Characterization of CmaA, an Adenylation-Thiolation Didomain Enzyme Involved in the Biosynthesis of Coronatine

Robin Couch,1 Sarah E. O’Connor,2† Heather Seidle,1 Christopher T. Walsh,2 and Ronald Parry1*

Department of Chemistry, Rice University, Houston, Texas 77005,1 and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 021152

Received 7 July 2003/Accepted 1 October 2003

Several pathovars of Pseudomonas syringae produce the phytotoxin coronatine (COR), which contains an unusual amino acid, the 1-amino-2-ethylcyclopropane carboxylic acid called coronamic acid (CMA), which is covalently linked to a polyketide-derived carboxylic acid, coronafacic acid, by an amide bond. The region of the COR biosynthetic gene cluster proposed to be responsible for CMA biosynthesis was resequenced, and errors in previously deposited cmaA sequences were corrected. These efforts allowed overproduction of P. syringae pv. glycinea PG4180 CmaA in P. syringae pv. syringae FFS as a FLAG-tagged protein and overproduction of P. syringae pv. tomato CmaA in Escherichia coli as a His-tagged protein; both proteins were in an enzymatically active form. Sequence analysis of CmaA indicated that there were two domains, an adenylation domain (A domain) and a thiolation domain (T domain). ATP-32PPi exchange assays showed that the A domain of CmaA catalyzes the conversion of branched-chain L-amino acids and ATP into the corresponding aminoacyl-AMP derivatives, with a kinetic preference for 1-allo-isoleucine. Additional experiments demonstrated that the T domain of CmaA, which is posttranslationally modified with a 4'-phosphopantetheinyl group, reacts with the AMP derivative of 1-allo-isoleucine to produce an aminoacyl thiolester intermediate. This covalent species was detected by incubating CmaA with ATP and 1-[G-32H]allo-isoleucine, followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis. It is postulated that the 1-allo-isoleucine covalently tethered to CmaA serves as the substrate for additional enzymes in the CMA biosynthetic pathway that catalyze cyclopropane ring formation, which is followed by thiolester hydrolysis, yielding free CMA. The availability of catalytically active CmaA should facilitate elucidation of the details of the subsequent steps in the formation of this novel cyclopropyl amino acid.

Coronatine (COR) (Fig. 1) is a novel phytotoxin that is produced by five distinct pathovars of Pseudomonas syringae, including P. syringae pv. atropurpurea, P. syringae pv. glycinea, P. syringae pv. maculicola, P. syringae pv. morsprunorum, and P. syringae pv. tomato, which infect ryegrass, soybean, crucifers, Prunus spp., and tomato, respectively. COR contributes to virulence in several host-pathogen interactions and elicits diffuse chlorosis in a wide variety of plant species (3, 13). COR also induces hypertrophy, inhibits root elongation, and stimulates ethylene production (11, 14, 17, 39). In several reports workers have noted the striking structural and functional homologies among COR, jasmonic acid, and 12-oxophytodienoic acid, suggesting that COR may function as a molecular mimic of the octadecanoid signaling molecules of higher plants (11, 14, 17, 39). COR consists of a bicyclic polyketide moiety, coronafacic acid (CFA), that is linked to an ethylcyclopropyl amino acid moiety, coronamic acid (CMA) (Fig. 1). These two moieties are produced by different biosynthetic pathways (26). Three acetate units, one butyrate unit, and a four-carbon unit that is derived from glutamic acid are combined to form CFA (27), while CMA is derived from 1-isoleucine via 1-allo-isoleucine which is oxidatively cyclized to form the cyclopropane ring (25). A ligase is then presumed to join CFA to CMA and produce COR via formation of an amide bond (2, 18). In the fermentation broth of P. syringae pv. glycinea, COR is accompanied by small amounts of norcoronatine, which contains norcoronamic acid (20) (Fig. 1), and by other congeners in which CMA is replaced by leucine, valine, isoleucine, or allo- isoleucine (20, 21).

