVT)\textsuperscript{36} using a commercially available kit\textsuperscript{37} (see the Methods section) and then reacted with TglHI for 2 h. No turnover was visible by MALDI-TOF MS (Figure 2b).

The observed activity of TglHI with TglACys\textsuperscript{[26–33]} but not with TglACys\textsuperscript{[30–37]}, TglACys\textsuperscript{[34–41]}, or TglACys\textsuperscript{[38–45]} suggested that residues C-terminal to Glu\textsubscript{34} of TglACys might be necessary for TglHI recognition. To determine if these residues are also sufficient for catalysis, short DNA templates encoding an initiator methionine followed by C-terminal fragments of TglACys beginning with residues Glu\textsubscript{29} (23mer) through Asp\textsubscript{40} (12mer) were constructed from synthetic oligonucleotides and the corresponding peptides were prepared by IVT (Figure S1; see Table S3 for sequences). TglHI activity was detectable in vitro toward the 17mer and increased with peptide length, nearly depleting the 20mer substrate in 2 h (Figure 3). The 19mer fragment (20 residues total including the N-terminal Met) was then used as a starting point for subsequent mutagenesis experiments. This peptide is shorter than the heterologously expressed 31mer that in a previous study was not converted by TglHI.\textsuperscript{1} The previous study was conducted at 100 \mu M peptide, whereas IVT results in much lower concentrations. It is possible that

Figure 2. TglHI requires the C-terminal region of TglACys for activity. (a) Activity of TglB and TglHI in a coupled assay with synthetic TglA peptide variants. The assay mixture also contained Cys and P. syringae CysRS and tRNA\textsuperscript{Cys}. Pairs of overlaid MALDI-TOF mass spectra depict each substrate (gray) and the product of its overnight reaction with TglB and TglHI (black). Formation of
TglACys (orange) from TglA (blue) indicates TglB activity; conversion of TglACys to TglACys—CH₂ (yellow) indicates TglHI activity. (b) Activity assay of TglHI with TglACys[38–45] generated by IVT, colored as in (a). Translation initiation with either methionine or N-formylmethionine (Met) produces two TglACys species with a mass difference of 28 Da.

Figure 3. TglHI activity toward C-terminal fragments of TglACys generated by IVT. (a) Sequence of the longest fragment tested (23mer) with the F[N/D]LD motif in blue and the N-terminal Met introduced during translation initiation in red. Start sites (excluding the required N-terminal Met for each peptide) of the longest and shortest fragments tested are marked with lines; the 19mer was used for subsequent mutagenesis. (b) Pairs of MALDI-TOF mass spectra depicting IVT-generated C-terminal fragments of TglACys (gray) and the products of 2 h in vitro reactions with TglHI (black). See Figure S1 for MS data with peptides shorter than the 16mer and longer than the 20mer.

Considering the importance of the F[N/D]LD sequence motif for LanB binding to class I lanthipeptide precursors,\(^8,38,39\) and the presence of this motif in the TglA sequence (Figure 1a), single and double alanine substitutions in Phe36–Asp39 of TglACys were generated by IVT and investigated for their effects on TglHI activity in vitro. Turnover was abolished on the F36A variant and barely detectable for L38A, but the D37A and D39A variants were converted at comparable efficiency to the original 19mer (Figure 4a). Of the six possible double variants, TglHI did not act on the three containing the F36A mutation; activity was barely detectable toward D37A/L38A, L38A/D39A, and D37A/D39A (Figure 4b). Taken together, these results indicate that Phe36 is required and Leu38 is important for TglACys recognition by TglHI, explaining the lack of TglHI activity toward the full-length 30–37A and 34–41A peptides (Figure 2). The presence of either Asp37 or Asp39 is also important, but substitution of one residue is moderately tolerated if the other is present. These findings reveal the similarity between the peptide recognition modes of TglHI and class I lanthipeptide synthetases (LanBs) such as NisB, which recognize the F[N/D]LD motifs of LanA precursor peptides primarily through hydrophobic interactions, especially with the Phe and Leu residues of the motif.\(^5,40–42\)

Notably, this recognition mode is not shared by TglB, which is closely related to LanBs, but accepts a 12mer TglA substrate that only contains the final Asp residue of the motif.\(^1\)

