Crystal Structure of the ECH$_2$ Catalytic Domain of CurF from Lyngbya majuscula

INSIGHTS INTO A DECARBOXYLASE INVOLVED IN POLYKETIDE CHAIN $\beta$-BRANCHING

Received for publication, May 14, 2007, and in revised form, September 14, 2007. Published, JBC Papers in Press, October 10, 2007. DOI 10.1074/jbc.M703921200

Todd W. Geders, Liancai Gu, Jonathan C. Mowers, Haichuan Liu, William H. Gerwick, Kristina Håkansson, David H. Sherman, and Janet L. Smith

From the Life Sciences Institute, Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109, the Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, the Life Sciences Institute, Departments of Medicinal Chemistry, Chemistry, Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan 48109, the Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, and the Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093

Curacin A is a mixed polyketide/nonribosomal peptide possessing anti-mitotic and anti-proliferative activity. In the biosynthesis of curacin A, the N-terminal domain of the CurF multifunctional protein catalyzes decarboxylation of 3-methylglutaconyl-acetyl carrier protein (ACP) to 3-methylcrotonyl-ACP, the postulated precursor of the cyclopropane ring of curacin A. This decarboxylase is encoded within an “HCS cassette” that is used by several other polyketide biosynthetic systems to generate chemical diversity by introduction of a $\beta$-branch functional group to the natural product. The crystal structure of the CurF N-terminal ECH$_2$ domain establishes that the protein is a crotonase superfamily member. Ala$^{78}$ and Gly$^{118}$ form an oxyanion hole in the active site that includes only three polar side chains as potential catalytic residues. Site-directed mutagenesis and a biochemical assay established critical functionalities, including a cyclopropyl ring, a cis-vinyl thiazole heterocycle, and a terminal alkene (Fig. 1A). Curacin A was first shown in 1994 to possess anti-mitotic activity by the inhibition of tubulin polymerization, cell cycle arrest at the G$_2$/M transition, and anti-proliferative activity against colon, renal, and breast cancer-derived cell lines (5). Further studies have demonstrated rapid, essentially irreversible binding of curacin A to tubulin with potent inhibition of colchicine binding (6).

The biosynthetic gene cluster (cur) for curacin A was recently identified and characterized, revealing that the metabolite is generated by a hybrid type I PKS/NRPS (7). The first six genes in the cur cluster (curA$\cdots$curF) have been implicated in the formation of the cyclopropyl ring as well as the thiazoline ring of curacin A. Genes within this region include an “HCS cassette” that encodes an acyl carrier protein (ACP) (CurB), a putative ketosynthase (CurC), a 3-hydroxy-3-methylglutaryl (HMG)-CoA synthase (HCS) (CurD), and two domains con-
taining similarity to enoyl-CoA hydratases (CurE (ECH1)) and the N-terminal ∼260 residues of CurF (ECH2)) (Fig. 2).

HCS cassettes are important sources of β-branches in polyketides and allow for the generation of wide chemical diversity at the branch point (Fig. 1). Among the functional groups generated are the cyclopropyl ring of curacin A (7, 8), the vinyl chloride of jamaicamide (9), the exomethylene group of the pederin family of polyketides (10), and the pendant
methyl (or ethyl) groups of bacillaene (11, 12), mupirocin (pseudomonic acid A) (13), myxoviricin A (14, 15), and virginiamycin M (16). Subsets of the HCS-containing gene cassette are found in gene clusters that specify production of the bryostatins (17) and leinamycin (18).

Among PKS gene clusters with HCS cassettes, at least six encode ECH1/ECH2-like enzyme pairs. Recently, the unique catalytic activities of CurE ECH1 and the CurF ECH2 domain have been ascertained (8). CurF ECH2 catalyzes conversion of (S)-HMG-ACP to 3-methylglutaconyl-ACP. Subsequently, the N-terminal ECH2 domain of CurF catalyzes decarboxylation to 3-methylcrotonyl-ACP, the postulated precursor of the cyclopropane ring of curacin A (Fig. 1A). The analogous activities of two other ECH1/ECH2-like enzyme pairs was demonstrated for the HCS cassettes of myxovirescin A and bacillaene (11, 14). Based on the structures of pathway end products, decarboxylation by ECH2-type enzymes are postulated to be followed by proton donation at C-4 leading to a $\Delta^2$ unsaturated product in curacin A, bacillaene, virginiamycin M, mupirocin, or myxovirescin A or at C-2 leading to a $\Delta^3$ unsaturated product in jamacamide and pederin (Fig. 1). Thus, ECH2 represents a critical point of divergence for the generation of correct downstream functionalities in the corresponding natural product. Understanding how these enzymes control product double bond regiochemistry will facilitate the prediction of enzymatic activity for this class of enzymes, as well as aid protein engineering attempts to tailor chemistry for the production of novel natural products.

