Alteration of reaction and substrate specificity of a bacterial type III polyketide synthase by site-directed mutagenesis

Nobutaka FUNA*, Yasuo OHNISHI*, Yutaka EBIZUKA† and Sueharu HORINOUCI*¹

*Department of Biotechnology, Graduate School of Agriculture and Life Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan, and †Department of Natural Product Chemistry, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan

RppA, which belongs to the type III polyketide synthase family, catalyses the synthesis of 1,3,6,8-tetrahydroxynaphthalene (THN), which is the key intermediate of melanin biosynthesis in the bacterium *Streptomyces griseus*. The reaction of THN synthesis catalysed by RppA is unique in the type III polyketide synthase family, in that it selects malonyl-CoA as a starter substrate. The Cys-His-Asn catalytic triad is also present in RppA, as in plant chalcone synthases, as revealed by analyses of active-site mutants having amino acid replacements at Cys¹³⁸, His²⁷⁰ and Asn³⁰³ of RppA. Site-directed mutagenesis of the amino acid residues that are likely to form the active-site cavity revealed that the aromatic ring of Tyr²²⁴ is essential for RppA to select malonyl-CoA as a starter substrate, since substitution of Tyr²²⁴ by amino acids other than Phe and Trp abolished the ability of RppA to accept malonyl-CoA as a starter, whereas the mutant enzymes Y224F and Y224W were capable of syn-

thesizing THN via the malonyl-CoA-primed reaction. Of the site-directed mutants generated, A305I was found to produce only a triketide pyrone from hexanoyl-CoA as starter substrate, although wild-type RppA synthesizes tetraketide and triketide pyrones in the hexanoyl-CoA-primed reaction. The kinetic parameters of Ala³⁰⁵ mutants and identification of their products showed that the substitution of Ala³⁰⁵ by bulky amino acid residues restricted the number of elongations of the growing polyketide chain. Both Tyr²²⁴ (important for starter substrate selection) and Ala³⁰⁵ (important for intermediate elongation) were found to be conserved in three other RppAs from *Streptomyces antibioticus* and *Streptomyces lividans*.

Key words: chalcone synthase, RppA, starter substrate selectivity, *Streptomyces griseus*, 1,3,6,8-tetrahydroxynaphthalene.

INTRODUCTION

The type III polyketide synthases (PKSs) are small homodimeric proteins (monomer size 40–45 kDa) that are responsible for catalysing the biosynthesis of aromatic polyketides in bacteria and plants [1,2]. RppA, the first PKS to be characterized as a type III PKS in bacteria, catalyses the synthesis of 1,3,6,8-tetrahydroxynaphthalene (THN) (Scheme 1A), which is the key intermediate in melanin biosynthesis in the bacterium *Streptomyces griseus* [3]. Chalcone synthase (CHS) is a plant-specific type III PKS that catalyses the formation of 4,2',4',6'-tetrahydrochalcone (naringenin chalcone) (Scheme 1B) [4], an essential precursor of flavonoids and isoflavonoids, which have attracted significant attention because of their wide range of pharmacological properties. A number of type III PKSs are present in plants that differ from CHSs; these include 2-pyrone synthase (PS), which yields 4-hydroxy-6-methyl-2-pyrone (triacetic acid lactone; TAL) (Scheme 1C) [5]. The discovery of type III PKSs in bacteria has opened up the field of investigation of the pathways of aromatic polyketide biosynthesis. Recently, Pfeifer et al. [6] revealed that the polyketide moiety of a vancomycin-type antibiotic, which is the last-resort antibiotic for treating methicillin-resistant *Staphylococcus aureus* infections, is synthesized by DpgA, a type III PKS sharing 22% sequence identity with RppA.

The first step in the formation of aromatic polyketides by type III PKSs involves the covalent binding of a starter moiety derived from CoA-linked starter substrates on to the thiol group

of the cysteine at the active site [7]. Then the acetyl-CoA carbanion, generated by decarboxylation of malonyl-CoA, is condensed stepwise to elongate the growing polyketide chain. A variety of patterns are observed in the final step of the reaction forming new aromatic rings by type III PKSs. The typical reaction is intramolecular Claisen condensation yielding phloroglucinols, which is observed for CHS, while PS catalyses pyrone formation by lactonization of a triketide intermediate [8]. Recent three-dimensional structural studies of the CHS from *Medicago sativa* (alfalfa) identified four amino acids (Cys¹⁶⁴, Phe²¹⁵, His³⁰³ and Asn³³⁶) involved in malonyl-CoA decarboxylation and condensation [9]. The difference between the products formed by CHS and RppA arises from the selection of starter substrate, the number of condensations of malonyl-CoA, and the pattern of final ring closure. CHS uses *p*-coumaroyl-CoA as a starter substrate and catalyses the condensation of three malonyl-CoAs to form an enzyme-bound tetraketide intermediate that is subsequently folded into naringenin chalcone by a Claisen condensation (Scheme 1B). In contrast, RppA catalyses THN synthesis by selecting malonyl-CoA as the starter substrate, carrying out four successive extensions of malonyl-CoA, and finally Claisen condensation for the ring closure that is presumably accompanied by decarboxylation of the carboxy group of the malonyl-CoA used as the starter substrate (Scheme 1A) [3].

Type III PKSs are known to utilize a wide range of starter substrates to yield phloroglucinols and α -pyrones. For example, CHS accepts aliphatic acyl-CoAs and *p*-coumaroyl-CoA as

Abbreviations used: CHS, chalcone synthase; LC-APCIMS, liquid chromatography/atmospheric-pressure chemical ionization MS; PKS, polyketide synthase; PS, 2-pyrone synthase; TAL, triacetic acid lactone; THN, 1,3,6,8-tetrahydroxynaphthalene.

¹ To whom correspondence should be addressed (e-mail asuhori@mail.ecc.u-tokyo.ac.jp).

The nucleotide sequence data reported have been deposited in the GenBank®, DDBJ, EMBL and GSDB Nucleotide Sequence Databases under the accession numbers AB084489, AB084490 and AB084491 for Sa-rppA, Sa-rppB and Sl-rppA respectively.

Scheme 1 Reactions catalysed by type III PKSs

(A) THN synthesis by RppA. The structure of the pentaketide intermediate is hypothetical. THN is readily oxidized to flavinolin non-enzymically. (B) Chalcone formation by CHS. (C) TAL formation by PS. (D) Substrate specificity of RppA. RppA yields tetra- and tri-ketide pyrones when hexanoyl-CoA (**1**) is incorporated as the starter substrate. Phenylacetyl-CoA (**2**) is not incorporated into wild-type RppA, but is incorporated into mutants Y224G, Y224L, Y224W, Y224F, Y224C, A305S and A305I. The ring closure catalysed by type III PKSs is supposed to proceed when polyketide intermediates are attached to CoA.

starter substrate at similar efficiencies [10]. Likewise, RppA can accept C₄–C₈ aliphatic acyl-CoAs as well as malonyl-CoA, and catalyses the sequential condensation of malonyl-CoA [11]. Despite this low substrate specificity of type III PKSs, the ability to use malonyl-CoA as a starter substrate is unique to RppA. However, the mechanism of this starter substrate specificity was not clear. We therefore investigated the mechanism of the starter substrate selectivity of RppA by site-directed mutagenesis of amino acid residues that are probably involved in starter substrate selectivity. We chose the candidate amino acids to be changed on the basis of the crystal structure of alfalfa CHS and amino acid alignment of RppAs from *Streptomyces*. For this alignment, three other *rppA* genes, which were later found to encode RppA proteins with the same substrate specificity as RppA from *S. griseus*, were cloned from *Streptomyces antibioticus* and *Streptomyces lividans*, and their amino acid sequences were deduced from the nucleotide sequences. In the present study, we first confirmed that the Cys-His-Asn triad in the active site is essential for malonyl-CoA decarboxylation and polyketide formation by RppA. Subsequent site-directed mutagenesis of amino acid residues probably forming the active-site pocket led to the identification of Tyr²²⁴ as an essential residue for the selection of malonyl-CoA as the starter substrate, and Ala³⁰⁵ as an important residue for determining the number of elongations of an acetate unit.

