Supporting Information

Morita et al. 10.1073/pnas.0909982107

SI Text

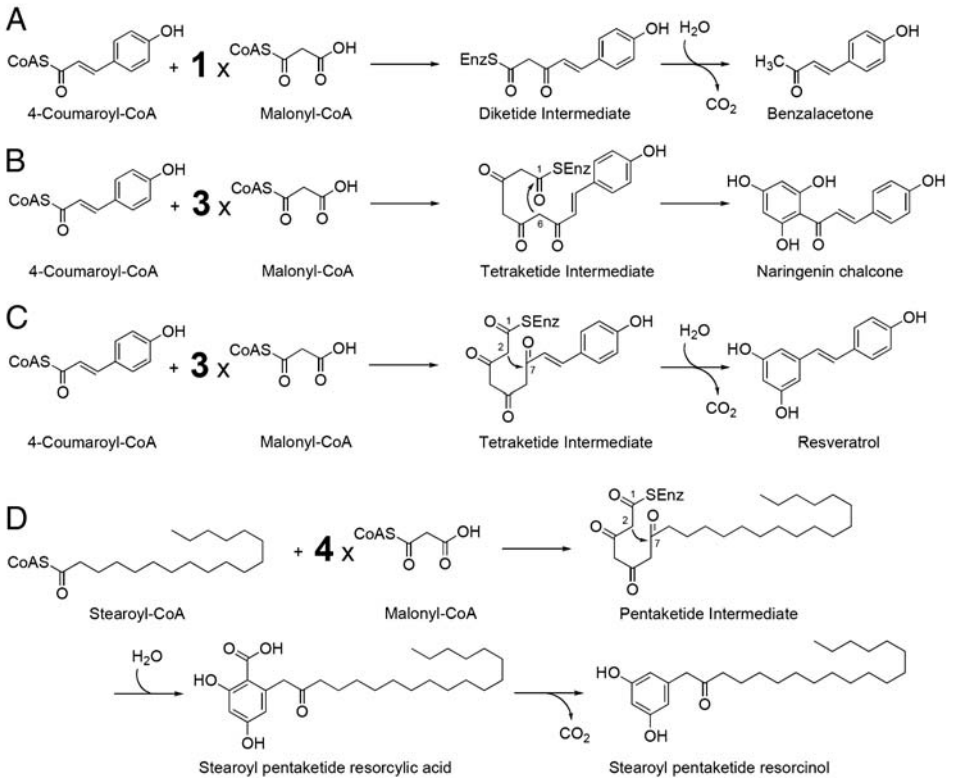

Fig. S1. Proposed synthetic mechanisms of polyketide. **(A)** benzalacetone from 4-coumaroyl-CoA and one molecule of malonyl-CoA by BAS; **(B)** naringenin chalcone from 4-coumaroyl-CoA and three molecules of malonyl-CoA by CHS; **(C)** resveratrol from and three molecules of malonyl-CoA by STS; and, **(D)** stearoyl pentaketide resorcylic acid and resorcinol from stearoyl-CoA and four molecules of malonyl-CoA by ORAS.