Enoyl-CoA hydratases belong to the crotonase superfamily, consisting of a wide variety of mechanistically diverse enzymes that exhibit various activities, such as hydratase (19), dehalogenase (20), decarboxylase (21, 22), isomerase (Refs. 22–25; Protein Data Bank code 2F6Q), hydrolyase (22), and carbon-carbon bond formation (22) or cleavage (27, 28). Because of this catalytic diversity, the low sequence identity superfamily lacks an invariant catalytic residue, and superfamily membership is established through crystal structures. The common mechanistic theme of the superfamily is the stabilization of an enolate anion intermediate of phosphopantetheine-linked substrates by two backbone amide groups forming an oxyanion hole (29). CurF ECH2 has less than 20% sequence identity to identified members of the superfamily. A unique feature of the ECH2-like enzymes involved in natural product biosynthesis is that they are thought to depend exclusively on ACP-linked substrates in vivo. Structural clues that enable these enzymes to discriminate between CoA and ACP-linked substrates will increase our understanding of the evolutionary changes that have occurred to favor one substrate over another, as well as to control functional group diversity in the final natural product.

In this paper we report the 1.85 Å crystal structure of the wild type, N-terminal ECH2 domain of CurF and the 1.65 Å structure of the corresponding Y82F variant. Modeling of the substrate of CurF ECH2 was performed, and Tyr82, Lys186, and His240 were identified as potential catalytic or substrate-binding residues. Site-directed mutagenesis in a coupled ECH1/ECH2 dehydration/decarboxylation assay with ACP-linked substrates demonstrated that CurF ECH2 His240 and Lys186 are critical to catalysis. Finally, we demonstrated that the 3-methylglutaconyl-ACP substrate is preferred over 3-methylglutaconyl-CoA, a preference that may be controlled by a conserved bulky amino acid residue in the HCS cassette class of decarboxylases.

**EXPERIMENTAL PROCEDURES**

**Cloning, Site-directed Mutagenesis, and Protein Expression**—The plasmid pMCSG7::CurFd17 was generated by PCR amplification of coding sequence corresponding to residues 17–257 of CurF from the pML9 cosmid DNA (7) and inserted into the vector pMCSG7 (30). The plasmid pMCSG7::CurFd17 was transformed into BL21(DE3) and grown at 37 °C in 2×YT medium to an A600 of 0.6–0.8 in 2 L baffled flasks. The cultures were adjusted to 18 °C, and isopropyl $\beta$-D-thiogalactopyranoside was added to final concentration of 0.4 mM and allowed to grow 8–16 h with shaking. The cells were harvested by centrifugation, and cell pellets were frozen immediately at −20 °C. Selenomethionine-labeled protein was produced in BL21(DE3) using selenomethionine minimal medium according to the protocol of Guerrero et al. (31). Site-directed mutagenesis was performed by the QuikChange method (Stratagene) (see supplemental materials). All of the constructs were verified by DNA sequencing. CurA (2057–2146)-ACP(II) with a C-terminal His$_8$ tag was constructed by inserting synthetic DNA (a kind gift from Christopher Calderone and Christopher T. Walsh, Harvard Medical School) into pET29a using NdeI and XhoI restriction sites. Protein overexpression and purification of CurA-ACP(II) were performed as described previously for CurB (8).

