

Characterization of the Substrate Specificity of PhlD, a Type III Polyketide Synthase from *Pseudomonas fluorescens**[§]

Received for publication, July 10, 2006, and in revised form, August 23, 2006 Published, JBC Papers in Press, August 24, 2006, DOI 10.1074/jbc.M606500200

Wenjua Zha, Sheryl B. Rubin-Pitel¹, and Huimin Zhao²

From the Departments of Chemical and Biomolecular Engineering, and Chemistry, Center for Biophysics and Computational Biology, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

PhlD, a type III polyketide synthase from *Pseudomonas fluorescens*, catalyzes the synthesis of phloroglucinol from three molecules of malonyl-CoA. Kinetic analysis by direct measurement of the appearance of the CoASH product ($k_{\text{cat}} = 24 \pm 4 \text{ min}^{-1}$ and $K_m = 13 \pm 1 \mu\text{M}$) gave a k_{cat} value more than an order of magnitude higher than that of any other known type III polyketide synthase. PhlD exhibits broad substrate specificity, accepting C₄–C₁₂ aliphatic acyl-CoAs and phenylacetyl-CoA as the starters to form C₆-polyoxoalkylated α -pyrones from sequential condensation with malonyl-CoA. Interestingly, when primed with long chain acyl-CoAs, PhlD catalyzed extra polyketide elongation to form up to heptaketide products. A homology structural model of PhlD showed the presence of a buried tunnel extending out from the active site to assist the binding of long chain acyl-CoAs. To probe the structural basis for the unusual ability of PhlD to accept long chain acyl-CoAs, both site-directed mutagenesis and saturation mutagenesis were carried out on key residues lining the tunnel. Three mutations, M21I, H24V, and L59M, were found to significantly reduce the reactivity of PhlD with lauroyl-CoA while still retaining its physiological activity to synthesize phloroglucinol. Our homology modeling and mutational studies indicated that even subtle changes in the tunnel volume could affect the ability of PhlD to accept long chain acyl-CoAs. This suggested novel strategies for combinatorial biosynthesis of unnatural pharmaceutically important polyketides.

Polyketide synthases (PKSs)³ from bacteria, fungi, and plants produce an amazing array of polyketides that have important biological and pharmacological activities (1–6). These PKSs are

divided into three classes. Type I PKSs, the so-called modular PKSs, are multifunctional enzymes and catalyze the synthesis of polyketide in a modular fashion (7–11). Type II PKSs, on the other hand, are multienzyme complexes with each domain located on a single protein (12, 13). Type III PKSs, the simplest PKSs, are homodimeric enzymes that iteratively carry out the polyketide synthesis at one single active site (14–16). This third type of PKSs shares a common mechanism for assembly of the polyketide scaffold: loading of the starter substrate (derived from an acyl-CoA) onto the active site cysteine, repetitive decarboxylative condensation reaction of the extension substrate (usually malonyl-CoA) to elongate the polyketide chain, and the final cyclization of polyketide product that is subsequently released from the PKS (5, 16). Although simple, type III PKSs produce a wide variety of polyketides. The diversity of the products arises from the choices of different starter units (aliphatic or aromatic acyl-CoAs) and extender units (malonyl or modified malonyl-CoAs), the number of elongation cycles to control the size of the polyketides, and the final cyclization pattern of the linear polyketides (Claisen condensation, aldol condensation, or heterocyclic lactonization).

Type III PKSs are traditionally isolated from higher plants, where they play important roles in the defense system against microorganisms, but recently a handful of type III PKSs have been discovered in prokaryotes. The first one is RppA from *Streptomyces griseus*, which catalyzes the condensation of five malonyl-CoA molecules to synthesize 1,3,6,8-tetrahydroxynaphthalene (THN), a key intermediate in the formation of melanin in *Streptomyces* (17, 18). Other members of the bacterial type III PKSs include DpgA from *Amycolatopsis orientalis*, which produces 3,5-dihydroxyphenylacetyl-CoA from four molecules of malonyl-CoA (19, 20); PKS10 and PKS18 from *Mycobacterium tuberculosis*, which are capable of accepting long chain acyl-CoAs (C₁₂–C₂₀) to produce α -pyrones (21); and PhlD from *Pseudomonas fluorescens* (22). PhlD was first discovered in the gene cluster responsible for the biosynthesis of 2,4-diacetylphloroglucinol in a variety of *P. fluorescens* strains (23, 24). However, its catalytic function was not defined until very recently (22). In a proposed mechanism (22), PhlD first condenses three molecules of malonyl-CoA to form a triketide. The cyclization of the triketide is then initiated by a decarboxylation reaction of the carboxyl group derived from the starter malonyl-CoA, leading to a C₆ carbanion that attacks C₁ in a Claisen condensation (Fig. 1A). A similar mechanism was also seen with RppA, where the Claisen condensation of the C₁₀ carbanion to C₁ cyclizes the linear pentaketide into THN (Fig. 1B) (18, 25). Although the plant type III PKSs typically share a high level of

* This work was supported by an Office of Naval Research Grant N00014-02-1-0725. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[§] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Table S1 and Fig. S1.

¹ Supported by the United States National Science Foundation Graduate Research Fellowship Program.

² To whom correspondence should be addressed: Depts. of Chemical and Biomolecular Engineering and Chemistry, Center for Biophysics and Computational Biology, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave., Urbana, IL 61801. Tel.: 217-333-2631; Fax: 217-333-5052; E-mail: zhao5@uiuc.edu.

³ The abbreviations used are: PKS, polyketide synthase; THN, 1,3,6,8-tetrahydroxynaphthalene; THNS, THN synthase; Ni-NTA, nickel-nitrilotriacetic acid; HPLC, high pressure liquid chromatography; LC, liquid chromatography; ESI, electrospray ionization; MS, mass spectrometry; TFA, trifluoroacetic acid.

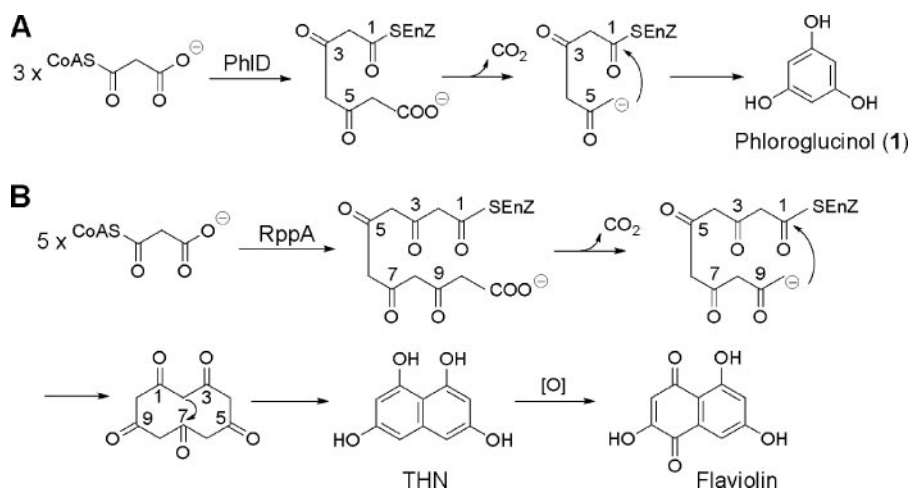


FIGURE 1. Proposed reaction mechanisms of PhlD (A) and THNS (or RppA; B).

sequence identity (70–95%), those of bacterial origin are more dissimilar. RppA and PhlD share about 40% sequence identity, the highest among all the bacterial type III PKSs, whereas the rest share only ~20%.

Type III PKSs have broad substrate specificity and utilize a variety of acyl-CoAs as the starter substrates (26, 27), as exemplified first by chalcone synthase, which reacts with C₄–C₆ aliphatic acyl-CoAs and benzoyl-CoA to produce acylphloroglucinols. Likewise, RppA can accept aliphatic acyl-CoAs with chain length of up to 8 carbons and catalyzes the synthesis of polyketides as long as hexaketide (27). The recently available crystal structures of several type III PKSs together with mutational studies have unveiled some structural bases for the formation of dissimilar polyketide products (28–32), namely the important role of the volume of the active site cavity in controlling the substrate selectivity as well as the length of the polyketides. This hypothesis, however, could not explain the reactivity of RppA with long chain acyl-CoAs and the extra extension step in polyketide synthesis. The crystal structure of RppA from *Streptomyces coelicolor* revealed a buried tunnel extending out from the active site cavity. The tunnel was proposed to dynamically increase the volume of the active site cavity and consequently assist in the binding of long chain acyl-CoAs and the polyketide extension (25). Such a buried tunnel was also found in PKS18 to be responsible for binding C₁₂–C₂₀ aliphatic acyl-CoAs as the starter substrates (32). To date, extensive studies have been conducted in plant type III PKSs to investigate the mechanism dictating the substrate, extension, and final cyclization specificity (16, 29, 30). However, the bacterial counterparts remain poorly characterized, presumably because of the low sequence homology among the latter category.

Here we report the biochemical, mutational, and structural characterization of the substrate specificity of PhlD. Enzyme kinetics was determined by direct measurement of the free CoASH produced by the reaction. The thermal stability of PhlD was also investigated. A starter substrate profiling employing various acyl-CoAs revealed an unusual substrate specificity of PhlD. A homology structural model accompanied by mutational analysis was further used to probe the structural basis for

this unusual property. A buried tunnel, similar to that seen in RppA, was identified adjacent to the active site in the modeled PhlD structure and was proposed to aid the acceptance of long chain acyl-CoAs. Surprisingly, it was shown that subtle changes in the tunnel volume can drastically affect the substrate specificity of PhlD.

