J|A|C|S
ARTICLES
Published on Web 08/17/2005

Engineered Biosynthesis of Plant Polyketides: Chain Length Control in an Octaketide-Producing Plant Type III Polyketide Synthase

Ikuro Abe,* Satoshi Oguro, Yoriko Utsumi, Yukie Sano, and Hiroshi Noguchi

Contribution from the School of Pharmaceutical Sciences and the COE21 Program, University of Shizuoka, 52-1 Yada, Shizuoka 422-8526, Japan

Received June 15, 2005; E-mail: abei@ys7.u-shizuoka-ken.ac.jp

Abstract: The chalcone synthase (CHS) superfamily of type III polyketide synthases (PKSs) produces a variety of plant secondary metabolites with remarkable structural diversity and biological activities (e.g., chalcones, stilbenes, benzophenones, acrydones, phloroglucinols, resorcinols, pyrones, and chromones). Here we describe an octaketide-producing novel plant-specific type III PKS from aloe (*Aloe arborescens*) sharing 50−60% amino acid sequence identity with other plant CHS-superfamily enzymes. A recombinant enzyme expressed in *Escherichia coli* catalyzed seven successive decarboxylative condensations of malonyl-CoA to yield aromatic octaketides SEK4 and SEK4b, the longest polyketides known to be synthesized by the structurally simple type III PKS. Surprisingly, site-directed mutagenesis revealed that a single residue Gly207 (corresponding to the CHS's active site Thr197) determines the polyketide chain length and product specificity. Small-to-large substitutions (G207A, G207T, G207M, G207L, G207F, and G207W) resulted in loss of the octaketide-forming activity and concomitant formation of shorter chain length polyketides (from triketide to heptaketide) including a pentaketide chromone, 2,7-dihydroxy-5-methylchromone, and a hexaketide pyrone, 6-(2,4-dihydroxy-6-methylphenyl)-4-hydroxy-2-pyrone, depending on the size of the side chain. Notably, the functional diversity of the type III PKS was shown to evolve from simple steric modulation of the chemically inert single residue lining the active-site cavity accompanied by conservation of the Cys-His-Asn catalytic triad. This provided novel strategies for the engineered biosynthesis of pharmaceutically important plant polyketides.

Introduction

The CHS superfamily of type III PKS enzymes is structurally and mechanistically distinct from the modular type I and dissociated type II PKSs of bacterial origin; the simple homodimer of 40−45 kDa proteins directly utilize CoA-linked substrates without involvement of the phosphopantetheine-armed acyl carrier protein to catalyze complete series of polyketide formation reactions with a single active site in an internal cavity.[1] The remarkable functional diversity of the CHS-superfamily enzymes derives from the differences of their selection of starter substrate, number of polyketide chain extensions, and mechanisms of cyclization reactions. For example, formation of 4,2′,4′,6′-tetrahydroxychalcone (naringenin chalcone) by CHS proceeds through sequential condensation of the C_6-C_3 unit of 4-coumaroyl-CoA as a starter with three C_2 units from malonyl-CoA, and the subsequent Claisen-type cyclization of the enzyme-bound tetraketide intermediate leads to formation of a new aromatic ring system (Figure 1A).[1] On the other hand, 2-pyrone synthase (2PS) from daisy (*Gerbera hybrida*) selects acetyl-CoA as a starter and carries out only two condensations with malonyl-CoA to produce a triketide, 6-methyl-4-hydroxy-2-pyrone (triacetic acid lactone) (Figure 1B).[2] Recent crystallographic and site-directed mutagenesis studies on plant[2−4] and bacterial[5,6] enzymes have begun to reveal structural and functional details of the type III PKSs that share almost the same three-dimensional overall fold and common

(1) For recent reviews, see: (a) Schröder, J. In *Comprehensive Natural Products Chemistry*; Elsevier: Oxford, 1999; Vol. 1, pp 749−771. (b) Austin, M. B.; Noel, J. P. *Nat. Prod. Rep.* **2003**, *20*, 79−110. (c) Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18*, 380−416.

(2) (a) Eckermann, S.; Schröder, G.; Schmidt, J.; Strack, D.; Edrada, R. A.; Helariutta, Y.; Elomaa, P.; Kotilainen, M.; Kilpeläinen, I.; Proksch, P.; Teeri, T. H.; Schröder, J. *Nature* **1998**, *396*, 387−390. (b) Jez, J. M.; Austin, M. B.; Ferrer, J.; Bowman, M. E.; Schröder, J.; Noel, J. P. *Chem. Biol.* **2000**, *7*, 919−930.

(3) (a) Ferrer, J. L.; Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Nat. Struct. Biol.* **1999**, *6*, 775−784. (b) Jez, J. M.; Ferrer, J. L.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Biochemistry* **2000**, *39*, 890−902. (c) Jez, J. M.; Noel, J. P. *J. Biol. Chem.* **2000**, *275*, 39640−39646. (d) Jez, J. M.; Bowman, M. E.; Noel, J. P. *Biochemistry* **2001**, *40*, 14829−14838. (e) Tropf, S.; Kärcher, B.; Schröder, G.; Schröder, J. *J. Biol. Chem.* **1995**, *270*, 7922−7928. (f) Suh, D. Y.; Fukuma, K.; Kagami, J.; Yamazaki, Y.; Shibuya, M.; Ebizuka, Y.; Sankawa, U. *Biochem. J.* **2000**, *350*, 229−235. (g) Suh, D. Y.; Kagami, J.; Fukuma, K.; Sankawa, U. *Biochem. Biophys. Res. Commun.* **2000**, *275*, 725−730. (h) Jez, J. M.; Bowman, M. E.; Noel, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5319−5324. (i) Austin, M. B.; Bowman, M. E.; Ferrer, J.-L.; Schröder, J.; Noel, J. P. *Chem. Biol.* **2004**, *11*, 1179−1194.

(4) (a) Abe, I.; Takahashi, Y.; Morita, H.; Noguchi, H. *Eur. J. Biochem.* **2001**, *268*, 3354−3359. (b) Abe, I.; Sano, Y.; Takahashi, Y.; Noguchi, H. *J. Biol. Chem.* **2003**, *278*, 25218−25226. (c) Abe, I.; Utsumi, Y.; Oguro, S.; Noguchi, H. *FEBS Lett.* **2004**, *562*, 171−176.

(5) (a) Funa, N.; Ohnishi, Y.; Fujii, I.; Shibuya, M.; Ebizuka, Y.; Horinouchi, S. *Nature* **1999**, *400*, 897−899. (b) Funa, N.; Ohnishi, Y.; Ebizuka, Y.; Horinouchi, S. *J. Biol. Chem.* **2002**, *277*, 4628−4635. (c) Funa, N.; Ohnishi, Y.; Ebizuka, Y.; Horinouchi, S. *Biochem. J.* **2002**, *367*, 781−789. (d) Izumikawa, M.; Shipley, P. R.; Hopke, J. N.; O'Hare, T.; Xiang, L.; Noel, J. P.; Moore, B. S. *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 510−515. (e) Austin, M. B.; Izumikawa, M.; Bowman, M. E.; Udwary, D. W.; Ferrer, J.-L.; Moore, B. S.; Noel, J. P. *J. Biol. Chem.* **2004**, *279*, 45162−45174.