**Protein Purification**—All of the steps were performed at 4 °C. The cell pellet from 1 liter of culture was resuspended in 35 ml of buffer A (20 mM Tris, pH 7.9, 500 mM NaCl, 10% glycerol, 20 mM imidazole). Cellytic Express (200–300 mg) was added prior to lysis by sonication (Sigma-Aldrich). The lysate was cleared by centrifugation at >38,000 × g for 45 min. Supernatant was filtered by a 0.45-μm filter, loaded onto a 5-ml HisTrap (GE Healthcare) nickel-nitrilotriacetic acid resin column, and
washed with 10 column volumes of buffer A. The protein was eluted with ~200 mM imidazole by a linear gradient of buffer B (20 mM Tris, pH 7.9, 500 mM NaCl, 10% glycerol, 750 mM imidazole). For His tag removal, the fractions were pooled, and buffer was exchanged with buffer A containing 1 mM dithiothreitol and incubated at 4 °C for 24–72 h with 2% (w/w) His-tagged tobacco etch virus protease. The reaction mixture was loaded onto the HiTrap column, and flow-through fractions were pooled, concentrated, and loaded onto a HiLoad 16/60 Superdex 75 (GE Healthcare) column equilibrated with 20 mM Tris, pH 7.9, 500 mM NaCl, and 10% glycerol. The fractions were then combined, concentrated to 20 mg/ml, and either flash frozen in liquid N2 or stored at 4 °C.

**Crystallization**—The crystals were grown in 24–72 h at 4 °C by microseeding in hanging drops using the vapor diffusion method. Equal volumes were mixed of protein solution and mother liquor containing 1.3–1.8 M sodium malonate, pH 7.0, 50 mM HEPES, pH 6.8, and 0–25 mM 2-ethanamidoethyl 3-methylbut-2-enethioate. The crystals were harvested in loops and directly frozen by plunging into liquid N2. Both trigonal (P321) and rhombohedral (R32) crystals grew in the above crystallization condition from the same seed stock.

**Data Collection and Structure Determination**—The diffraction data were collected at 100 K on GM/CA-CAT beamlines 23ID-B and 23ID-D at the Advanced Photon Source in the Argonne National Laboratory (Argonne, IL). The data were processed using the HKL2000 suite (32). Initial phasing was performed using a three-wavelength multiple wavelength anomalous dispersion data set from a single trigonal wild type selenomethionine-labeled protein crystal. SOLVE was used to find the four selenium sites (average figure of merit (FOM) = 0.47, score = 25.07) and for multiple wavelength anomalous dispersion phasing (overall FOM = 0.57) (33). RESOLVE was used for density modification (overall FOM = 0.69) and partial automated model building (34, 35). In both crystal forms, one ECH2 subunit was present in the asymmetric unit. Modeling was completed manually using COOT (36). Refinement was performed using REFMAC5 of the CCP4 suite with TLS (37–40). The refined model was used for density modification (overall FOM = 0.69) and partial automated model building (34, 35). In both crystal forms, one ECH2 subunit was present in the asymmetric unit. Modeling was completed manually using COOT (36). Refinement was performed using REFMAC5 of the CCP4 suite with TLS (37–40). The refined model was used as a probe structure for molecular replacement using PHASER with data from rhombohedral crystals (40, 41).

**Sequence and Structural Analysis**—Sequence alignments were performed by T_COFFEE (42). The figures and structural alignments were generated with PyMOL (43). Surface area calculations were calculated using AREAIMOL (44). Identification of structural neighbors was performed by a DALI search (45).

**Structure Alignment and Substrate Modeling**—Ca atoms of the crotonase core domain of CurF ECH2 (residues 18–222) were aligned with the analogous core region of liganded crotonase superfamily members carboxymethylproline synthase (Ref. 22; Protein Data Bank code 2A81, RMSD = 2.3 Å for 167 Cα), methylmalonyl CoA decarboxylase (Ref. 21; Protein Data Bank code 1EF9, RMSD = 1.7 Å for 153 Cα), 1,4-dihydroxy-2-naphthoyl-CoA synthase (Ref. 46; Protein Data Bank code 1Q51, RMSD = 1.6 Å for 138 Cα), rat enoyl CoA hydratase (Ref. 19; Protein Data Bank code 1DUB, RMSD = 1.8 Å for 156 Cα), and 4-chlorobenzoyl coenzyme A dehalogenase (Ref. 20; Protein Data Bank code 1NZY, RMSD = 2.0 Å for 163 Cα) (supplemental Fig. S2). Initial atomic coordinates and topology files for the substrate 3-methylglutaconyl moiety attached to 4-phosphopantetheinic acid were generated using the PRODRG2 server (47). Using the conformation of 4-phosphopantetheinic acid bound to rat ECH as a guide (Protein Data Bank code 1DUB), the substrate thioester oxygen was fixed at the position of a water molecule in the oxyanion hole of wild type CurF ECH2. Three water molecules overlapping the modeled substrate were removed. The model was refined by energy minimization using the program CNS (48) in 500 steps of conjugate gradient minimization with no experimental energy terms, no crystallographic symmetry restraints, and fixed main chain positions after the addition of polar hydrogens.