The genes required for COR biosynthesis were first identified in P. syringae pv. glycinea PG4180, in which the 32.8-kb COR gene cluster is borne on a 90-kb plasmid designated p4180A (3). Investigations have shown that the structural genes for CFA and CMA biosynthesis are located at opposite ends of the gene cluster (40). An intermediate region, located between the biosynthetic regions, encodes the three regulatory proteins involved in transcriptional control of the other two regions (29, 31, 38). Nucleotide sequence analysis of each of the biosynthetic regions revealed open reading frames which indicate that CFA is biosynthesized by monofunctional and multifunctional polyketide synthase proteins (30, 32), whereas CMA appears to be biosynthesized by a thiotemplate mechanism that resembles nonribosomal peptide synthetases (37). Protein overexpression and function assays are required to confirm these predictions based on in silico data.

To begin to decipher the mechanism of CMA biosynthesis, we report here the initial characterization of CmaA, a protein encoded by the CMA region that appears to be a didomain protein containing an adenylation domain (A domain) and a thiolation domain (T domain). We found that CmaA catalyzes...
the adenylation of L-αllo-isoleucine and the attachment of L-αllo-isoleucine to the CmaA T domain. We postulate that the enzyme-bound L-αllo-isoleucine serves as the substrate for the later stages of CMA biosynthesis.

**MATERIALS AND METHODS**

**Bacterial strains and plasmids.** *Pseudomonas* strains were routinely cultured on King’s medium B (16) at 28 or 18°C, while *Escherichia coli* cultures were grown on Luria-Bertani medium at 37°C (33). Ampicillin (100 μg/ml) and kanamycin (30 μg/ml) were used for plasmid selection in both organisms.

**Reagents.** Unless indicated otherwise, all chemicals were purchased from Sigma Chemical Co. (St. Louis, Mo.). Restriction enzymes were obtained from NEB Life Technologies (Beverly, Mass.), and *Pfu* Turbo polymerase was purchased from Stratagene (La Jolla, Calif.). Protein concentrations were determined with the Advanced protein assay reagent from Cytoskeleton Inc. (Denver, Colo.) or with Bio-Rad protein assay reagent (Bio-Rad Laboratories, Hercules, Calif.).

**DNA manipulations.** Agarose gel electrophoresis, restriction enzyme digestion, and purification of DNA fragments were performed by standard procedures (33). Plasmids were prepared by using a QIAprep Spin miniprep kit (Qiagen, Chatsworth, Calif.). Electrococmpetent *P. syringae* FF5 cells were prepared as described previously (12), and electroporation was conducted by using a Bio-Rad Gene Pulser II at 200 μF, 2.5 kV, and 25 μF. Creation of the *Pseudomonas* expression vectors pSFLAG-CTC and pSFLAG-MAC has been described previously (9). *P. syringae* strains harboring the tagged-expression vectors were selected by ampicillin resistance. The nucleotide sequences of all expression vectors pSFFLAG-CTC and pSFFLAG-MAC has been described previously (9).

**Cloning of CmaA.** (i) FLAG-tagged proteins. Two sets of PCR primer pairs were used to amplify the *cmaA* gene from plasmid pSAY10 (40) (underlined) and CmaA-R-Smal (5′-AACCCCGGGTCTCAGTATTTCCATGTTG-3′), and the second set (primer pair 2) consisted of CmaA-F-EcoRI and CmaA-RNS-Smal (5′-AACCCCGGGTCTCAGTATTTCCATGTTG-3′). The two sets of primer pairs resulted in a PCR product containing unique EcoRI and Smal restriction sites (underlined) at the 5′ and 3′ ends, respectively. After high-fidelity PCR amplification with *Pfu* Turbo DNA polymerase and either primer pair 1 or primer pair 2, the ~1.8-kb PCR product was purified with a Quiaquick PCR purification kit (Qiagen), digested with EcoRI and Smal, and ligated into multiple cloning sites of expression vector pSFFLAG-MAC (for the product of primer pair 1) or pSFLAG-CTC (for the product of primer pair 2). The resulting constructs were electroporated into *E. coli* by using a Gene Pulser II electroporator (Bio-Rad Laboratories) according to the manufacturer’s instructions. The transformants were confirmed by ampicillin selection, restriction mapping, and DNA sequencing and were designated pSFLAG-MAC-ct and pSFLAG-CTC-mac.

