Chemistry & Biology
Article

Structural Insight into Chain-Length Control and Product Specificity of Pentaketide Chromone Synthase from *Aloe arborescens*

Hiroyuki Morita,[1] Shin Kondo,[2] Satoshi Oguro,[3] Hiroshi Noguchi,[3] Shigetoshi Sugio,[2,*] Ikuro Abe,[3,4,*] and Toshiyuki Kohno[1,*]
[1] Mitsubishi Kagaku Institute of Life Sciences (MITILS), 11 Minamiooya, Machida, Tokyo 194-8511, Japan
[2] ZOEGENE Corporation, 1000 Kamoshida, Aoba, Yokohama, Kanagawa 227-8502, Japan
[3] School of Pharmaceutical Science and the COE21 Program, University of Shizuoka, Shizuoka 422-8526, Japan
[4] PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan
*Correspondence: ssugio@rc.m-kagaku.co.jp (S.S.), abei@ys7.u-shizuoka-ken.ac.jp (I.A.), tkohno@mitils.jp (T.K.)
DOI 10.1016/j.chembiol.2007.02.003

SUMMARY

The crystal structures of a wild-type and a mutant PCS, a novel plant type III polyketide synthase from a medicinal plant, *Aloe arborescens*, were solved at 1.6 Å resolution. The crystal structures revealed that the pentaketide-producing wild-type and the octaketide-producing M207G mutant shared almost the same overall folding, and that the large-to-small substitution dramatically increases the volume of the polyketide-elongation tunnel by opening a gate to two hidden pockets behind the active site of the enzyme. The chemically inert active site residue 207 thus controls the number of condensations of malonyl-CoA, solely depending on the steric bulk of the side chain. These findings not only provided insight into the polyketide formation reaction, but they also suggested strategies for the engineered biosynthesis of polyketides.

INTRODUCTION

The chalcone synthase (CHS) (EC 2.3.1.74) superfamily, which consists of type III polyketide synthases (PKSs), produces an array of structurally diverse, biologically active plant polyphenols [1, 2]. The simple homodimeric type III PKSs, consisting of 40–45 kDa subunits, are structurally and mechanistically distinct from the modular type I and dissociated type II PKSs of bacterial origin. The enzymes catalyze iterative decarboxylative condensations of malonyl-CoA with a CoA-linked starter molecule, without the involvement of an acyl carrier protein. Recent crystallographic and site-directed mutagenesis studies have begun to reveal the structural and functional details of the type III PKSs that share a common three-dimensional overall fold and an active-site architecture with a conserved Cys-His-Asn catalytic triad [3–11]. The polyketide formation reaction is initiated by a starter molecule loading onto the active-site Cys, which is followed by malonyl-CoA decarboxylation, polyketide chain elongation, and cyclization/aromatization of the enzyme-bound polyketide intermediate. For example, plant-specific CHS, the pivotal enzyme for flavonoid biosynthesis, produces naringenin chalcone through the sequential condensation of coumaroyl-CoA with three C_2 units from malonyl-CoA [1] (Figure 1A). On the other hand, 2-pyrone synthase (2PS) from *Gerbera hybrida* selects acetyl-CoA as a starter and carries out only two condensations with malonyl-CoA to produce triacetic acid lactone (TAL) [12] (Figure 1B). In principle, the molecular diversity of the polyketide products derives from the differences in (1) the selection of the starter molecule, (2) the number of polyketide chain elongations, and (3) the mechanism of the cyclization and aromatization reaction. Notably, only a small modification of the active-site architecture generates the remarkable functional diversity of the type III PKSs [2, 5].

Pentaketide chromone synthase (PCS) and octaketide synthase (OKS), two novel plant-specific type III PKSs recently obtained from a medicinal plant, *Aloe arborescens*, share 50%–60% amino acid sequence identity with those of the other CHS-superfamily enzymes of plant origin [13, 14]. PCS retains the Cys-His-Asn catalytic triad as well as most of the CHS active-site residues (Figure 2). However, CHS's conserved residues, Thr197, Gly256, and Ser338 (numbering according to *Medicago sativa* CHS), are replaced with Met207, Leu266, and Val351 (numbering according to *A. arborescens* PCS), respectively. Although many of the type III PKSs condense up to three molecules of malonyl-CoA to a starter substrate, PCS catalyzes condensation of five molecules of malonyl-CoA to produce 5,7-dihydroxy-2-methylchromone [13] (Figure 1C), which is a biosynthetic precursor of kehellin, one of the famous antiasthmatic furochromones of medicinal plants [15]. In contrast, as reported in our previous work, the PCS M207G mutant efficiently produced a 1:4 mixture of SEK4:SEK4b from eight molecules of malonyl-CoA [13] (Figures 1D and 1E). The pentaketide-producing PCS was thus functionally transformed into an OKS by the single amino acid substitution; the chemically inert single residue determines the polyketide chain length and the product