EXPERIMENTAL PROCEDURES

Materials—The pET-26b(+) expression vector and BL21(DE3) *Escherichia coli* strain were obtained from Novagen (Madison, WI). *Phusion* DNA polymerase, T4 DNA ligase, and restriction endonucleases were purchased from New England Biolabs (Beverly, MA). QIAprep spin plasmid miniprep kit, QIAquick gel extraction kit, QIAquick PCR purification kit, and Ni-NTA-agarose resins were obtained from Qiagen. Oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA). [2-¹⁴C]Malonyl-CoA (1.74 GBq/mmol) was purchased from PerkinElmer Life Sciences. Nonlabeled acyl-CoAs and all of the other reagents were obtained from Sigma-Aldrich.

Cloning, Expression, and Purification—The *phlD* gene was PCR-amplified from the genomic DNA of *P. fluorescens* strain Pf-5 with the specific forward primer (PhlDfor) and reverse primer (PhlDrev) designed according to the putative *phlD* sequence of that strain (GenBank™ accession number CP000076): PhlDfor, GGATCGGTCATATGTCTACACTTTGCCTTCCACAC (the NdeI restriction site is underlined and the ATG start codon is italicized); PhlDrev: TCATTACTCGAGGGCGGTCCACTCGCCAC (the XhoI restriction site is underlined). The PCR amplification was carried out using *Phusion* DNA polymerase (New England Biolabs). The amplified product was then cloned into the expression vector pET-26b(+) using the restriction enzymes NdeI and XhoI. The stop codon was not included in the reverse primer to add a His₆ tag to the C terminus of PhlD. After confirming the accuracy of the cloned *phlD* gene by DNA sequencing, the plasmid pET26-PhlD was transformed into *E. coli* BL21 (DE3) for protein overexpression. The cells harboring the plasmid were initially cultured in LB medium at 37 °C to an A₆₀₀ of 0.6. Then 0.25 mM isopropyl-β-D-thiogalactoside inducer was added, and the cells were further grown at 20 °C overnight to achieve maximal protein expression. The cells were harvested by centrifugation (3,000 × g, 4 °C, 20 min), resuspended in lysis buffer (50 mM sodium phosphate buffer, 300 mM NaCl, 10 mM imidazole, pH 8.0), supplemented with protease inhibitor mixture (Novagen), and stored at –70 °C when needed. PhlD protein with a C-terminal His₆ tag was purified to near homogeneity by one-step affinity chromatography using Ni-NTA resins. Briefly, the cells were lysed by two passes through the French press at 16,000 p.s.i. After removal of the cell debris by centrifugation (5,000 × g, 4 °C, 20 min), the cleared lysate was incubated with the Ni-NTA resins at 4 °C for 1 h. The mixture was then loaded onto a

Substrate Specificity of PhlD

polypropylene column, and the resins were washed extensively (50 mM sodium phosphate buffer, 300 mM NaCl, 20 mM imidazole, pH 8.0). PhlD protein was eluted with elution buffer (50 mM sodium phosphate buffer, 300 mM NaCl, 100 mM imidazole, pH 8.0). Finally, the purified PhlD protein was buffer-exchanged and concentrated to ~2–3 mg/ml using dry polyethyleneglycol 20,000 sprinkled around the dialysis bag (10,000 MWCO).

Determination of Enzyme Kinetics and Stability—The kinetic parameters of PhlD for its substrate malonyl-CoA were determined based on the method used by Tseng *et al.* (20), with some modifications. The experiments were carried out in triplicate using a range of concentrations of malonyl-CoA (5–100 μM). For each reaction, substrate was incubated with PhlD (4–16 $\mu\text{g}/\text{ml}$) in buffer (50 mM Tris-HCl, pH 7.5, 500 μl) at 25 °C using a refrigerated water bath. Buffer and malonyl-CoA were preincubated at 25 °C for 20 min before addition of enzyme. At 10 s, 1 min, and 2 min, 100- μl aliquots of the reaction were quenched by the addition of 20 μl of 4% trifluoroacetic acid (TFA) (final concentration 0.667% (TFA)) and kept at room temperature until analysis by HPLC. The reactions were analyzed using an Agilent 1100 Series HPLC and ZORBAX SB-C18 reverse-phase column (3.0 \times 150 mm, 3.5 μm) (Agilent Technologies, Palo Alto, CA) monitoring at 258 nm (0.5 ml/min; 25 °C; 0–1 min, 3% B; 1–15 min, 3–15% B; where A = H₂O, 0.1% TFA and B = acetonitrile, 0.1% TFA). Peaks at 7.5, 8.9, and 10 min correspond to authentic CoASH, malonyl-CoA, and acetyl-CoA, respectively. The initial rates of CoASH production determined from these reactions were fitted to the Michaelis-Menten equation using nonlinear least squares regression analysis in Microcal Origin 5.0 (Microcal Software, Northampton, MA) to calculate k_{cat} and K_m .

To determine thermal inactivation, samples of PhlD enzyme (1 mg/ml concentration) were incubated at 25, 30, and 37 °C in a MJ Research PTC-200 thermocycler (± 0.3 °C). The aliquots were taken over a range of times and cooled on ice, and the enzymatic activity was subsequently determined using the two-enzyme system utilized by Achkar *et al.* (22), with some slight adjustments. Each reaction contained 50 μg of PhlD, 150 milliunits of α -ketoglutarate dehydrogenase enzyme, 2 mM α -ketoglutarate, 0.3 mM β -NAD⁺, and 160 μM malonyl-CoA in potassium phosphate buffer (50 mM, pH 7.0). The initial reaction rate was determined at 25 °C by following the appearance of NADH at 340 nm ($\epsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$) in a Varian Cary 100 Bio UV-visible spectrophotometer (Palo Alto, CA). Typically inactivation was monitored until activity was diminished by at least 70%.

Substrate Specificity Characterization—The standard reaction was carried out in 50 mM Tris-HCl buffer, pH 7.5, containing 1 mM Tris-(2-carboxyethyl)-phosphine hydrochloride, 200 μM starter acyl-CoA, 20 μM malonyl-CoA (including 1.7 μM [^{2-¹⁴C}]malonyl-CoA), and 60 μg of purified wild type or mutant PhlD protein in a total volume of 600 μl . The reaction was incubated at 30 °C for 30 min and quenched with 20 μl of 6 M HCl. The products were extracted twice with 600 μl of ethyl acetate, and the combined extracts were dried under N₂ flush. Finally, the residue material was dissolved in 50 μl of methanol for HPLC analysis. A ZORBAX SB-C18 reverse-phase column

(3.0 \times 150 mm, 3.5 μm ; Agilent) was run at room temperature on a SpectraSYSTEM LC system (Thermo, Waltham, MA) with a flow rate of 0.5 ml/min and UV absorbance detection at 280 nm (UV₂₈₀). Gradient elution was carried out with H₂O (solvent A) and acetonitrile (solvent B), both containing 0.1% TFA. The mobile phase linear gradient was set for reactions A, B, and C (Fig. 2): 5% B to 40% B in 30 min; for reactions D, E, F, and H: 20% B to 80% B in 30 min; and for reaction G: 40% B to 100% B in 30 min. The elution was collected every minute, and the radioactivity was measured by a LS 6500 Multi-Purpose Scintillation Counter (Beckman Coulter, Fullerton, CA), which was plotted on the same HPLC profile as UV₂₈₀. For direct detection of the products under UV₂₈₀, the substrates in the reaction were increased to 0.5 mM for the starter acyl-CoA and 0.2 mM for malonyl-CoA without the addition of radiolabeled malonyl-CoA.

Product Characterization—The polyketide products were characterized by liquid chromatography-electrospray ionization mass spectrometry (LC/ESI-MS) on a LCQ DECAXP ion trap (Thermo, Waltham, MA). Reactions for LC/ESI-MS were scaled up 10-fold compared with those prepared for UV detection, and only nonradioactive malonyl-CoA was used.

Homology Modeling—The crystal structure of 1,3,6,8-tetrahydroxynaphthalene synthase (THNS) from *S. coelicolor* was downloaded from the Protein Data Bank (accession code 1U0M). The dimer coordinates of the structure were prepared by PyMol (pymol.sourceforge.net/). The amino acid sequence of Pf-5 PhlD was then aligned with THNS using the sequence alignment tool provided by Insight II software (Insight II, version 2000; Accelrys Inc., San Diego, CA), followed by manual inspection and correction to make sure that the conserved residues were well aligned. This sequence alignment was subjected to the automatic model building module of Insight II, MODELER, using default parameters with moderate refinement for the loop regions to build the homology structure model of PhlD. Of nine models created, the best model was selected based on visual inspection, such as the proper orientation of the key active site residues, the score from the Profile3D estimation, and the values of the ϕ and ψ angles calculated by ProStat. The reaction intermediate, lauroyl-primed triketide, was generated and manually docked into the active site cavity of the created model using Molecular Operating Environment (Chemical Computing Group Inc., Montreal, Canada), according to the configuration of the polyethyleneglycol molecule cocrystallized with THNS. Using forcefield MMFF94s (33), the whole structure was subjected to energy minimization to relieve steric and torsional artifacts from the modeling and docking processes.