After high-fidelity PCR amplification with primers cmaA-NdeI (5′-GGGAATTCCATATGACCTCCTACCATTCA-3′) and cmaA-EcoRI (5′-AAAAAGAATTCTCAGTATTTCCATGTTG-3′), the PCR product was purified as described above and ligated into the multiple cloning site of *NdeI*-EcoRI plasmid vector pET28b (Novagen, Madison, Wis.). The constructs were transformed into *E. coli* as described above and were selected by kanamycin resistance. DNA sequencing confirmed the identity of the insert. For overproduction of CmaA in *E. coli*, this construct was transformed into *E. coli* BL21(DE3) along with plasmid pSU20-Sfp containing the gene encoding Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase with broad substrate specificity (22), pSU20-Sfp was constructed by removal of the sfp gene, along with its promoter and ribosome binding site, from plasmid pUC8-Sfp (24) as an *EcoRI*-BamHI fragment and ligation of the fragment into plasmid pSU20 (1) digested with *EcoRI* and *BamHI*.

For overproduction of the His-tagged CmaA protein (His-CmaA), BL21(DE3) cells harboring the desired plasmids were grown in Luria-Bertani medium supplemented with 30 μg of kanamycin per ml and 34 μg of chloramphenicol per ml. One liter of medium was inoculated with 10 ml of an overnight starter culture and incubated at 25°C until an optical density at 595 nm of 0.55 was reached. Protein expression was then induced by addition of 100 μM isopropyl-β-D-thiogalactopyranoside (IPTG), and cells were allowed to grow for an additional 15 h at 15°C. Cells were harvested by centrifugation (10 min at 6,000 × g) and resuspended in lysis buffer (25 mM Tris [pH 8], 500 mM NaCl, 10% glycerol) and lysed by two passages through a French press at 10,000 lb/in². Cell extracts were clarified by centrifugation (30 min at 15,000 × g) and applied to nickel-nitriotriacetic acid resin (1 ml of resin for 3 liters of culture) (Qiagen). Cell lysate was allowed to bind in batch to the resin for 2 h at 4°C and then was decanted into a column. The resin was washed with 15 column volumes of lysis buffer and was eluted with a step gradient of lysis buffer containing increasing amounts of imidazole (5, 30, 60, 100, and 500 mM). CmaA eluted in the 100 and 500 mM imidazole fractions. Fractions containing the desired protein were dialyzed against 25 mM Tris (pH 7.5)–200 mM NaCl–10% glycerol and frozen at −80°C.

**A C-terminal His-tagged version of CmaA was also constructed by amplification of the gene with primers cmaA-Neol (5′-AAAAAACCTTAGGCTGAGCTCCTAACATCACT-3′; restriction site is underlined) and cmaA-Xhol (5′-AAAAACCTTAGGCTGAGCTCCTAACATCACT-3′).** The resulting product was ligated into Neol-Xhol-digested pET28b. Selection, transformation, and protein expression were performed as described above.

**ATP-PPi exchange assays.** ATP-PPi exchange reactions were carried out at 25°C in 100-μl mixtures that contained 75 mM Tris-HCl (pH 7.5), 10 mM MgCl₂, 1 mM tris-(2-carboxyethyl)phosphine (Molecular Probes, Eugene, Ore.), 1 mM [32P]adenosine monophosphate (5 Ci/mol; Dupont NEN, Boston, Mass.), 5 mM ATP, 0.2 nmol of FLAG-tagged CmaA protein (CmaA-FLAG) or 0.14 nmol of His-CmaA, and various concentrations of amino acid substrate. The reactions were initiated by addition of enzyme, were allowed to proceed for 10 min, and then were quenched by addition of a 1.6% activated charcoal–4.46% tetrasodium pyrophosphate–3.5% perchloric acid mixture in water. The charcoal was collected by centrifugation or filtration, washed twice with a 4.46% tetrasodium pyrophosphate–3.5% perchloric acid solution, and then, if collected by centrifugation, resuspended in 0.5 ml of water, combined with liquid scintillation cocktail (ScintiVerse; Fisher Scientific, Pittsburgh, Pa.), and placed in a liquid scintillation counter. If filtered, the charcoal was placed directly into the scintillation vial after the wash step, mixed vigorously with scintillation cocktail, and counted. Each reaction was performed...
at least in duplicate. The amount of charcoal-bound radioactivity was converted to reaction velocity by using the specific activity of the $^{38}P$P. A nonlinear regression fit of the plot of velocity versus substrate concentration provided the $V_{\text{max}}$ and $K_{m}$ values.