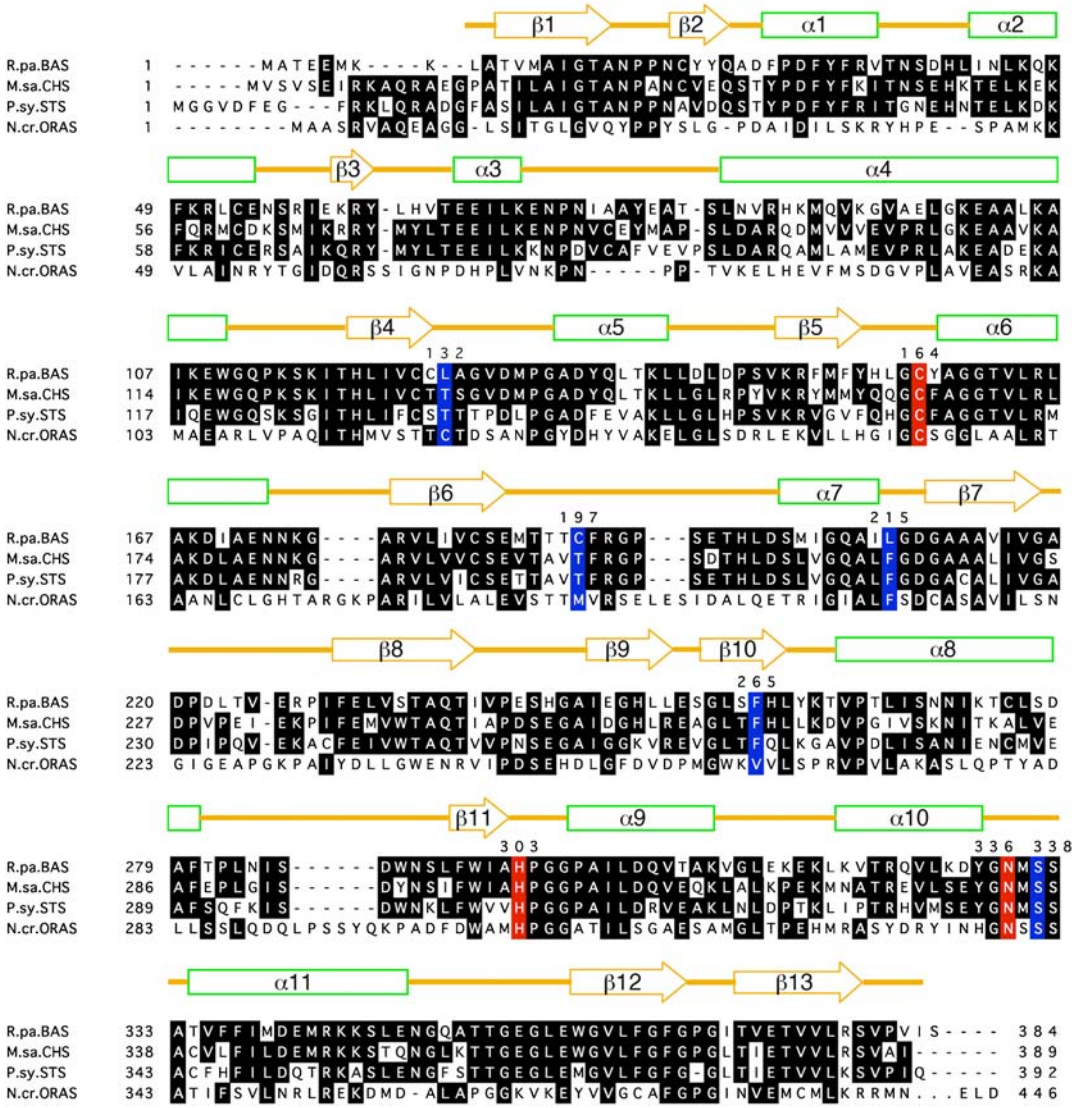

Fig. S2. Comparison of the primary sequence of BAS and other type III PKSs. The secondary structures of BAS are also delineated: α-helices (green rectangles); β-strands (orange arrows); and loops (orange, bold lines), which are diagrammed. The catalytic triad Cys-His-Asn residues are highlighted in red. The residues thought to be crucial for the functional diversity of type III PKSs are colored in blue. Abbreviations (GeneBank accession numbers): R.pa.BAS, *Rheum palmatum* BAS (AAK82824); M.sa.CHS, *M. sativa* CHS (P30074); P.sy. STS, *Pinus sylvestris* STS (AAB24341).