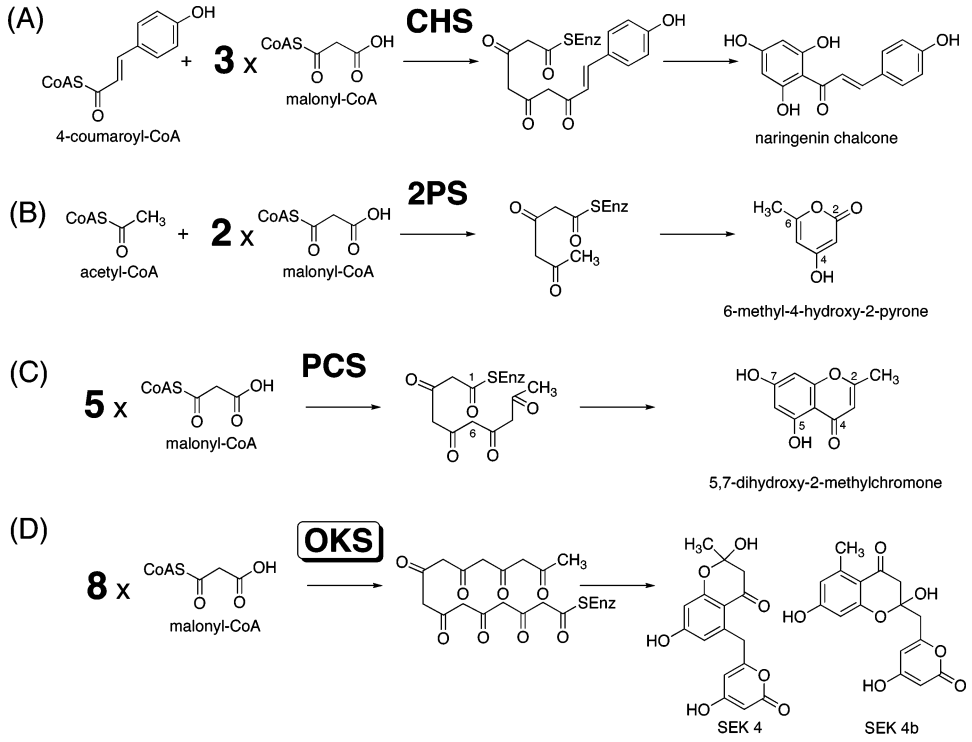

Figure 1. Proposed mechanism for the formation of (A) naringenin chalcone from 4-coumaroyl-CoA and three molecules of malonyl-CoA by CHS, (B) triacetic acid lactone from acetyl-CoA and two molecules of malonyl-CoA by 2PS, (C) 5,7-dihydroxy-2-methylchromone from five molecules of malonyl-CoA by PCS, and (D) SEK4 and SEK4b from eight molecules of malonyl-CoA by *A. arborescens* octaketide synthase (OKS).

active-site architecture with an absolutely conserved Cys-His-Asn catalytic triad. The polyketide formation reaction is thus initiated by starter molecule loading at the active site Cys, which is followed by malonyl-CoA decarboxylation, polyketide chain elongation, and final cyclization of the polyketide intermediate.[1]

Aloe (*Aloe arborescens*) is a medicinal plant rich in aromatic polyketides such as pharmaceutically important aloenin (hexaketide), aloesin (heptaketide), and barbaloin (octaketide) (Figure 2A). Therefore, in addition to regular CHSs involved in the flavonoid biosynthesis, the presence of functionally distinct multiple PKS enzymes that catalyze the initial key reactions in the biosynthesis of the secondary metabolites was expected.[7] Indeed, we recently succeeded in cloning a novel type III pentaketide chromone synthase (PCS) from the medicinal plant; PCS catalyzed condensation of five molecules of malonyl-CoA to produce 5,7-dihydroxy-2-methylchromone (Figure 1C), a biosynthetic precursor of the antiasthmatic furochromones, kehellin and visnagin.[7] Here we now report another novel *A. arborescens* type III PKS that produces aromatic octaketides SEK4 and SEK4b (Figure 1D), the longest polyketides known to be synthesized by the structurally simple homodimeric type III PKS.[7] Notably, the octaketides are the products of the minimal Type II PKS (heterodimeric complex of ketosynthase and chain length factor) for the benzoisochromanequinone actinorhodin (*act* from *Streptomyces coelicolor*).[8] Since the aloe plant does not produce SEK4/SEK4b and their metabolites, it is tempting to speculate that the enzyme is originally involved in the biosynthesis of anthrones and anthraquinones in the medicinal plant (Figure 2B).[9] The physiological role of the enzyme in the medicinal plant remains to be elucidated. Most importantly, it was demonstrated that a single residue determines the polyketide chain length and product specificity; the functional diversity of the type III PKS was shown to evolve from simple steric modulation of the chemically inert single residue Gly207 (corresponding to the CHS's active site thr197) lining the active-site cavity.

Experimental Section

Chemicals. [2-^{14}C]Malonyl-CoA (48 mCi/mmol) and [1-^{14}C]acetyl CoA (47 mCi/mmol) were purchased from Moravek Biochemicals (CA). 4-Coumaroyl-CoA, cinnamoyl-CoA, and benzoyl-CoA were chemically synthesized as described previously.[10] Malonyl-CoA, acetyl-CoA, and other aliphatic CoA esters were purchased from Sigma. Authentic samples of SEK4/SEK4b[7] and aloesone[4c] were obtained in our previous works.

cDNA Cloning. As described before,[7] total RNA was extracted from young roots of *A. arborescens* and reverse-transcribed using oligo dT primer (RACE 32 = 5′-GGC CAC GCG TCG ACT AGT ACT TTT

(6) (a) Saxena, P.; Yadav, G.; Mohanty, D.; Gokhale, R. S. *J. Biol. Chem.* **2004**, *278*, 44780−44790. (b) Sankaranarayanan, R.; Saxena, P.; Marathe, U.; Gokhale, R. S.; Shanmugam, V. M.; Rukmini, R. *Nat. Struct. Mol. Biol.* **2004**, *11*, 894−900. (c) Pfeifer, V.; Nicholson, G. J.; Ries, J.; Recktenwald, J.; Schefer, A. B.; Shawky, R. M.; Schröder, J.; Wohlleben, W.; Pelzer, S. *J. Biol. Chem.* **2001**, *276*, 38370−38377. (d) Tseng, C. C.; McLoughlin, S. M.; Kelleher, N. L.; Walsh, C. T. *Biochemistry* **2004**, *43*, 970−980.

(7) Abe, I.; Utsumi, Y.; Oguro, S.; Morita, H.; Sano, Y.; Noguchi, H. *J. Am. Chem. Soc*. **2005**, *127*, 1362−1363.