EXPERIMENTAL

Materials

[2-¹⁴C]Malonyl-CoA (1.96 GBq/mmol) was purchased from Amersham Pharmacia Biotech. All unlabelled CoA esters were purchased from Sigma.

Identification and functional characterization of RppAs from *S. lividans* and *S. antibioticus*

Southern hybridization, colony hybridization and nucleotide sequencing were carried out as described by Maniatis et al. [12]. The 1.1 kb *rppA* gene from *S. griseus*, excised from pET16b-RppA [3], was used as a ³²P-labelled hybridization probe to clone *rppA* homologues from the chromosomal DNA of *S. antibioticus* IFO13271 and *S. lividans* TK21. A gene (Sl-*rppA*) from *S. lividans* and two genes (Sa-*rppA* and Sa-*rppB*) from *S. antibioticus* were cloned and their nucleotide sequences were determined. An *Nde*I site was introduced at the start codon of each *rppA* gene by PCR with pairs of primers, as follows: Sa-RppA, 5'-GCG-GAATTCCAAGAGAAGTGGATTCCACATATGCCAGG-3' and 5'-GCGGGATCCTTTCTCCGCTGTCCGTCAGGG-GCGC-3'; Sa-RppB, 5'-CGCGAATTCGGCACGGCATATG-GCGACTT-3' and 5'-GCGGGATCCGTGATCGTCTCTTC-GGTCAT-3'; Sl-RppA, 5'-GCCAAGCTTCATATGGCGAC-

TTTGTGCAGACC-3' and 5'-GCCGAATTCCTCGAGTTC-AGTCATGCCTGCCTCAC-3' (italic letters indicate an *NdeI* site and the underlined bases represent restriction sites used for further cloning). After amplification by PCR under standard conditions, the *EcoRI*–*Bam*HI fragment containing the *Sa-rppA* coding sequence, the *EcoRI*–*Bam*HI fragment containing the *Sa-rppB* coding sequence and the *Hind*III–*EcoRI* fragment containing the *Sl-rppA* coding sequence were cloned in pUC19. The absence of undesired alterations during PCR was confirmed by nucleotide sequencing. The *Sa-rppA*, *Sa-rppB* and *Sl-rppA* coding sequences were then excised as *NdeI*–*Bam*HI, *NdeI*–*Bam*HI and *NdeI*–*XhoI* fragments respectively, and inserted between the corresponding cloning sites of pET16b (Novagen), resulting in pET16-Sa-RppA, pET16-Sa-RppB and pET16-Sl-RppA respectively. Each of the expression plasmids would direct the synthesis of a protein with the structure Met-Gly-His₁₀-Ser₂-Gly-His-Ile-Glu-Gly-Arg-His-RppA/B. The histidine-tagged RppA/B proteins were produced in *Escherichia coli* BL21 (DE3) and purified with a His-binding resin, as described in [3].

Construction of RppA mutants

Mutations were introduced into the *rppA* gene of *S. griseus* using the Takara *in vitro* mutagenesis kit (Mutan-Super Express K_m). The *NdeI*–*Bam*HI fragment excised from pET-RppA was cloned into pKF19k, resulting in pKF19k-RppA. The oligonucleotides used for PCR are shown in Table 1. PCR amplification using pKF19k-RppA as the template was performed according to the supplier's instructions. The entire coding sequences of the mutant constructs were confirmed by nucleotide sequencing. For protein purification, the mutant *rppA* genes were cloned between the *NdeI* and *Bam*HI sites of pET16b. The mutant RppAs were produced and purified as described above.

PKS assay

The standard reaction mixture contained [2-¹⁴C]malonyl-CoA, 100 mM Tris/HCl (pH 7.5) and 45 µg of RppA in a total volume of 500 µl. Reactions were incubated at 30 °C and stopped by adding 100 µl of 6 M HCl, and the products were extracted with 600 µl of ethyl acetate. The organic layer was dried by flushing with N₂ and the residual material was dissolved in 15 µl of methanol for HPLC or TLC analysis. Reverse-phase HPLC conditions were as follows: ODS-80Ts (C₁₈) column (4.6 mm × 150 mm; Tosoh), maintained at 40 °C, using gradient solvent systems of solvent A [2% (v/v) acetic acid in acetonitrile] and solvent B [2% (v/v) acetic acid in water]. The eluate was collected every 1 min and radioactivity was measured directly by liquid scintillation counting. Reverse-phase TLC was performed by using RP-18 (Merck). The polyketides produced were identified by LCQ (Thermo Quest), liquid chromatography/atmospheric-pressure chemical ionization MS (LC-APCIMS).

Products of the malonyl-CoA-primed reaction

(HPLC was performed in the mobile-phase system described in the legend to Figure 2.) (i) THN: retention time was 15 min on LC. LC-APCIMS: positive, *m/z* 193 (*M*+H)⁺; negative, *m/z* 191 (*M*–H)[–]. (ii) 2,5,7-Trihydroxy-1,4-naphthoquinone (flavolin): retention time was 22 min on LC. LC-APCIMS: positive, *m/z* 207 (*M*+H)⁺; negative, *m/z* 205 (*M*–H)[–]. (iii) TAL: retention time was 9.5 min on LC. LC-APCIMS: positive, *m/z* 127 (*M*+H)⁺; negative, *m/z* 125 (*M*–H)[–]. (iv) 6-Acetonil-4-hydroxy-2-pyrone (tetraketide lactone): retention time was 8 min on LC. LC-APCIMS: positive, *m/z* 169 (*M*+H)⁺; negative, *m/z* 167 (*M*–H)[–].

Table 1 Primer sequences used for site-directed mutagenesis

Codons encoding the mutations are underlined.