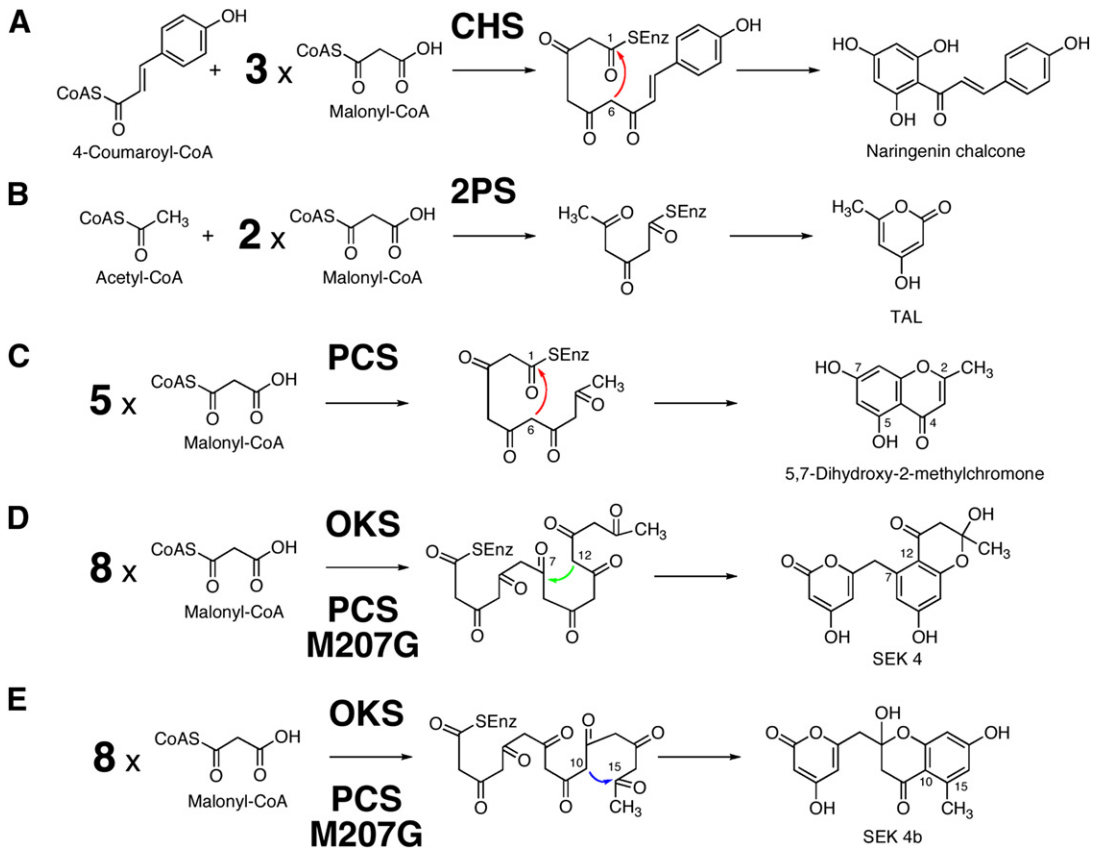

Figure 1. Formation of Polyketides by Type III Polyketide Synthases

(A–E) Proposed synthetic mechanisms of (A) naringenin chalcone from 4-coumaroyl-CoA and three molecules of malonyl-CoA by CHS, (B) triacetic acid lactone (TAL) from acetyl-CoA and two molecules of malonyl-CoA by 2PS, (C) 5,7-dihydroxy-2-methylchromone from five molecules of malonyl-CoA by PCS, (D) SEK4 from eight molecules of malonyl-CoA by OKS and the M207G mutant, and (E) SEK4b from eight molecules of malonyl-CoA by OKS and the M207G mutant. The red arrow depicts a C-6/C-1 Claisen-type cyclization reaction. The green arrow shows a C-12/C-7 aldol-type cyclization reaction, while the blue arrow indicates a C-10/C-15 aldol-type cyclization reaction.

specificity of the enzyme. The octaketides SEK4 and SEK4b are known as the shunt products of the minimal type II PKS for the benzoisochromanequinone actinorhodin (*act* from *Streptomyces coelicolor*) [16] and are the longest polyketides generated by the structurally simple, homodimeric type III PKS [14, 17, 18].

In this paper, we now present the crystal structures of the pentaketide-producing wild-type *A. arborescens* PCS and its octaketide-producing M207G mutant. The three-dimensional structures at 1.6 Å resolution clearly revealed that residue 207 lining the active-site cavity indeed occupies a crucial position for the polyketide chain-elongation reactions, and provided a structural basis for polyketide chain-length control in the type III PKS enzymes.

RESULTS

Overall Structure of PCS

The homodimeric structure of *A. arborescens* PCS complexed with CoA-SH was solved by molecular replacement, by using *G. hybrida* 2PS [4] as a search model, and was refined to 1.6 Å resolution. The experimental map allowed the modeling of residues 1–402, with 3 additional residues from the protease cleavage site of monomer A, and residues 8–402 of monomer B (Figure 3A). The pentaketide-producing PCS shares 58% identity with *M. sativa* CHS and 57% identity with *G. hybrida* 2PS (Figure 2). As expected from the sequence homology, the overall structure of PCS is nearly identical to those of *M. sativa* CHS [3] and *G. hybrida* 2PS, with rms deviations of 0.7 Å and 0.9 Å, respectively, in the Cα atoms (Figure 3B). Each monomer consists of 403 amino acids with an αβαβα pseudo symmetric motif and forms a symmetric dimer with a two-fold axis. Upon dimerization, each monomer buries 3400 $Å^2$ of its surface, and Met147 (numbering in *A. arborescens* PCS) forms cross-subunit interactions by protruding into another monomer, forming a partial wall of the active-site cavity. The conserved catalytic triad (Cys174-His316-Asn349) is buried inside, in a location and orientation very similar to those of *M. sativa* CHS and *G. hybrida* 2PS, and is structurally maintained in each monomer. The internal active site is connected to the protein surface by the characteristic

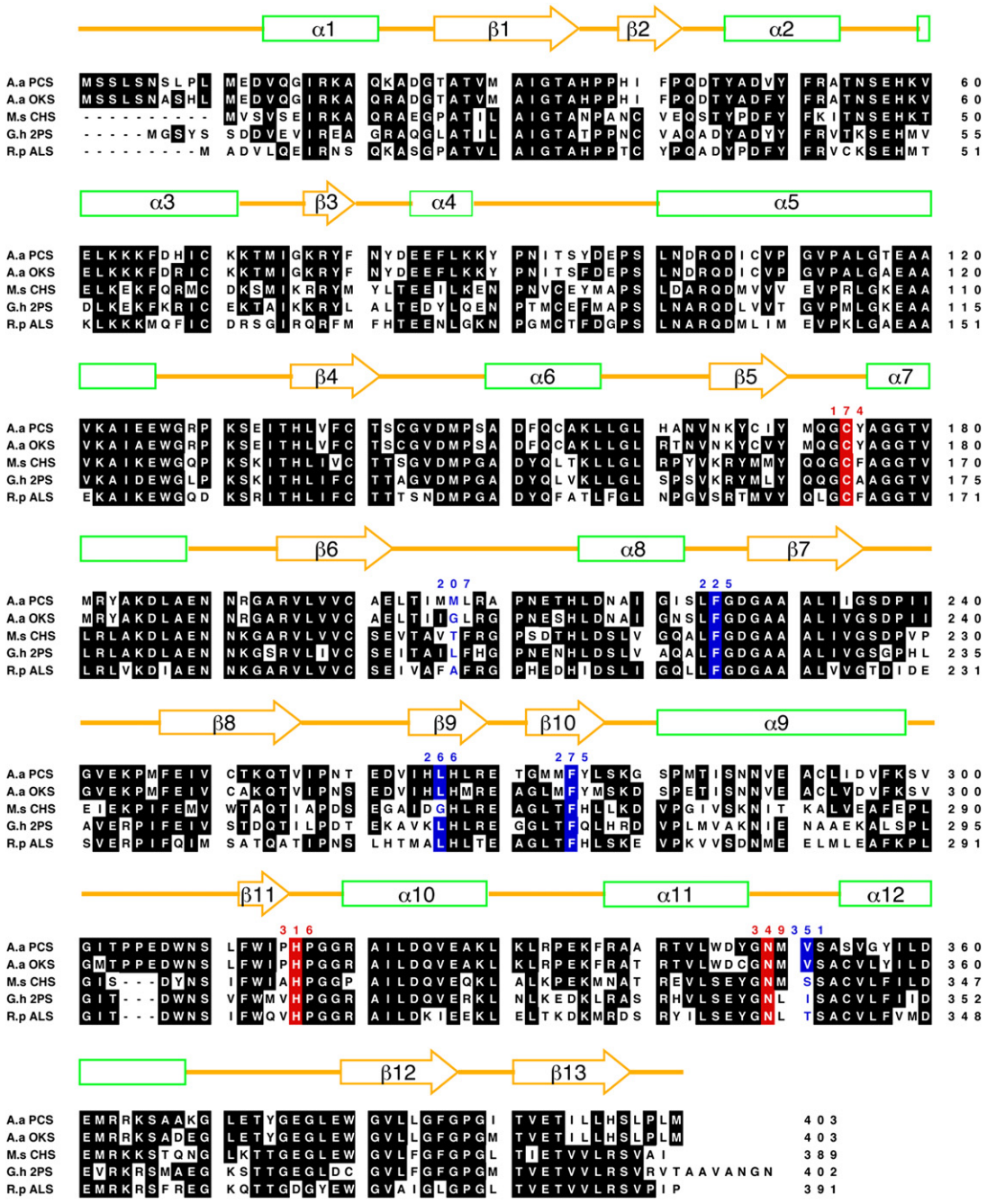

Figure 2. Comparison of the Primary Sequences of PCS and Other Plant Type III PKSs

A.a PCS, *A. arborescens* PCS; A.a OKS, *A. arborescens* OKS; M.s CHS, *M. sativa* CHS; G.h 2PS, *G. hybrida* 2PS; R.p ALS, *R. palmatum* ALS. The secondary structures of PCS are also delineated: α helices (green rectangles), β strands (orange arrows), and loops (orange, bold lines) are diagrammed. The catalytic triad of Cys-His-Asn is colored red. The residues thought to be crucial for the steric modulation of the active site in a number of divergent type III PKSs are highlighted in blue.