(8) (a) Fu, H.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1994**, *116*, 4166−4170. (b) Fu, H.; Hopwood, D. A.; Khosla, C. *Chem. Biol*. **1994**, *1*, 205−210.

(9) Dewick, P. M. In *Medicinal Natural Products, A Biosynthetic Approach*, 2nd ed.; Wiley: West Sussex, 2002.

(10) (a) Abe, I.; Morita, H.; Nomura, A.; Noguchi, H. *J. Am. Chem. Soc*. **2000**, *122*, 11242−11243. (b) Morita, H.; Takahashi, Y.; Noguchi, H.; Abe, I. *Biochem. Biophys. Res. Commun.* **2000**, *279*, 190−195. (c) Morita, H.; Noguchi, H.; Schröder, J.; Abe, I. *Eur. J. Biochem.* **2001**, *268*, 3759−3766. (d) Abe, I.; Takahashi, Y.; Noguchi, H. *Org. Lett.* **2002**, *4*, 3623−3626. (e) Abe, I.; Takahashi, Y.; Lou, W.; Noguchi, H. *Org. Lett.* **2003**, *5*, 1277−1280. (f) Abe, I.; Watanabe, T.; Noguchi, H. *Phytochemistry* **2004**, *65*, 2447−2453. (g) Oguro, S.; Akashi, T.; Ayabe, S.; Noguchi, H.; Abe, I. *Biochem. Biophys. Res. Commun.* **2004**, *325*, 561−567.

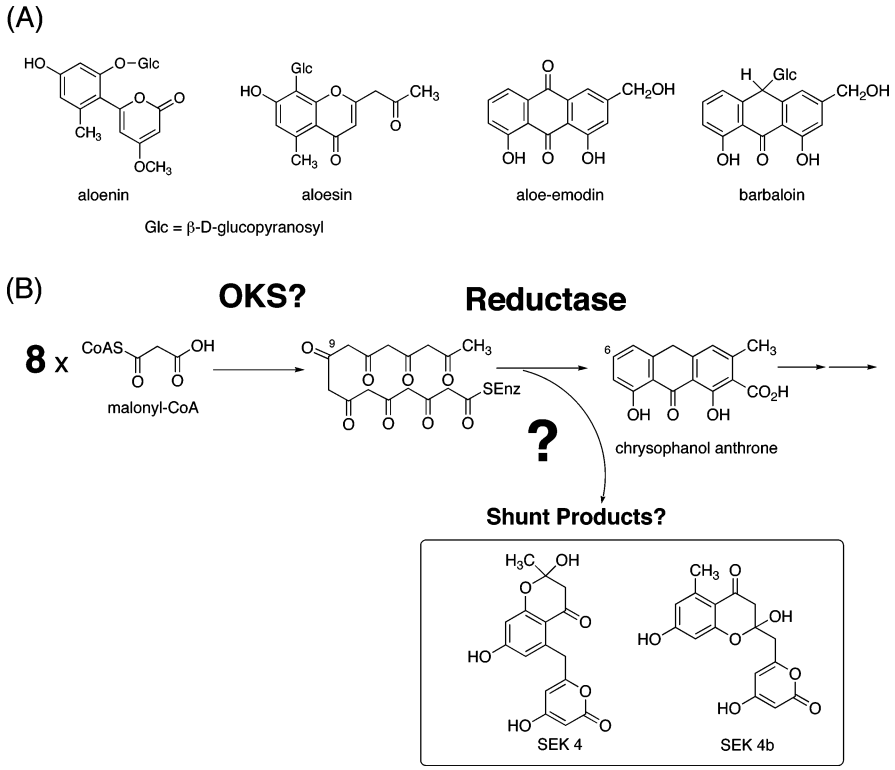

Figure 2. (A) Structures of aromatic polyketides produced by aloe (*A. arborescens*). (B) Hypothetical scheme[9] for involvement of OKS and a yet unidentified ketoreductase in the biosynthesis of anthrones and anthraquinones in *A. arborescens*. In the absence of interactions with the tailoring enzyme, OKS just affords SEK4/SEK4b as shunt products as in the case of the minimal type II PKS.

TTT TTT TTT TTT T-3′). The obtained cDNA mixture was used as a template for the PCR reactions with inosine-containing degenerate oligonucleotide primers based on the conserved sequences of known CHSs: 112S = 5′-(A/G)A(A/G) GCI ITI (A/C)A(A/G) GA(A/G) TGG GGI CA-3′, 174S = 5′- GCI AA(A/G) GA(T/C) ITI GCI GA(A/G) AA(T/C) AA-3′, 368A = 5′-CCC (C/A)(A/T)I TCI A(A/G)I CCI TCI CCI GTI GT-3′, and 380A = 5′-TCI A(T/C)I GTI A(A/G)I CCI GGI CC(A/G) AA-3′. The number of primer indicates the amino acid number of corresponding alfalfa (*Medicago sativa*) CHS. Nested PCR was carried out with the primer sets of 112S and 380A and then with 174S and 368A to amplify a 544-bp DNA fragment. For the PCR 30 cycles of reactions (94 °C for 0.5 min, 42 °C for 0.5 min, and 72 °C for 1 min) were performed each time with a 10 min final extension. The gel-purified PCR product was ligated into pT7Blue T-Vector (Novagen) and sequenced.

For the 3′-end amplification two specific primers (278S = 5′-GGA GCT CAC CAT CAT GAT GC-3′ and 327S = 5′-CAT CTC GAC AAT GCC ATC GG-3′) were designed based on the obtained core sequence. First RT-PCR was carried out with the primer set of 278S and RACE32 and the second PCR with 327S and RACE32 to amplify a 359-bp DNA fragment. 5′-End amplification was carried out using the Marathon cDNA Amplification kit (CLONTECH) and two specific primers (331A = 5′-CCA ATG ATC AGT GCA GCA GC-3′ and 232A = 5′-CCG ATG GCA TTG TCG AGA TG-3′) to amplify a 702-bp DNA fragment.

Expression of cDNA. A full-length cDNA was obtained using N- and C-terminal PCR primers; 5′-CCG CC*G AAT TC*A TGA GTT CAC TCT CCA AC-3′ (sense, the *EcoR* I site is italic) and 5′-ATA ATA *CTC GAG* CAT GAG AGG CAG GCT GTG-3′ (antisense, the *Xho* I site is italic). The amplified DNA was digested with *EcoR* I/*Xho* I and cloned into the *EcoR* I/*Xho* I site of pET-22b(+) (Novagen). Thus, the recombinant enzyme contains an additional hexahistidine tag at the C terminal. After confirmation of the sequence, the plasmid was transformed into *E. coli* BL21(DE3)pLysS. The cells harboring the plasmid were cultured to an A_{600} of 0.6 in Luria-Bertani medium containing 100 μg/mL of ampicillin at 23 °C. Then, 1.0 mM isopropyl thio-β-D-galactoside was added to induce protein expression, and the culture was incubated further at 23 °C for 16 h.