Mutant	Sequence
C106A	5'-TCGTCTACGTATCCGCCACGGGTTTCATG-3'
T107A	5'-TCGTCTACGTATCTGCGGGGTTTCATG-3'
C138A	5'-GCTCGGGCGCGGGCGGGCGGGCGGGCGATCAAC-3'
C138Q	5'-GCTCGGGCGAGCGGGCGGGCGGGCGGGCGATCAAC-3'
C168A	5'-TCGTGTCTGCGAGTTCGCGAGCCTGTGC-3'
C171A	5'-CTGGGCTACACGCCACCGATATCGGGGTC-3'
C171L	5'-CTGCTCTACACGCCACCGATATCGGGGTC-3'
C171T	5'-CTGACTACACGCCACCGATATCGGGGTC-3'
F188A	5'-CGGACTCGCGGGCGAGCGGCTCTCCGCGGGCGTC-3'
F188S	5'-CTCTCGGGCGAGCGGCTCTCCGCGGGCGTC-3'
F188Y	5'-CTCTACGGCGAGCGGCTCTCCGCGGGCGTCGA-3'
I222A	5'-CTGGGCTCTACGCGGTTGCGACACC-3'
I222F	5'-GGACTGGTTCTCTACGCGGTTGCGGACACC-3'
I222M	5'-GGACTGGATGCTCTACGCGGTTGCGGACACC-3'
Y224A	5'-TCTCGCGCGGGTTGCGGACACC-3'
Y224C	5'-TCTCTCGCGGGTTGCGGACACC-3'
Y224F	5'-TCTCTCTCGCGGGTTGCGGACACC-3'
Y224G	5'-TCTCTCGGGCGGTCGCGACACC-3'
Y224H	5'-TCTCTCGCGGGTTGCGGACACC-3'
Y224L	5'-TCTCTCTCGCGGGTTGCGGACACC-3'
Y224S	5'-TCTCTCTCGCGGGTTGCGGACACC-3'
Y224W	5'-TCTCTCTGGCGGGTTGCGGACACC-3'
F231A	5'-CGCGGTTGCGGACACC-3'
F233A	5'-CGCGGTTGCGGACACC-3'
H270A	5'-TCGTGCGCGGGCGGACCGCGCATCTTGGAC-3'
H270N	5'-TCGTCAACGCGGGCGGACCGCGCATCTTGGAC-3'
H270Q	5'-TCGTCTAGCGGGCGGACCGCGCATCTTGGAC-3'
N303A	5'-CCGAACGCGGGCGGACCGAGCTCCGTCGCTTCGAC-3'
N303D	5'-CCGAACGCGGGCGGACCGAGCTCCGTCGCTTCGAC-3'
N303H	5'-CCGAACGCGGGCGGACCGAGCTCCGTCGCTTCGAC-3'
N303Q	5'-CCGAACGCGGGCGGACCGAGCTCCGTCGCTTCGAC-3'
A305D	5'-CGCGGCAACATCGAGCTCCGTCGCTTCGAC-3'
A305F	5'-CGCGGCAACATCTTACGCTCCGTCGCTTCGAC-3'
A305G	5'-CGCGGCAACATCGGCGAGCTCCGTCGCTTCGAC-3'
A305I	5'-CGCGGCAACATCATGAGCTCCGTCGCTTCGAC-3'
A305L	5'-CGCGGCAACATCTGAGCTCCGTCGCTTCGAC-3'
A305R	5'-CGCGGCAACATCGGCGAGCTCCGTCGCTTCGAC-3'
A305S	5'-TCTCGAGCTCCGTCGCTTCGAGCGGCTGGCCGCTC-3'
A305T	5'-CGCGGCAACATCAGGCTCCGTCGCTTCGAC-3'

Products of the hexanoyl-CoA-primed reaction

Hexanoyl-CoA is structure (1) in Scheme 1. (i) 4-Hydroxy-6-pentyl-2-pyrone (3): retention time was 12.5 min on LC. LC-APCIMS: positive, *m/z* 183 (*M*+H)⁺; negative, *m/z* 181 (*M*–H)[–]. LC-APCIMS/MS (precursor ion at *m/z* 181): *m/z* 137 (*M*–CO₂–H)[–]. (ii) 4-Hydroxy-6-(2'-oxoheptyl)-2-pyrone (5): retention time was 11 min on LC. LC-APCIMS: positive, *m/z* 225 (*M*+H)⁺; negative, *m/z* 223 (*M*–H)[–]. LC-APCIMS/MS (precursor ion at *m/z* 223): *m/z* 179 (*M*–CO₂–H)[–], 125 (C₆H₅O₃)[–].

Products of the phenylacetyl-CoA (2)-primed reaction

(i) 6-Benzyl-4-hydroxy-2-pyrone (4): retention time was 9.5 min on LC. LC-APCIMS: positive, *m/z* 203 (*M*+H)⁺; negative, *m/z* 201 (*M*–H)[–]. (ii) 4-Hydroxy-6-(2'-oxo-3'-phenylpropyl)-2-pyrone (6): retention time was 11 min on LC. LC-APCIMS: positive, *m/z* 245 (*M*+H)⁺; negative, *m/z* 243 (*M*–H)[–].

Malonyl-CoA decarboxylation assay

The reaction mixture contained 51.8 µM [2-¹⁴C]malonyl-CoA, 100 mM Tris/HCl (pH 7.5) and 4.5 µg of wild-type or mutant

RppA in a total volume of 50 μ l. Reactions were incubated at 30 °C for 15 min and stopped by adding 1 μ l of 6 M HCl, and a portion (20 μ l) was applied to Merck no. 5744 TLC plate. The TLC plate was developed with a solvent system of isopropyl alcohol/water/28 %-satd ammonium hydroxide (80:5:15, by vol.). The R_f of acetyl-CoA was determined by using the authentic compound as a standard.

Determination of the kinetic parameters of Tyr²²⁴ and Ala³⁰⁵ mutant enzymes

The standard reaction, in a total volume of 100 μ l, contained 100 mM Tris/HCl (pH 7.5), 52.6 μ M [2-¹⁴C]malonyl-CoA (264000 d.p.m.), 5–50 μ M hexanoyl-CoA and 2.9 μ g of wild-type or mutant RppA. After the reaction mixture had been preincubated at 30 °C for 3 min, the reaction was initiated by adding the substrates and then continued for 3–30 min. The reaction was stopped with 20 μ l of 6 M HCl and the material in the mixture was extracted with 200 μ l of ethyl acetate. The organic layer was combined with non-labelled authentic products and dried by flushing with N₂ for HPLC analysis. After separation by HPLC, the amount of polyketides produced was quantified by means of [2-¹⁴C]malonyl-CoA incorporation.

RESULTS

Characterization of RppA from *S. antibioticus* and *S. lividans*

A large number of CHS genes have been isolated from plants and sequenced, providing a reliable consensus sequence that can be used to predict target amino acid residues for mutational analysis [13]. However, only two genes, one from *S. coelicolor* A3(2) (accession number CAC01488) and the other from *S. avermitilis* (BAB69299), encoding proteins highly similar to RppA were

deposited in the DNA databases. In the present study we isolated three genes, Sl-rppA from *S. lividans* TK21 and Sa-rppA and Sa-rppB from *S. antibioticus* IFO13271, by standard DNA manipulation, including colony hybridization with the *S. griseus* rppA gene as probe. The nucleotide sequences of the three cloned genes were determined and their deduced amino acid sequences were compared with those of CHS, PS and all available RppAs (Figure 1). The overall amino acid identity in alignment of Sl-RppA, Sa-RppA and Sa-RppB with *S. griseus* RppA is 70 %, 68 % and 72 % respectively.