Site-directed Mutagenesis—Site-directed mutagenesis was carried out using the megaprimer PCR method (34), which consists of two sequential PCRs, both employing the wild type *phlD* gene as the template. Briefly, the first PCR was carried out with the mutagenesis primer containing the desired mutation with 20–21 additional nucleotides on either side (5' and 3') and one of the flanking primers, PhlDfor or PhlDrev, previously used for amplification. The first PCR product was then used as the megaprimer in the second PCR in conjunction with the other flanking primer to amplify the full-length product. The mutant

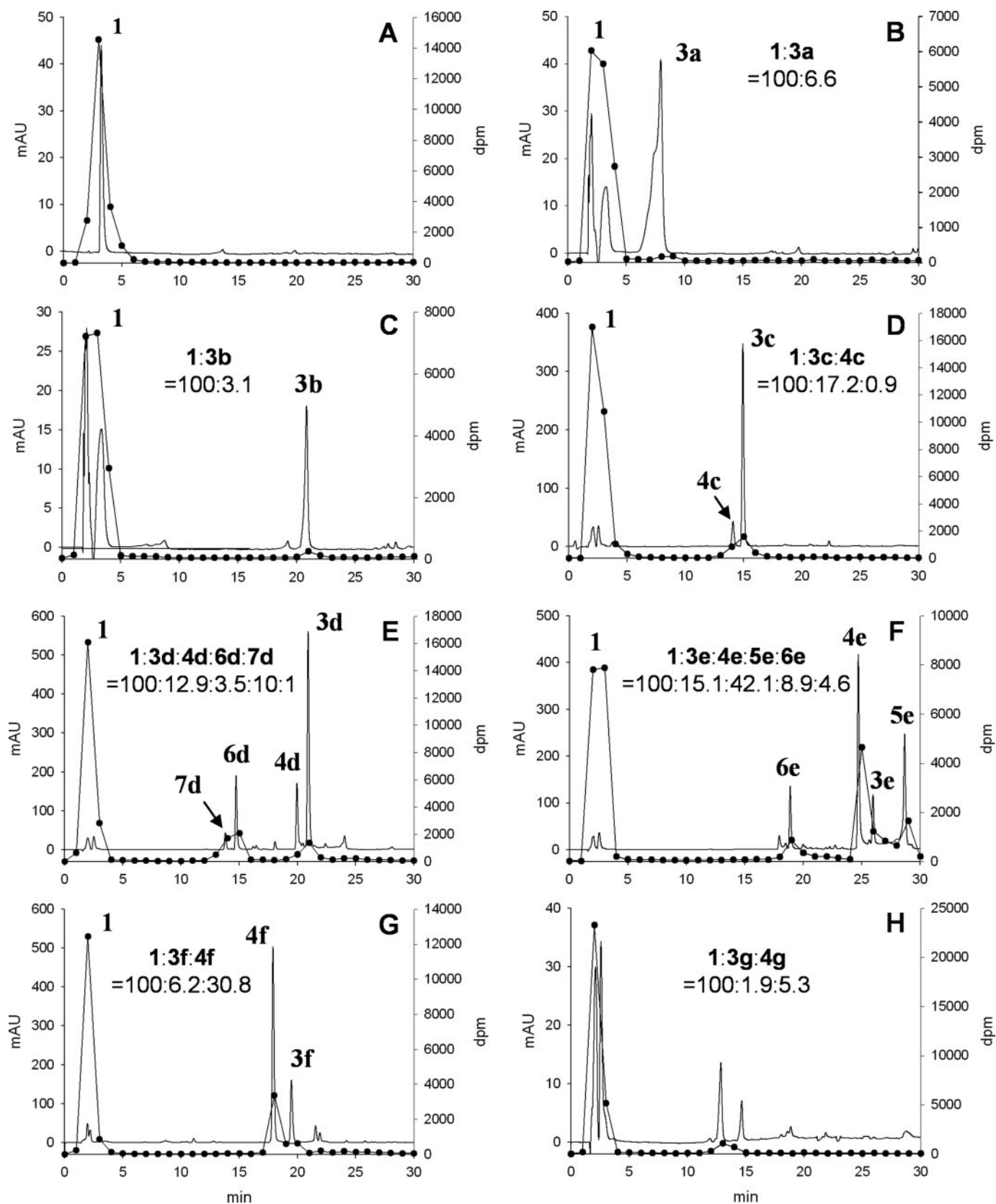


FIGURE 2. HPLC analysis of the products synthesized by PhID from various acyl-CoA starter substrates. The reactions for radiolabeling HPLC contain 20 μM malonyl-CoA including 1.7 μM [$2\text{-}^{14}\text{C}$]malonyl-CoA and 200 μM acyl-CoA starters. A, malonyl-CoA; B, acetoacetyl-CoA (**2a**); C, butyryl-CoA (**2b**); D, hexanoyl-CoA (**2c**); E, octanoyl-CoA (**2d**); F, decanoyl-CoA (**2e**); G, lauroyl-CoA (**2f**); H, phenylacetyl-CoA (**2g**). For substrate numbering, see also Fig. 4. The reactions for UV₂₈₀ HPLC contain 200 μM nonradioactive malonyl-CoA and 500 μM of each acyl-CoA starters. The UV₂₈₀ HPLC for A was prepared by injection of pure phloroglucinol. The solid lines represent the UV absorbance at 280 nm (milli-absorbance unit (mAU), left y axis), and the solid lines with filled circles represent radioactivity measurement (disintegrations/min (dpm), right y axis).

Substrate Specificity of PhlD

genes were subcloned into pET-26b(+) using the restriction enzymes NdeI and XhoI and sequenced to confirm that only the desired mutations were incorporated.

Creation and Screening of the Saturation Mutagenesis Library—Saturation mutagenesis was carried out using an overlap extension PCR-based method as described elsewhere (35). PhlD_{for} and PhlD_{rev} were used as the flanking primers to amplify the mutant *phlD* gene. One of the mutagenesis primers has the same backbone nucleotide composition as those used in site-directed mutagenesis but contains the sequence NNS to replace the codon encoding the residue to be mutated (N = A, C, G, or T, and S = G or C; supplemental Table S1). As a result, the substitution NNS allows the incorporation of all 20 amino acids while reducing the total number of possible codons to 32. The second mutagenesis primer is exactly complementary with the 5' 20–21 nucleotides flanking the NNS codon in the first mutagenesis primer. For construction of a saturation library for each selected residue, two separate PCRs were performed, using the wild type *phlD* gene as the template and each containing one flanking primer and one mutagenic primer. The two PCR products were purified from the agarose gel using QIA-Quick gel extraction kit after DNA electrophoresis and treated with the restriction enzyme DpnI to remove the methylated template. The DpnI-digested PCR products were then combined and reassembled into the full-length gene in an overlap extension PCR without primers and subsequently amplified with the two flanking primers. Products of the correct size were purified from the gel and cloned into pET-26b(+) as described previously. The libraries were then transformed into BL21(DE3) and screened for the clones that still retained the wild type activity to produce phloroglucinol. The screening was based on the specific colorimetric reaction of cinnamaldehyde with phloroglucinol in the presence of HCl to produce a yellow-colored adduct ($\lambda_{\max} = 466$ nm). For each saturation mutagenesis library, around 20 active clones were randomly picked in addition to a wild type PhlD clone and a clone containing the empty pET-26b(+) vector that would be used as the positive and negative control, respectively. The expression of PhlD-c(His)₆ was induced as described previously in 5 ml of terrific broth. All of the following experiments were performed in 96-well microplates. The cell pellets were first lysed by the addition of 1 mg/ml of lysozyme and freeze-thaw. Then the recombinant proteins were purified using Ni-NTA resins (Qiagen) according to the manufacturer's protocol. The purified proteins (≈ 0.1 mg/ml) were assayed with 0.5 mM lauroyl-CoA and 0.2 mM nonradiolabeled malonyl-CoA as described in the enzyme assay section in a total reaction volume of 200 μ l. After ethyl acetate extraction, the products were analyzed on an Agilent 1100 Series HPLC system with the ZORBAX SB-C18 column to assess the incorporation of lauroyl-CoA by the mutants as compared with the wild type PhlD reaction. The HPLC program was modified to shorten the analysis time: 70–80% acetonitrile in H₂O (both with 0.1% TFA) in 10 min. At the same time, the ability of the mutant PhlD to synthesize phloroglucinol was also assayed with 0.2 mM of the malonyl-CoA substrate, and the amount of phloroglucinol produced was estimated using the colorimetric assay employing cinnamaldehyde

on a SpectraMax 190 plate reader (Molecular Devices, Sunnyvale, CA) with the wavelength detection set at 466 nm.

RESULTS

Kinetics and Stability of PhlD—The recent characterization of native PhlD by Achkar *et al.* (22) reported significantly different kinetic values ($k_{\text{cat}} = 10$ min⁻¹, $K_m = 5.6$ μ M, $k_{\text{cat}}/K_m = 1787$ mM⁻¹ min⁻¹) than other chalcone synthase superfamily enzymes, for example the related bacterial enzymes RppA from *S. griseus* ($k_{\text{cat}} = 0.77$ min⁻¹, $K_m = 0.93$ μ M, $k_{\text{cat}}/K_m = 828$ mM⁻¹ min⁻¹) (27) and DpgA from *A. orientalis* ($k_{\text{cat}} = 0.81$ min⁻¹, $K_m = 15$ μ M, $k_{\text{cat}}/K_m = 54$ mM⁻¹ min⁻¹) (20), which condense and cyclize five and four malonyl-CoAs, respectively, whereas PhlD utilizes three. We determined the kinetic parameters of PhlD to be $k_{\text{cat}} = 24 \pm 4$ min⁻¹, $K_m = 13 \pm 1$ μ M, and $k_{\text{cat}}/K_m = 1883$ mM⁻¹ min⁻¹. Interestingly, these results do corroborate the finding that, in terms of k_{cat} value, PhlD is highly unusual among the chalcone synthase superfamily of type III PKS enzymes. Additionally, the inactivation of PhlD over time was analyzed at 25, 30, and 37 °C. Loss of activity did not follow the typical first order kinetics, so results are not reported as the half-life of thermal inactivation. The times required to reduce the activity of the PhlD enzyme by 50% at 25, 30, and 37 °C were found to be 128, 74, and 7.2 min, respectively, showing that PhlD is relatively unstable even at lower temperatures.