Enzyme Purification. The *E. coli* cells were harvested by centrifugation and resuspended in 40 mM potassium phosphate buffer, pH 7.9, containing 0.1 M NaCl. Cell lysis was carried out by sonication and centrifuged at 15 000 *g* for 40 min. The supernatant was passed through a column of Pro-Bond resin (Invitrogen) containing Ni^{2+} as an affinity ligand. After washing with 20 mM potassium phosphate buffer, pH 7.9, containing 0.5 M NaCl and 40 mM imidazole, the recombinant OKS was finally eluted with 15 mM potassium phosphate buffer, pH 7.5, containing 10% glycerol and 500 mM imidazole. Protein concentration was determined by the Bradford method (Protein Assay, Bio-Rad) with bovine serum albumin as standard. Finally, to determine subunit composition, the purified enzyme was applied to HPLC gel filtration column (TSK-gel G3000SW, 7.5 × 600 mm, TOSOH), which was eluted with 0.1 M KPB, pH 6.8, containing 10% glycerol and 0.2 M KCl at a flow rate of 1.0 mL/min. The standard molecular weight markers used were β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome *c* (12.4 kDa).

Site-Directed Mutagenesis. Gly207 mutants were constructed using the QuickChange Site-Directed Mutagenesis Kit (Stratagene) and a pair of primers as follows (mutated codons are italic): G207M sense 5′-GAG CTC ACC ATA ATC *ATG* CTT CGA GGC CCT-3′, anti sense 5′-AGG GCC TCG AAG *CAT* GAT TAT GGT GAG CTC-3′; G207T sense 5′-GAG CTC ACC ATA ATC *ACG* CTT CGA GGC CCT-3′, anti sense 5′-AGG GCC TCG AAG *CGT* GAT TAT GGT GAG CTC-3′; G207A sense 5′-GAG CTC ACC ATA ATC *GCG* CTT CGA GGC CCT-3′, anti sense 5′-AGG GCC TCG AAG *CGC* GAT TAT GGT GAG CTC-3′; G207L sense 5′-GAG CTC ACC ATA ATC *CTG* CTT CGA GGC CCT-3′, anti sense 5′-AGG GCC TCG AAG *CAG* GAT TAT GGT GAG CTC-3′; G207F sense 5′-GAG CTC ACC ATA ATC *TTC* CTT CGA GGC CCT-3′, anti sense 5′-AGG GCC TCG AAG *GAA* GAT TAT GGT GAG CTC-3′; G207W sense 5′-GAG CTC ACC

ATA ATC *TGG* CTT CGA GGC CCT-3′, anti sense 5′-AGG GCC TCG AAG *CCA* GAT TAT GGT GAG CTC-3′.

Enzyme Reaction. As described before, the standard reaction mixture contained 54 nmol of malonyl-CoA (and 27 nmol of other CoA ester) and 280 pmol of the purified recombinant enzyme in a final volume of 500 μL of 100 mM potassium phosphate buffer, pH 7.0.[7] Incubations were carried out at 30 °C for 2 h and stopped by adding 50 μL of 20% HCl. The products were then extracted twice with 1000 μL of EtOAc and analyzed by HPLC and LC-ESIMS on a TSK-gel ODS-80Ts column (4.6 × 150 mm, TOSOH) with a flow rate of 0.8 mL/min. For the standard assay gradient elution was performed with H_2O and MeOH, both containing 0.1% TFA: 0−5 min, 30% MeOH; 5−17 min, 30−60% MeOH; 17−25 min, 60% MeOH; 25−27 min, 60−70% MeOH. Retention time (min): SEK4 (19.3), SEK4b (20.6), aloesone (20.6), 6-(2,4-dihydroxy-6-methylphenyl)-4-hydroxy-2-pyrone (16.9), 2,7-dihydroxy-5-methylchromone (22.7), tetraacetic acid lactone (4.4), and triacetic acid lactone (6.0). For separation of SEK4b and aloesone, gradient elution was performed with H_2O and CH_3CN, both containing 0.1% TFA: 0−40 min, 15−30% CH_3CN. Retention time (min): SEK4 (17.8), SEK4b (20.3), and aloesone (23.6). On-line HPLC-ESIMS spectra were measured with a Hewlett-Packard HPLC 1100 series (Wilmington, DE) coupled to a Finnigan MAT LCQ ion trap mass spectrometer (San Jose, CA) fitted with an ESI source.

For large-scale enzyme reaction, 20 mg of purified enzyme was incubated with malonyl-CoA (20 mg, 30 mmol) in 100 mL of 100 mM phosphate buffer, pH 7.5, containing 1 mM EDTA, at 30 °C for 18 h. The reaction was quenched by addition of 20% HCl (10 mL) and extracted with ethyl acetate (200 mL × 3). The enzyme reaction products (ca. 0.5−1.0 mg) were purified by reverse-phase HPLC. Spectroscopic data of the products were as follows. 2,7-Dihydroxy-5-methylchromone: ESIMS R_t = 22.7 min, *m/z* 193 $[M + H]^+$. UV: λ_{max} 308 nm. ^{1}H NMR (400 MHz, DMSO-d_6): δ 6.53 (1H, d, J = 2.0 Hz, H-6), 6.48 (1H, d, J = 2.0 Hz, H-8), 5.32 (1H, s, H-3), 2.56 (3H, s, CH_3). ^{13}C NMR (100 MHz, DMSO-d_6): δ 182.3 (C-4), 164.7 (C-2), 161.8 (C-7), 160.3 (C-1a), 143.3 (C-5), 117.7 (C-4a), 114.5 (C-3), 112.5 (C-6), 100.3 (C-8), 22.7 (CH_3). 6-(2,4-Dihydroxy-6-methylphenyl)-4-hydroxy-2-pyrone: ESIMS R_t = 16.9 min, *m/z* 233 $[M + H]^+$. UV: λ_{max} 304 nm. ^{1}H NMR (400 MHz, DMSO-d_6): δ 6.21 (1H, d, J = 2.1 Hz, H-3′), 6.14 (1H, d, J = 2.1 Hz, H-5′), 6.02 (1H, d, J = 2.1 Hz, H-5), 5.27 (1H, d, J = 2.1 Hz, H-3), 2.06 (3H, s, CH_3). ^{13}C NMR (100 MHz, DMSO-d_6): δ 170.5 (C-4), 166.5 (C-2), 159.1 (C-4′), 157.8 (C-6), 156.7 (C-2′), 138.6 (C-6′), 116.8 (C-1′), 113.1 (C-5′), 105.3 (C-5), 100.1 (C-3′), 88.7 (C-3), 19.7 (CH_3). The NMR assignments were performed by comparison with those of 2-acetonyl-7-hydroxy-5-methylchromone[4c] (aloesone) and 6-(2,4-dihydroxy-6-methylphenyl)-4-methoxy-2-pyrone[11] (the aglycone of aloenin).