The putative RppAs contained all of the signature sequences for CHS-related enzymes (Cys¹³⁸, Phe¹⁸⁸, His²⁷⁰ and Asn³⁰³) (Figure 1) [9], which suggested that the protein would catalyse the condensation of an acyl-CoA substrate with malonyl-CoA. To test this prediction, Sl-RppA, Sa-RppA and Sa-RppB were produced in *E. coli* as histidine-tagged proteins, purified and characterized. All of the recombinant proteins had the same mobility on SDS/PAGE, migrating with a monomer molecular size of approx. 40 kDa (results not shown). TLC and HPLC analysis of the reaction products revealed that these proteins catalysed THN formation solely from malonyl-CoA (results not shown), which clearly indicates that the three RppA homologues are THN synthases that belong to type III PKSs.

Activities of active-site mutants

The amino acid sequence alignment of all available RppAs and CHS (Figure 1) revealed the presence of four highly conserved residues; these are Cys¹³⁸, Phe¹⁸⁸, His²⁷⁰ and Asn³⁰³ in RppA from *S. griseus*. Structural and functional studies of CHS showed that Cys¹⁶⁴ serves as the nucleophile that initiates the reaction and as the attachment site for the enzyme-bound polyketide intermediates required for chalcone formation [4]. The imidazolium ion of His³⁰³ has a property to stabilize the thiolate

<i>S. griseus</i> RppA (Sg)	MATLCRPAIAVPEHVITMQQTDLDA-RETHAGH--PORDLVRLIQTNGVQTRHLV-----QPIEKTLAHPGFVRNQVYEAERKT	78
<i>S. antibioticus</i> RppA (Sa A)	MPRLCKPAMPAEHIITVEETLEFA-EKAHAGK--POLPLALRLIINTGVQKRHHV-----QPIEQTLRHPGFEEERNRIYEVESKK	78
<i>S. antibioticus</i> RppB (Sa B)	MATLCRPSVSVPEHVITMEETLELA-RSLHADH--POLPLALRLIINTGVQTRHHV-----QPIEETLKHGPFEEERNKVVYEAERKA	78
<i>S. lividans</i> RppA (Sl)	MATLCRPSVSVPEHVITMEETLELA-RRCHTDH--POLPLALRLIINTGVQTRHHV-----QPIEDTLRHPGFEDRNKVVYEAERKA	78
<i>S. coelicolor</i> RppA (Sc)	MATLCRPSVSVPEHVITMEETLELA-RRCHTDH--POLPLALRLIINTGVQTRHHV-----QPIEDTLRHPGFEDRNKVVYEAERKA	78
<i>S. avermitilis</i> RppA (Sav)	MATLCRPAVSVPEHVITMEETLELA-RSRHPDH--POLPLALRLIINTGVQTRHHV-----QPIEETLKHGPFEEERNHVVYEAERKA	78
CHS	MVSVSEIRKAQRAEPATILAIPTANPANCVEQSTYDFYFKITNSEHKTLEKFKFORMCDKSMIKRMYLTTEEILKENPNVCEYMAPSLDARDQMVVVEVPR	104
PS	MGYSDDVEIREAGRAQGLATILAIPTATPNCVAQADYADYFRVTRKSEHMDLKEKFKRICETAIKKRYLALTEDYLQENPTMCFMAPSLNARQDLVVTGVPM	109
138		
<i>Sg</i>	RVPEVRRALANAETEPSEIDLIVYVCTGFMMPSLTAWLINSMGFRPETROLPIAQLGCAAGGAAINRAHDFCVAYPDSNVLVSCFESLICYQ-PTDIGVGSLLSNG	186
<i>Sa A</i>	RTPEVIERALANAELRARDIDAIIVYVCTGFMMPSLTAWLINLKGFRPDTRQPIAQLGCAAGGAAINRAHDFCVAHGPSNVLVSCFESLICYQ-PEADDIGSLLSDG	186
<i>Sa B</i>	RVPAVVRALDDAEVLASDIDVIVYVCTGFMMPSLTAWLINEMDFSTTRQPIAQLGCAAGGAAINRAHDFCTAYPEANALIVACEFSLICYQ-PTDLGIGNLLSNG	186
<i>Sl</i>	RVPVAVRQALDDAELLATDIDVIVYVCTGFMMPSLTAWLINEMGFDSTTRQPIAQLGCAAGGAAINRAHDFCTAYPEANALIVACEFSLICYQ-PTDLGVSLLSNG	186
<i>Sc</i>	RVPVAVRQALDDAELLATDIDVIVYVCTGFMMPSLTAWLINEMGFDSTTRQPIAQLGCAAGGAAINRAHDFCTAYPEANALIVACEFSLICYQ-PTDLGVSLLSNG	186
<i>Sav</i>	RVPVAVRQALDDAELLATDIDVIVYVCTGFMMPSLTAYLINSMDFSDTRQPIAQLGCAAGGAAINRAHDFCTAYPEANALIVACEFSLICYQ-PTDLGVSLLSNG	186
CHS	LGKEAAVKAIKEWGPKSKITHLICTTGVDMPGADYQLTKLLGLRPYVRYMYQQGCFAGGTVLRLAKDLAENNKGRVLCVSEVAVLIFRGPSTHLDLSLVQA	213
PS	LGKEAAVKAIKEWGLPKSKITHLICTTGVDMPGADYQLVLLGLSPVRYMYLQQGCAAGGTVLRLAKDLAENNKGRVLCVSEVAVLIFRGPSTHLDLSLVQA	218
188		
<i>Sg</i>	LFGDALSAAVVRGGGTGMR-----LERNGSHLVPDTEWISAVVDTGHEFFOLDKRRVPGTMEPLAPVLLDVLH-LGWSVPNMDFFIIVHAGGPRILDDLCHFLDLPE	289
<i>Sa A</i>	LFGDAVAASAVVRGIGGGIE-----LERNASYLIPNTEDWISAVVDTGHEFFOLDKRRVPGTMEPLAPVLLRELAKDH-SWDAGKLDFFIIVHAGGPRILDDLAKFLEVDRE	289
<i>Sa B</i>	LFGDGIAASAVVRGGGGEGIA-----LERNGSYLIPKTEDWIMDVATGHEFFOLDKRRVPATMEPLAPALQELAGRH-GWDAADLDFFIIVHAGGPRILDDLKFLQVDPH	289
<i>Sl</i>	LFGDGIAASAVVRGGGTGVR-----LERNGSYLIPKTEDWIMDVKATGHEFFOLDKRRVPATMEPLAPALKEAGEH-GWDASDLDFIIVHAGGPRILDDLSTFLVDPH	289
<i>Sc</i>	LFGDGIAASAVVRGGGTGVR-----LERNGSYLIPKTEDWIMDVKATGHEFFOLDKRRVPATMEPLAPALKEAGEH-GWDASDLDFIIVHAGGPRILDDLSTFLVDPH	289
<i>Sav</i>	LFGDGIAASAVVRGGGTGIT-----LERNASYLIPKTEDWISVAVRATGHEFFOLDKRRVPGTMEPLAPALQELASQH-GWDASDLDFIIVHAGGPRILDDLKFLRVDPH	289
CHS	LFGDGAAALIVGSDPVEIEKPIFEMVWTAQTIAPDSEGAIDGHLREAGITFHLKLDVPGIVSKNITKALVEAFELGSDYNSIFWIAHGGPAILDQVEQRLAKPE	322
PS	LFGDGAAALIVGSGPHLAVERPIFIVSTDQITLPDTEKAMHLHREGGLTEQLHRDVFPLMVAKNIENAEKALSPLGITDWSVFWMVHGGRAILDQVERKLNKED	327
224		
<i>Sg</i>	MFRYSRATLTERGNIPSSVVF-----DALARLFDDGGAESAQGL-----IAGFGPGITAEVAVGSWAGEGLGADVGRDLELTAGVALSG	372
<i>Sa A</i>	VFRHWSLTSEYGNIPSAVVL-----EAARRLFEEDAPAGATGL-----IAGFGPGITAEAMALGTWAGSDRSAS	355
<i>Sa B</i>	AFRFSRATLTERGNIPSAVVL-----DALRRLFDAGGAPDRARGL-----LAGFGPGITAEMSLGRWRRTDET	353
<i>Sl</i>	AFRFSRATLTERGNIPSAVVL-----DALRRLFDAGGVEEGARGL-----LAGFGPGITAEMSLGCWQTADVRRGIQDVTAAARGVSRVRQA	374
<i>Sc</i>	AFRFSRATLTERGNIPSAVVL-----DALRRLFDAGGVEEGARGL-----LAGFGPGITAEMSLGCWQTADVRRGIQDVTAAARGVSRVRQA	374
<i>Sav</i>	AFRFSRATLTERGNIPSAVVL-----DALRRLFDAGGAEHAARGM-----LAGFGPGITAEMSLGRWRHTDEA	352
CHS	KMNATREVLSEYGNIPSAVLFILDEMRRKSTQNGLKTGEGLEWGVLFGFGPGITETVTVLRSVAI	389
PS	KLRSRHLVSEYGNIPSAVLFILDEVRRKSMAGKSTTGELDGCGLVFGFGPGITVETVTVLRSVRVTAAVANGN	402
303		
<i>Sg</i>	MFRYSRATLTERGNIPSSVVF-----DALARLFDDGGAESAQGL-----IAGFGPGITAEVAVGSWAGEGLGADVGRDLELTAGVALSG	372
<i>Sa A</i>	VFRHWSLTSEYGNIPSAVVL-----EAARRLFEEDAPAGATGL-----IAGFGPGITAEAMALGTWAGSDRSAS	355
<i>Sa B</i>	AFRFSRATLTERGNIPSAVVL-----DALRRLFDAGGAPDRARGL-----LAGFGPGITAEMSLGRWRRTDET	353
<i>Sl</i>	AFRFSRATLTERGNIPSAVVL-----DALRRLFDAGGVEEGARGL-----LAGFGPGITAEMSLGCWQTADVRRGIQDVTAAARGVSRVRQA	374
<i>Sc</i>	AFRFSRATLTERGNIPSAVVL-----DALRRLFDAGGVEEGARGL-----LAGFGPGITAEMSLGCWQTADVRRGIQDVTAAARGVSRVRQA	374
<i>Sav</i>	AFRFSRATLTERGNIPSAVVL-----DALRRLFDAGGAEHAARGM-----LAGFGPGITAEMSLGRWRHTDEA	352
CHS	KMNATREVLSEYGNIPSAVLFILDEMRRKSTQNGLKTGEGLEWGVLFGFGPGITETVTVLRSVAI	389
PS	KLRSRHLVSEYGNIPSAVLFILDEVRRKSMAGKSTTGELDGCGLVFGFGPGITVETVTVLRSVRVTAAVANGN	402