16 Å-long CoA-binding tunnel, as in the case of other type III PKSs [3–7]. Unexpectedly, the crystal contained a CoA-SH molecule in each monomer, which may have been co-purified with the PCS protein from *Escherichia coli*. The CoA-SH is bound to PCS in a similar manner as in the case of *M. sativa* CHS and *G. hybrida* 2PS [4] (Figure 3C).

Furthermore, the electron density map indicated the oxidation of the catalytic Cys174 into sulfic acid (-SO_2H) during crystallization, which was also observed in the structures of *M. sativa* CHS [3] and *G. hybrida* 2PS [4].

A comparison of the structures of *A. arborescens* PCS, *M. sativa* CHS, and *G. hybrida* 2PS reveals a significant

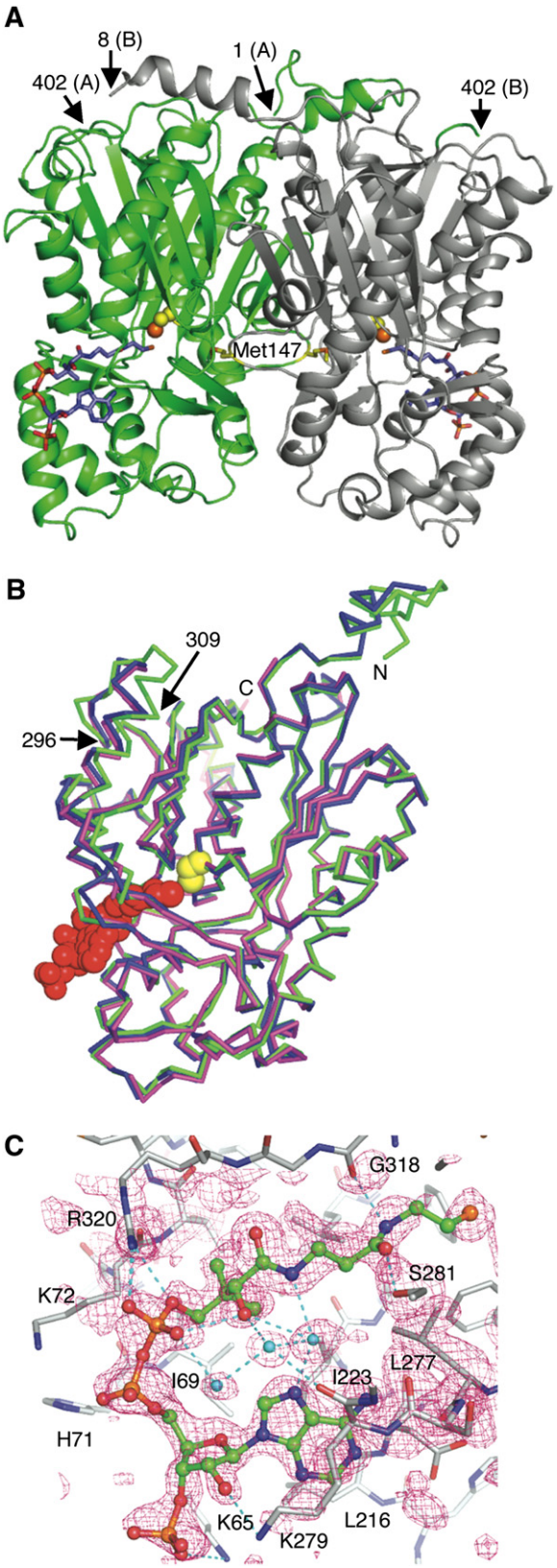

Figure 3. Overall Structure of PCS Complexed with CoA-SH

(A) Ribbon representation of the PCS homodimer. The monomers are colored green and silver, and the CoA-SH molecules are shown as blue stick models. The catalytic Cys174 and Met147, which form a partial wall of the active-site cavity of another monomer, are highlighted as yellow CPK and stick models, respectively.

(B) Comparison of PCS (green), *M. sativa* CHS (blue), and *G. hybrida* 2PS (purple). The catalytic Cys174 and the bound CoA-SH in PCS are also shown as yellow and red CPK molecules, respectively.

(C) CoA-SH binding to the PCS structure. The CoA-SH (green) and the SIGMA-weighted $|2F_o - F_c|$ electron density (0.8σ, red cage) for CoA-SH are shown. The water molecules (light-blue spheres) and hydrogen bonds (dotted lines) are also indicated.

backbone change in residues 296–309 of PCS (corresponding to residues 286–296 of *M. sativa* CHS and 291–301 of *G. hybrida* 2PS) (Figure 3B). This is due to the characteristic insertion of 3 residues, 304–306, and the addition of 10 residues at the N-terminus in PCS (Figure 2). These residues not only allow the formation of the expanded surface-exposed loop (residues 296–309), but also contribute to the larger dimerization surface area, compared to those of *M. sativa* CHS and *G. hybrida* 2PS. However, these conformational differences in the regions far from the traditional active site are not thought to be crucial for catalytic activity.

Active-Site Architecture of PCS

The pentaketide-forming *A. arborescens* PCS and the chalcone-forming *M. sativa* CHS share identical catalytic residues and the highly conserved CoA binding site, but they produce chemically distinct polyketides. A structural comparison revealed that most of the active site residues of *A. arborescens* PCS and *M. sativa* CHS are superimposable in nearly identical positions. The only exceptions are Cys143, Thr204, Met207, Leu266, and Val351 of *A. arborescens* PCS (corresponding to Ser133, Thr194, Thr197, Gly256, and Ser338, respectively, of *M. sativa* CHS), which may account for the different catalytic activities of the two enzymes (Figures 4A and 4B). The latter 3 residues, Met207, Leu266, and Val351, are altered in a number of functionally divergent type III PKSs and are thought to be critical for the substrate and product specificities of the enzymes [4, 13, 14]. Indeed, we have previously demonstrated that residue 207 plays a decisive role in controlling the product chain length of the enzyme reaction: the pentaketide-producing PCS was functionally transformed into an OKS by the single amino acid substitution M207G [13]. Further, small-to-large substitutions (G207A, G207T, G207M, G207L, G207F, and G207W) in OKS from *A. arborescens* resulted in loss of octaketide-forming activity and the concomitant formation of shorter chain length polyketides depending on the size of the side chain [14]. Interestingly, the PCS crystal displays a broad electron density in the position of Met207, suggesting that the residue is fairly flexible and adopts at least two distinct conformations, as shown in Figures 4A and 4B. In contrast, the structural analyses of *M. sativa* CHS and *G. hybrida* 2PS indicated that Thr197 of *M. sativa* CHS and Leu202 of *G. hybrida* 2PS, corresponding to Met207 of PCS, are not so flexible [3, 4].