Determination of Starter Substrate. Acetyl-CoA, resulting from decarboxylation of malonyl-CoA, was also accepted as a starter substrate as in the case of *A. arborescens* PCS[7] but not so efficiently as in the case of *R. palmatum* ALS.[4c] This was confirmed by the ^{14}C incorporation rate from [1-^{14}C]acetyl CoA in the presence of cold malonyl-CoA, while the yield of the octaketides SEK4/SEK4b from [2-^{14}C]malonyl-CoA was almost at the same level in the presence or absence of cold acetyl-CoA in the reaction mixture. Theoretical ^{14}C-specific incorporation from [1-^{14}C]acetyl-CoA should be 12.5% of those from [2-^{14}C]-malonyl-CoA if acetyl CoA serves as a starter unit of the octaketides forming reaction, which was largely matched with the observed incorporation rate.

Enzyme Kinetics. Steady-state kinetic parameters were determined using [2-^{14}C]malonyl-CoA (1.8 mCi/mmol) as a substrate. The experiments were carried out in triplicate using five concentrations of substrate (from 6.5 to 117.8 μM) in the assay mixture, containing 10 μg of purified enzyme, 1 mM EDTA, in a final volume of 500 μL of 100 mM K-phosphate buffer, pH 7.5. Incubations were carried out at 30 °C for 30 min. The reaction products were extracted and separated by Si-gel TLC (Merck Art. 1.11798; ethyl acetate/hexane/AcOH = 63:27:5, v/v/v). Radioactivities were quantified by autoradiography using a bioimaging analyzer BAS-2000II (FUJIFILM). Lineweaver−Burk plots of data were employed to derive the apparent K_M and k_{cat} values (average of triplicates) using EnzFitter software (BIOSOFT).

Homology Modeling. As described before,[4b] the model was produced by the SWISS-MODEL package (http://expasy.ch/spdbv/) provided by the Swiss-PDB-Viewer program.[12] A standard homology modeling procedure was applied based on the sequence homology of residues 6−403 of *A. arborescens* OKS and the X-ray crystal structures of CHS including *M. sativa* CHS (1BI5A.pdb, 1BQ6A.pdb, 1CGKA.pdb, 1CGZA.pdb, 1CHWA.pdb, 1CHWB.pdb, 1CMLA.pdb), *M. sativa* CHS C164A mutant (1D6FA.pdb), *M.* sativa CHS N336A mutant (1D6HA.pdb), *M. sativa* CHS H303Q mutant (1D6IA.pdb, 1D6IB.pdb), *M. sativa* CHS G256A mutant (1I86A.pdb), *M. sativa* CHS G256V mutant (1I88A.pdb, 1I88B.pdb), *M. sativa* CHS G256L mutant (1I89A.pdb, 1I89B.pdb), *M. sativa* CHS G256F mutant (1I8BA.pdb, 1I8BB.pdb), and *M.* sativa CHS F215F mutant (1JWXA.pdb). The corresponding Ramachandran plot was also created with Swiss PDB-Viewer software to confirm that the majority of residues grouped in the energetically allowed regions. Calculation of cavity volumes (Connolly's surface volumes) was then performed with the CASTP program (http://cast.engr.uic.edu/cast/).

Results and Discussion

A cDNA encoding the octaketide synthase (OKS) (GenBank accession no. AY567707) was cloned and sequenced from young roots of aloe (*A. arborescens*) by RT-PCR using degenerate primers based on the conserved sequences of known CHSs.[7] A 1441-bp full-length cDNA contained a 60-bp 5′ noncoding region, a 1,212-bp open reading frame encoding a M_r 44 568 protein with 403 amino acids, and a 169-bp of 3′ noncoding region. The deduced amino acid sequence showed 50−60% identity to those of other type III PKSs of plant origin (Figure 3): 91% identity (368/403) with *A. arborescens* PCS,[7] 60% identity (240/403) with alfalfa (*Medicago sativa*) CHS[3a], and 54% identity (216/403) with a heptaketide-producing 2-acetonyl-7-hydroxy-5-methylchromone (aloesone) synthase (ALS) from *Rheum palmatum*.[4c] In contrast, it showed only 23% identity (93/403) with a bacterial 1,3,6,8-tetrahydroxynaphthalene synthase (RppA) from *Streptomyces griseus*.[5a]

Comparison of the sequence revealed conservation of the catalytic triad (Cys164, His303, and Asn336) and most of the CHS active-site residues (Met137, Gly211, Gly216, Phe215, Phe265, and Pro375) (numbering in *M. sativa* CHS);[3] however, CHS's conserved Thr197, Gly256, and Ser338, sterically altered in a number of divergent type III PKSs, are uniquely replaced with Gly, Leu, and Val, respectively (Figure 3). The three residues are also missing in the heptaketide-producing *R. palmatum* ALS (T197A/G256L/S338T),[4c] the pentaketide-producing *A. arborescens* PCS (T197M/G256L/S338V),[7] and the triketide-producing *G.* hybrida 2PS (T197L/G256L/S338I).[2] These chemically inert residues lining the active-site cavity have been proposed to control starter substrate selectivity and polyketide chain length by steric modulation of the initiation/elongation cavity.[2,7] The CHS-based homology modeling predicted that *A. arborescens* OKS has the same three-dimensional overall fold as *M. sativa* CHS,[3a] with the total cavity volume (1124 Å^3) slightly larger than that of the chalcone ($C_{15}H_{12}O_5$)-

(11) Suga, T.; Hirata, T. *Bull. Chem. Soc. Jpn* **1978**, *51*, 872−877.

(12) Guex, N.; Peitsch, M. C. *Electrophoresis* **1977**, *18*, 2714−2723.