Additional recognition determinants in the TglACys sequence were identified in the same manner through a systematic alanine replacement scan of the 19mer as well as insertions and deletions in residues 46–50. No single-alanine substitution outside the F[N/D]LD motif completely abolished modification by TglHI, but only the E42A, V43A, E45A, and S46A variants were well tolerated; mutations in the remaining residues caused a significant reduction in activity (Table 1 and Figure S2). Alanine insertions at positions 46 and 47 resulted in partial turnover, whereas insertions at positions 49 and 50 abolished modification and only slight modification was observed with the Ins48A mutant; deletions at any of the five C-terminal positions were not tolerated (Table 1 and Figure S3). The identity of Ala50, the residue flanking the Cys that TglHI modifies, was also important for recognition, with replacements with Ser, Gly, or Val leading to low levels of modification and substitutions with Phe, Lys, or Asp preventing modification entirely (Figure 5). Thus, it appears that the recognition of TglACys by TglHI involves both the Phe and Leu of the F[N/D]LD motif remote from the site of modification and additional residues nearer the Cys that is modified.

We also investigated whether the binding site of TglACys is predominantly in the RRE-containing TglI protein, or the active site of the TglH protein, or both. Because TglH and TglI could not be expressed individually and could not be separated after expression and purification, native mass spectrometry (nMS) was used to observe the intact TglHI complex and partially dissociate the subunits. Analysis of TglHI by nMS in the presence of the substrate TglACys led to the detection of the TglHI complex bound to what appears to be the product of the reaction (Figure 6a) as the increase in mass compared to the apo complex was 7125.65 Da (average mass of N-terminally His₆-tagged TglACys—CH₂ = 7126.58 Da). When the TglHI complex was incubated with TglA (i.e., lacking the C-terminal Cys), the analogous complex was observed (Figure 6b). Hence, the Cys does not appear to be required for binding.
to the TglHI heterodimer. The monomers corresponding to TglH and TglI were also observed in the spectrum, but they were not bound to the scaffold peptide. Therefore, it appears that peptide binding to TglHI involves extensive interactions with both TglH and TglI even in the absence of Cys51.

To provide a visual approximation of the TglHI complex and its interaction with the substrate, the AlphaFold-Multimer algorithm was used to predict the structure of a 1:1 heterodimer of TglHI both with and without the TglACys 19mer (Figure S4a). In the models, TglH did not contain iron, but in the trimeric complex, the C-terminal Cys of the 19mer was still located in the vicinity of the iron ligands (based on the structure of another DUF692 family member, PDB 3BWW, and a recent study on the DUF692 enzyme MbnB). The predicted structure of apo-TglH in the complex with TglI is similar to that of MbnB, illustrating the capability of AlphaFold (the prediction was performed before the MbnB structure was reported). MbnB contains three irons in the crystal structure, and the MbnABC complex shows threading of the substrate MbnA from MbnC to the active site of MbnB where a cysteine in the core peptide makes a direct contact with one of the irons. In the predicted TglHI complex with the 19mer peptide of TglA, a very similar interaction is seen that starts at TglI and ends in the active site of TglH. In the AlphaFold model, the C-terminal domain of TglI adopts an RRE fold (three antiparallel \( \beta \)-strands followed by three \( \alpha \)-helices), and the substrate TglA binds to the RRE; MbnC does not contain a canonical RRE fold, but its C-terminal domain contains a \( \beta\)-
sheet that makes an antiparallel β-sheet interaction with the MbnA leader peptide similar to the way substrate binds to β3 in RREs in other structurally characterized RiPP systems. Thus, the predicted model of the 19mer binding to TglHI recaptures several of the features seen in the crystal structure of the related MbnABC protein–substrate complex.

The TglHI model (with or without the substrate) predicts a distance of 50−55 Å from the putative metal-binding site in TglH to the predicted “hydrophobic cage” formed between β3, α1, and α3 of the RRE that interacts with the Phe residue of the F[N/D]LD motif in the NisB/NisA system (Figure S4b). If these features function similarly in TglHI, the sequence of TglACys that should span this distance, Phe36 to Cys51, would have a maximum extended length of about 56−64 Å (based on contour lengths of 3.5−4.0 Å per residue). Therefore, the predicted structure of the TglHI heterodimer is consistent with a binding mode in which (1) the F[N/D]LD motif of TglACys binds to the RRE of TglI; (2) the C-terminal cysteine binds to the iron center of TglH; and (3) intervening residues such as Asp40, Ile41, and Ile44 make additional key contacts across the TglH–TglI interface, as indicated by sluggish TglHI-catalyzed modification of alanine variants at these positions (Table 1 and Figure S2).