Figure 1 Amino acid sequence comparison of type III PKSs from *Streptomyces* and plants

Active-site residues are marked by ▼ (RppA numbering). The residues proposed for the determination of substrate specificity are highlighted with boxes. The CHS sequence is from *Medicago sativa* (accession number P30074) and the PS sequence is from *Gerbera hybrida* (CAA86219).

Figure 2 HPLC analysis of the products synthesized from [^{14}C]malonyl-CoA by wild-type (WT) and mutant RppAs

Under the conditions of the standard RppA assay, 51.4 μM [^{14}C]malonyl-CoA (88000 d.p.m.) was incubated for 1 h (C138A, F188A, H270A and N303A) or 15 min (wild-type and H270Q). HPLC gradient conditions: linear from 5% (v/v) solvent A [2% (v/v) acetic acid in acetonitrile] to 40% (v/v) solvent A in solvent B [2% (v/v) acetic acid in water] over 30 min, then 100% solvent A for 10 min. The flow rate was 0.8 ml/min. The enzyme used in each reaction is indicated at the top right of the chromatogram. Symbols connected by broken lines represent radioactivity measurements (right axis), and those connected by solid lines represent the UV absorbance detected at 280 nm (left axis).

anion of Cys¹⁶⁴ by ionic interaction. Both His³⁰³ and Asn³³⁶ of CHS interact with the incoming malonyl-CoA to facilitate the decarboxylation of the substrate and the formation of an acetyl-CoA carbanion by stabilizing the developing negative charge on the thioester carbonyl oxygen. In addition, Phe²¹⁵ could provide a non-polar environment that favours the decarboxylation of malonyl-CoA [4].

The roles of the conserved active-site residues in the reaction were investigated using site-directed mutants. The mutants were assayed for their ability to catalyse polyketide formation from [^{14}C]malonyl-CoA, and their product profiles were compared

Figure 3 TLC analysis of the decarboxylation of [^{14}C]malonyl-CoA by wild-type (WT) and mutant RppAs

The positions of malonyl-CoA (M-CoA; R_f 0.43) and acetyl-CoA (A-CoA; R_f 0.83) are indicated.

with that of the wild-type enzyme (Figure 2). The reaction of wild-type RppA with malonyl-CoA results in THN with a small amount of TAL, a derailment product that is formed as a result of fewer condensations of malonyl-CoA. Cys¹³⁸ mutants, Phe¹⁸⁸ mutants, H270A, N303D, N303H or N303Q did not form any polyketide (for alanine-substituted mutants, the results are shown in Figure 2). N303A (Figure 2) and H270N (results not shown) produced TAL and flaviolin, an auto-oxidized product of THN, at a reduced rate. H270Q retained high activity to catalyse polyketide formation, whereas the derailment products TAL and 6-acetyl-4-hydroxy-2-pyrone (tetraketide lactone) were observed in increased amounts (Figure 2).