PPi release assay. PP, levels were measured by monitoring the appearance of NADH (at 340 nm) generated by a coupled, continuous spectrophotometric assay (8). Reactions were carried out at 25°C, and the reaction mixtures (500 μl) contained 1 M Tris-HCl (pH 8), 50 mM UDP-glucose (Sigma-Aldrich, St. Louis, Mo.), 50 mM glucose 1,6-bisphosphate (Sigma-Aldrich), 10 mM NaF, 100 mM MgCl$_2$, 1 U of UMP-glucose pyrophosphorylase (Sigma-Aldrich), 5 U of phosphoglucosemutase (Sigma-Aldrich), 5 U of phosphoglucosemutase (Sigma-Aldrich), 5 U of phosphoglucosemutase (Sigma-Aldrich). Samples were analyzed by electrophoresis with a SDS bromophenol blue) and boiling for 5 min. Samples were analyzed by electrophoresis. Thus, based upon these sequence analyses, it appeared reasonable to propose that CmaA is an adenylation activation enzyme that covalently loads its substrate by first adenylylating it (via the A domain) and then transferring it onto the 4'-phosphopantetheine arm located within the T domain (7, 19). The presence of signature motifs by using the Expasy ScanProsite algorithm resulted in identification of both an AMP-binding domain signature sequence motif (A domain) and several residues that are characteristic of a T domain (Fig. 2). Both of these domains are common in adenylation activation proteins. Thus, based upon these sequence analyses, it appeared reasonable to propose that CmaA is an adenylation activation enzyme that covalently loads its substrate by first adenylylating it (via the A domain) and then transferring it onto the 4'-phosphopantetheine arm located within the T domain (7, 19). The presence of A and T domains in CmaA was recognized previously 

RESULTS

Sequence revisions. In 1998, a corrected version of the P. syringae pv. glycinea PG4180 cmA nucleotide sequence was deposited in the GenBank database (4). In order to characterize the CmaA protein, it was critical to confirm the accuracy of this sequence. Therefore, the cmA region of P. syringae pv. glycinea PG4180 was resequenced. A comparison of the new sequence, which has at least twofold coverage in the forward and reverse directions of the entire gene, with the 1998 sequence revealed a number of differences between the two sequences. To clarify these differences, the P. syringae pv. tomato DC3000 cmA gene sequence was retrieved from The Institute for Genomic Research (TIGR) website (www.tigr.org) and compared with the other two cmA gene sequences. The TIGR P. syringae pv. tomato sequence exhibited 97% identity with the newly obtained cmA nucleotide sequence but only 93% identity with the 1998 cmA gene sequence. At the protein level, the TIGR P. syringae pv. tomato amino acid sequence exhibited 96% identity with the new PG4180 cmA sequence (Fig. 2) but only 79% identity with the 1998 cmA sequence (data not shown). The differences are due to reading frame changes caused by inserted or deleted bases in the 1998 cmA sequence. Because of the repeated coverage of the resequenced cmA region and the high degree of similarity between the new cmA sequence and the TIGR sequence, we assumed that the correct sequence for the PG4180 cmA gene has been obtained (GenBank accession number AJ386681).

This sequence was used as a basis for overproduction of CmaA in P. syringae FF5.

Because of the errors discovered in the previously deposited PG4180 cmA sequence, the entire CMA region of PG4180 was also resequenced, and the resequenced region was compared with the sequence reported for P. syringae pv. tomato. A significant number of errors were discovered in the original PG4180 sequence. Furthermore, both the new PG4180 sequence (GenBank accession number AJ391839) and the P. syringae pv. tomato sequence appeared to contain at least three previously unrecognized open reading frames, which we designated cmaC, cmaD, and cmaE (Fig. 3).

CmaA sequence analysis. An analysis of either the P. syringae pv. glycinea or P. syringae pv. tomato cmA nucleotide sequence revealed a 1,788-bp open reading frame encoding a 595-amino-acid CmaA protein. A BLAST analysis of each CmaA protein revealed that the greatest similarities were with adenylation activation enzymes and domains that play a role in nonribosomal peptide biosynthesis (results not shown). Furthermore, screening the CmaA amino acid sequence for the presence of signature motifs by using the Expasy ScanProsite algorithm resulted in identification of both an AMP-binding domain signature sequence motif (A domain) and several residues that are characteristic of a T domain (Fig. 2). Both of these domains are common in adenylation activation proteins. Thus, based upon these sequence analyses, it appeared reasonable to propose that CmaA is an adenylation activation enzyme that covalently loads its substrate by first adenylylating it (via the A domain) and then transferring it onto the 4'-phosphopantetheine arm located within the T domain (7, 19). The presence of A and T domains in CmaA was recognized previously, despite the errors present in the cmA nucleotide sequence (37).