Starter Substrate Specificity of PhlD—The starter acyl-CoA specificity of PhlD was investigated using [2-¹⁴C]malonyl-CoA mixed with various acyl-CoAs. The reactivity of PhlD toward each acyl-CoA was then deduced from the amount of [2-¹⁴C]malonyl-CoA incorporation as the extension unit into the final polyketide products. Efficient acceptance of unnatural acyl-CoAs could only be seen when they were in at least 2-fold excess over malonyl-CoA (data not shown). This indicates a clear preference of PhlD to select malonyl-CoA over other acyl-CoAs as the starter substrates, similar to what was seen with RppA that also uses malonyl-CoA as the sole substrate (27). When nonradiolabeled malonyl-CoA was used in place of unnatural acyl-CoA in the reaction mixture, only one radioactive product was observed, which co-eluted with authentic phloroglucinol (Fig. 2A). This confirmed the function of PhlD as a phloroglucinol synthase from solely the malonyl-CoA substrate.

To further probe the substrate specificity of PhlD, we first tested aliphatic acyl-CoAs with different chain length. PhlD did not react very well with short chain acyl-CoAs. It showed barely detectable activity toward acetyl-CoA. When butyryl-CoA (**2b**) was used as the starter substrate, beside phloroglucinol (**1**), a small amount of product was also detected by both the UV absorbance at 280 nm and the radioactivity measurement (Fig. 2C). The product (**3b**) was identified by LC/ESI-MS to be a triketide ($M_r = 154$) from the butyryl-primed reaction. The amount of the triketide and phloroglucinol were calculated from their radioactivity as a ratio of 3.1:100. Interestingly, PhlD showed a greater ability to accept long chain aliphatic acyl-CoAs. With hexanoyl-CoA (**2c**) as the starter, PhlD produced a triketide (**3c**, $M_r = 182$) and a tetraketide (**4c**, $M_r = 224$) as determined by LC/ESI-MS (Fig. 2D). The product amount ratio

was 100 (phloroglucinol):17.2 (triketide):0.9 (tetraketide). When octanoyl-CoA (**2d**) was incubated with PhlD as the starter, in addition to the triketide (**3d**, $M_r = 210$) and tetraketide (**4d**, $M_r = 252$) products as observed with the hexanoyl-CoA reaction, two more products were identified (Fig. 2E). Product **6d** had a M_r of 336 by LC/ESI-MS, which corresponds to a hexaketide resulted from five condensation cycles of the octanoyl moiety with malonyl-CoA. Product **7d** was a heptaketide ($M_r = 378$) from six cycles of condensation. The product amount from octanoyl-CoA reaction had a ratio of 100 (phloroglucinol):12.9 (triketide):3.5 (tetraketide):10 (hexaketide):1 (heptaketide). Decanoyl-CoA (**2e**) was accepted by PhlD with even greater efficiency, leading to the formation of a triketide (**3e**, $M_r = 238$), tetraketide (**4e**, $M_r = 280$), pentaketide (**5e**, $M_r = 322$), and hexaketide (**6e**, $M_r = 364$) (Fig. 2F). The product amount from the decanoyl-CoA starter was significantly increased compared with that from hexanoyl and octanoyl-CoA, with the ratio of 100 (phloroglucinol):15.1 (triketide):42.1 (tetraketide):8.9 (hexaketide):4.6 (heptaketide). PhlD could still react with lauroyl-CoA (**2f**) to produce a triketide (**3f**, $M_r = 266$) and a tetraketide (**4f**, $M_r = 308$) (Fig. 2G), whose amounts were 6.2 and 30.8, respectively, compared with the amount of phloroglucinol that was set to 100. However, when the aliphatic chain in an acyl-CoA was further elongated, such as in myristoyl-CoA and palmitoyl-CoA, no more activity was detected with PhlD. Therefore, PhlD could accept aliphatic acyl-CoA with carbon chain length of up to 12.

Five additional acyl-CoAs were also investigated for their reactivity with PhlD. Acetoacetyl-CoA (**2a**), whose structure mimics an intermediate in the polyketide elongation pathway, was not used very well by PhlD, yielding only a triketide (**3a**, $M_r = 126$; Fig. 2B) in addition to phloroglucinol (**1**). Under the assay conditions, PhlD poorly accepted acyl-CoAs with a branched β position, such as isobutyryl-CoA and could not accept acyl-CoAs with a rigid β position, such as crotonoyl-CoA. Regarding acyl-CoAs with aromatic side chains, benzoyl-CoA was not accepted by PhlD, but phenylacetyl-CoA (**2g**) could be incorporated by PhlD with moderate efficiency, generating a triketide (**3g**, $M_r = 202$) and a tetraketide (**4g**, $M_r = 244$) (Fig. 2H). The product amount was distributed as 100 (phloroglucinol):1.9 (triketide):5.3 (tetraketide). Therefore, PhlD exhibits very broad substrate specificity and produces a wide range of products (see Fig. 4). The amount of phloroglucinol produced in the malonyl-CoA reaction was set as the 100% standard for calculation of the relative incorporation efficiency of other acyl-CoAs by PhlD as the starter substrates, and the data are summarized in Table 1. The remarkable ability of PhlD to react with long chain and bulky acyl-CoA substrates and to carry out extensive polyketide elongation is not reflected by its physiological activity, which only produces a triketide from the condensation of three molecules of malonyl-CoA. These results could not be plausibly interpreted by the prevailing theory for type III PKS catalysis, which states that the size of the polyketide product, influenced by the choice of the starter molecule and the number of extension steps, is controlled by the volume of the active site cavity (16, 29, 30).

Products from the Unnatural Starter Substrates—To identify the products formed from the unnatural acyl-CoA starters, we

TABLE 1
The reactivity of PhlD with various acyl-CoAs as the starter substrates

Starter substrates	Incorporation efficiency by PhlD ^a
	%
Malonyl-CoA	100
Acetyl-CoA	<0.1
Butyryl-CoA (2b)	0.2 ± 0.1
Hexanoyl-CoA (2c)	8 ± 1
Octanoyl-CoA (2d)	7 ± 2
Decanoyl-CoA (2e)	18 ± 2
Lauroyl-CoA (2f)	7.6 ± 0.4
Acetoacetyl-CoA (2a)	1.6 ± 0.1
Isobutyryl-CoA	<0.1
Benzoyl-CoA	<0.1
Phenylacetyl-CoA (2g)	3.3 ± 0.8
Crotonoyl-CoA	<0.1

^a For each acyl-CoA starter, its incorporation was calculated as the total sum of all its products, whose amounts were determined by the division of the measured dpm by the number of [²⁻¹⁴C]malonyl-CoA incorporation associated with each product.

performed LC/ESI tandem MS analysis. For triketides **3a–g**, the $[M-CO_2-H]^-$ fragment was the major fragment ion on the spectra, giving the structure of a C6-alkylated α -pyrone. The spectra for tetraketides **4a–g** showed an additional fragment ion at m/z of 125 corresponding to $[C_6H_5O_3]^-$, indicating a C6-2'-oxoalkylated α -pyrone structure. The spectrum of the pentaketide from the decanoyl-CoA starter (**5e**) had a fragment ion of 167 besides a fragment ion of 125 corresponding to $[C_6H_5O_3]^-$ and a fragment ion of 277 corresponding to $[M-CO_2-H]^-$, which agrees with a structure of 4-hydroxyl-6-(2',4'-dioxotridecyl)-2-pyrone (Fig. 3A). Both hexaketides from the octanoyl-CoA starter and decanoyl-CoA starter (**6d** and **6e**) had the signature fragment ions for α -pyrones, 125 corresponding to $[C_6H_5O_3]^-$ and 291 (**6d**) or 319 (**6e**) corresponding to $[M-CO_2-H]^-$, respectively. Two additional fragment ions corresponding to the cleavage at the 2-oxo and 4-oxo groups in the side chain were seen in the spectrum of octanoyl-primed hexaketide (**6d**), giving a structure of 4-hydroxyl-6-(2',4',6'-trioxotridecyl)-2-pyrone (Fig. 3B). These two equivalent fragment ions somehow were missing in the decanoyl-hexaketide spectrum (**6e**). Nevertheless, the two signature α -pyrone fragment ions together with the determined molecular weight implied **6e** with the structure of 4-hydroxyl-6-(2',4',6'-trioxopentadecyl)-2-pyrone (Fig. 3C). The spectrum of the heptaketide product from octanoyl-CoA reaction (**7d**) gave five fragment ions, including $[C_6H_5O_3]^-$, $[M-CO_2-H]^-$, and the other three generated by the cleavage at the internal keto groups in the side chain. The fragmentation pattern of **7d** was consistent with the proposed structure of 4-hydroxyl-6-(2',4',6',8'-tetraoxopentadecyl)-2-pyrone (Fig. 3D). As a result, the LC-MS/MS analysis demonstrated that the products identified from the unnatural acyl-CoA reaction by radiolabeling and UV HPLC were all single ring α -pyrones. The product formation is summarized in Fig. 4. Note that previous studies showed that RppA was able to produce phloroglucinol type products when primed with hexanoyl-CoA and octanoyl-CoA (27). However, with PhlD, no phloroglucinols were observed as major products from the unnatural starter substrates.