Remarkably, both the locations and angles of Thr197, Gly256, and Ser338 of *M. sativa* CHS are well conserved

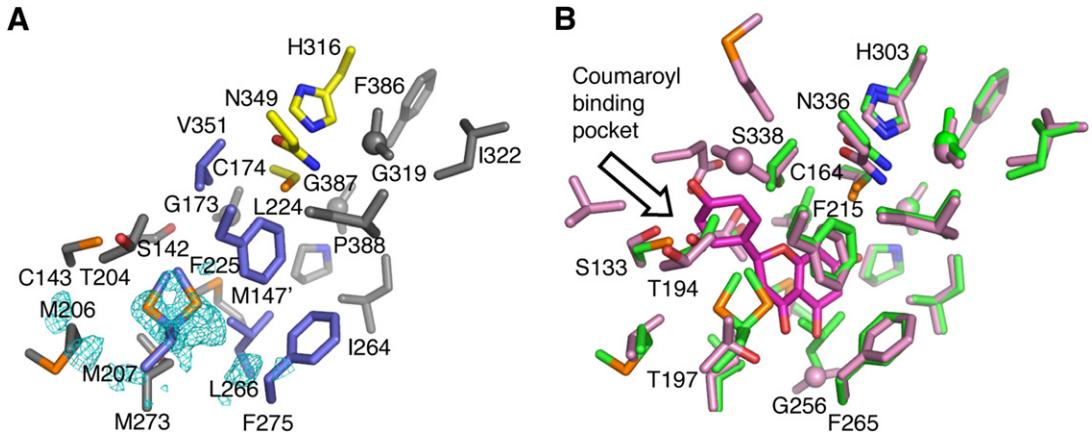

Figure 4. The Active-Site Geometry of PCS and *M. sativa* CHS

(A) The PCS active-site residues with an omit map of the SIGMA-weighted $|F_o - F_c|$ electron density (2.0σ, light-blue cage) for Met207. Both conformers of Met207 are shown. The Cys-His-Asn catalytic triad is represented in yellow. The 3 residues that appeared to be most important for distinguishing PCS activity from those of other type III PKSs are highlighted in blue. Two phenylalanine residues that appear to function as "gatekeepers," in the case of *M. sativa* CHS, are also colored blue.

(B) Superposition of the PCS (green) and *M. sativa* CHS active-site residues (pink). The bound naringenin molecule in the *M. sativa* CHS structure (PDB entry 1CGK) is also indicated as a purple stick model.

in Met207, Leu266, and Val351, respectively, of *A. arborescens* PCS. The only difference is the steric bulk of each residue, which causes a substantial steric contraction of the active-site cavity of PCS. The similar case of the steric contraction by the 3 residues has also been observed in the crystal structure of *G. hybrida* 2PS (T197L/G256L/S338I, numbering according to *M. sativa* CHS). As a result, the total cavity volume of the pentaketide ($C_{10}H_8O_4$)-forming PCS is about 247 Å^3, which is much smaller than that of the chalcone ($C_{15}H_{12}O_5$)-forming *M. sativa* CHS (754 Å^3) and is almost as large as that of the triketide ($C_6H_6O_3$)-forming *G. hybrida* 2PS (253 Å^3). Interestingly, the small-to-large Gly256L substitution (numbering according to *M. sativa* CHS) is also observed in *G. hybrida* 2PS, *A. arborescens* OKS, and *Rheum palmatum* ALS, which each select acetyl- or malonyl-CoA as a starter substrate [4, 14, 17]. Unexpectedly, further conformational differences are observed for the other 2 residues, Thr204 and Cys143. The backbone torsion angles of Thr204 (−62, 1) and Cys143 (−117, 94), compared with Thr194 (−77, 6) and Ser133 (−119, 104), are shifted by ϕ angles of −15° and by ψ angles of +10° toward Val351 in PCS and toward Ser338 in *M. sativa* CHS, respectively. This causes the loss of the CHS's coumaroyl-binding pocket [3] from the PCS active-site cavity. As a result, PCS no longer accepts coumaroyl-CoA as a starter substrate for the chalcone-forming reaction.

Finally, it should be noted that the PCS crystal structure revealed neither a hydrogen-bond network, as in the case of *Pinus sylvestris* stilbene synthase (STS) [5] (as discussed later), nor any additional catalytic Cys residues that appeared to be crucial for the enzyme reaction. In contrast, it was recently reported that *S. coelicolor* 1,3,6,8-tetrahydroxynaphthalene synthase (THNS), the bacterial type III PKS catalyzing a malonyl-primed pentaketide-forming reaction, features 3 possible catalytic Cys residues in addition to the conventional Cys164 (numbering according to *M. sativa* CHS) [7].

Structure of the PCS M207G Mutant with Additional Novel, Buried Pockets

Besides the wild-type enzyme, the crystal structure of the octaketide-producing PCS M207G mutant was also determined to 1.6 Å resolution. Unlike the wild-type enzyme, the crystal of the point mutant was obtained without a CoA-SH molecule. The electron density map clearly shows the substitution of Met207 with Gly and the oxidation of Cys174 to sulfic acid. The overall molecular structure was nearly identical to that of the wild-type enzyme, except for residues 211–220, the region thought to be involved in CoA-SH binding (Figures 3C and 5A). To determine whether this change is due to either the point mutation or the absence of CoA-SH binding, the M207G mutant was once again cocrystallized with CoA-SH. The newly obtained crystal, including CoA-SH, is an almost perfect analog of that of wild-type PCS and does not indicate any backbone changes for residues 211–220, confirming that the region is perturbed by CoA-SH binding. The rmsd between wild-type and the M207G mutant complexed with CoA-SH is only 0.2 Å in the Cα atoms. Thus, further structural analyses of the M207G mutant were carried out using the structure complexed with the CoA-SH molecule.

A comparison of the wild-type and M207G mutant structures revealed that the conformations of the residues lining the active-site cavity are perfectly conserved in the mutant, and that the active site of the mutant contains two additional novel, buried pockets (pockets A and B) that extend into the traditional active-site cavity (Figures 5B–5E). Thus, Met207 blocks the entrance of pockets A and B in wild-type PCS, but the replacement of M207