In an attempt to identify the substrate specificity of CmaA, key residues that comprise the A domain specificity-conferring code (5, 35) were examined (underlined residues in Fig. 2). Of the 10 residues that make up the specificity-conferring code, 7 are found in A domains that are specific for l-isoleucine, L-leucine, and L-valine. Both l-isoleucine and l-allo-isoleucine have been shown to be incorporated into CMA, but l-allo-isoleucine, whose specificity-conferring code has not been defined yet, is a much more efficient precursor (25). The timing of the epimerization of l-isoleucine to l-allo-isoleucine is currently unknown. By analogy with the conversion of L-allo-isoleucine into CMA, it appears that norcoronacic acid should be derived from L-valine. For these reasons, l-isoleucine, L-allo-isoleucine, and L-valine each appeared to be a potential substrate for CmaA.

Protein overexpression and visualization. Two complementary approaches were taken to overexpress CmaA for substrate specificity assays. In one approach, the P. syringae pv. glycinea CmaA protein was expressed in P. syringae pv. syringae FF5, a Pseudomonas strain that lacks the COR gene cluster (36). Additionally, the P. syringae pv. tomato CmaA protein was overexpressed in E. coli. To generate the appropriate constructs, the cmA gene was PCR amplified from P. syringae pv. glycinea and P. syringae pv. tomato, cloned into expression vectors that produced FLAG-tagged and His-tagged proteins, respectively, and transformed into the appropriate hosts. The soluble protein yields were compared with N-terminal and...
C-terminal tag locations, and a C-terminal FLAG-tagged protein and an N-terminal His-tagged protein were selected for subsequent functional assays. Figure 4A shows the SDS-PAGE results for the N-terminal His-CmaA and the C-terminal CmaA-FLAG. As shown in the Fig. 4, denatured CmaA-FLAG electrophoresed faster than expected, since the protein migrated at a molecular weight $\approx 66,000$, whereas the predicted molecular weight is actually $\approx 68,000$. Denatured His-CmaA electrophoresed at approximately the same position as CmaA-FLAG. Native PAGE of CmaA-FLAG indicated that the protein exists as a dimer (Fig. 4B).

**PPi exchange assay.** The substrate specificity of CmaA was assayed by using the amino acid-dependent exchange of radioactivity from $^{32}$PPi into ATP (10). This assay measures the reversible formation of the aminoacyl-AMP derivative and allows determination of amino acid selectivity. His-CmaA and FLAG-CmaA were each assayed with variable concentrations of substrate in a buffered solution containing ATP and $^{32}$PPi. After a brief incubation period, newly formed, radiolabeled ATP was collected by using activated charcoal and subsequently was quantified by liquid scintillation counting. A nonlinear regression analysis of reaction velocity versus substrate concentration provided the $K_m$ and $V_{max}$ values, the latter of which was used to calculate $k_{cat}$.

L-Leucine, L-valine, and L-isoleucine were suggested to be possible substrates by examining the specificity-conferring code of CmaA. A series of qualitative assays with His-CmaA indicated that L-allo-isoleucine is the preferred substrate, and a small amount of activation was also observed with L-leucine, L-valine, or L-isoleucine (Fig. 5). Kinetic parameters for the

![FIG. 2. Alignment of CmaA amino acid sequences. A comparison of the CmaA sequence from *P. syringae* pv. glycinea with the CmaA sequence from *P. syringae* pv. tomato DC3000 indicates that these two sequences are 96% identical. The residues enclosed in the rectangle create an AMP-binding domain signature sequence, whereas the residues enclosed in the ellipses are common to a T domain, as determined by using the ScanProsite program at the Expasy web site. The arrow indicates the putative 4'-phosphopantetheine attachment site. The underlined residues are the residues that contribute to the specificity-conferring code of A domains in nonribosomal peptide synthetases. See the text for further discussion.](http://jb.asm.org/)

![FIG. 3. Gene organization in the CMA region of the COR biosynthetic gene cluster of *P. syringae* pv. glycinea PG4180.](http://jb.asm.org/)

![1 kb](http://jb.asm.org/)
four isoleucine diastereomers were measured (Table 1). A more detailed kinetic analysis was performed with the more active CmaA-FLAG.