However, the binding pose of TglACys interacting with TglHI predicted by AlphaFold-Multimer deviates significantly from the anticipated interaction. Although the predicted structure shows the TglACys leader peptide making antiparallel β strand interactions with β1 of the TglI RRE as observed in other systems, Phe36 of the F[N/D]LD motif does not interact with the hydrophobic cage, in contrast to the NisA–NisB complex. Instead, Phe36 interacts with an extended loop between β1 and β2 of the RRE that is closer to the TglH active site, and Phe33 of TglACys interacts with the hydrophobic cage (Figure S4d,e). The length of this loop is unique to TglI among structurally characterized RREs (Figure S4c). The predicted binding model does not explain several of our experimental observations. First, Phe33 is less critical than Phe36 for modification of TglACys by TglHI (Table 1 and Figure 4). Second, the interaction of Phe36 with the loop in TglI rather than the hydrophobic cage means that the distance between Phe36 and the active site would be only 35−40 Å (Figure S4b). If this prediction is correct, then the observation that TglHI did not process the deletion mutants cannot be due to the inability of the C-terminal Cys to reach the active site and instead must reflect specific interactions in the substrate tunnel that cannot be made in these variants. Future structural biology studies will be needed to resolve these questions.

To further characterize the role of Cys51 in TglACys recognition by TglHI, a set of TglACys analogs with noncanonical amino acid residues replacing Cys51 was generated by expressed protein ligation (EPL). TglA was expressed as a C-terminal fusion with a temperature-dependent intein and chitin-binding domain (CBD), intein catalysis was induced to generate IVT-generated TglACys 19mer peptide variants (gray) and the products of 2 h in vitro reactions with TglHI (black).

### Table 1. Modification of TglACys 19mer Point Mutants by TglHIA

| variant   | activityb | variant   | activityb |
|-----------|-----------|-----------|-----------|
| F49A      | +         | Ins50A    | −         |
| V48A      | +         | Ins49A    | −         |
| K47A      | +         | Ins48A    | +         |
| S46A      | +++       | Ins47A    | ++        |
| E45A      | +++       | Ins46A    | +         |
| I44A      | +         | ΔA50      | −         |
| V43A      | +++       | ΔF49      | −         |
| E42A      | +++       | ΔV48      | −         |
| I41A      | +         | ΔK47      | −         |
| D40A      | +         | ΔS46      | −         |
| E35A      | +         |           |           |
| E34A      | ++        | wild-type | +++       |
| F33A      | +         |           |           |

“For MS data, see Figure S2. Activity values reflect estimated substrate conversion based on the intensity of MALDI-TOF MS signals: −, no product visible; +, <33% conversion; ++, 33−67% conversion; ++++, >67% conversion.

https://doi.org/10.1021/acschembio.2c00087
ACS Chem. Biol. 2022, 17, 930−940
homocysteine (Hcy), D-cysteine, or L-penicillamine (Pen) other PEARL enzymes share a common ancestor with LanBs, βTglHI-catalyzed carbon excision from TglACys is strongly dependent on the F[N/D]LD motif comprising residues Phe36−Asp39, which has been shown in class I lanthipeptide systems to mediate substrate recognition by LanBs.7 TglB and other PEARL enzymes share a common ancestor with LanBs, whereas the TglII sequence (including its RRE) is not homologous to any protein with known function. The identity of the C-terminal Cys residue is also critical for catalysis by TglHI; the analogs TglA (lacking the Cys), TglAHcy, TglADCys, and TglAPen are not modified by TglHI in vitro and do not inhibit TglHI-catalyzed modification of TglACys peptides in competition experiments. However, the substrate analog TglASec nearly completely inhibits TglHI at equimolar concentrations with the substrate, though TglASec itself does not undergo Cβ excision.

This study also demonstrates that a synthetic 19mer peptide corresponding to the C-terminus of TglACys is a minimal substrate for TglHI to catalyze its carbon excision reaction. This peptide is significantly shorter than the previously reported length requirement and shows that both TglB and TglHI can act on peptides that consist of the C-terminus of the TglA sequence. For TglB, the minimal requirement is the final 12 amino acids, whereas TglHI requires the final 19 amino acids for catalysis, indicating overlapping but not identical substrate binding requirements. Structure prediction by AlphaFold-Multimer appears to predict the structures of TglHI well based on a recent crystal structure of MbnBC,18 but the predicted mechanism of substrate engagement either suggests a new type of interaction between the RRE and substrate or illustrates shortcomings of the current capabilities for predicting interactions of short peptides with their modifying enzymes.