**ECH1/ECH2-coupled Enzymatic Assay**—The activities of the CurF ECH2 wild type and variants were measured in the ECH1–ECH2 coupled assay, as previously reported (8). In brief, 50 μM (R,S)-HMG-CurA-ACP(II) was incubated with 2 μM ECH1 and ECH2 (wild type or variants) in 50 mM Tris-HCl, pH 7.5, at 37 °C for 1 h. The reactions were terminated by 10% formic acid immediately before loading the reaction mixture on the Jupiter C4 (5 μ, 300 A) reverse phase column (Phenomenex), and the proteins were eluted with CH3CN (0.05% HCOOH and 0.05% CF3COOH)/H2O (0.05% HCOOH and 0.05%CF3COOH). The chromatogram peaks were normalized by 32 Karat software (Beckman Coulter) and subjected to base-line subtraction before peak area calculations.

**Preparation of 3-Methylglutaconyl-CoA and ACP**—To separate the dehydration and decarboxylation steps catalyzed by ECH1 and ECH2, 3-methylglutaconyl-CoA was prepared by enzymatic dehydration of (R,S)-HMG-CoA. (R,S)-HMG-CoA was incubated with ECH1 at 37 °C for 5 h, and the dehydration product was isolated using XBridge Prep C18 column (Waters, 10 × 250 mm, 5 μm) under the similar HPLC conditions reported (8). The fractions were pooled and lyophilized. ~0.5 mg of 3-methylglutaconyl-CoA was generated from 6 mg of (R,S)-HMG-CoA. 3-Methylglutaconyl-ACP was prepared with the Sfp protocol (8).

**ECH1 and ECH2 Assays for ACP and CoA Substrates**—(R,S)-HMG-CoA/ACP and 3-methylglutaconyl-CoA/ACP were employed to test the ECH1 and ECH2 activities. 2 μM ECH1 or ECH2 was incubated with 50 μM ACP or CoA substrates in 50 mM Tris-HCl, pH 7.5, at 37 °C for 1 h. The CoA samples were analyzed by XBridge Prep C18 column (Waters, 4.6 × 250 mm, 5 μm) and eluted with MeOH/H2O (10 mM CH3COONH4), and ACP samples were analyzed by Jupiter C4 column and eluted with CH3CN (0.05% HCOOH and 0.05% CF3COOH)/H2O (0.05% HCOOH and 0.05%CF3COOH).

**Mass Spectrometry Analysis**—Multiplex protonated CurAACP(II) with or without treatment with ECH1 and/or ECH2 was generated by electrospray ionization (ESI) at 70 μl/h (Apollo ion source; Bruker Daltonics, Billerica, MA) of a solution containing 2 μM CurA-ACP(II) (55:45 CH3CN:H2O with 0.05% HCOOH and 0.05% CF3COOH). All of the mass spectra were collected with an actively shielded 7 Tesla Fourier transform ion cyclotron resonance (FTICR) mass spectrometer with a quadrupole front end (APEX-Q, Bruker Daltonics). Ions produced by ESI were externally accumulated in a hexapole for 1 s,
transferred via high voltage ion optics, and captured in an ICR cell by gated trapping. This accumulation sequence was looped six times. The ESI capillary voltage was set to −3.8 kV. Nitrogen drying gas (200–250 °C) was employed to assist desolvation of ESI droplets. All of the data were acquired with XMASS software (version 6.1; Bruker Daltonics) in broadband mode from m/z = 200–2000 with 512,000 data points and summed over 10 scans. The mass spectra were analyzed with MIDAS analysis software (49).

RESULTS

Structure Determination—Initial crystal screening with a polypeptide including residues 1–257 of CurF produced crystals that diffracted to only ~3.8 Å, but an N-terminal truncation including residues 17–257 yielded crystals that diffracted to beyond 2 Å. The structure was solved by multiple wavelength anomalous dispersion phasing using selenomethionyl CurF ECH2 (Table 1). The resulting model was refined against the 2.0 Å data set in crystal form I and used to solve the structures by molecular replacement in crystal form II (Table 2). The values in parentheses are for the outer shell.