Homology Modeling—To elucidate the structural basis for the unusual ability of PhlD to accept long chain acyl-CoAs, we built a homology structural model for PhlD based on the only

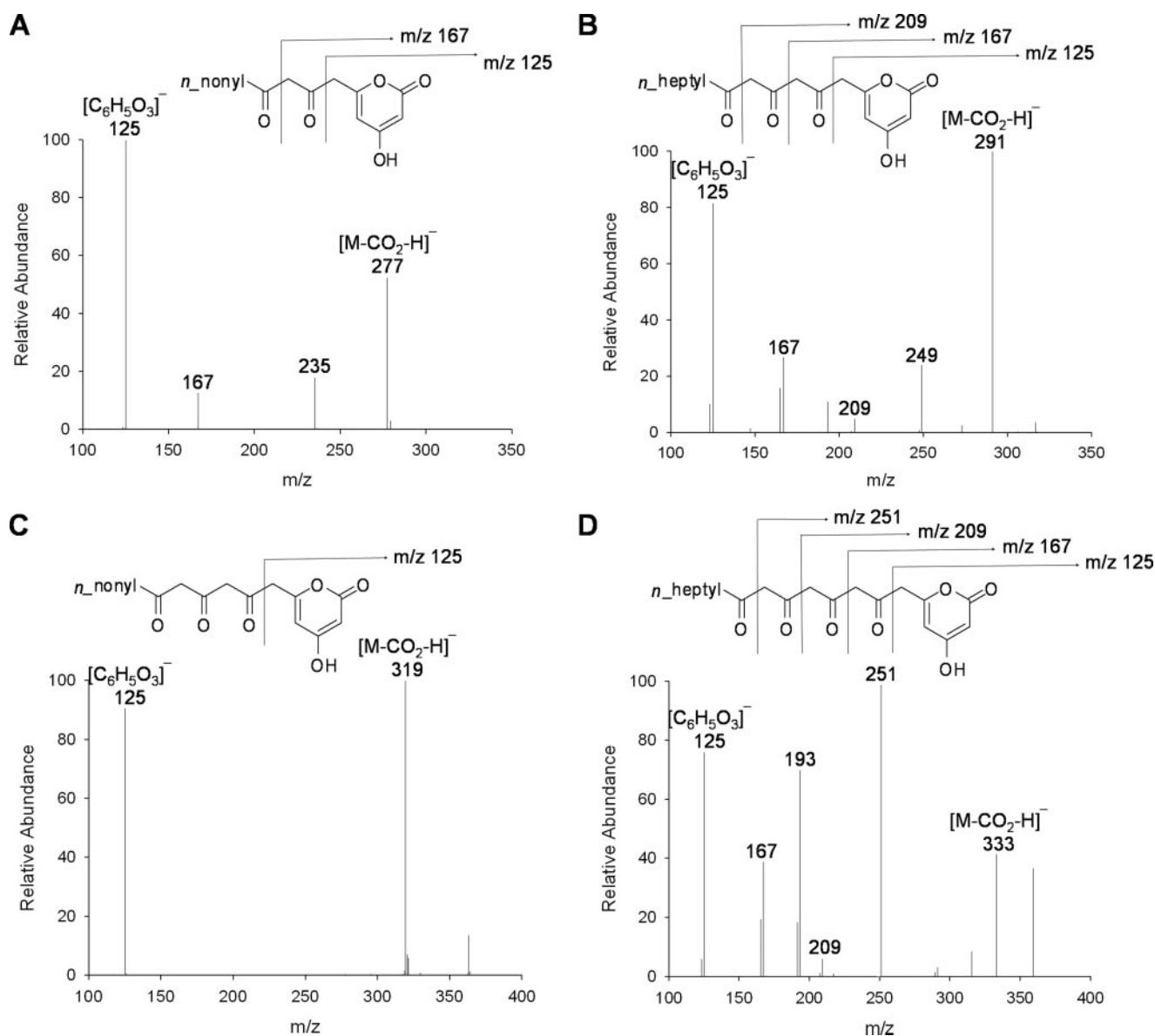


FIGURE 3. Selected MS/MS spectra of the products from unnatural acyl-CoAs. A, 4-hydroxyl-6-(2',4'-dioxotridecyl)-2-pyrone (**5e**), precursor ion at $m/z = 321$. B, 4-hydroxyl-6-(2',4',6'-trioxotridecyl)-2-pyrone (**6d**), precursor ion at $m/z = 335$. C, 4-hydroxyl-6-(2',4',6'-trioxopentadecyl)-2-pyrone (**6e**), precursor ion at $m/z = 363$. D, 4-hydroxyl-6-(2',4',6',8'-tetraoxopentadecyl)-2-pyrone (**7d**), precursor ion at $m/z = 377$.

crystal structure that showed high sequence identity with PhlD: THNS from *S. coelicolor* (Protein Data Bank accession code 1U0M; 43.4% identity). When the modeled PhlD structure was superimposed onto the structure of THNS, all of the residues lining the active site cavity were superimposed well; in particular, the catalytic triads, Cys-His-Asn (located at position 138, 272, and 305, respectively in PhlD), overlapped excellently.

It was reported that subtle change of the structure elements underneath the active site cavity of THNS formed an extra cavity, which extended out as a buried tunnel from the traditional active site cavity and was occupied by a polyethyleneglycol heptamer in the crystal structure of THNS (25). This tunnel was proposed to assist the binding of long chain acyl-CoAs and the extra extension steps beyond tetraketide in THNS (25). A similar tunnel was also found in the modeled structure of PhlD. Thus, we hypothesized that the tunnel in PhlD might also be responsible for the ability of PhlD to bind long chain acyl-CoA

starters. Note that the closest homolog to PhlD, RppA from *S. griseus*, can only accept C_4 – C_8 acyl-CoAs, whereas PhlD is capable of accepting acyl-CoAs with C_4 – C_{12} aliphatic chains. Comparison of the residues lining the tunnel between PhlD and RppA should cast some light regarding how the tunnel affects the acceptance of long chain acyl-CoAs. About 13 residues delineated the tunnel in the modeled structure of PhlD and were picked to compare with those in RppA by sequence alignment (supplemental Fig. S1). Among them, only six residues were different (highlighted by boxes): Met²¹, His²⁴, Leu²⁸, Leu⁵⁹, Ala⁶⁰, and Ala¹⁸⁵ in PhlD were replaced by Thr²¹, Val²⁴, Thr²⁸, Thr⁵⁹, Leu⁶⁰, and Asn¹⁸⁵ in RppA. These residues were located in the bottom half of the buried tunnel and might be important for differentiating the ability to accept long chain acyl-CoAs between PhlD and RppA (see Fig. 6A).

Mutational Analysis—To investigate the role of the above six different residues between PhlD and RppA on substrate bind-

the better ability of PhID to accept long chain acyl-CoAs than that of RppA.

To further probe the functional role of these six sites, we carried out saturation mutagenesis on each of them. The resulting libraries were prescreened for the clones active to produce phloroglucinol. The library at site 21 (site 21 library) had only 38% active fraction as assessed by the colorimetric assay on 200 randomly picked colonies, site 24 library had 48%, site 28 library had 53%, site 59 library had 61%, site 60 library had 36%, and site 185 library had 36%. To obtain a reasonable representation of the entire saturation mutagenesis library, about 20 active clones were randomly picked from each site followed by protein purification. The purified mutant proteins were first assayed *in vitro* for the retained activity to produce phloroglucinol from malonyl-CoA alone. Only those that showed more or less the same phloroglucinol forming activity as the wild type PhID were further analyzed on HPLC for their products from the lauroyl-CoA primed reaction as described under "Experimental Procedures." Mutants M21I, H24V, and L59M were identified from the libraries of site 21, site 24, and site 59 to have reduced ability to accept lauroyl-CoA, whereas the libraries at site 28, site 60, and site 185 did