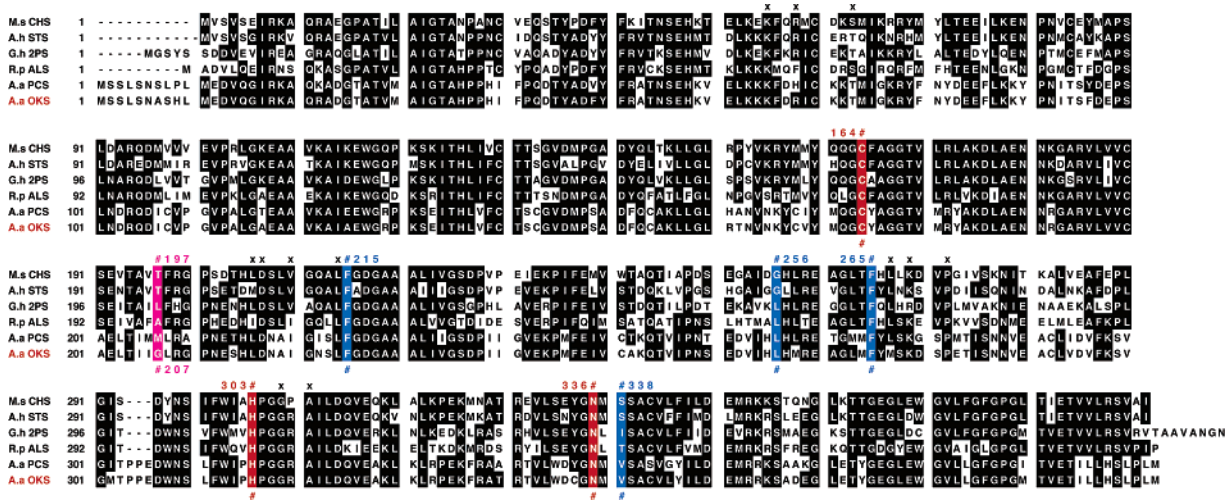

Figure 3. Comparison of the deduced amino acid sequences of *A. arborescens* OKS and other CHS-superfamily type III PKSs. M.s CHS, *M. sativa* CHS; A.h STS, *A. hypogaea* stilbene synthase; G.h 2PS, *G. hybrida* 2PS; R.p ALS, *R. palmatum* ALS; A.a PCS, *A. arborescens* PCS; A.a OKS, *A. arborescens* OKS. The critical active-site residue 197 (in pink), the catalytic triad (Cys164, His303, and Asn336) (in red), and the residues lining the active site (Phe215, Gly256, F265, Ser338) (in blue) are marked with # (numbering in *M. sativa* CHS) and residues for the CoA binding with +.

forming CHS (1019 Å^3) but much larger than that of the triketide ($C_6H_6O_3$)-producing *G. hybrida* 2PS (298 Å^3). This suggested that the active site of OKS is large enough to perform the seven successive condensation reactions and accommodate the octaketide products ($C_{16}H_{14}O_7$).

A. arborescens OKS was heterologously expressed in *E. coli* with an additional hexahistidine tag at the C-terminal. The purified enzyme gave a single band with a molecular mass of 45 kDa on SDS−PAGE, while the native OKS appeared to be a homodimer since it had an apparent molecular mass of 90 kDa as determined by gel filtration. Despite the 91% amino acid sequence identity with the pentaketide-producing PCS, the recombinant OKS did not produce 2,7-dihydroxy-2-methylchromone (Figure 1C) but instead efficiently accepted malonyl-CoA as a sole substrate to yield a 1:4 mixture of SEK4 and SEK4b (Figure 1D and 5A), the longest polyketides known to be generated by the structurally simple homodimeric type III PKS.[7] The octaketides SEK4/SEK4b are the shunt products of the minimal type II PKS (heterodimeric complex of ketosynthase and chain length factor) when expressed in the absence of the ketoreductase.[8] Since the medicinal plant does not produce SEK4/SEK4b and their metabolites but instead produces a significant amount of anthrones and anthraquinones (octaketides) (Figure 2A), it is tempting to speculate that the enzyme is originally involved in the biosynthesis of anthrones/anthraquinones. However, maybe because of misfolding of the protein in the *E. coli* expression system or because of the absence of interactions with tailoring enzymes such as a yet unidentified ketoreductase[9] the recombinant OKS just afforded SEK4/SEK4b as shunt products as in the case of the minimal type II PKS (Figure 2B). The physiological role of OKS in the medicinal plant remains to be elucidated.

The recombinant OKS showed K_M = 95.0 μM and k_{cat} = 94.0 × 10^{-3} min^{-1} for malonyl-CoA (SEK4b forming activity) with a pH optimum at 7.5 at 30 °C, which was comparable with PCS (K_M = 71.0 μM and k_{cat} = 445 × 10^{-3} min^{-1}).[7] As in the case of other type III PKSs,[3h,10] *A. arborescens* OKS exhibited unusually broad substrate tolerance; the enzyme also accepted aromatic (4-coumaroyl, cinnamoyl, and benzoyl) and

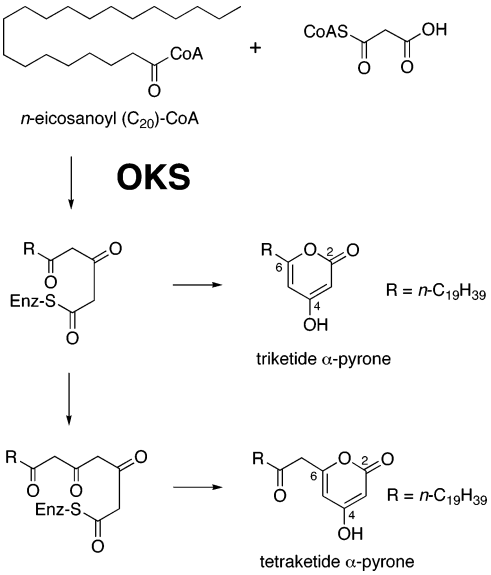

Figure 4. Formation of a triketide and a tetraketide α-pyrone from *n*-eicosanoyl-CoA and malonyl-CoA by OKS.

long-chain fatty acyl (*n*-hexanoyl, *n*-octanoyl, *n*-decanoyl, *n*-dodecanoyl, *n*-tetradecanoyl, *n*-hexadecanoyl, *n*-octadecanoyl, and *n*-eicosanoyl) CoA esters as a starter and carried out sequential condensations with malonyl-CoA to produce triketide and tetraketide α-pyrones without formation of a new aromatic ring system (Figure 4). In particular, it is noteworthy that OKS readily accepted the C_{20} chain length ester, much longer than the previously reported *Scutellaria baicalensis* CHS that only accepted up to the C_{12} ester,[10f] suggesting a difference in the active-site structure between the two enzymes. Finally, acetyl-CoA, resulting from decarboxylation of malonyl-CoA, was also accepted as a starter substrate just as in the case of the previously reported *A. arborescens* PCS[7] (data not shown) but not so efficiently as in the case of *R. palmatum* ALS.[4c] This was confirmed by the ^{14}C incorporation rate from [1-^{14}C]acetyl CoA; the yield of SEK4/SEK4b from [2-^{14}C]malonyl-CoA was almost at the same level in the presence or absence of cold acetyl-CoA in the reaction mixture.

As mentioned above, the octaketide-producing OKS and the pentaketide-producing PCS share 91% amino acid sequence identity, and the CHS's active site Thr197[3a] is uniquely substituted with small Gly207 in OKS while with bulky Met207 in PCS (Figure 3). Further, as we reported in a previous paper, a PCS mutant in which Met207 was replaced by Gly efficiently yielded the octaketides SEK4/SEK4b instead of the pentaketide chromone.[7] This suggested that the single residue determines the polyketide chain length and product specificity by simple steric modulation of the active-site cavity. To further test the hypothesis we constructed a series of OKS mutants bearing small-to-large changes in place of Gly207 (G207A, G207T, G207M, G207L, G207F, and G207W mutant) and investigated the mechanistic consequences of the point mutations.