| Parameter                     | Crystal form I  | Crystal form II |
|-------------------------------|-----------------|-----------------|
| Space group                   | P321            | P321            |
| Dimensions (Å) a, c           | 105.8, 46.3     | 105.8, 46.2     |
| X-ray source                  | 23ID-B          | 23ID-B          |
| Wavelength λ (Å)              | 1.0332          | 0.95373         |
| d_iso (Å)                     | 2.0             | 3.8             |
| Unique observations           | 20,390          | 13,394          |
| Rmerge(%)                     | 6.3 (68)        | 8.8 (67)        |
| Completeness (%)              | 99.9 (100)      | 18.9 (2.3)      |
| Average redundancy(Å²)        | 11.0 (10.4)     | 29.9 (5.6)      |

| Data collection               |           |               |
|-------------------------------|-----------|---------------|
| Source                        | X-ray     | ESI droplets  |
| Dimensions (Å)                | 1.0332    | 0.95373       |
| Space group                   | P321      | P321          |
| Data range                    | 50–2.0    | 50–1.85       |
| RMSD bond length (Å)          | 0.011     | 0.012         |
| Average protein B-factor (Å²) | 29.6      | 22.8          |
| Average solvent B-factor (Å²) | 44.6      | 41.0          |
| Wilson B (Å²)                 | 33.2      | 25.6          |

Structure of the N-terminal ECH2 Domain of CurF—The CurF ECH2 domain possesses the crotonase core fold (residues 17–209) consisting of a central mostly parallel β sheet flanked by α helices (Fig. 3A). Roughly perpendicular to the main β sheet (β1, β2, β3, β5, and β7) are short β strands (β4, β6, and β8) that are a hallmark of the crotonase superfamily. Crotonase superfamily members can be classified into one of three structural classes based on the conformation of the helical C terminus (22, 29). The helical C terminus (α9, α10, and α11) of CurF ECH2 has the “self-association” fold seen in at least eight other crotonase superfamily members (Refs. 23–25, 28, and 50; Protein Data Bank code 2F6Q), including the two with biotin-dependent decarboxylase activities, methylmalonyl CoA decarboxylase (MMCD) (21) and carboxymethylproline synthase (CarB) (22).

Quaternary Structure of the ECH2 Domain of CurF—The tri- mer structure fundamental to the crotonase superfamily occurs in both crystal forms of ECH2 (Fig. 3B). Subunit contacts are virtually identical in the two crystal forms. A total of 15% of the solvent-accessible surface area of the monomer is buried in the ECH2 trimer (total buried surface area per monomer = 1050 Å²). The extensive buried surface and the lack of water in the subunit interface together indicate that the trimeric association observed in the crystal structure reflects a true quaternary structure for the protein. However, in solution CurF ECH2 also displayed concentration-dependent dissociation (supplemental Fig. S1), indicating a dynamic equilibrium between trimeric and lower oligomeric states.

Active Site—Despite less than 20% sequence identity with other crotonase superfamily members, strong structural similarity of the CurF ECH2 domain to crotonase superfamily members enabled us to align the structures and to identify critical elements of the active site, including the substrate-binding tunnel, active site oxyanion hole, and active site chamber. Structural alignments with several other crotonase superfamily members (supplemental Fig. S2) clearly indicate that the backbone amides of residues Ala77 and Gly117 form the oxyanion hole. These residues follow conserved Gly77 and Gly117, which have backbone conformations only accessible to glycine. In this manner, the peptide planes of residues 77–78 and 117–118 are oriented so their amides can stabilize the proposed enolate anion intermediate by hydrogen bonding. In the crystal structures, a water molecule occupies the oxyanion hole (Fig. 4A). Attempts to obtain crystal structures of complexes with product analogues were unsuccessful. The substrate of CurF ECH2 was then modeled into the active site (see "Experimental Procedures"). Only three polar side chains

**Table 2**

| Refinement statistics | Form I       | Form II      |
|-----------------------|--------------|--------------|
| Space group           | P321         | R32          |
| Data range            | 50–2.0       | 50–1.85      |
| RMSD bond length (Å)  | 0.011        | 0.012        |
| Average protein B-factor (Å²) | 29.6      | 22.8          |
| Average solvent B-factor (Å²) | 44.6      | 41.0          |
| Wilson B (Å²)         | 33.2         | 25.6         |

| Ramachandran plot* | Form I       | Form II      |
|--------------------|--------------|--------------|
| Favored            | 98.8%        | 99.2%        |
| Allowed            | 1.2%         | 0.8%         |
| Disallowed         | 0.0%         | 0.0%         |
| Protein atoms      | 1899         | 1895         |
| Water molecules    | 145          | 184          |
| Protein Data Bank code | 2F6Q        | 2Q34         |

* R = Σ|Fo| − |Fc|/Σ|Fo|, where |Fo| is the observed structure factor, and |Fc| is the calculated structure factor used in the refinement.