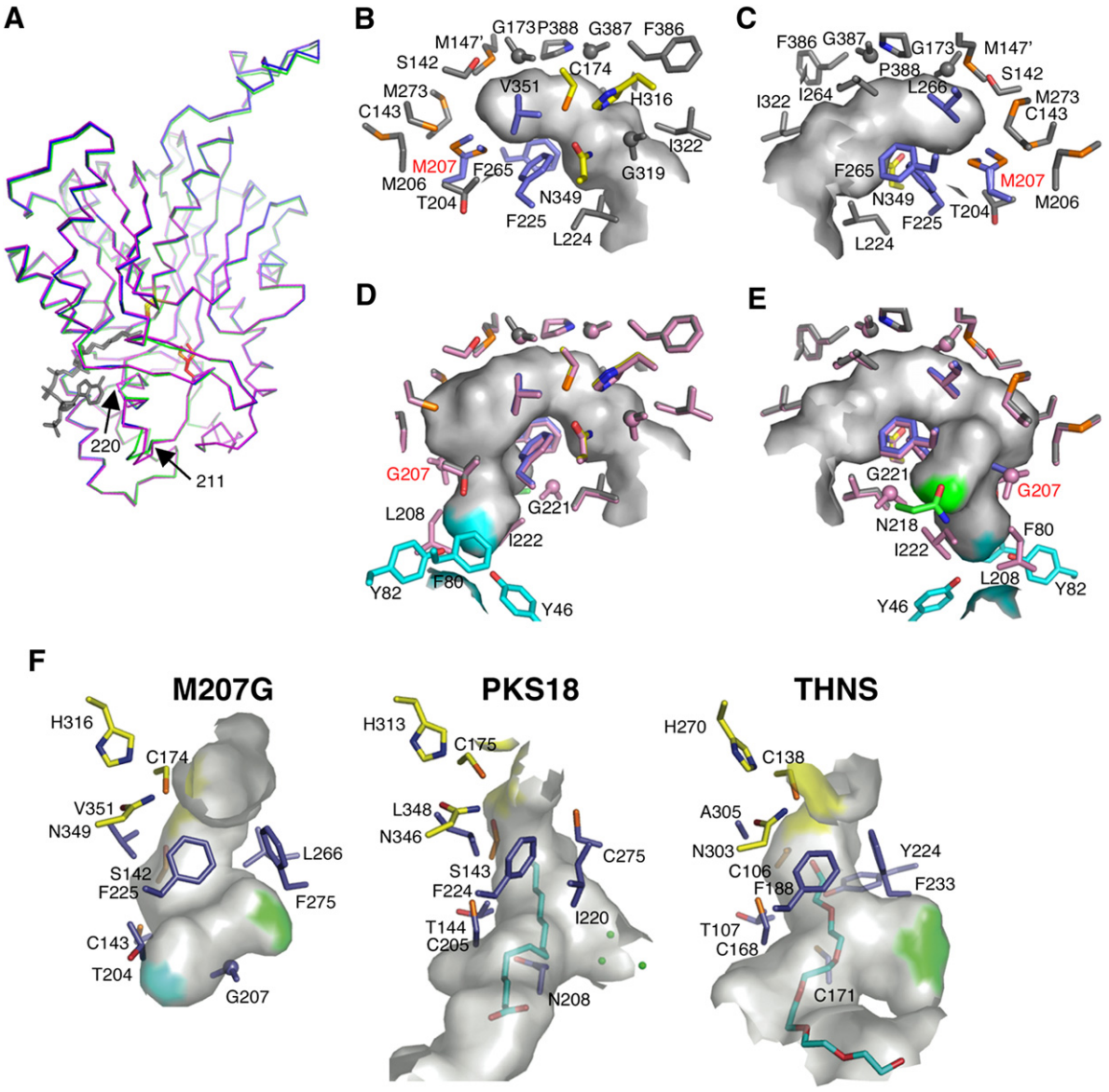

Figure 5. Crystal Structure of the PCS M207G Mutant

(A) Superposition of the overall structures of wild-type PCS (green), the M207G mutant (purple), and the M207G mutant complexed with CoA-SH (blue). The catalytic Cys174 (yellow), the two conformers of Met207 (red), and the CoA-SH molecules (gray) are also indicated as stick models.

(B and C) (B) Front and (C) rear views of the active-site cavity surface of wild-type PCS. The residue number of Met207 is shown in red. The PCS active-site residues are colored as in Figure 4A.

(D and E) (D) Front and (E) rear views of the active-site cavity surface of the PCS M207G mutant. The residue number of Gly207 (red) and the bottoms of pockets A (light blue) and B (green) are highlighted. The PCS active-site residues are also superimposed in (D) and (E). The only residue numbers that are found in the active-site cavity of the PCS M207G mutant, compared with wild-type PCS, are shown in (D) and (E). The M207G mutant clearly shows the expanded cavity, consisting of the PCS active-site cavity and the two novel, buried pockets, compared to that of wild-type PCS.

(F) Comparison of the active-site cavities of the M207G mutant, *M. tuberculosis* PKS18, and *S. coelicolor* THNS. The only residues (blue and orange) that appear to determine the respective activities of these enzymes are shown with the catalytic triad (yellow). Cys106 of *S. coelicolor* THNS is the reactive residue (orange) as well as a member of the catalytic triad in the case of *S. coelicolor* THNS. The bottoms of pockets A and B of the M207G mutant are highlighted as light-blue and green surfaces, respectively. A myristic acid and a polyethylene glycol heptamer, bound to the respective active-site cavities of *M. tuberculosis* PKS18 and *S. coelicolor* THNS, are shown as light-blue stick models, to indicate the acyl-binding tunnel of their active-site cavities. The regions of the bacterial type III PKSs corresponding to pocket B are colored green.

with the less bulky Gly opens the gate to the buried pockets, thereby expanding the putative polyketide chain-elongation tunnel. Amazingly, the single amino acid substitution dramatically expands the active-site cavity; the total cavity volume of the M207G mutant is 649 Å^3, which is 2.6 times larger than that of wild-type

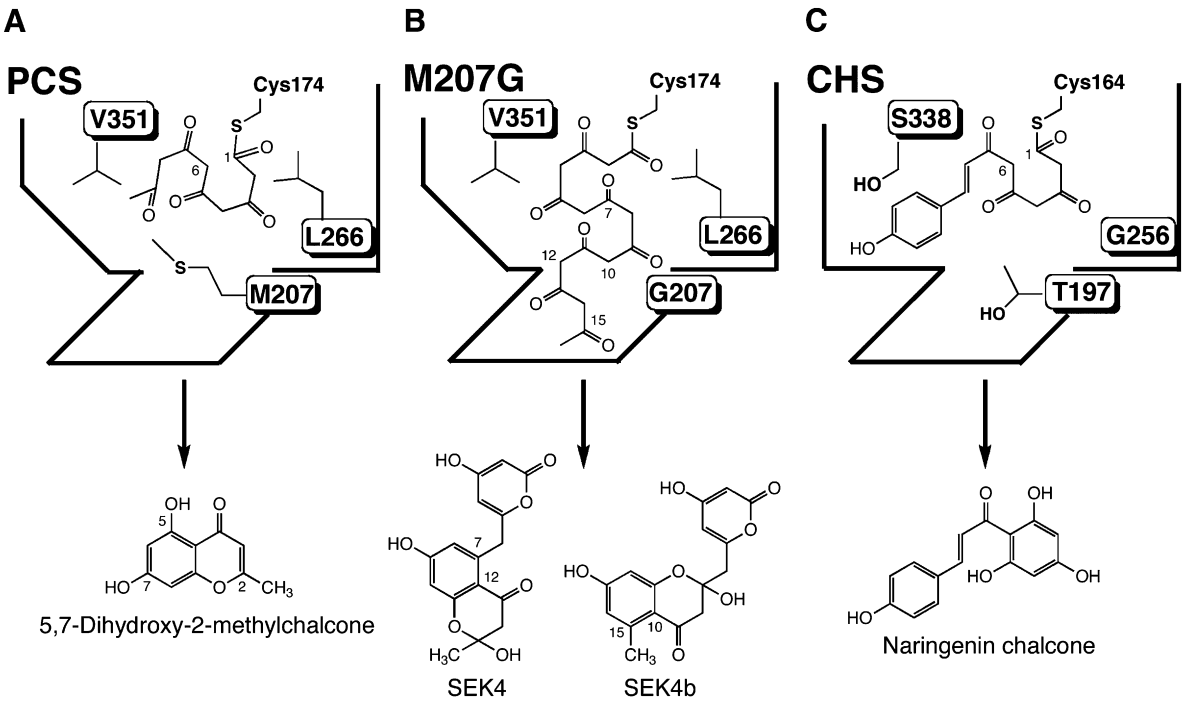

Figure 6. Schematic Representation of the Active-Site Architecture of Wild-Type PCS, the M207G Mutant, and *M. sativa* CHS

(A–C) The M207G substitution opens a gate to the buried pocket A that extends into the "floor" of the active-site cavity, resulting in a 4:1 mixture of SEK4b:SEK4 instead of 5,7-dihydroxy- 2-methylchormone. PCS locks the methyl end of its linear pentaketide intermediate between Met207 and Val351, as in the case in which *M. sativa* CHS locks the aromatic ring derived from 4-coumaroyl-CoA with the coumaroyl-binding pocket.