First, when Gly207 was substituted with bulky Met as in the case of PCS, OKS G207M mutant completely lost the octaketide-forming activity but instead efficiently afforded an unnatural pentaketide (*m/z* 193 $[M + H]^+$) as a single product (Figure 5B). The NMR (^{1}H NMR and ^{13}C NMR) and MS spectra of the product obtained from a large-scale enzyme reaction (1.0 mg from 20 mg of malonyl-CoA) were characteristic of those of chromones and showed good agreement with those of 2-acetonyl-7-hydroxy-5-methylchromone[4c] (aloesone), except the signals due to the terminal acetonyl group. Accordingly, the structure of the novel pentaketide was determined to be 2,7-dihydroxy-5-methylchromone,[13] which was uniquely consistent with both biogenetic reasoning and the spectroscopic data. The octaketide-producing OKS was thus converted to a pentaketide synthase by single amino acid substitution. Interestingly, the pentaketide is a regioisomer of the PCS product, 5,7-dihydroxy-2-methylchromone (R_t = 23.5 min; UV, λ_{max} 292 nm), which is formed by a C-1/C-6 Claisen-type cyclization (Figure 1B), while in OKS G207M a C-4/C-9 aldol-type cyclization of the polyketide intermediate folded in a different conformation yielded 2,7-dihydroxy-5-methylchromone (R_t = 22.7 min, UV, λ_{max} 308 nm) (Figure 6A). Despite the high sequence identity, OKS and PCS are not functionally interconvertible by the point mutation, suggesting further subtle structural differences of the active site between the two enzymes. In addition, other bulky substitutions G207L and G207F also resulted in production of the pentaketide 2,7-dihydroxy-5-methylchromone along with a trace amount of SEK4/SEK4b (Figure 6B).

Interestingly, unlike the wild-type OKS that efficiently converted the C_{20} fatty acyl CoA into the triketide and tetraketide α-pyrones (Figure 4), the pentaketide-producing G207M mutant only accepted shorter chain length CoAs (up to the C_{12} ester) as a starter substrate for the pyrone formation reactions (data not shown). It is thus likely that the fatty acyl-CoAs are loaded into the same active-site cavity where the sequential decarboxylative condensations of the growing polyketide chain take place. Steric contraction of the cavity by the G207M bulky substitution rejected the longer chain CoAs as well as octaketide formation. Further intriguing, the G207M mutant did not afford any reaction products from the C_{14} or C_{16} ester, while it efficiently yielded the pentaketide chromone from malonyl-CoA as a starter when incubated with the C_{18} and the longer chain CoAs. This suggested that the C_{14} and C_{16} ester just blocked up the active-site cavity and totally inhibited the condensation reactions, while the C_{18} and larger CoAs could be no longer loaded into the active site of the G207M mutant.

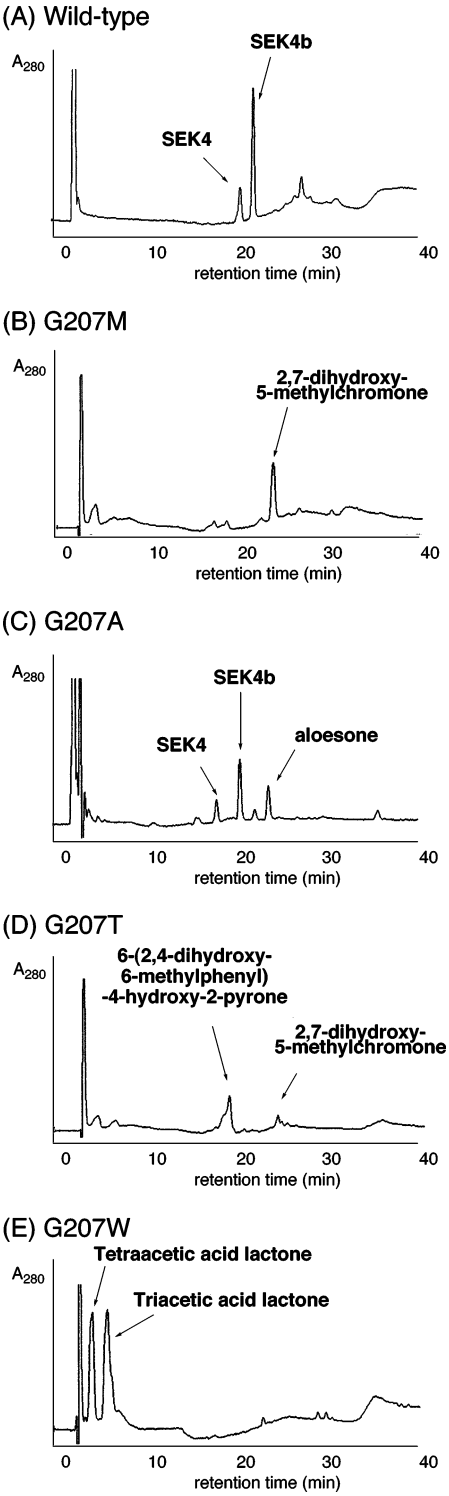

Figure 5. HPLC elution profiles of enzyme reaction products of (A) wild-type OKS, (B) OKS G207M mutant, (C) OKS G207A mutant, (D) OKS G207T mutant, and (E) OKS G207W mutant. HPLC separation conditions were as described in the Experimental Section. Note that only C was eluted with the gradient program with H_2O and CH_3CN for separation of SEK4b and aloesone.

Next, when Gly207 was replaced with Ala as in the case of the heptaketide-producing ALS, OKS G207A mutant indeed yielded aloesone in addition to SEK4/SEK4b (Figures 5C and 6A). Notably, the heptaketide is a biosynthetic precursor of

(13) Stockinger, H.; Schmidt, U. *Liebigs Ann. Chem.* **1976**, 1617−1625.