**Table 1**

| Data collection |            |               |
|-----------------|-------------|---------------|
| Source          | X-ray       | ESI droplets  |
| Dimensions (Å)  | 1.0332      | 0.95373       |
| Space group     | P321        | P321          |
| Data range      | 50–2.0      | 50–1.85       |
| RMSD bond length (Å) | 0.011      | 0.012        |
| Average protein B-factor (Å²) | 29.6      | 22.8          |
| Average solvent B-factor (Å²) | 44.6      | 41.0          |
| Wilson B (Å²)   | 33.2        | 25.6         |
(Tyr\textsuperscript{82}, Lys\textsuperscript{86}, and His\textsuperscript{240}) are present within the primarily hydrophobic active site chamber (Fig. 4B).

**Decarboxylase Activity of Active Site Variants with an ACP-linked Substrate**—Based on the active site structure and substrate modeling, the three polar residues in the active site chamber, Tyr\textsuperscript{82}, Lys\textsuperscript{86}, and His\textsuperscript{240}, were tested by site-directed mutagenesis. Proteins possessing Y82F, K86A, K86Q, H240A, or H240Q were produced and activity toward ACP-linked substrates was evaluated in a coupled ECH1/ECH2 assay (8). FTICR-MS confirmed that the ECH1 and ECH2 products were produced from the ACP-linked substrate (Fig. 5C and supplemental Fig. S3) 8). The effect of each amino acid substitution was identical in assays with ACP- and CoA-linked substrates (Ref. 8 and data not shown). Substitution of Tyr\textsuperscript{82} by Phe led to 2-fold reduction in product formation in comparison to wild type (Fig. 5), indicating that Tyr\textsuperscript{82} plays little or no role in catalysis. Consistent with this conclusion, the Tyr\textsuperscript{82} phenyl ring in the 1.65 Å crystal structure of the Y82F
variant had a similar position to the Phe\textsuperscript{82} phenyl ring in the wild type enzyme. On the other hand, product formation was reduced more than 20-fold in the K86A, K86Q, H240A, and H240Q variants compared with wild type. Thus, we conclude that Lys\textsuperscript{86} and His\textsuperscript{240} play an important role in substrate binding or catalysis. Although the proteins are stable, the CurF ECH\textsubscript{2} K86A, K86Q, H240A, and H240Q variants have not yielded crystals.

Decarboxylase Preference for ACP-linked Substrates over CoA-linked Substrates—CoA- and ACP-linked substrates for ECH\textsubscript{2} were prepared by enzymatic synthesis to evaluate the substrate preference of the decarboxylase without the complication of a coupled ECH\textsubscript{1}/ECH\textsubscript{2} assay. CurF ECH\textsubscript{2} had a 20-fold preference for the ACP-linked substrate under our standard assay conditions. CurF ECH\textsubscript{2} decarboxylated >70\% of 3-methylglutaconyl-ACP but only 3\% of 3-methylglutaconyl-CoA (Fig. 6). A preference for ACP-linked substrates has not been reported for any member of the crotonase superfamily. Although it presumably has an ACP-linked substrate \textit{in vivo}, CurE ECH\textsubscript{1} differed from CurF ECH\textsubscript{2} and showed no substrate preference by providing similar yields of 3-methylglutaconyl-ACP from (R,S)-HMG-ACP and of 3-methylglutaconyl-CoA from (R,S)-HMG-CoA (Fig. 6).