PCS (247 Å^3) and almost as large as that of *M. sativa* CHS (754 Å^3). The active-site cavity of the point mutant is now large enough to perform the iterative condensations of eight molecules of malonyl-CoA and to accommodate the octaketide product. These observations clearly support the proposal that the octaketide-producing activity of the PCS M207G mutant is totally dependent upon the presence of the additional buried pockets, and that the growing polyketide intermediates extend into at least one of the two pockets.

To test which pocket is crucial for the polyketide chain-elongation reaction, we constructed a series of PCS M207G/N218X double mutants, in which Asn218, at the bottom of pocket B (Figure 5E), was simultaneously substituted with Asp, Glu, Gln, Lys, and Ala, and tested the enzyme activity. As a result, all of the double mutants turned out to be functionally almost identical to the M207G single mutant (data not shown), suggesting that pocket B is not crucial for the octaketide-producing activity of the PCS M207G mutant. Furthermore, as discussed later, the location and orientation of pocket A overlap those of acyl-binding tunnels of the bacterial type III PKSs, *Mycobacterium tuberculosis* PKS18 [6], and *S. coelicolor* THNS [7] (Figure 5F).

DISCUSSION

The structural analyses of the malonyl-primed pentaketide-forming wild-type PCS and the octaketide-forming PCS M207G mutant clearly demonstrated that residue 207 (corresponding to Thr197 in *M. sativa* CHS) lining the active-site cavity indeed occupies a crucial position for the polyketide chain-elongation reactions. Met207 of wild-type PCS blocks the entrances of novel, buried pockets that extend into the "floor" of the active-site cavity. It is quite remarkable that the single amino acid replacement causes such a dramatic increase in the active-site cavity volume. The large-to-small Met207G substitution opens the gate to the pockets, thereby substantially expanding the putative polyketide chain-elongation tunnel, which leads to the formation of the longer octaketides SEK4 and SEK4b instead of the pentaketide 5,7-dihydroxy-2-methylchromone (Figures 6A and 6B). Interestingly, the malonyl-primed, octaketide-forming *A. arborescens* OKS and the heptaketide-forming *R. palmatum* ALS also feature the similar large-to-small Thr197A and Thr197G substitutions, respectively (numbering according to *M. sativa* CHS) (Figure 2). As we have previously demonstrated, in OKS and ALS, the chemically inert residue also determines the polyketide chain length and the product specificity [14, 17, 19]. For example, small-to-large substitutions (G207A, G207T, G207M, G207L, G207F, and G207W) in OKS resulted in the loss of octaketide-forming activity and the concomitant formation of shorter chain length polyketides, depending on the steric bulk of the side chain. These long-chain polyketide-producing plant type III PKSs seem to share a common active-site architecture by

combining the large-to-small Thr197(G/A) and the small-to-large Gly256L substitutions (numbering according to *M. sativa* CHS), as in the case of the PCS M207G mutant.

Recent structural analyses of *M. tuberculosis* PKS18 [6] and *S. coelicolor* THNS [7] have revealed that these bacterial type III PKSs incorporate an acyl-binding tunnel extending into the conventional active-site cavity floor (Figure 5F), and that the acyl-binding tunnel is used to produce the long-chain triketide and tetraketide pyrones (C_{12}-C_{22}) from a fatty acyl starter unit. We propose that these long-chain polyketides may be regarded as analogs of the octaketide (C_{16}) intermediate of the *A. arborescens* PCS M207G mutant. Interestingly, the location and orientation of pocket A of the PCS M207G mutant overlap with those of the acyl-binding tunnel of the bacterial type III PKSs. Since *A. arborescens* OKS readily accepts a long-chain (C_{12}-C_{20}) fatty acyl starter to produce triketide and tetraketide pyrones [14], it is possible that the plant and bacterial type III PKSs share a similar active-site architectural strategy to produce the long-chain polyketides, including pyrones, phloroglucinols, and resorcinols [20, 21]. With the results of the functional analyses of pocket B in PCS, these observations support an idea that the PCS M207G mutant uses the novel, buried pocket A to accommodate the octaketide intermediate, as in the case of the bacterial type III PKSs. Indeed, we have very recently succeeded in generating a novel polyketide-producing PCS mutant enzyme by introducing mutations to the residues lining pocket A. This mutant enzyme not only increased polyketide chain length, compared with the wild-type and PCS M207G mutant enzymes, but it also altered the mechanism of the cyclization reaction, thereby resulting in the formation of a novel, unnatural polyketide, which will be discussed elsewhere. On the basis of the crystal structure of the naphthalene-producing *S. coelicolor* THNS, the possible involvement of additional catalytic Cys residues has been proposed for the formation of the aromatic pentaketide in the bacterial type III PKS [7]. However, in both wild-type PCS and the PCS M207G mutant, no catalytic Cys residues were observed in addition to the conventional Cys that serves as the covalent attachment site for the growing polyketide intermediates.

The M207G point mutation also substantially alters the mechanism of the cyclization reaction: wild-type PCS catalyzes the C-1/C-6 Claisen-type cyclization of a pentaketide intermediate to produce 5,7-dihydroxy-2-methylchromone (Figure 1C), whereas the M207G mutant performs the C-10/C-15 and C-12/C-7 aldol-type cyclization of an octaketide intermediate to form the 4:1 mixture of SEK4b:SEK4 (Figures 1D and 1E). Recently, structural analyses of *P. sylvestris* STS led to the proposal that the so-called ''aldol-switch'' hydrogen-bond network involving Ser338-H_2O-Thr132-Glu192 (numbering according to *M. sativa* CHS) plays a critical role in the determination of product specificity [5]. Thus, Noel and Schröder proposed that electronic effects, rather than steric factors, balance the competing cyclization specificities in CHS (Claisen-type) and STS (aldol-type) from a common tetraketide intermediate [5]. However, in *A. arborescens* PCS, the CHS's active-site Ser338 is substituted with the hydrophobic Val, and such a hydrogen-bond network is not observed in either wild-type or the M207G mutant. It should be also noted that the CHS active site can be extended to facilitate the synthesis of larger polyketides, including the octaketides SEK4 and SEK4b, by the S338V single amino acid mutation [18]. These observations suggest that residue 338, located in the proximity of the catalytic Cys164 (numbering according to *M. sativa* CHS), which is the covalent attachment site for the growing polyketide intermediates, provides steric guidance so that the linear intermediate extends into the additional buried pocket, thereby leading to the formation of the longer polyketides [18]. Again, the formation of the 4:1 mixture of SEK4b:SEK4, not 100% conversion to a single product, has been shown in the PCS M207G mutant. In addition, the PCS M207G mutant still retains ~80% of the catalytic efficiency compared with wild-type PCS. The formation of the 4:1 SEK4b:SEK4 mixture might be controlled by steric modulation rather than the hydrogen network. It is remarkable that the functional conversions are simply based on the steric modulation of a chemically inert single residue lining the active-site cavity.