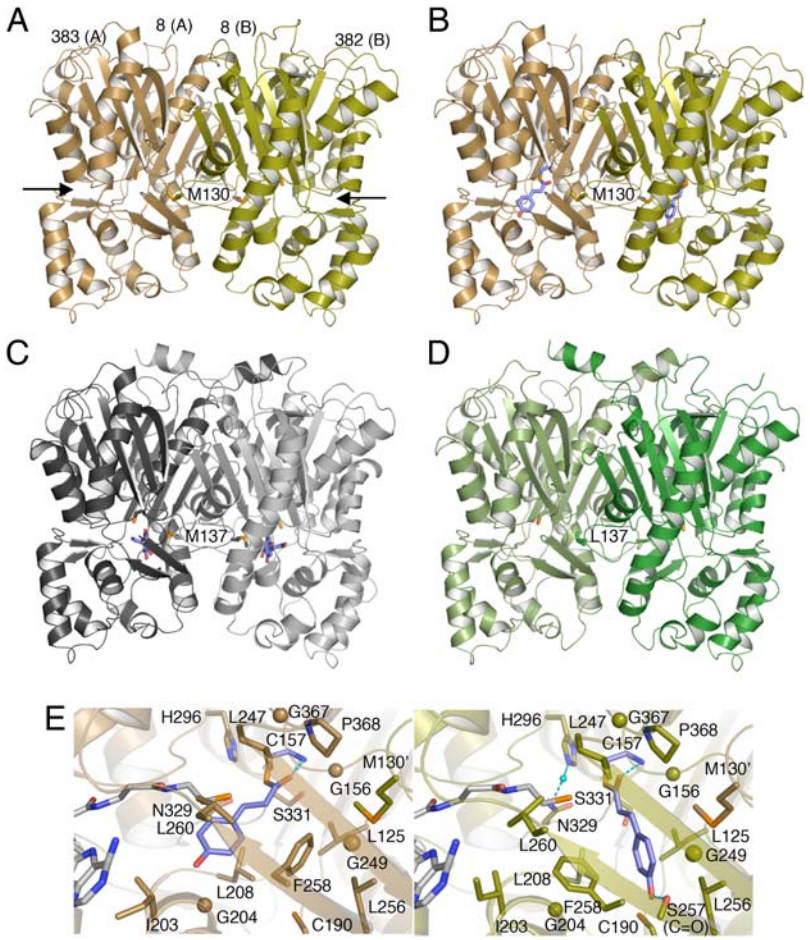

Fig. S3. Schematic representations of: (*A*) BAS_apo structure; (*B*) BAS complexed with the monoketide intermediate (4-coumaroyl thioester); (*C*) *M. sativa* CHS complexed with naringenin (PDB entry 1CGK); and, (*D*) *P. sylvestris* STS (PDB entry 1U0U). Arrows indicate the substrate entrance in each monomer. The catalytic cysteine and methionine (or leucine), which form a partial wall of the active-site cavity of another monomer, are indicated in stick model. The bound monoketide intermediate in BAS and naringenin in *M. sativa* CHS are highlighted as blue stick model. (*E*) Comparison of the active-site in (*left*) monomer A and (*right*) monomer B of the BAS structure complexed with the monoketide intermediate. A CoA-SH molecule bound in the *M. sativa* CHS structure (PDB entry 1BQ6) was also shown as a gray stick model to indicate the CoA-binding tunnel.

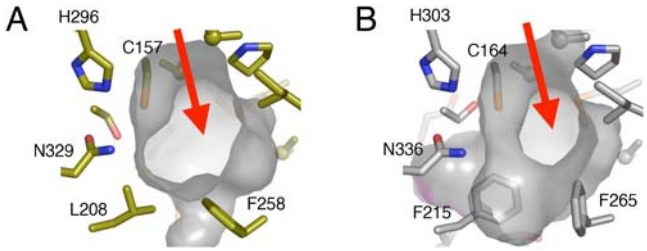

Fig. S4. Comparison of the active-site entrance of (*A*) BAS with that of (*B*) *M. sativa* CHS. Arrows indicate the substrate entrance in each monomer.

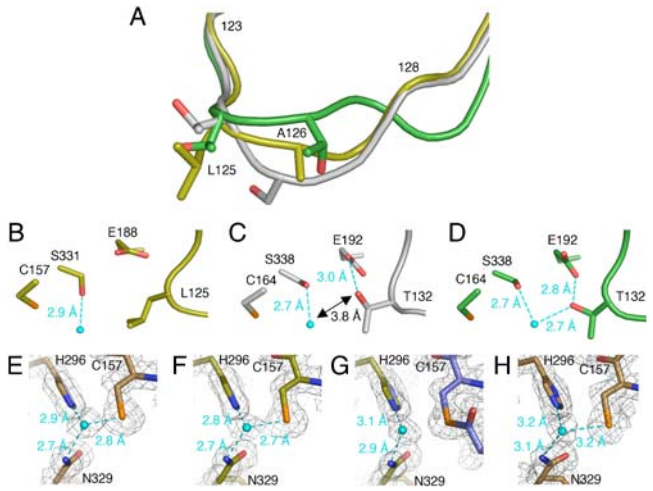

Fig. S5. Close-up view of the putative nucleophilic water molecule. (*A*) Superimposition of the loop between residues 123–128 of BAS with equivalent loops of *M. sativa* CHS and *P. sylvestris* STS. Only BAS residue numbers are marked. Close-up view of the residues and the water molecule in (*B*) BAS and (*C*) *M. sativa* CHS equivalent to those involving the aldol switch in (*D*) *P. sylvestris* STS. The $2F_o$-F_c density map of the catalytic triad and the putative nucleophilic water molecule in (*E*) monomer A, and (*F*) monomer B of wild-type BAS apo structure, (*G*) monomer B of BAS structure complexed with the monoketide intermediate, and (*H*) monomer A of the BAS I207L/L208F mutant structure, countered at 1.0 σ. The water molecules and the hydrogen bonds are indicated as light-blue sphere and dotted lines.