**DISCUSSION**

The N-terminal ECH\textsubscript{1} domain of CurF possesses the crotonase superfamily fold. As is seen with the two other structurally characterized crotonase superfamily members possessing biotin-independent decarboxylase activity
The active site chamber enhances the reactivity of these groups. The hydrophobic environment of CurF ECH\textsubscript{2} and Tyr\textsubscript{140} of MMCD are not in analogous parts of the protein structures. Significantly, Tyr\textsubscript{82} is not conserved among sequences of decarboxylases in the crotonase superfamily. Nevertheless, the close structural similarity with other crotonase superfamily members suggests strongly that Ala\textsubscript{78} and Lys\textsubscript{86} is the only invariant polar residue within the active site chamber. Lys\textsubscript{86} resides on helix \textit{A} of the CurF ECH\textsubscript{2} monomer, and we propose that it serves to stabilize the substrate carboxylate here and in other ECH\textsubscript{2}-like decarboxylases of HCS cassettes (Fig. 3\textsubscript{S2}). His\textsubscript{240} is not a candidate proton donor in CurF ECH\textsubscript{2}, given the relatively fixed position to assist catalysis by hydrogen bonding with the substrate carboxylate, perhaps orienting the carboxyl optimally to be a leaving group (Fig. 7\textsubscript{A}). In liganded CarB, His\textsuperscript{229} is in an identical position to His\textsubscript{240} of CurF ECH\textsubscript{2} and also was proposed to function by stabilizing the substrate carboxylate (22). The importance of His\textsubscript{240} in catalysis was demonstrated in the H240A and H240Q variants of CurF ECH\textsubscript{2}, in which catalysis was severely impaired (Fig. 5). His\textsubscript{240} of CurF ECH\textsubscript{2} is the only invariant polar residue within the active site chamber, and we propose that it serves to stabilize the substrate carboxylate here and in other ECH\textsubscript{2}-like decarboxylases of HCS cassettes (Fig. 3\textsubscript{C}). An analogous His does not exist in MMCD. The equivalent helix to \textalpha\textsubscript{10} of CurF ECH\textsubscript{2}, which contains His\textsubscript{240}, is further from the active site chamber in MMCD and forms extensive trimer contacts with the adjacent monomer.

After decarboxylation, proton donation to the substrate C-4 carbon to form isopentenyl-ACP is necessary for enolate collapse to the \Delta^2 unsaturated product (Fig. 7\textsubscript{B}). Lys\textsuperscript{86} is the most likely proton donor in CurF ECH\textsubscript{2}, given the relatively fixed position of the enolate anion in the oxyanion hole during substrate binding and the hydrophobic nature of the active site chamber. Lys\textsuperscript{86} resides on helix \textalpha\textsubscript{2} and has the necessary flexibility and reach to accommodate proton donation at the C-4 carbon of the substrate (Fig. 4\textsubscript{B}). His\textsuperscript{240} is not a candidate pro-
ton donor because it appears to be unable to reach the C-4 carbon of the decarboxylated intermediate, as required for subsequent product formation, and also unable to move closer to the C-4 carbon. Its backbone is held in place as part of the longer helix α2, and its side chain is positioned by a hydrogen bond with the carbonyl of Thr147 (Fig. 4A) deep within the active site chamber. In contrast, Lys86 is located on helix α2, the most mobile part of the CurF ECH2 structure as seen in distinctly higher crystallographic temperature factors. This region of the crotonase fold is generally mobile and is completely disordered in some structures.

Proton addition to the C-4 carbon following decarboxylation should produce a $\Delta^2$ unsaturated product, whereas proton addition to C-2 should yield the $\Delta^1$ product (Fig. 7). Examples of both reaction routes exist, based on the structures of the PKS products (Fig. 1). Thus, the mechanism of ECH2 is a key determinant of the regiochemistry of the ultimate product of each pathway. The active site chambers of other ECH2-like decarboxylases appear equally as hydrophobic as the CurF ECH2 active site with the exception of the loop between helices α2 and α3 (residues 89–95 of CurF; Fig. 3C). This region covers the active site (Fig. 4B) but differs in length and sequence among the ECH2 decarboxylases (Fig. 3C). It is uncertain whether all residues aligned with CurF 89–95 constitute the α2-α3 loop in these decarboxylases or whether some of them are part of a longer helix α3, as in CarB and MMCD (supplemental Fig. S2). Variability in the α2-α3 loop is expected among the ECH2 decarboxylases. Each enzyme must possess a unique conformation to accommodate its substrate, often bulkier than the methyl group of the CurF ECH2 substrate (Fig. 1). In addition, a unique proton donor must be positioned to generate specifically a $\Delta^2$ or $\Delta^1$ unsaturated product. Candidate proton donors include Lys80, Asp83, or Asp84 of PksI; His68, Asp73, or Asp77 of VirE; Cys74 or Asp75 of TaY; Asp79 or Asp82 of MupK; Glu88, Glu91, Lys92, or Asp95 of JamJ; and Asp2949, Lys2954, Asp2962, or Glu2963 of Ped3 (Fig. 3C).