Since Phe¹⁸⁸, His²⁷⁰ and Asn³⁰³ are thought to contribute to acetyl-CoA carbanion formation, we investigated the malonyl-CoA decarboxylation activity of mutants (Figure 3). Amino acid replacement at Phe¹⁸⁸ and His²⁷⁰ caused complete loss (F188A and F188S) or a marked decrease (F188Y, H270A and H270N) of activity towards malonyl-CoA, with the notable exception of H270Q, which was active at a rate comparable with that of the wild-type enzyme. Asn³⁰³ mutants were still able to decarboxylate malonyl-CoA, but with substantially diminished activity when compared with the wild-type, indicating that this residue is important, but not essential, for malonyl-CoA decarboxylation. Mutation of Cys¹³⁸ to Ser did not influence the decarboxylation of malonyl-CoA. However, mutants C138A and C138Q lost the ability to decarboxylate malonyl-CoA, presumably because of

Figure 4 TLC analysis of products synthesized from various starter substrates and [^{14}C]malonyl-CoA by wild-type (WT) and mutant RppAs

The total volume of the reaction was 100 μl . For the extender substrate, 51.8 μM [^{14}C]malonyl-CoA (88000 d.p.m.) was incubated for 15 min. Substrate(s) used were 50 μM hexanoyl-CoA plus 51.8 μM [^{14}C]malonyl-CoA (A), 50 μM phenylacetyl-CoA plus 51.8 μM [^{14}C]malonyl-CoA (B), and 51.8 μM [^{14}C]malonyl-CoA (C). TLC was developed in acetonitrile/water/acetic acid (50:50:1, by vol.). The positions of THN (R_f 0.71), tetraketide pyrones (5, R_f 0.43; 6, R_f 0.52; see Scheme 1D) and triketide pyrones (3, R_f 0.34; 4, R_f 0.46) are indicated.

Figure 5 HPLC analysis of products synthesized from various starter substrates and [^{14}C]malonyl-CoA by wild-type (WT) and mutant RppAs

Under the conditions for the standard RppA assay, $51.4 \mu\text{M}$ [^{14}C]malonyl-CoA (88 000 d.p.m.) and hexanoyl-CoA (left panels) or phenylacetyl-CoA (right panels) were incubated for 15 min. HPLC gradient conditions: linear from 20% (v/v) solvent A [2% (v/v) acetic acid in acetonitrile] to 80% (v/v) solvent A in solvent B [2% (v/v) acetic acid in water] over 20 min, then 100% solvent A for 5 min; the flow rate was 1.0 ml/min. The enzyme used in each reaction is indicated at the top right of the chromatogram. Symbols connected by broken lines represent radioactivity measurements (right axes), and those connected by solid lines represent the UV absorbance detected at 280 nm (left axes). The existence of the relatively small peaks indicated by arrows was further confirmed by selected ion monitoring at the molecular mass of corresponding polyketides using LC-APCIMS.

the instability of these two mutants, which accumulated predominantly as an insoluble form in *E. coli* cells.

Screening for residues that determine the substrate specificity of RppA

The crystal structure of alfalfa CHS2 was used to predict the residues in the active-site cavity of RppA. Jez et al. [14] proposed that residues Thr¹³², Ser¹³³, Thr¹⁹⁴, Thr¹⁹⁶, Ile²⁵⁴, Gly²⁵⁶, Leu²⁶³, Phe²⁶⁵ and Ser³³⁸ determine the substrate specificity of CHS and the size of the product. In RppA, these residues are replaced by Cys¹⁰⁶, Thr¹⁰⁷, Cys¹⁶⁸, Cys¹⁷¹, Ile²²², Tyr²²⁴, Phe²³¹, Phe²³³ and Ala³⁰⁵ respectively, which are completely conserved among the RppA sequences (Figure 1). We therefore chose these nine residues for mutagenesis analysis to determine their contribution

to the starter substrate selectivity of RppA. Since RppA accepts both hydrophilic malonyl-CoA and hydrophobic hexanoyl-CoA as a starter substrate, specific residues seem to interact with each substrate. Therefore some of the mutants generated may be inactive only in the malonyl-CoA-primed reaction. We tested the ability of the mutants to accept as a starter substrate not only malonyl-CoA, but also hexanoyl-CoA and phenylacetyl-CoA. The selectivity of the starter substrate was assessed by TLC using [^{14}C]malonyl-CoA as the extender substrate (Figure 4). When incubated with hexanoyl-CoA and malonyl-CoA, wild-type RppA catalysed the formation of tetra- and tri-ketide pyrones besides THN; phenylacetyl-CoA was a poor substrate of wild-type RppA, as previously observed [11]. Mutants C106A, T107A, C168A, C171A, C171T, I222F, F231A and F233A did not form polyketides, probably because replacement of the amino acid

Table 2 Steady-state kinetic parameters of wild-type and mutant RppAs

Results are means $\pm$ S.D. ($n = 3$). All k_{cat} and apparent K_{m} values were determined using a Hanes plot.

Enzyme	Triketide (3)			Tetraketide (5)			Hexanoyl-CoA			3/5
	$10^3 \times k_{\text{cat}}$ (min^{-1})	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \cdot \text{M}^{-1}$)	$10^3 \times k_{\text{cat}}$ (min^{-1})	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \cdot \text{M}^{-1}$)	$10^3 \times k_{\text{cat}}$ (min^{-1})	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \cdot \text{M}^{-1}$)	
Wild type	25.1 $\pm$ 1.2	15.0 $\pm$ 2.9	27 782	11.1 $\pm$ 0.7	10.2 $\pm$ 2.4	18 101	36.0 $\pm$ 1.9	13.2 $\pm$ 2.7	45 409	1.5
Y224G	39.4 $\pm$ 10.1	56.8 $\pm$ 15.3	11 553	1.6 $\pm$ 0.0	19.6 $\pm$ 2.6	1355	39.2 $\pm$ 9.5	50.8 $\pm$ 12.9	12 854	8.5
Y224L	83.4 $\pm$ 3.5	14.9 $\pm$ 1.1	93 064	4.4 $\pm$ 0.1	7.6 $\pm$ 0.6	9699	87.0 $\pm$ 3.4	14.3 $\pm$ 1.1	101 735	9.6
A305S	379 $\pm$ 110	16.0 $\pm$ 5.9	563 195	157 $\pm$ 37	14.5 $\pm$ 4.8	180 571	540 $\pm$ 152	15.7 $\pm$ 5.7	574 668	3.1
A305I	124 $\pm$ 13	8.7 $\pm$ 0.9	238 863	—*	—*	—*	124 $\pm$ 13	8.7 $\pm$ 0.9	238 863	—*

* A305I synthesized only a triketide (**3**).

residues diminished appropriate binding of the substrates. C171L, I222A, I222M, A305G and A305S were active in hexanoyl-CoA-, phenylacetyl-CoA- and malonyl-CoA-primed reactions with a similar or reduced rate when compared with that of the wild-type enzyme, suggesting that these residues are not crucial for determination of the starter substrate selectivity of RppA. Exceptionally, mutant A305I preferentially catalysed hexanoyl-CoA- and phenylacetyl-CoA-primed reactions, and each of the reactions gave one major product with little THN, which was further confirmed by HPLC analysis. These findings suggest that the replacement of Ala³⁰⁵ alters the size of the polyketide produced. On the other hand, Tyr²²⁴ mutants substituted with Ala, Gly or Leu showed no detectable activity in the malonyl-CoA-primed reaction, but were still active in hexanoyl-CoA- and phenylacetyl-CoA-primed reactions, implying that the substitution of Tyr²²⁴ results in the complete loss of selectivity of malonyl-CoA as the starter substrate. We therefore generated and characterized more mutants with amino acid replacements at Ala³⁰⁵ and Tyr²²⁴, as shown below.