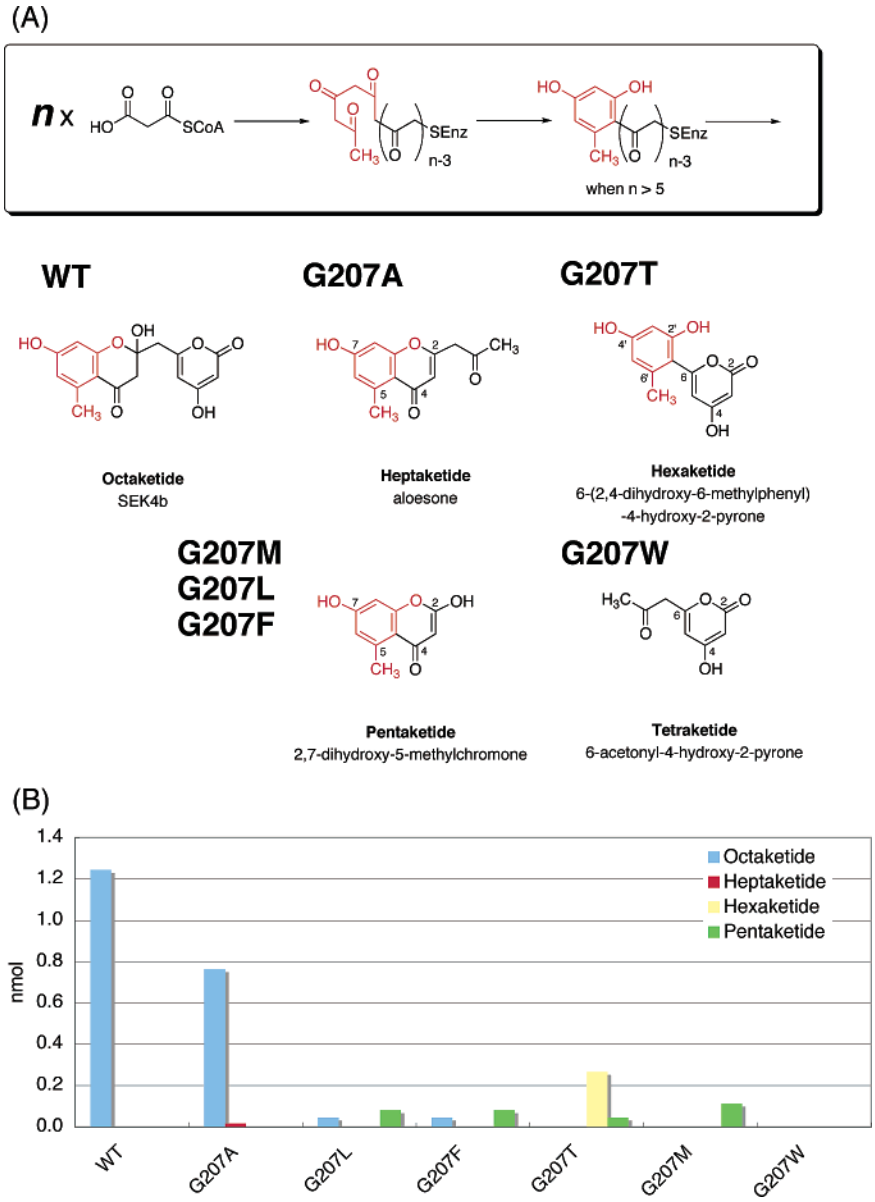

Figure 6. (A) Proposed mechanism for the formation of aromatic polyketides by OKS and its mutants. The enzymes catalyze chain initiation and elongation, possibly initiating the first aromatic ring formation reaction at the methyl end of the polyketide intermediate (in red). The partially cyclized intermediates are then released from the active site and undergo subsequent spontaneous cyclizations, leading to formation of the fused ring systems. (B) Distribution pattern of aromatic polyketides produced by OKS and its mutants. Quantification of the products was calculated from the ^{14}C incorporation rate from [2-^{14}C]malonyl-CoA under the standard assay condition.

aloesin (Figure 2A), the antiinflammatory agent of the medicinal plant.[4c] On the other hand, when Gly207 was substituted with Thr as in the case of CHS and most other plant type III PKSs[3], OKS G207T mutant completely lost the octaketide-forming activity and instead efficiently afforded a hexaketide (*m*/*z* 235 $[M + H]^+$) along with the pentaketide chromone (Figures 5D and 6B). Spectroscopic data (^1H NMR, ^{13}C NMR, MS, and UV) of the enzyme reaction product showed good agreement with those of 6-(2,4-dihydroxy-6-methylphenyl)-4-methoxy-2-pyrone[11] (the aglycone of aloenin), except the signals due to the 4-methoxy group. The structure of the hexaketide was thus determined to be 6-(2,4-dihydroxy-6-methylphenyl)-4-hydroxy-2-pyrone (Figure 6A), which is a biosynthetic precursor of aloenin (Figure 2A), the antihistamic agent of the medicinal plant.[11] Finally, when Gly207 was substituted with the most bulky Trp, OKS G207W mutant just afforded a tetraketide pyrone, 6-acetonyl-4-hydroxy-2-pyrone (tetraacetic acid lactone)[14] (*m*/*z* 169 $[M + H]^+$), and triacetic acid lactone without formation of an aromatic ring system (Figures 5E and 6A).

The small-to-large substitutions in place of Gly207 in *A. arborescens* OKS (corresponding to CHS's active site Thr197) resulted in loss of the octaketide-forming activity and concomitant formation of shorter chain length polyketides (from triketide to heptaketide) including the hexaketide pyrone and the pentaketide chromone depending on the size of the side chain (Figure 6B). It is conceivable that the point mutations cause steric contraction of the active-site cavity where the polyketide chain elongations take place, consequently shortening the product chain length. Remarkably, the functional diversity of

(14) (a) Bentley, R.; Zwitkowits, P. M. *J. Am. Chem. Soc*. **1967**, *89*, 676−680. (b) Samappito, S.; Page, J.; Schmidt, J.; De-Eknamkul, W.; Kutchan, T. M. *Planta* **2002**, *216*, 64−71.

the type III PKS was shown to evolve from the simple steric modulation of the active-site cavity accompanied by conservation of the Cys-His-Asn catalytic triad. An analogous result for the chain length control has been recently described for the minimal type II PKS (the heterodimeric complex of ketosynthase and chain length factor).[15] Presumably, OKS and the G207 mutant enzymes catalyze chain initiation and elongation and possibly initiate the first aromatic ring formation reaction at the methyl end of the polyketide intermediate (Figure 6A). The partially cyclized aromatic intermediates would be then released from the active site and undergo subsequent spontaneous cyclizations, thereby completing the formation of the fused ring systems.

Very recently we succeeded in crystallization of the pentaketide-producing *A. arborescens* PCS, both wild type and the octaketide-producing M207G mutant enzyme (studies in progress). The crystal structures at 1.6 Å resolution revealed that PCS and CHS share the same three-dimensional overall fold, including the CoA binding tunnel and the geometry of the catalytic triad. Remarkably, it was clearly demonstrated that the Met207 lining the active-site cavity indeed occupies a crucial position for the polyketide chain elongation reactions. Substitution of Met207 with Gly opens a gate to a novel buried pocket and thus expands a putative polyketide chain elongation tunnel, which lead to formation of the longer octaketides. It is very likely that *A. arborescens* OKS shares the similar active-site architecture and machinery.

In conclusion, *A. arborescens* OKS is a novel plant type III PKS that produces octaketides SEK4/SEK4b. We demonstrated the mechanistic consequences of substitution of the crucial active-site residue involved in the polyketide chain length control and provided the structural basis for understanding the structure−function relationship of type III PKS enzymes. These findings revolutionized our concept for the catalytic potential of the structurally simple plant type III PKSs and suggest novel strategies for the engineered biosynthesis of pharmaceutically important plant polyketides.

Acknowledgment. This work was supported in part by the COE21 Program, Grant-in-Aid for Scientific Research (nos. 16510164, 17310130, and 17035069), by the Cooperation of Innovative Technology and Advanced Research in Evolutional Area (CITY AREA) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by a Health and Labor Sciences Research Grant from the Ministry of Health, Labour and Welfare, Japan.

JA053945V

(15) An analogous result in Type II PKS has been reported, see: (a) Tang, Y.; Tsai, S.-C.; Khosla, C. *J. Am. Chem. Soc.* **2003**, *125*, 12708−12709. (b) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. *Nat. Struct. Mol. Biol.* **2004**, *11*, 888−893.