Interestingly, both PCS and CHS catalyze the C-6/C-1 Claisen-type cyclization of the enzyme-bound polyketide intermediate (Figures 1A and 1C). Although the PCS active-site cavity lacks CHS's coumaroyl-binding pocket, which locks the aromatic moiety of the coumaroyl starter (Figures 4A and 4B), PCS and CHS share almost identical active-site architectures, with the exceptions of Cys143, Thr204, Met207, Leu266, and Val351 in PCS. Considering the functional and structural similarities between the two enzymes, it is possible that the acetyl moiety of the pentaketide intermediate is locked by the conformationally flexible Met207, and that the Claisen-type cyclization/aromatization reaction of the enzyme-bound intermediate takes place with common structural machinery, as in the case of CHS (Figures 6A and 6C). This reaction would be followed by the probable spontaneous formation of the fused ring system, leading to the production of 5,7-dihydroxy-2-methylchromone.

On the other hand, one of the interesting points for the formation of SEK4b and SEK4 by the PCS M207G mutant is the timing of the thioester-bond cleavage of the enzyme-bound polyketide intermediate and the cyclization/aromatization reaction. Although the detailed structural mechanism remains uncertain, it is likely that the mutant enzyme catalyzes polyketide chain elongation, and the aromatic ring formation through the C-10/C-15 and C-7/C-12 aldol-type cyclizations of the octaketide intermediate. The partially cyclized intermediate may be released from the active site, and then undergo a spontaneous, nonenzymatic cyclization reaction, leading to the formation of SEK4b and SEK4, as in the case of the bacterial minimal type II PKS.

Table 1. Data Collection and Refinement Statistics

	PCS CoA-SH	M207G	M207G CoA-SH
Data Collection			
Space group	$P2_1$	$P2_1$	$P2_1$
Unit cell			
a, b, c (Å)	73.2, 88.4, 70.0	74.3, 89.1, 70.7	74.1, 89.2, 70.8
α, β, γ (°)	90.0, 95.6, 90.0	90.0, 95.5, 90.0	90.0, 95.7, 90.0
Resolution (Å)	30.0–1.6 (1.66–1.60)	30.0–1.6 (1.66–1.60)	30.0–1.6 (1.66–1.60)
Unique reflections	116,521	117,664	120,438
Redundancy	3.7 (3.8)	3.5 (3.3)	3.8 (3.6)
Completeness (%)	99.9 (100)	97.7 (95.3)	99.8 (98.1)
$<I/(\sigma I)>$	33.5 (8.4)	33.2 (5.0)	27.4 (3.9)
R_{sym} (%)[a]	6.9 (23.0)	4.6 (24.1)	5.7 (28.7)
Refinement			
Resolution (Å)	1.6	1.6	1.6
R_{cryst}/R_{free} (%)[b]	19.8/20.7	19.2/20.9	19.3/20.8
Number of atoms			
Protein	6,204	6,193	6,193
Water	566	592	545
Ligand	96	-	96
B factors (Å^2)			
Protein	12.7	16.8	14.0
Water	21.6	25.0	20.8
Ligand	34.6	-	23.0
Rms deviations			
Bond lengths (Å)	0.011	0.011	0.011
Bond angles (°)	1.5	1.5	1.5

Values in parentheses are for the highest-resolution shell.
[a] $R_{sym} = \Sigma_h \Sigma_i |I(h)_i - < I(h) >| / \Sigma_h \Sigma_i I(h)_i$, where $I(h)$ is the intensity of refraction h, Σ_h is the sum over all reflections, and Σ_i is the sum over i measurements of reflection h.
[b] R_{free} was calculated with 5% of data excluded from refinement.

SIGNIFICANCE

In the sequential condensation reactions of type III PKSs, polyketide chain-length control is one of major determining factors of the molecular diversity of polyketides. The crystal structures revealed that the pentaketide-producing wild-type and the octaketide-producing M207G mutant shared almost the same overall folding, and that the large-to-small substitution dramatically increases the volume of the polyketide-elongation tunnel by opening a gate to two novel, hidden pockets behind the active site of the enzyme. Residue 207, a chemically inert active-site residue, thus controls the number of condensations of malonyl-CoA solely based on the steric bulk of the side chain. This is interesting not only from the point of view of molecular evolution of the enzymes, but also for engineering the catalytic potential of the functionally diverse type III PKSs.

EXPERIMENTAL PROCEDURES

Structure Determination

Expression, purification, and crystallization of the recombinant PCS were carried out as previously reported [22]. Diffraction data were collected at 100 K at BL24XU of SPring-8 (wavelength, 0.82656 Å), by using a Rigaku R-AXIS V imaging plate. The crystals belonged to space group $P2_1$, with unit cell parameters of a = 73.2 Å, b = 88.4 Å, c = 70.0 Å, $\alpha = \gamma = 90.0°$, and $\beta = 95.6°$, and contained two molecules in the asymmetric unit. Data were indexed, integrated, and scaled with HKL2000 [23]. Crystal data and intensity statistics are summarized in Table 1.

The crystal structure was determined by the molecular replacement method, and the crystal structure of *G. hybrida* 2PS (PDB entry 1EE0) [4] was used as a search model. The molecular replacement was carried out with EPMR [24]. Crystallographic refinement and model building were performed by using CNS [25] and XTALVIEW [26],

respectively. Since a broad range of electron density corresponding to methionine at position 207, which has been shown to determine the polyketide chain length in PCS, was observed in the preliminary structure refinement, the initial refinements and model building were done by using a model with the methionine replaced by alanine. After several rounds of refinement and model building, the methionine was replaced, and further refinements and model building were carried out. The electron density map also indicated the presence of a small molecule in the traditional CoA-binding tunnel. The small molecule was determined to be CoA-SH, from a comparison with the *M. sativa* CHS structure complexed with CoA-SH (PDB entry 1BQ6B) [3]. The determination was also confirmed by the M207G mutant structure complexed with CoA-SH, as described later. Refinement statistics are listed in Table 1. The quality of the model was checked with PROCHEK [27]. The cavity volume was calculated on the basis of the residues composing the active-site cavity wall by using CASTP (http://cast.engr.uic.edu/cast/).

Expression and Purification of the M207G Mutant

The M207G mutant expression plasmid was constructed with the QuickChange Site-Directed Mutagenesis Kit (Stratagene), according to the manufacturer's protocol, by using 5′-GAGCTCACCATAATC ATGCTTCGAGGCCCT-3′ as a sense primer, 5′-AGGGCCTCGAAGC ATGATTATGGTGAGCTC-3′ as an antisense primer, and the PCS expression plasmid as the template. The mutant protein was expressed and purified according to the same procedure as in the case of wild-type PCS, and was concentrated to 10 mg/ml in 20 mM HEPES-NaOH (pH 7.0) buffer, containing 100 mM NaCl and 2 mM DTT.