### CONCLUSIONS

The experiments described herein used IVT for rapid high-throughput generation and characterization of variant TglACys peptide sequences, allowing the systematic identification of substrate residues important for in vitro recognition and β-carbon excision by TglHI. The results of the coupled TglB-TglHI activity assays corroborate previous observations that TglB-catalyzed cysteinylation of the TglA C-terminus does not require the N-terminal 38 residues of TglA.1 In contrast, TglHI-catalyzed Cβ excision from TglACys is strongly dependent on the F[N/D]LD motif comprising residues Phe36−Asp39, which has been shown in class I lanthipeptide systems to mediate substrate recognition by LanBs.7 TglB and other PEARL enzymes share a common ancestor with LanBs, but we cannot rule out a covalent modification during expression. The formation of a TglHI heterodimer and TglHI bound to the substrate suggests that the proteins maintain a folded conformation under the conditions employed.

Figure 6. Binding of His_{50}-TglA-Cys—CH₂ or TglA to purified TglHI shown by nMS. (a) Binding of TglHI to the product of its reaction with His_{50}-TglACys. (b) Binding of TglHI to TglA. The calculated average mass of His_{50}-apo-TglHI is 36,702.4 Da, and the observed mass after deconvolution is 36,701.3 Da. The calculated average mass of TglII is 31,014.3 Da, and the observed mass after deconvolution is 31,143.4 Da. Higher masses of proteins than their calculated molecular weights have been observed previously in nMS and have been attributed to incomplete desolvation, but we cannot rule out a covalent modification during expression. The formation of a TglHI heterodimer and TglHI bound to the substrate suggests that the proteins maintain a folded conformation under the conditions employed.
short, synthetically accessible substrate and a potent inhibitor for TglHI should serve as a springboard for more detailed mechanistic and structural characterization of this remarkable reaction.

### METHODS

For materials, expression and purification of enzymes and substrates, and MS procedures, see the Supporting Information.

**TglB-Coupled TglHI Activity Assays with Synthetic Substrates.** TglA (10 μM) was mixed with 5 μM TglHI, 0.5 μM His₂-TglB, 50 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 5 mM ATP, 2 mM L-cysteine HCl, 1 mM TCEP–HCl, 2.5 μM *P. syringae* tRNA⁵⁵⁵, and 0.5 μM *P. syringae* CysRS in a total volume of 100 μL. Reactions were initiated with His₂-TglB and incubated overnight at 30 °C. After 18 h, reaction mixtures were acidified by addition of trifluoroacetic acid (TFA) to a final concentration of 0.3% (v/v), desalted using C18 ZipTips, and analyzed by MALDI-TOF MS.

**Transcription-Coupled In Vitro Translation (IVT) and TglHI Activity Assays with IVT-Generated Peptides.** DNA templates for IVT (except for TglACys[38−45A], see the Supporting Information) were generated from single-stranded oligonucleotides (Table S4) by two rounds of 16-cycle overlap extension PCR (one round for the TglACys 12mer template) with Taq polymerase. Each first-round PCR mixture served as the template for the second round of PCR by direct 1:100 dilution into the second-round reaction. Primers and annealing temperatures for each template are listed in

---

**Figure 7.** TglACys analogs with noncanonical C-terminal residues. (a) MALDI-TOF mass spectra of in vitro reaction mixtures containing 100 μM TglASec (black), TglAHcy (red), TglADCys (green), or TglAPen (blue) after 4 h at room temperature with or without 10 μM TglHI. The TglASec sample also contained TglA as a major impurity, which is not shown in the spectral window (see also Figure S5a). (b) Inhibition of TglHI by TglASec. MALDI-TOF MS spectra of 50 μM test substrate (top traces; His₂-TglACys at left, TglACys Ac-19mer at right) and room-temperature reactions containing 50 μM substrate, 5 μM TglHI, and either no inhibitor, 50 μM TglASec, or 50 μM TglA are shown. TglHI-catalyzed modification of the substrate is nearly complete after 30 min; however, the addition of TglASec prevents any appreciable turnover during the same time. TglA (lacking Cys) at the same concentration does not inhibit turnover (bottom traces).
In a typical IVT reaction, dsDNA template (3 μM, approx. 6 μg) was mixed with NEB PURExpress in vitro Protein Synthesis Kit reagents A (4 μL) and B (3 μL) on ice in a prelubricated RNase-free 0.6 mL microcentrifuge tube. The reaction was incubated at 37 °C for 1 h on a prewarmed heat block and divided into 2.5 μL aliquots. One aliquot was immediately desalted and analyzed by MALDI-TOF MS; each remaining aliquot was used in a 10 μL activity assay containing 5 μM TglHI in 0.7X reaction buffer (1X: 50 mM Na2HPO4, 300 mM NaCl, 10% (v/v) glycerol, pH 7.6). Assay mixtures were incubated at room temperature in ambient air for 2 h, desalted, and analyzed by MALDI-TOF MS.