Activities of Tyr²²⁴ and Ala³⁰⁵ mutants

To investigate the influence of Tyr²²⁴ with regard to starter substrate selectivity, we generated Tyr²²⁴ mutants using Cys, His, Phe, Ser and Trp replacements at this position, in addition to the Y224A, Y224G and Y224L mutants studied previously. Y224H and Y224S were produced predominantly as an insoluble form in *E. coli*, which prevented further study. The product profiles of the mutants were analysed by HPLC using [¹⁴C]malonyl-CoA as the extender substrate, and the products were identified by LC-APCIMS analysis (Figure 5). Consistent with the results of TLC analysis, mutants Y224G, Y224L and Y224C were found not to produce THN from malonyl-CoA. However, these mutants incorporated hexanoyl-CoA and phenylacetyl-CoA as a starter substrate, both yielding a tetraketide α -pyrone (**5**, **6**) and triketide α -pyrone (**3**, **4**) (Scheme 1D; Figure 5). Interestingly, those mutants with the aromatic amino acids Trp or Phe at the Tyr²²⁴ position retained their ability to accept malonyl-CoA as a starter substrate, which is indicated by the presence of THN in the products.

Mutants A305D, A305F, A305L, A305R and A305T were also generated, in addition to A305G, A305S and A305I, to test the possibility that Ala³⁰⁵ contributes to the determination of product size. A305D, A305F, A305L and A305R were not analysed because they were produced as an inclusion body in *E. coli* cells. Although A305S and A305T showed a similar product profile, yielding a tetra- and tri-ketides in both the hexanoyl-CoA- and phenylacetyl-CoA-primed reactions, as in the case of

wild-type (results shown in Figure 5 for A305S), A305I produced only triketide pyrones (**3**, **4**; Scheme 1D).

Steady-state kinetic analysis of Tyr²²⁴ and Ala³⁰⁵ mutants with hexanoyl-CoA as the starter substrate

Tyr²²⁴ and Ala³⁰⁵ of RppA correspond to Gly²⁵⁶ and Ser³³⁸ respectively in CHS, and these residues of CHS are believed to form the coumaroyl-binding and cyclization pockets in chalcone formation [14,15]. In PS, the bulky residues Leu²⁶¹ and Ile³⁴³ correspond to Gly²⁵⁶ and Ser³³⁸ respectively of CHS. These replacements would reduce the volume of both of the pockets in PS, so that the number of elongations of an acetate unit is restricted to two, as implied from the three-dimensional structure of PS [14]. Simultaneous amino acid replacement at three positions in CHS (Thr¹⁹⁷, Gly²⁵⁶ and Ser³³⁸) by the 'PS-type' residues functionally converts CHS into PS, due to the shrinking of the active-site cavity volume, which governs the choice of starter substrate and controls the final length of the polyketide [14]. We determined the steady-state kinetic parameters of the reactions of wild-type and mutant RppAs using hexanoyl-CoA as a starter substrate (Table 2). In terms of the product ratio of tetraketide versus triketide, A305S was kinetically similar to wild-type RppA, whereas A305I produced a triketide exclusively. This suggests that the replacement of Ala³⁰⁵ with the bulky 'PS-type' Ile results in a decrease in the volume of the cyclization pockets in RppA, as a result of which the size of the polyketide formed is restricted.

All of the mutants, except for Y224G, consumed hexanoyl-CoA with 2–12-fold improvements in catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$) compared with the reaction catalysed by the wild-type enzyme, presumably because of less competition in occupying Cys¹³⁸ against malonyl-CoA. Although monopoly of the catalytic centre by hexanoyl-CoA would also occur in the reaction of Y224G, this mutant showed reduced activity towards hexanoyl-CoA. We suppose that hexanoyl-CoA takes up different conformations in the mutants.

DISCUSSION

On the basis of the reaction mechanism of CHS, the initial step of THN formation by RppA would involve the loading of malonyl-CoA on to the thiol group of Cys¹³⁸, resulting in formation of enzyme-bound malonate via the thioester bond. Apparently, this transfer reaction is a distinct step in the determination of the starter substrate in type III PKSs. RppA mutants of Tyr²²⁴, such as Y224A, Y224C, Y224G and Y224L, cannot synthesize polyketides solely from malonyl-CoA (Figure

5), whereas these mutants are active in the hexanoyl-CoA- and phenylacetyl-CoA-primed reactions, implying that they accept malonyl-CoA as the elongation substrate, but not as the starter substrate. This suggests the existence of two distinct substrate-binding pockets for starter and elongation substrates in the active-site pocket of RppA. It is possible that some malfunction occurs at the loading step of malonyl-CoA on to Cys¹³⁸ in the initial reaction in the Tyr²²⁴ mutants. In contrast, mutants having an aromatic amino acid, Trp or Phe, at this position are still active with malonyl-CoA as starter substrate, and produce THN. These results clearly indicate that the aromatic ring of Tyr²²⁴ is essential for recruiting malonyl-CoA to the thiol group of Cys¹³⁸. The electronic X-ray crystallographic structure of CHS has revealed that Gly²⁵⁶, the residue corresponding to Tyr²²⁴ of RppA, is present at the surface of the active-site pocket [9]. The substitution of Gly²⁵⁶ of CHS with a bulky residue reduces the size of the active-site cavity, which causes alterations in both the number of condensation reactions for chain extension and the conformational changes of the triketide and tetraketide intermediates during the cyclization reaction [16]. Considering the high amino acid sequence conservation between RppA and CHS, we believe that the overall backbone architecture of the active-site cavity of RppA is similar to that of CHS, and therefore that Tyr²²⁴ of RppA lies in the active-site pocket, permitting direct interaction with malonyl-CoA. The structural similarity without sequence identity between CHS and ketoacyl acyl carrier protein synthase III [17,18], which initiates fatty acid biosynthesis and is believed to have a common ancestor with type III PKSs, supports this idea. A minor difference in architecture leads to the mechanistic difference; Tyr²²⁴ of RppA contributes to starter substrate selectivity, while Gly²⁵⁶ of CHS determines the size of the active-site cavity and the terminal chain lengths of resultant polyketides. It is worth noting that Tyr²²⁴ of RppA is also conserved in DpgA, a type III PKS involved in the biosynthesis of balhimycin (a vancomycin derivative) that also utilizes malonyl-CoA as a starter substrate [6].

The substitution of Ala³⁰⁵ with the bulky 'PS-type' Ile causes the mutant RppA to condensate an acetate unit twice in the hexanoyl-CoA- and phenylacetyl-CoA-primed reactions. However, a small amount of THN was detected in the same reaction mixture, indicating that the mutation neither eliminates the ability to adopt malonyl-CoA as a starter substrate nor affects the size of the product of the malonyl-CoA-primed reaction. It appears that, in the active-site pocket, the conformation of the nascent chain of the malonyl-CoA-primed polyketide might differ from that of the hexanoyl-CoA-primed reaction. No observed influence on the size of the products by amino acid replacement at position Ala³⁰⁵ can be explained in terms of the conformational difference in the nascent chains. Recently, Jez et al. [16] reported that a decrease in the size of the active-site cavity of CHS, generated by substitution of Gly²⁵⁶ by bulkier residues, decreased the number of extensions in the aromatic-starter-primed reaction, but did not affect that in the hexanoyl-CoA-primed reaction. Given that a single PKS uses various starter substrates and synthesizes products with a wide variety in size [11], it is likely that the structure of the starter substrate moiety, which correlates with the conformation of the nascent polyketide chain accommodated in the active-site cavity, contributes more towards the determination of the length of the polyketide produced than does the physical capacity determined by the volume of the active-site pocket.