Crystallization and Structure Refinement of the M207G Mutant

The M207G mutant crystals were obtained at 20°C, in 100 mM Tris-HCl (pH 8.5) buffer containing 12% (w/v) PEG 8,000, 320 mM trisodium citrate, and 3% (v/v) MPD with 5 mg/ml of the protein solution, by using the sitting-drop vapor-diffusion method. On the other hand, the M207G mutant crystals, complexed with CoA-SH, were obtained by the same crystallization methods as those used for wild-type PCS, except for the cocrystallization with 2 mM CoA-SH. Both crystals were independently transferred into the reservoir solution with 10% (v/v) glycerol as a cryoprotectant, and they were then flash cooled at 100 K in a nitrogen-gas stream. Both X-ray diffraction data sets were collected at NW12 of Photon Factory-AR (wavelength, 1.00000 Å) by using an ADSC Quantum 210 CCD detector. Crystal data and intensity statistics are summarized in Table 1. Both crystals belonged to space group $P2_1$, and the unit cell dimensions were very similar to those of wild-type PCS. The position and orientation of each molecule were readily obtained by using the molecule found in the wild-type PCS structure with CNS [25]. The final models of the M207G mutant structure, which consists of the same number of residues as wild-type, were obtained by additional cycles of refinement. Refinement statistics are given in Table 1.

Mutagenesis Studies and Enzyme Assay of PCS M207G/N218X

All of the PCS M207G/N218X expression plasmids were directly constructed from the M207G mutant expression plasmid by using the QuickChange Site-Directed Mutagenesis Kit according to the manufacturer's protocol. The mutant proteins were purified by affinity chromatography on Glutathione Sepharose 4B (Amersham Biosciences), as previously reported [22], and were then used in the enzyme assay. The enzyme assay and the product identification were carried out as previously reported [13], except for an elution program with a linear gradient of CH_3CN from 15% to 30% in H_2O containing 0.1% TFA for 40 min to separate the reaction products.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the National Project on Protein Structural and Functional Analyses. The authors declare that they have no competing financial interests.

Received: December 10, 2006
Revised: January 24, 2007
Accepted: February 6, 2007
Published: April 27, 2007

REFERENCES

1. Schröder, J. (1999). The chalcone/stilbene synthase-type family of condencing enzymes. In Comprehensive Natural Reports, D. Barton, K. Nakanishi, O. Meth-Cohn, and U. Sankawa, eds. (Oxford, UK: Elsevier), pp. 749–772.
2. Austin, M.B., and Noel, J.P. (2003). The chalcone synthase superfamily of type III polyketide synthases. Nat. Prod. Rep. *20*, 79–110.
3. Ferrer, J.L., Jez, J.M., Bowman, M.E., Dixon, R.A., and Noel, J.P. (1999). Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat. Struct. Biol. *6*, 775–784.
4. Jez, J.M., Austin, M.B., Ferrer, J., Bowman, M.E., Schröder, J., and Noel, J.P. (2000). Structural control of polyketide formation in plant-specific polyketide synthases. Chem. Biol. *7*, 919–930.
5. Austin, M.B., Bowman, M.E., Ferrer, J.L., Schroder, J., and Noel, J.P. (2004). An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. Chem. Biol. *11*, 1179–1194.
6. Sankaranarayanan, R., Saxena, P., Marathe, U.B., Gokhale, R.S., Shanmugam, V.M., and Rukmini, R. (2004). A novel tunnel in mycobacterial type III polyketide synthase reveals the structural basis for generating diverse metabolites. Nat. Struct. Mol. Biol. *11*, 894–900.
7. Austin, M.B., Izumikawa, M., Bowman, M.E., Udwary, D.W., Ferrer, J.L., Moore, B.S., and Noel, J.P. (2004). Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates. J. Biol. Chem. *279*, 45162–45174.
8. Jez, J.M., Ferrer, J.L., Bowman, M.E., Dixon, R.A., and Noel, J.P. (2000). Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. Biochemistry *39*, 890–902.
9. Lanz, T., Tropf, S., Marner, F.J., Schroder, J., and Schroder, G. (1991). The role of cysteines in polyketide synthases. Site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways. J. Biol. Chem. *266*, 9971–9976.
10. Jez, J.M., and Noel, J.P. (2000). Mechanism of chalcone synthase. pKa of the catalytic cysteine and the role of the conserved histidine in a plant polyketide synthase. J. Biol. Chem. *275*, 39640–39646.
11. Austin, M.B., Saito, T., Bowman, M.E., Haydock, S., Kato, A., Moore, B.S., Kay, R.R., and Noel, J.P. (2006). Biosynthesis of *Dictyostelium discoideum* differentiation-inducing factor by a hybrid type I fatty acid-type III polyketide synthase. Nat. Chem. Biol. *2*, 494–502.
12. Eckermann, S., Schröder, G., Schidt, J., Strack, D., Edrada, R.A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpeläinen, I., Proksch, P., et al. (1998). New pathway to polyketides in plants. Nature *396*, 387–390.
13. Abe, I., Utsumi, Y., Oguro, S., Morita, H., Sano, Y., and Noguchi, H. (2005). A plant type III polyketide synthase that produces pentaketide chromone. J. Am. Chem. Soc. *127*, 1362–1363.
14. Abe, I., Oguro, S., Utsumi, Y., Sano, Y., and Noguchi, H. (2005). Engineered biosynthesis of plant polyketides: chain length control in an octaketide-producing plant type III polyketide synthase. J. Am. Chem. Soc. *127*, 12709–12716.
15. Dewick, P.M. (2002). Medicinal Natural Products, A Biosynthetic Approach, Second Edition (West Sussex, UK: Wiley).
16. Fu, H., Hopwood, D.A., and Khosla, C. (1994). Engineered biosynthesis of novel polyketides: evidence for temporal, but not

regiospecific, control of cyclization of an aromatic polyketide precursor. Chem. Biol. *1*, 205–210.

17. Abe, I., Watanabe, T., Lou, W., and Noguchi, H. (2006). Active site residues governing substrate selectivity and polyketide chain length in aloesone synthase. FEBS J. *273*, 208–218.
18. Abe, I., Watanabe, T., Morita, H., Kohno, T., and Noguchi, H. (2006). Engineered biosynthesis of plant polyketides: manipulation of chalcone synthase. Org. Lett. *8*, 499–502.
19. Abe, I., Utsumi, Y., Oguro, S., and Noguchi, H. (2004). The first plant type III polyketide synthase that catalyzes formation of aromatic heptaketide. FEBS Lett. *562*, 171–176.
20. Abe, I., Watanabe, T., and Noguchi, H. (2004). Enzymatic formation of long-chain polyketide pyrones by plant type III polyketide synthases. Phytochemistry *65*, 2447–2453.
21. Funa, N., Ozawa, H., Hirata, A., and Horinouchi, S. (2006). Phenolic lipid synthesis by type III polyketide synthases is essential for cyst formation in *Azotobacter vinelandii*. Proc. Natl. Acad. Sci. USA *103*, 6356–6361.
22. Morita, H., Kondo, S., Abe, T., Noguchi, H., Sugio, S., Abe, I., and Kohno, T. (2006). Crystallization and preliminary crystallographic analysis of a novel plant type III polyketide synthase that produces pentaketide chromone. Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun. *62*, 899–901.
23. Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.
24. Kissinger, C.R., Gehlhaar, D.K., and Fogel, D.B. (1999). Rapid automated molecular replacement by evolutionary search. Acta Crystallogr. D Biol. Crystallogr. *55*, 484–491.
25. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.
26. McRee, D.E. (1992). A visual protein crystallographic software system for X11/Xview. J. Mol. Graph. *10*, 44–46.
27. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. *26*, 283–291.

Accession Numbers

Coordinates have been deposited in the Protein Data Bank with accession codes 2D3M for PCS complexed with CoA-SH, 2D51 for the M207G mutant, and 2D52 for the M207G mutant complexed with CoA-SH.