**TglHI Activity Assays with Purified Substrates.** In a typical assay, 10 μM TglHI was added to 100 μM of substrate peptide in 1X reaction buffer in a final volume of 100 μL. Inhibition assays contained 50 μM of inhibitor peptide, 50 μM of substrate peptide, and 5 μM TglHI. After initiation with TglHI, reactions were incubated at room temperature (23 °C); at various time points, 20 μL aliquots were withdrawn, desalted, and analyzed by MALDI-TOF MS.

**ASSOCIATED CONTENT**

**Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.2c00087.

Sequence of the codon-optimized TglACys[38–45A] gene, materials used, detailed procedures for protein expression, purification, and activity assays, Figures S1–S7, and Tables S1–S5 (PDF)

**Accession Codes**

NCBI Protein accessions: TglH, WP_007252215.1; TglI, WP_007252216.1; TglIA, WP_007252214.1; TglIB, WP_007252217.1. Nucleotide accessions: *P. syringae* pv. *maculicola* str. ES4326 genome, NZ_CP047260.1.

**AUTHOR INFORMATION**

**Corresponding Author**

Wilfred A. van der Donk — Department of Chemistry and Carl R. Woese Institute for Genomic Biology and Howard Hughes Medical Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States; orcid.org/0000-0002-5467-7071; Email: vddonk@illinois.edu

**Authors**

Martin I. McLaughlin — Department of Chemistry and Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States; Present Address: Department of Bioengineering, Stanford University, Stanford, California 94305, United States; orcid.org/0000-0003-4410-0786

Yue Yu — Department of Chemistry and Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States; orcid.org/0000-0001-8306-4648

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acschembio.2c00087

**Funding**

This study was supported by the National Institutes of Health (Grant R37GM058822 to W.A.v.d.D.). The Bruker UltraflexTreme MALDI-TOF/TOF mass spectrometer was purchased in part with a grant from the National Center for Research Resources, National Institutes of Health (S10 RR027109 A). W.A.v.d.D. is an Investigator of the Howard Hughes Medical Institute (HHMI).

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The authors thank Z. Li and J. Arrington of the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign for performing LC-MS analysis of TglAsEc using a Thermo Q Exactive instrument (Z.L.) and for assistance in acquiring native mass spectrometry data (J.A.). The authors thank D. Nguyen for sharing key expertise in IVT methods. This study is subject to HHMI’s Open Access to Publications policy. HHMI laboratory heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

**ABBREVIATIONS**

4-MPAA, 2-(4-mercaptophenyl)acetic acid; ATP, adenosine 5′-triphosphate; CBD, chitin-binding domain; DTT, DL-dithiothreitol; dsDNA, double-stranded DNA; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); EDTA, ethylenediaminetetraacetic acid; EPL, expressed protein ligation; ESI, electrospray ionization; fMet, 1-[(N-formyl)methylionine; HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; HPLC, high-performance liquid chromatography; IMAC, immobilized metal affinity chromatography; IVT, in vitro transcription-coupled translation; LB, Luria-Bertani medium; MALDI, matrix-assisted laser desorption ionization; MWCO, molecular weight cutoff; nMS, native mass spectrometry; PEARL, peptide aminoacyl-tRNA ligase; PMSF, phenylmethanesulfonyl fluoride; RiPP, ribosomally translated and posttranslationally modified peptide; RRE, RiPP recognition element; SAH, S-adenosyl-L-homocysteine; SAM, S-adenosyl-L-methionine; SPE, solid-phase extraction; ssDNA, single-stranded DNA; TB, Terrific Broth; TCEP, tris(carboxymethyl)phosphine; TFA, trifluoroacetic acid; TOF, time of flight; Tris, tris(2-hydroxyethyl)aminomethane; tRNA 5′ GTP, cysteine transfer RNA