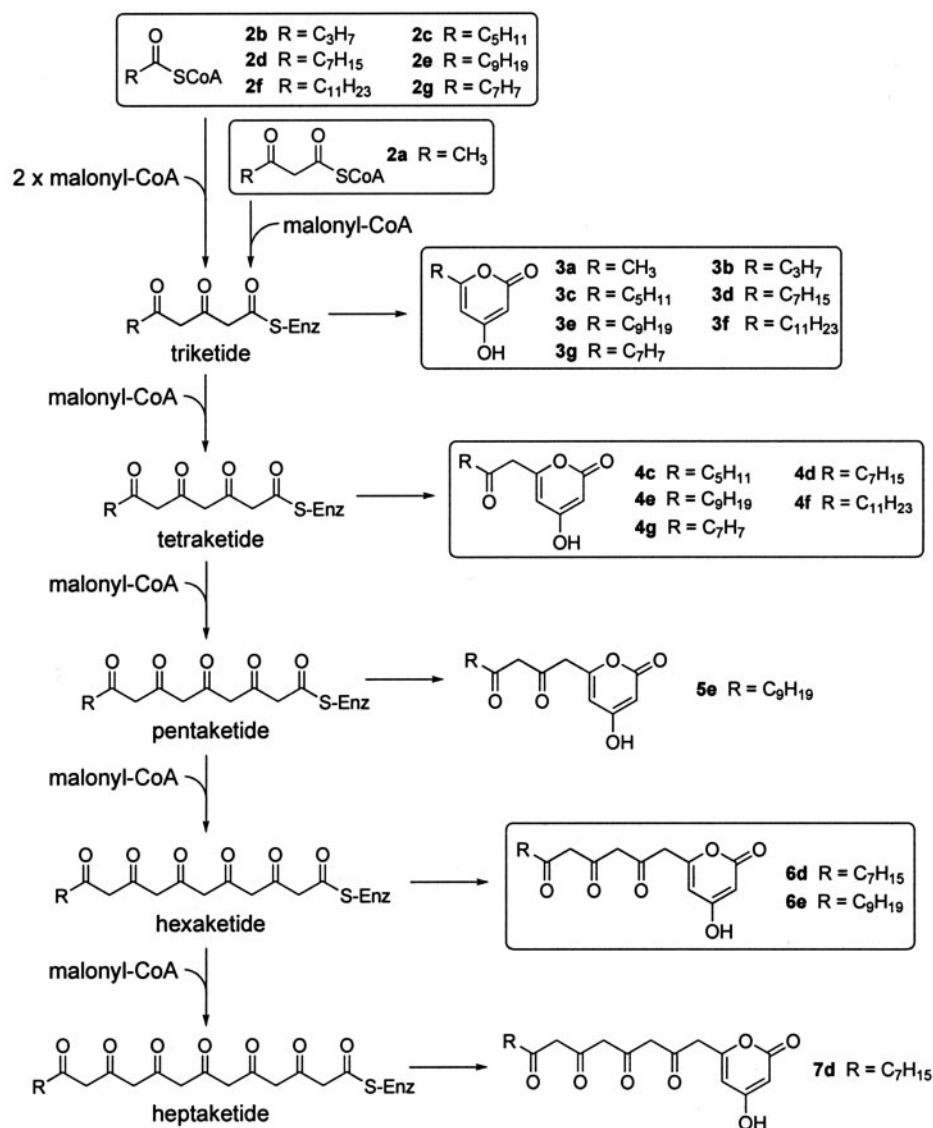


FIGURE 4. Summary of PhID reactions with various acyl-CoA starter substrates.

ing individually, each of the six residues in PhID was changed, respectively, into the corresponding residue in RppA by site-directed mutagenesis. The ability of these mutants to accept long chain acyl-CoAs was probed using lauroyl-CoA, the substrate used by PhID but not by RppA. None of the six mutants had the expected effect on the acceptance of long chain acyl-CoAs: L59T did not produce any soluble protein; M21T and A60L still retained the wild type activity to react with lauroyl-CoA; H24L, although it showed reduced reactivity toward lauroyl-CoA, had lost the ability to produce phloroglucinol to the same extent as the wild type PhID; and A185D was completely inactive. Unlike the successful examples from plant type III PKSs, where one PKS could be converted into a functional equivalent of another PKS by just mutating a few active site residues (29–31), the low sequence homology between bacterial type III PKSs, such as PhID and RppA, made such rational design difficult. Note that the function of residues lining the tunnels of PhID and RppA could be affected by their interacting residues farther away. However, given their important location, these six selected sites might still be important in determining

not yield any mutants with such reduced reactivity toward lauroyl-CoA. These three mutants were purified in large scale and were further verified using the assay condition for direct detection of the products under UV₂₈₀ as described under "Experimental Procedures." The ability of M21I, H24V, and L59M to synthesize phloroglucinol was well maintained, showing 91, 118, and 106% reactivity with malonyl-CoA substrate, respectively, as compared with the wild type PhID (Fig. 5). On the other hand, M21I and L59M lost half of the ability to accept lauroyl-CoA, and more interestingly, H24V retained only around 11% reactivity toward lauroyl-CoA as compared with the wild type PhID (Fig. 5).

DISCUSSION

The closest homolog of PhID in the type III PKS family is RppA. Notably, PhID from *P. fluorescens* Pf-5 and RppA from *S. griseus* have ~45% identity in the amino acid sequence, the highest among all known bacterial type III PKSs. PhID and RppA share several unique features in catalysis. They use solely malonyl-CoA as their physiological substrates with the differ-

Substrate Specificity of PhlD

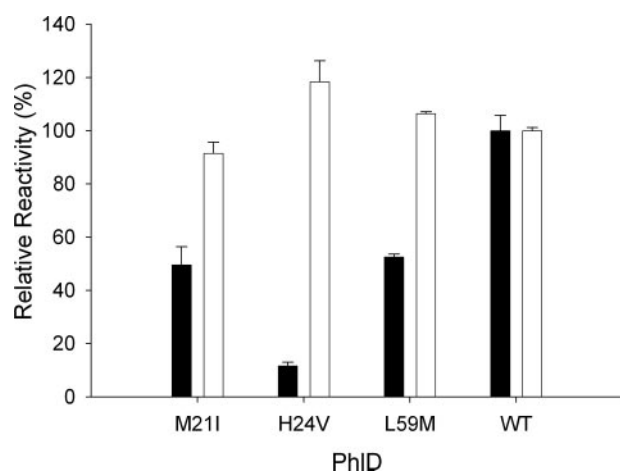


FIGURE 5. Comparison between the mutants and wild type PhlD of the ability to react with lauroyl-CoA (■) and the retained activity to synthesize phloroglucinol (□). The lauroyl-CoA reaction was carried out as described for the direct detection of products under UV₂₈₀, containing 500 μ M lauroyl-CoA, 200 μ M malonyl-CoA, and 0.2 mg/ml of each protein. The products after ethyl acetate extraction were analyzed on HPLC, and their relative amounts were compared by the integrated peak area. The reaction for phloroglucinol synthesis contained only 200 μ M malonyl-CoA and 0.2 mg/ml of each protein. The relative amount of phloroglucinol produced was estimated by the cinnamaldehyde colorimetric assay monitored at 466 nm. WT, wild type.

ence being that PhlD condenses three molecules of malonyl-CoA, whereas RppA condenses five. Another unique feature of PhlD and RppA is that their polyketide products are cyclized through a Claisen condensation of the end carbanion group (generated via the decarboxylation of the starter malonyl-CoA derived carboxyl group) to C1 (linked to the thiol group of the active site cysteine). RppA was shown to have broad starter substrate specificity (27). This study demonstrated that PhlD can also utilize a wide variety of acyl-CoAs as its starter substrate. To fairly compare the reactivity of PhlD toward unnatural acyl-CoAs with that of RppA, we used the previously reported reaction conditions for *S. griseus* RppA with some slight modification (27). Our data revealed a distinct substrate specificity profile for PhlD.

PhlD accepted acyl-CoAs with aliphatic chains of C₄–C₁₂. When octanoyl-CoA was incubated with PhlD as the starter, in addition to the triketide, tetraketide, and hexaketide, the same products as with RppA, a heptaketide product, resulting from six cycles of condensation with malonyl-CoA, was observed. The incorporation efficiency of octanoyl-CoA as the starter was 7% compared with that of malonyl-CoA (Table 1). The substrate specificity study on RppA reported a 6% incorporation efficiency for RppA to react with octanoyl-CoA (27). The rates in that study, however, were all estimated as the comparison with acetoacetyl-CoA. Thus, the actual incorporation rate for RppA, if malonyl-CoA was used as the standard, should be much lower. Surprisingly, decanoyl-CoA, a starter that was not reactive with RppA, was accepted by PhlD, yielding a triketide, tetraketide, pentaketide, and hexaketide with the total incorporation rate of 18% relative to malonyl-CoA (Table 1). PhlD could also react with lauroyl-CoA to produce a triketide and tetraketide. Thus, whereas the upper limit of the aliphatic chain length of the starter acyl-CoAs is 8 carbons for RppA, PhlD is able to accept those with up to 12 carbons. The remarkable

abilities of PhlD to react with longer chain acyl-CoA substrates and carry out extensive polyketide extension over that of RppA were not reflected by their physiological activities, where PhlD only produces a triketide from the condensation of three molecules of malonyl-CoA, whereas RppA condenses five malonyl-CoA into a pentaketide. These results contradict with the well accepted volume controlling model for type III PKS catalysis that the volume of the active site cavity dictates the size of the starter molecule as well as the number of extension steps (16, 29, 30).

To understand the structural basis of the unusual ability of PhlD to accept long chain acyl-CoAs, we built a homology model for PhlD. This modeled structure of PhlD revealed a similar buried tunnel as found in RppA and PKS18, which was believed to accommodate the binding of long chain acyl-CoAs (25, 32). The comparison of the residues lining the tunnel between PhlD and *S. griseus* RppA identified six sites that are occupied by different residues. These sites were hypothesized to contribute to the different binding capabilities of long chain acyl-CoAs between PhlD and RppA. However, direct replacement of the residues in PhlD by the corresponding ones in RppA affected not only PhlD binding of long chain acyl-CoAs but also its physiological activity to synthesize phloroglucinol. This might not be surprising considering the low sequence homology between PhlD and RppA. The function of the residues lining the tunnel could not be accurately predicted because they might be affected by the surrounding interacting residues on the second shell or even the third shell of the tunnel. Therefore, saturation mutagenesis was carried out on the six sites of PhlD individually to more comprehensively explore their effects on the acceptance of long chain acyl-CoAs. Interestingly, the six saturation mutagenesis libraries yielded only 30–60% active fraction that produced phloroglucinol. Although located toward the bottom of the buried tunnel and far away from the actual active site cavity, these six sites still impinged on the physiological activity of PhlD to a great extent.