The CurF ECH2 domain is the first crotonase superfamily member shown to act preferentially upon ACP-linked substrates. Product formation in vitro with an ACP-linked substrate was 20-fold greater than with a CoA-linked substrate (Fig. 6). Despite the strong preference for ACP-linked substrates by CurF ECH2, structure alignment with other crotonase superfamily enzymes having CoA ligands (supplemental Fig. S2) revealed remarkably few structural changes in the CoA-binding region. Structural motifs that recognize the CoA adenine ring in other crotonase enzymes and the overall backbone conformation that forms the CoA adenosine-binding site are identical in CurF ECH2. One notable difference in CurF ECH2 is the position of the Tyr73 side chain. In the CoA-dependent enzymes, a basic side chain occupies this space and interacts with the 4-phosphate of pantothetic acid and the 5′-phosphate of the CoA nucleotide. CurF Tyr73 is conserved as a bulky phenylalanine or tyrosine residue in the ECH2-like decarboxylases from HCS cassettes with presumed specificity for ACP-linked substrates (Fig. 3C), whereas the analogous position is always a smaller alanine, serine, or valine residue in the CoA-dependent enzymes. Thus, it appears that CoA is a poorer substrate of CurF ECH2 in part because the Tyr73 side chain blocks a basic side chain (Arg38 in CurF) from entering the CoA-binding site. CurF ECH2 containing an alanine substitution at the Tyr73 site was insoluble and has frustrated efforts to test this hypothesis directly. The structural conservation in this region implies a relatively recent evolution of protein function toward ACP-linked substrates.

An intriguing outcome of the CurF ECH2 structure is the apparent symmetry mismatch between its trimeric structure and the fundamentally dimeric structure of several downstream domains in the CurF polypeptide (enoyl reductase, ketosynthase, and dehydratase) (7, 51). The CurF ECH2 domain possesses the crotonase self-association fold with the active site fully formed by each monomer and also demonstrates a capacity for trimer dissociation in solution (supplemental Fig. S1). Thus, the ECH2 domain may be monomeric in the context of full-length CurF, similar to the bacterial fatty acid β-oxidation multienzyme complex in which an N-terminal monomeric crotonase domain is fused to a dimeric dehydrogenase domain (50). Nevertheless, existence of the classical crotonase trimer in the isolated ECH2 domain (Fig. 3B) is strongly suggestive of a trimeric ECH2 within full-length CurF. This could be accomplished in either of two ways. CurF could be a hexamer containing two ECH2 trimers flexibly tethered to three dimers of the downstream domains. Alternatively, CurF could be a dimer in which ECH2 forms a heterotrimer with ECH1 (CurE). Determination of the oligomeric organization of CurF is an ongoing investigation.

In summary, the crystal structure of the N-terminal ECH2 domain of CurF PKS-NRPS multifunctional protein from L. majuscula possesses a crotonase fold. The backbone amides of Ala78 and Gly118 form an oxyanion hole. The hydrophobic active site chamber includes only three polar side chains. Of these, Lys86 and His240 are critical for catalytic activity, but Tyr82 is not. Based on assay results and comparisons with other crotonase decarboxylases, His240 is proposed to stabilize the substrate carboxylate, and Lys286 is proposed to donate a proton to the C-4 position of product. These structural features enable CurF ECH2 to generate specifically the $\Delta^2$ alkene regiochemistry in formation of the key isopentenyl-ACP product during curacin A biosynthesis. CurF ECH2 is highly selective for ACP-linked substrates, whereas CurE ECH1 is nonselective. Strong sequence conservation of both the ECH1 dehydratases (35–95% identity) and the ECH2 decarboxylases (35–60% identity) encoded by HCS cassettes suggests that these domains have similar substrate preferences to the curacin ECH1 and ECH2.

Acknowledgments—We thank David Akey for assistance with domain assignments. We are grateful for the synthetic DNA of CurA-ACP(II) from Christopher Calderone and Christopher T. Walsh. The country of Curacão and the CARMABI Research Station are acknowledged for access to the source organism, the curacin A-producing L. majuscula. We are indebted to the staff of the GM/CA beamlines (supported by the NIGMS and NCI, National Institutes of Health) at the Advanced Photon Source (supported by the United States Department of Energy).