L-Isoleucine appears to be preferred over L-leucine and L-valine, as implied by the relatively small $K_m$ value and the relatively large $k_{cat}/K_m$ value. Furthermore, CmaA is quite selective with respect to the absolute configuration of isoleucine; that is, the $k_{cat}$ and $k_{cat}/K_m$ values of the D stereoisomer of isoleucine are roughly one-third those of the L stereoisomer, clearly indicating that there is a preference for the latter. However, after examination of the L-allo-isoleucine data, it became very apparent that this amino acid is by far the preferred substrate, as it yields the smallest $K_m$, the greatest $k_{cat}$, and the largest $k_{cat}/K_m$ value of all the substrates tested. The preference for the L isomer of allo-isoleucine was also apparent, as D-allo-isoleucine resulted in a threefold increase in the $K_m$, an 85-fold decrease in the $k_{cat}$, and consequently a 272-fold decrease in the $k_{cat}/K_m$ value. Interestingly, CMA, the ethylcyclopropyl amino acid component of COR which is thought to be the product of the cyclization of allo-isoleucine, was the second best substrate tested, as judged by the $k_{cat}/K_m$ value. Norcoronamic acid, which is thought to be the product of cyclization of valine, was the worst substrate tested.

PP$_i$ release assay. A continuous, spectrophotometric assay was also employed to determine the substrate specificity of CmaA (8). This assay measures the release of PP$_i$, by means of a coupled enzyme system in which PP$_i$ drives the reduction of NAD$^+$ to NADH, a process that is conveniently monitored at 340 nm. Continuous release of PP$_i$ must be accompanied by the release of the aminoacyl-AMP derivative to regenerate the free A domain for catalysis of another reaction cycle. Thus, this assay is in effect a measure of the frequency of loss of the aminoacyl-AMP derivative from the A domain active site. Following adenylation, preferred amino acid substrates are expected to be held tightly by the enzyme, presumably to await transfer to the 4'-phosphopantetheine arm of the T domain. On the other hand, less preferred substrates have a tendency to leak from the adenylation site following the adenylation reaction. With this assay, therefore, highly preferred substrates should generate relatively small $K_m$ and $k_{cat}$ values (i.e., tight binding with little leakage), while the substrates that are less preferred should yield larger $K_m$ and $k_{cat}$ values.

Table 2 shows the kinetic values obtained in the PP$_i$ release assay when CmaA-FLAG was used in conjunction with various substrates. As observed with the PP$_i$ exchange assay (Table 1), the PP$_i$ release assay results illustrated the preference of CmaA for L-isoleucine over L-leucine and L-valine. Furthermore, the specificity of CmaA for L-isoleucine over D-isoleucine was also apparent when this assay was used. Most importantly, it is clear that L-allo-isoleucine was the most preferred substrate for CmaA of all the substrates tested, as it displayed the smallest $K_m$ and $k_{cat}$ values. The lowest $K_m$, and the largest $k_{cat}/K_m$ value (Table 2). These results are in good agreement with those of the PP$_i$ exchange assay.

Covalent loading of L-allo-isoleucine by CmaA. With the preferred substrate of the adenylation reaction identified, the ability of CmaA to catalyze loading of its T domain with L-allo-isoleucine was examined. CmaA-FLAG was incubated with L-[G-3H]allo-isoleucine in the presence or absence of ATP. The proteins were subsequently denatured, electrophoresed, and analyzed by autoradiography. Figure 6 shows the autoradiograph obtained, which verified that there was ATP-dependent acylation of CmaA with L-allo-isoleucine. Identical results were obtained with His-CmaA (data not shown).

**DISCUSSION**

Bioynthesis of the plant phytotoxin COR by *P. syringae* appears to proceed by formation of CMA and CFA, which are then linked to produce COR (Fig. 1). A set of three genes, *cmaA*, *cmaB*, and *cmaT*, in the CMA region of the biosynthetic cluster (Fig. 3) have previously been implicated in CMA formation from the proteogenic amino acid L-isoleucine. To begin to decipher the mechanism of cyclopropane ring formation, we have undertaken biochemical characterization of the CmaA protein. This protein is predicted to be a 68-kDa, two-domain
(A domain-T domain) protein that resembles the amino acid activation modules observed in nonribosomal peptide biosynthetic pathways.