Three mutations identified from the library screening, which significantly reduced the reactivity of PhlD with lauroyl-CoA but not its physiological activity to synthesize phloroglucinol, were mapped onto the structural model of PhlD. At site 21, the original methionine residue and the new mutation isoleucine both have hydrophobic side chains of similar size. Nevertheless, the branched side chain in isoleucine increased the contacting surface on one side of the tunnel and consequently pushed this side slightly into the tunnel, causing a small reduction in the volume (Fig. 6C) compared with the wild type enzyme (Fig. 6B). Indeed, using the cavity volume calculation function provided by Molecular Operating Environment, we found that the change of Met to Ile decreased the volume of the tunnel slightly from 1227 to 1199 Å³. The new narrowed tunnel might not accommodate the large aliphatic moiety of lauroyl-CoA as well as the wild type PhlD, resulting in its loss of ability to accept this acyl-CoA substrate by more than half. At site 24, the imidazole plane of histidine in the wild type PhlD did not directly point into the tunnel because of the presence of surrounding bulky residues such as Tyr¹⁷² (Fig. 6A, shown as a *stick*). Valine, on the contrary, contains a small side chain that can be positioned toward the tunnel. As a result, the substitution of His with Val

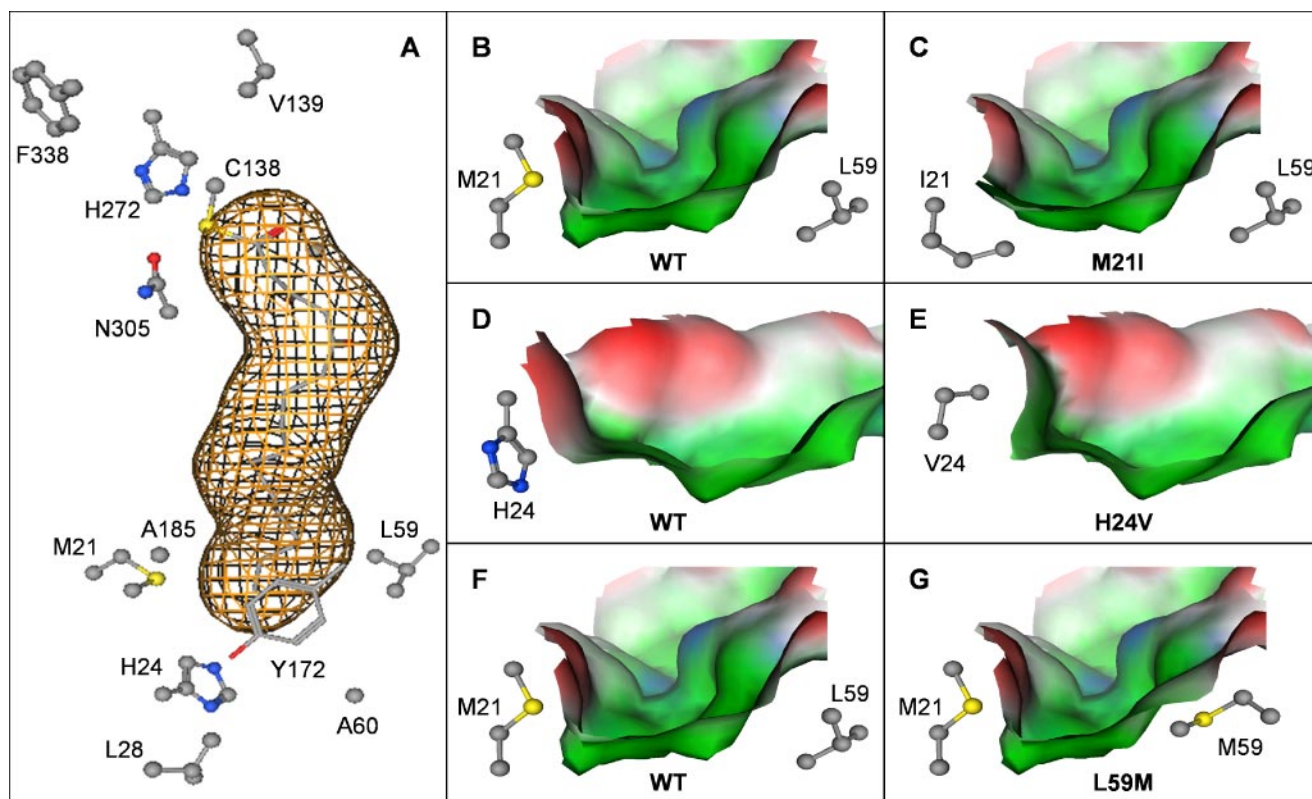


FIGURE 6. Modeling and mutational studies of the tunnel for binding of long chain acyl-CoAs. *A*, schematic representation of the key residues lining the tunnel and the active site cavity. The lauroyl-primed triketide intermediate, covalently linked to the active site cysteine (Cys¹³⁸), was modeled in the PhlD active site cavity and was shown with an electron density cloud around it (generated by Molecular Operating Environment). *B–G*, molecular surface at the tunnel. The residues lining the front side of the tunnel in *A* were removed for a clear view of the tunnel surface. *B* and *F*, wild type PhlD (WT) showing Met²¹ and Leu⁵⁹; *C*, mutant M21I; *D*, WT showing His²⁴; *E*, mutant H24V; *G*, mutant L59M. *B*, *C*, *F*, and *G* were viewed from the bottom of the tunnel, whereas *D* and *E* were viewed from the right side of the tunnel as shown in *A*. Green, blue, and red represent hydrophobic, hydrophilic, and exposed (due the removal of the front residues) surfaces, respectively.

also caused a reduction in the tunnel volume by 78 Å³. The bottom of the tunnel in H24V was raised up high enough to hinder the binding of lauroyl-CoA, giving only 1/10 of the reactivity compared with that of the wild type enzyme (Fig. 6, *D* and *E*). The mutation of Leu⁵⁹ to Met made the most obvious change on the tunnel volume as the long side chain of methionine extended into the tunnel, pushing the side of the tunnel inwards (Fig. 6, *F* and *G*). Although a reduction of 94 Å³ in the tunnel volume was obtained with L59M, the most significant of all the three identified mutations, this mutation only rendered 50% decrease in the ability to accept lauroyl-CoA.

In summary, several interesting findings may be drawn from the saturation mutagenesis study. First, it seems that subtle changes in the residues lining the tunnel can significantly affect the substrate specificity. All three mutations identified in our study did not result in significant alteration in the side chain size or the hydrophobicity, and yet they significantly affected the binding of lauroyl-CoA. It is not atypical in protein engineering that small changes in active site residues play a significant role in altering protein substrate specificity (36, 37). In the case of PhlD, more drastic changes of tunnel residues than what was found in this study might not even be desired because they could affect the correct folding of the tunnel and consequently damage the overall activity of PhlD. Second, among the three mutations, H24V had the biggest impact on the reactivity of PhlD toward lauroyl-CoA, but this mutation reduced the vol-

ume of the tunnel less than L59M. Mutation H24V occurred at the bottom of the tunnel, whereas M21I and L59M were both located on the side of the tunnel. As a result, H24V shortened the length of the tunnel, whereas the other two only influenced the diameter. The positioning of the mutations along the tunnel might explain why H24V affected the acceptance by PhlD of lauroyl-CoA more drastically than the other two mutations. Finally, it is apparent that all three mutations identified in this study that affected the acceptance of lauroyl-CoA could not be readily predicted by rational design alone. This demonstrated the power of the combination of random mutagenesis and rational design in altering protein substrate specificity. Very likely, there might be other residues along the tunnel that could be tuned to affect PhlD binding of long chain acyl-CoAs. Nevertheless, our findings shed some light on the structural features controlling the selection of long chain acyl-CoAs in type III PKSs, namely that even subtle changes in the cavity volume can significantly affect the acceptance of these substrates.

Acyl-CoAs with shorter chains were incorporated by PhlD much less efficiently than by RppA. Butyryl and hexanoyl-CoA as the starter substrates gave only 0.2 and 8% incorporation rates, respectively (Table 1). The tunnel in PhlD seemed to be capable of accommodating larger acyl chain moieties than that in RppA. On the other hand, acyl-CoAs with short chains, such as butyryl and hexanoyl-CoA might not access, bind to, and thus stabilize the tunnel in PhlD as well as the RppA tunnel.