The results of mutagenesis analysis of the active-site residues of RppA are almost identical with those for CHS. For example, of the His³⁰³ mutants of CHS, only H303Q retains high activity for chalcone formation and decarboxylation of malonyl-CoA when

compared with wild-type CHS [4]. A minor difference was observed in the reaction of the Phe¹⁸⁸ mutants; the three Phe¹⁸⁸ mutants of RppA cannot efficiently decarboxylate malonyl-CoA as the extender, whereas the Phe²¹⁵ mutants of CHS were capable of forming polyketides. Our findings confirm that the mechanism of decarboxylation of malonyl-CoA proposed for plant CHSs also works for RppA, a bacterial type III PKS. The diversity of reactions catalysed by the type III PKSs is ascribed to changes in certain residues during evolution without changing the principal machinery of the active site. RppA mutant H270Q retains decarboxylation activity at a near wild-type level (Figure 3), but increased amounts of derailment products were observed (Figure 2). Perhaps the impaired catalytic machinery cannot provide a sufficient amount of acetyl-CoA carbanion, which is required for chain extension, and derailment occurs before the elongation reaction has been completed.

RppA mutant H270Q produces large amounts of tri- and tetra-ketide derailment products, but never yields any pentaketide derailment product. We had expected the production of pentaketide polyketides, such as α -acetylsorsellinic acid or its isomer, as derailment products. The predicted mechanism for the final ring closure of the pentaketide intermediate to yield THN, i.e. Claisen condensation accompanied by decarboxylation of the malonate moiety incorporated as the starter substrate [19], would resemble the decarboxylative chain elongation reaction. We had therefore thought that this reaction would be catalysed by the same machinery used for generating an acetyl-CoA carbanion. However, the absence of any pentaketide derailment product in the reaction products of H270Q, despite the presence of THN, suggests that the impaired machinery is still active in catalysing ring closure, although the possibility of a spontaneous nucleophilic attack of the active methylene and subsequent decarboxylation during aromatization cannot be excluded. Alternatively, an unidentified residue(s) may be responsible for the final ring closure.

It is surprising that a single amino acid replacement results in a change in the substrate specificity of RppA. Identification of the residues critical for substrate specificity and amino acid replacement at these positions could facilitate the tailoring of this catalyst for the production of novel polyketides. As demonstrated in the present study, for example, Tyr²²⁴ mutants uses phenylacetyl-CoA, instead of malonyl-CoA, as a starter substrate, since the mutation diminishes the competition against malonyl-CoA.

N.F. was supported by the Japan Society for the Promotion of Science. This work was supported by the Bio Design Program of the Ministry of Agriculture, Forestry, and Fisheries of Japan (BDP-02-VI-2-5), and by the Industrial Technology Research Grant Program in 00 Program of the New Energy and Industrial Technology Development Organization of Japan (00A03004).

REFERENCES

- Moore, B. S. and Hopke, J. N. (2001) Discovery of a new bacterial polyketide biosynthetic pathway. *Chem. Biochem.* **2**, 35–38.
- Schröder, J. (1997) A family of plant-specific polyketide synthases: facts and predictions. *Trends Plant Sci.* **2**, 373–378.
- Funa, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y. and Horinouchi, S. (1999) A new pathway for polyketide synthesis in microorganisms. *Nature (London)* **400**, 897–899.
- 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.
- Eckermann, S., Schröder, G., Schmidt, 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 (London)* **396**, 387–390.

- 6 Pfeifer, V., Nicholson, G. J., Ries, J., Recktenwald, J., Schefer, A. B., Shawky, R. M., Schröder, J., Wohlleben, W. and Pelzer, S. (2001) A polyketide synthase in glycopeptide biosynthesis. *J. Biol. Chem.* **276**, 38370–38377
- 7 Lanz, T., Tropsch, S., Marner, F.-J., Schröder, J. and Schröder, G. (1991) The role of cysteines in polyketide synthases. *J. Biol. Chem.* **266**, 9971–9976
- 8 Schröder, J. (1999) Probing plant polyketide biosynthesis. *Nat. Struct. Biol.* **6**, 714–716
- 9 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
- 10 Schütz, R., Heller, W. and Hahlbrock, K. (1983) Substrate specificity of chalcone synthase from *Petroselinum hortense*. *J. Biol. Chem.* **258**, 6730–6734
- 11 Funa, N., Ohnishi, Y., Ebizuka, Y. and Horinouchi, S. (2002) Properties and substrate specificity of RppA, a chalcone synthase-related polyketide synthase in *Streptomyces griseus*. *J. Biol. Chem.* **277**, 4628–4635
- 12 Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- 13 Schröder, G. and Schröder, J. (1992) A single change of histidine to glutamine alters the substrate preference of a stilbene synthase. *J. Biol. Chem.* **267**, 20558–20560
- 14 Jez, J. M., Austin, M. B., Ferrer, J.-L., Bowman, M. E., Schröder, J. and Noel, J. P. (2000) Structural control of polyketide formation in plant-specific polyketide synthases. *Chem. Biol.* **40**, 1–12
- 15 Jez, J. M., Ferrer, J.-L., Bowman, M. E., Austin, M. B., Schröder, J., Dixon, R. A. and Noel, J. P. (2001) Structure and mechanism of chalcone synthase-like polyketide synthases. *J. Ind. Microbiol. Biotechnol.* **27**, 393–398
- 16 Jez, J. M., Bowman, M. E. and Noel, J. P. (2001) Structure-guided programming of polyketide chain-length determination in chalcone synthase. *Biochemistry* **40**, 14829–14838
- 17 Qiu, X., Janson, C. A., Konstantinidis, A. K., Nwagwu, S., Silverman, C., Smith, W. W., Khandekar, S., Lonsdale, J. and Abdel-Meguid, S. S. (1999) Crystal structure of β -ketoacyl-acyl carrier protein synthase III. *J. Biol. Chem.* **274**, 36465–36471
- 18 Davies, C., Heath, R. J., White, S. W. and Rock, C. O. (2000) The 1.8 Å crystal structure and active-site architecture of β -ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure* **8**, 185–195
- 19 Fujii, I., Mori, Y., Watanabe, A., Kubo, Y., Tsuji, G. and Ebizuka, Y. (2000) Enzymatic synthesis of 1,3,6,8-tetrahydroxynaphthalene solely from malonyl-coenzyme A by a fungal iterative type I polyketide synthase PKS1. *Biochemistry* **39**, 8853–8858

Received 20 June 2002/24 July 2002; accepted 24 July 2002

Published as BJ Immediate Publication 24 July 2002, DOI 10.1042/BJ20020953