Discrete A domain-T domain proteins in biosynthetic pathways often capture a proteogenic amino acid and modify it. The modified amino acid is then released to create a dedicated fi

cration of an aminoacyl-AMP derivative, indicating that the covalently sequestered isoleucine and to link it to CmaA in the T domain with a phosphopantetheinylation moiety in vivo, either by the heterologously expressed Bacillus phosphopantetheinyl transferase Sfp in the E. coli expression system or by an endogenous Pseudomonas transferase. The kinetic data indicate that the CmaA that is produced in Pseudomonas is substantially more active than the enzyme that is produced in E. coli. The N-terminally tagged version produced in E. coli has a mutation in the protein sequence (Ser521Phe) which may contribute to the decrease in activity. However, a wild-type C-terminally Histagged construct also exhibited substantially reduced activity compared to the activity of the Pseudomonas protein (data not shown). Therefore, we suspect that the differences in the kinetic data are due to the fact that E. coli is a poor expression vehicle for the COR biosynthetic proteins. CmaA protein folding may not occur properly in E. coli, or alternatively, the B. subtilis phosphopantetheinyl transferase may not modify CmaA efficiently. Incubation of purified protein from either background with \( ^{38}\)PP, exchange data, which measured the enzyme-dependent reversible formation of an aminoacyl-AMP derivative, indicated that \( l\)-allo-isoleucine is strongly favored as the substrate over all the other amino acids examined. The data suggest that free \( l\)-allo-isoleucine is biosynthesized by the phytopathogenic P. syringae strains prior to this step of CMA biosynthesis. How this unusual isoleucine diastereomer is synthesized remains

Table 1. Kinetic parameters for CmaA-FLAG and His-CmaA as determined by the PP\(_i\) exchange assay

<table>
<thead>
<tr>
<th>Substrate</th>
<th>CmaA-FLAG</th>
<th></th>
<th>His-CmaA</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( K_m ) (mM)</td>
<td>( k_{cat} ) (min(^{-1}))</td>
<td>( k_{cat}/K_m ) (min(^{-1})mM(^{-1}))</td>
<td>( K_m ) (mM)</td>
</tr>
<tr>
<td>( l)-allo-Isoleucine</td>
<td>0.26 ± 0.02</td>
<td>17 ± 0.3</td>
<td>65.4</td>
<td>0.00 1.6</td>
</tr>
<tr>
<td>( l)-Isoleucine</td>
<td>1.0 ± 0.1</td>
<td>2.8 ± 0.1</td>
<td>2.8</td>
<td>ND 2.0</td>
</tr>
<tr>
<td>( l)-Leucine</td>
<td>7.4 ± 1.1</td>
<td>16 ± 3</td>
<td>2.2</td>
<td>ND 2.0</td>
</tr>
<tr>
<td>( l)-Valine</td>
<td>8.2 ± 1.3</td>
<td>12 ± 1</td>
<td>1.5</td>
<td>ND 2.0</td>
</tr>
<tr>
<td>( d)-allo-Isoleucine</td>
<td>0.84 ± 0.07</td>
<td>0.20 ± 0.06</td>
<td>0.24</td>
<td>ND 2.0</td>
</tr>
<tr>
<td>( d)-Isoleucine</td>
<td>1.1 ± 0.2</td>
<td>1.1 ± 0.2</td>
<td>1.0</td>
<td>ND 2.0</td>
</tr>
<tr>
<td>dl-Coronamic acid</td>
<td>1.5 ± 0.3</td>
<td>6.7 ± 0.4</td>
<td>4.5</td>
<td>ND 2.0</td>
</tr>
<tr>
<td>dl-Norcoronamic acid</td>
<td>11 ± 3</td>
<td>1.1 ± 0.2</td>
<td>0.10</td>
<td>ND 2.0</td>
</tr>
</tbody>
</table>

\( ^{38}\)ND, not determined.