Substrate Specificity of PhlD

Interestingly, the reactivity of PhlD toward acetoacetyl-CoA was only 1.6% (Table 1), less than 1/20 of that of RppA if the amount of the triketide product, 6-methyl-4-hydroxyl-pyrone (3a), relative to the malonyl-CoA product in the reaction was compared (6.6 triketide:100 phloroglucinal for PhlD versus 188 triketide:100 THN for RppA) (Fig. 2B) (27). The acetoacetyl moiety, after being loaded onto the thiol group of the active site cysteine, imitated the structure of the diketide in both PhlD and RppA polyketide elongation pathways. According to the reaction mechanism proposed for the phloroglucinal synthesis by PhlD, a carboxyl group or its decarboxylated product, carbanion group, was necessary to be retained at the end position in the linear polyketide intermediate to carry out the final Claisen condensation for ring closure (Fig. 1A). However, the end position in the acetoacetyl starter-derived polyketide was occupied by a methyl group. The environment of the PhlD active site cavity that favors the presence of a negatively charged end group must hinder the binding of the acetoacetyl moiety. On the other hand, RppA produces a much longer polyketide intermediate, and thus its active site cavity might be more flexible to accommodate the end methyl group in the acetoacetyl-primed reaction. PhlD was unable to accept the acyl-CoAs with a branched β position, exemplified by isobutyryl-CoA, whereas RppA could still react with such substrates to a certain extent. The comparison between the active site cavities of the two proteins revealed two different residues near the catalytic triad: Val¹³⁹ and Phe³³⁸ in PhlD versus Ala¹³⁹ and Ile³³⁷ in RppA (Fig. 6A). The two bulkier substitutions in PhlD might result in a steric repulsion to the side group of isobutyryl-CoA. Neither PhlD nor RppA reacted with crotonoyl-CoA or benzoyl-CoA, presumably because of the unfavorable electronic or steric environment conferred by the active site residues to such substrates with a rigid β position. However, phenylacetyl-CoA could be incorporated by PhlD with moderate efficiency (3.3% incorporation rate; Table 1) but not by RppA. Because the more extended phenyl group in the phenylacetyl moiety was very likely to reach to the buried tunnel, the different reactivity seen in this case might imply a bigger and more flexible tunnel in PhlD than in RppA.

Interestingly, when long chain acyl-CoAs, such as octanoyl-CoA or decanoyl-CoA, were incubated with PhlD, extension steps beyond tetraketide were observed, whereas this did not occur with acyl-CoAs containing shorter chains. The same trend was also observed in RppA, where only the longest acceptable acyl-CoA, octanoyl-CoA, was able to produce polyketides with more than three extension steps (27). The crystal structure of *S. coelicolor* THNS (RppA) showed that the three helices that formed the buried polyethyleneglycol binding tunnel underwent dynamic motions, which assisted the active site cavity to selectively increase its volume in response to long chain acyl-CoA starters to form polyketide products with further extension than tetraketide (25). Given the relatively high sequence identity between THNS and PhlD, the same mechanism might be responsible for the ability of PhlD to catalyze multiple extension steps when bound with long chain acyl-CoA starters.

The substrate specificity study of RppA from *S. griseus* reported the formation of phloroglucinols from tetraketides when hexanoyl-CoA and octanoyl-CoA were used as the start-

ers (27). However in our product analysis by LC/ESI-MS, we did not observe the formation of any phloroglucinol type products from the unnatural acyl-CoAs by PhlD within the detection limits of our UV absorbance and radioactivity based HPLC. It is clear that PhlD produces α -pyrones as its major products from the unnatural acyl-CoA starters. Even for RppA, only trace amounts of phloroglucinols were detected from hexanoyl and octanoyl-CoA, which were barely distinguishable from the noise on the HPLC chromatograms (27). The formation of α -pyrones by PhlD is likely to follow the mechanism by which 2-pyrone synthase produces 6-methyl-4-hydroxyl-pyrone (29). The electron distribution and geometry of the PhlD active site cavity stimulate the formation of the C5 oxyanion in the unnatural acyl-CoA starter-initiated linear polyketides and make it adopt a configuration favorable for the lactonizational attack of C5 oxyanion to C1, leading to α -pyrones.

In summary, PhlD exhibits relatively broad substrate specificity compared with other members of the type III PKS family. In addition to its ability to produce phloroglucinal from malonyl CoA, it accepts C₄–C₁₂ aliphatic acyl-CoAs and phenylacetyl-CoA as the starter substrates to form from tri- to heptaketide α -pyrones. The diverse products produced by PhlD from unnatural acyl-CoA starters greatly expand the existing reservoir of polyketides. Homology modeling of PhlD indicated the presence of a buried tunnel extending out from the active site cavity to accommodate the binding of long chain acyl-CoAs. Subsequent saturation mutagenesis guided by the structural knowledge from modeling on the tunnel successfully altered the substrate specificity of PhlD, which demonstrated that even subtle changes in the tunnel volume could affect the ability of PhlD to accept long chain acyl-CoAs. The powerful combination of random mutagenesis and rational design can be generally applied to tailor other properties of polyketide synthases beside the substrate specificity. This provides novel strategies for combinatorial biosynthesis of unnatural pharmaceutically important polyketides.

Acknowledgment—We thank Mike Nickels for help in radioactive HPLC analysis.

REFERENCES

1. Edwards, M. L., Stemreick, D. M., and Sunkara, P. S. (1990) *J. Med. Chem.* **33**, 1948–1954
2. Hopwood, D. A. (1997) *Chem. Rev.* **97**, 2465–2497
3. Jang, M. (1997) *Science* **275**, 218–220
4. Katz, L., and Donadia, S. (1993) *Annu. Rev. Microbiol.* **47**, 875–912
5. Keating, T. A., and Walsh, C. T. (1999) *Curr. Opin. Chem. Biol.* **3**, 598–606
6. Khosla, C., Gokhale, R. S., Jacobsen, J. R., and Cane, D. E. (1999) *Annu. Rev. Biochem.* **68**, 219–253
7. Donadia, S., Staver, M. J., McAlpine, J. B., Swanson, S. J., and Katz, L. (1991) *Science* **252**, 675–679
8. Shen, B. (2000) *Biosynthesis: Aromatic Polyketides, Isoprenoids, Alkaloids* **209**, 1–51
9. Shen, B., and Hutchinson, C. R. (1993) *Science* **262**, 1535–1540
10. Rawlings, B. J. (2001) *Nat. Prod. Rep.* **18**, 190–227
11. Rawlings, B. J. (2001) *Nat. Prod. Rep.* **18**, 231–281
12. Staunton, J., and Weissman, K. J. (2001) *Nat. Prod. Rep.* **18**, 380–416
13. Rawlings, B. J. (1999) *Nat. Prod. Rep.* **16**, 425–484

14. Tropf, S., Kaercher, B., Schroeder, J., and Schroeder, G. (1995) *J. Biol. Chem.* **270**, 7922–7928
15. Moore, B. S., and Hopke, J. N. (2001) *Chembiochem.* **2**, 35–38
16. Austin, M. B., and Noel, J. P. (2003) *Nat. Prod. Rep.* **20**, 79–110
17. Ueda, K., Kim, K. M., Beppu, T., and Horinouchi, S. (1995) *J. Antibiot. (Tokyo)* **48**, 638–646
18. Funai, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y., and Horinouchi, S. (1999) *Nature* **400**, 897–899
19. Chen, H., Tseng, C. C., Hubbard, B. K., and Walsh, C. T. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14901–14906
20. Tseng, C. C., McLoughlin, S. M., Kelleher, N. L., and Walsh, C. T. (2004) *Biochemistry* **43**, 970–980
21. Saxena, P., Yadav, G., Mohanty, D., and Gokhale, R. S. (2003) *J. Biol. Chem.* **278**, 44780–44790
22. Achkar, J., Xian, M., Zhao, H. M., and Frost, J. W. (2005) *J. Am. Chem. Soc.* **127**, 5332–5333
23. Bangera, M. G., and Thomashow, L. S. (1996) *Mol. Plant Microbe. Interact.* **9**, 83–90
24. Bangera, M. G., and Thomashow, L. S. (1999) *J. Bacteriol.* **181**, 3155–3163
25. Austin, M. B., Izumikawa, M., Bowman, M. E., Udvary, D. W., Ferrer, J. L., Moore, B. S., and Noel, J. P. (2004) *J. Biol. Chem.* **279**, 45162–45174
26. Schuz, R., Heller, W., and Hahlbrock, K. (1983) *J. Biol. Chem.* **258**, 6730–6734
27. Funai, N., Ohnishi, Y., Ebizuka, Y., and Horinouchi, S. (2002) *J. Biol. Chem.* **277**, 4628–4635
28. Ferrer, J. L., Jez, J. M., Bowman, M. E., Dixon, R. A., and Noel, J. P. (1999) *Nat. Struct. Biol.* **6**, 775–784
29. Jez, J. M., Austin, M. B., Ferrer, J. L., Bowman, M. E., Schroeder, J., and Noel, J. P. (2000) *Chem. Biol.* **7**, 919–930
30. Abe, I., Oguro, S., Utsumi, Y., Sano, Y., and Noguchi, H. (2005) *J. Am. Chem. Soc.* **127**, 12709–12716
31. Austin, M. B., Bowman, M. E., Ferrer, J. L., Schroeder, J., and Noel, J. P. (2004) *Chem. Biol.* **11**, 1179–1194
32. Sankaranarayanan, R., Saxena, P., Marathe, U. B., Gokhale, R. S., Shanmugam, V. M., and Rukmini, R. (2004) *Nat. Struct. Mol. Biol.* **11**, 894–900
33. Halgren, T. A. (1999) *J. Comput. Chem.* **20**, 720–729
34. Sarkar, G., and Sommer, S. S. (1990) *BioTechniques* **8**, 404–407
35. Georgescu, R., Bandara, G., and Sun, L. (2003) *Methods Mol. Biol.* **231**, 75–83
36. Chockalingam, K., Chen, Z. L., Katzenellenbogen, J. A., and Zhao, H. M. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 5691–5696
37. Lee, J., Ang, E. L., and Zhao, H. M. (2006) *J. Bacteriol.* **188**, 6179–6183