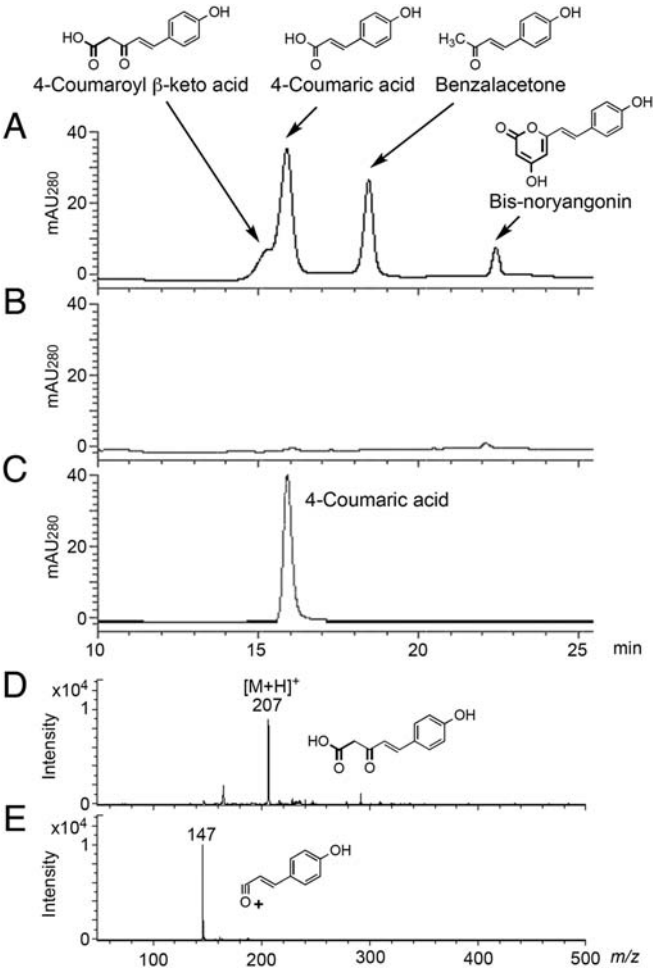

Fig. S6. HPLC elution profiles of the enzyme reaction products with 4-coumaroyl-CoA and malonyl-CoA: (*A*) wild-type BAS; (*B*) boiled wild-type BAS; (*C*) HPLC elution profile of authentic 4-coumaric acid; (*D*) MS; and, (*E*) MS/MS (precursor ion at *m*/*z* 207) spectra of 4-coumaroyl diketide β-keto acid produced by *R. palmatum* BAS.

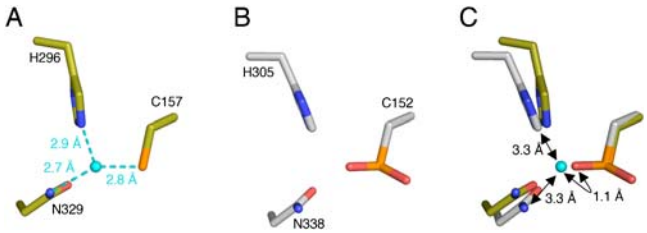

Fig. S7. Comparison of the catalytic triad of (*A*) *R. palmatum* BAS; and, (*B*) *N. crassa* ORAS. (*C*) Superimposition of the catalytic triad.

Table S1. Data collection and refinement statistics

Data collection	BAS	BAS + intermediate	I207L/L208F
Space group	$P2_1$	$P2_1$	$P2_1$
Unit-cell			
a, b, c (Å)	54.6, 89.6, 81.1	54.6, 89.9, 81.0	74.1, 89.2, 70.8
α,β,γ (°)	90.0, 100.5, 90.0	90.0, 100.5, 90.0	90.0, 95.7, 90.0
Resolution (Å)	30.0–1.8 (1.86–1.80)	30.0–1.6 (1.66–1.60)	30.0–1.8 (1.86–1.80)
Unique reflections	69,782	100,844	74,255
Redundancy	3.8 (3.8)	3.7 (3.6)	3.7 (3.5)
Completeness (%)	98.3 (97.3)	99.8 (99.3)	100 (99.9)
$<I/(\sigma I)>$	25.4 (7.5)	32.1 (8.3)	19.7 (4.4)
R_{sym} (%)*	9.6 (26.7)	7.1 (22.9)	7.3 (27.5)
Refinement			
Resolution (Å)	1.8	1.6	1.8
R_{cryst}/R_{free} (%)†	18.4/20.6	19.5/21.5	18.4/20.4
Number atoms			
Protein	5,796	5,796	5,802
Water	459	554	456
Ligand	-	22	-
B-factors (Å^2)			
Protein	12.8	13.0	12.9
Water	17.2	19.6	17.4
Ligand	-	21.4	-
r.m.s deviations			
Bond lengths (Å)	0.005	0.011	0.005
Bond angles (°)	1.2	1.5	1.2

Values in parentheses are for the highest resolution shell.

*$R_{sym} = S_h S_i |I(h)_i - < I(h) > | / S_h S_i I(h)_i$, where $I(h)$ is the intensity of refraction h, S_h is the sum over all reflections and S_i is the sum over i measurements of reflection h.

†R_{free} was calculated with 5% of data excluded from refinement.