Table 2. Kinetic parameters for CmaA-FLAG as determined by the forward PP\(_i\) assay

<table>
<thead>
<tr>
<th>Substrate</th>
<th>( K_m ) (mM)</th>
<th>( k_{cat} ) (min(^{-1}))</th>
<th>( k_{cat}/K_m ) (min(^{-1})mM(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( l)-allo-Isoleucine</td>
<td>0.04 ± 0.01</td>
<td>0.09 ± 0.00</td>
<td>2.3</td>
</tr>
<tr>
<td>( l)-Isoleucine</td>
<td>0.50 ± 0.06</td>
<td>0.71 ± 0.03</td>
<td>1.4</td>
</tr>
<tr>
<td>( l)-Leucine</td>
<td>1.5 ± 0.1</td>
<td>1.8 ± 0.1</td>
<td>1.2</td>
</tr>
<tr>
<td>( l)-Valine</td>
<td>1.2 ± 0.1</td>
<td>1.4 ± 0.1</td>
<td>1.2</td>
</tr>
<tr>
<td>( d)-allo-Isoleucine</td>
<td>5.6 ± 1.0</td>
<td>0.51 ± 0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>( d)-Isoleucine</td>
<td>5.3 ± 0.9</td>
<td>0.79 ± 0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>dl-Coronamic acid</td>
<td>1.0 ± 0.2</td>
<td>1.3 ± 0.1</td>
<td>1.3</td>
</tr>
<tr>
<td>dl-Norcoronamic acid</td>
<td>4.5 ± 1.0</td>
<td>1.8 ± 0.4</td>
<td>0.40</td>
</tr>
</tbody>
</table>

FIG. 6. Covalent labeling of CmaA-FLAG by \( l\)-[\( ^{38}\)H]allo-isoleucine in the presence of ATP.
unknown. The kinetic parameters for L-valine indicate that it is a rather poor substrate. This is somewhat surprising given the natural occurrence of norcoronamic acid, which is probably derived from L-valine by a process similar to CMA formation. However, norcoronatine is a minor constituent of P. syringae pv. glycinea fermentation, and so formation of norcoronamic acid may be a relatively inefficient process.

The A domains of nonribosomal peptide synthetases typically exhibit a consensus sequence that determines the amino acid substrate specificity. While the consensus sequence of 1-isoleucine-activating domains is very similar to that of CmaA (Fig. 2), there are three changes in the consensus sequence that may be responsible for the preference for 1-allo-isoleucine. Although other 1-allo-isoleucine-containing natural products exist (23, 34), this is the first example of an A domain specific for the (2S,3R) stereochemistry of 1-allo-isoleucine.

In this paper we describe the first characterization of the CMA biosynthetic pathway at the enzymatic level. The data reveal that there is a discrete A domain-T domain whose function is to tether L-allo-isoleucine, presumably for subsequent cyclization and hydrolysis by CmaT to produce CMA. Previous studies have shown that CmaT exhibits thioesterase activity with model substrates (28). In addition to the CmaA and CmaT genes, five other genes are present in the CMA biosynthetic region of P. syringae PG4180 (Fig. 3). The sequence of CmaB exhibits similarities to the sequences of 1-ketoglutarate-dependent dioxygenases, BarB1/Bar2 encoded by the barbamide gene cluster (6), and SyrB2 encoded by the syringomycin gene cluster (15, 41). These similarities suggest that CmaB is a nonheme iron dioxygenase that may carry out hydroxylation or chlorination of the CmaA-bound L-allo-isoleucine at C-6 (Fig. 7). CmaC exhibits similarity to methylmalonyl coenzyme A mutases, which suggests that its role may be to deprotonate and cyclize CmaA-bound 6-hydroxy-allo-isoleucine or 6-chloro-allo-isoleucine to produce CmaA-bound CMA (Fig. 7). CmaD exhibits similarities to acyl carrier proteins, while CmaE shows similarities to proteins with an α/β hydrolase fold. The role played by these two proteins is unclear. Characterization of the later steps in the CMA biosynthetic pathway is under way.

ACKNOWLEDGMENTS

We acknowledge support of this study by the National Institutes of Health (grant GM26564 to R.J.P. and grant GM20011 to C.T.W.), by the Robert A. Welch Foundation (grant C-0729 to R.J.P.), and by an Irving S. Sigal postdoctoral fellowship (to S.E.O.).

Construction of vector pSU20-sfp by M. G. Thomas is also gratefully acknowledged.

REFERENCES

male sterile, insensitive to methyl jasmonate, and resistant to a bacterial pathogen. Plant Cell 6:751–759.