**REFERENCES**

(1) Ting, C. P.; Funk, M. A.; Halaby, S. L.; Zhang, Z.; Gonen, T.; van der Donk, W. A. Use of a scaffold peptide in the biosynthesis of amino acid-derived natural products. *Science* 2019, 365, 280–284.
(2) Baltrus, D. A.; Nishimura, M. T.; Romanchuk, A.; Chang, J. H.; Mukhtar, M. S.; Cherks, K.; Roach, J.; Grant, S. R.; Jones, C. D.; Dangl, J. L. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog.* 2011, 7, e1002132.
(3) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; et al. New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* 2021, 38, 130–239.
(4) Reimer, D.; Hughes, C. C. Thiol-based probe for electrophilic natural products reveals that most of the ammosamides are artifacts. *J. Nat. Prod.* 2017, 80, 126–133.
(5) Nagata, H.; Ochiai, K.; Aotani, Y.; Ando, K.; Yoshida, M.; Takahashi, I.; Tamaoki, T. Lymphostin (LK6-A), a novel immunosuppressant from Streptomyces sp. *KY11783: taxonomy of the producing organism, fermentation, isolation and biological activities. J. Antibiot.* 1997, 50, 537–542.
(43) Evans, R. J.; O’Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Židek, A.; Bates, R.; Blackwell, S.; Yim, J. et al. Protein complex prediction with AlphaFold-Multimer bioRxiv 2021, DOI: 10.1101/2021.10.04.463034.

(44) Chekan, J. R.; Ongpipattanakul, C.; Nair, S. K. Steric complementarity directs sequence promiscuous leader binding in RiPP biosynthesis. Proc. Natl. Acad. Sci. U.S.A. 2019, 116, 24049−24055.

(45) Ainavarapu, S. R.; Brujic, J.; Huang, H. H.; Wiita, A. P.; Lu, H.; Li, L.; Walther, K. A.; Carrión-Vázquez, M.; Li, H.; Fernandez, J. M. Contour length and refolding rate of a small protein controlled by engineered disulfide bonds. Biophys. J. 2007, 92, 225−233.

(46) Koehne, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lehl, T.; Houssen, W. E.; Jaspers, M.; Naismith, J. H. Structural analysis of leader peptide binding enables leader-free cyanobactin processing. Nat. Chem. Biol. 2015, 11, 558−563.

(47) Davis, K. M.; Schramma, K. R.; Hansen, W. A.; Bacik, J. P.; Khare, S. D.; Seyedsayamdost, M. R.; Ando, N. Structures of the peptide-modifying radical SAM enzyme SuiB elucidate the basis of substrate recognition. Proc. Natl. Acad. Sci. U.S.A. 2017, 114, 10420−10425.

(48) Evans, R. L., 3rd; Latham, J. A.; Xia, Y.; Klinman, J. P.; Wilmot, C. M. Nuclear magnetic resonance structure and binding studies of PqqD, a chaperone required in the biosynthesis of the bacterial dehydrogenase cofactor pyrroloquinoline quinone. Biochemistry 2017, 56, 2735−2746.

(49) Grove, T. L.; Himes, P. M.; Hwang, S.; Yumerefendi, H.; Bonanno, J. B.; Kuhlman, B.; Almo, S. C.; Bowers, A. A. Structural insights into thioether bond formation in the biosynthesis of sactipeptides. J. Am. Chem. Soc. 2017, 139, 11734−11744.

(50) Sumida, T.; Dubley, S.; Wilcox, B.; Severinov, K.; Tagami, S. Structural basis of leader peptide recognition in lasso peptide biosynthesis pathway. ACS Chem. Biol. 2019, 14, 1619−1627.

(51) Weerasinghe, N. W.; Habibi, Y.; Uggowitzer, K. A.; Thibodeaux, C. J. Exploring the conformational landscape of a lanthipeptide synthetase using native mass spectrometry. Biochemistry 2021, 60, 1506−1519.

(52) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed protein ligation: a general method for protein engineering. Proc. Natl. Acad. Sci. U.S.A. 1998, 95, 6705−6710.

(53) Chong, S.; Mersha, F. B.; Comb, D. G.; Scott, M. E.; Landry, D.; Vence, L. M.; Perler, F. B.; Benner, J.; Kucera, R. B.; Hirvonen, C. A.; et al. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene 1997, 192, 271−281.