

Structure of polyhydroxyalkanoate (PHA) synthase PhaC from *Chromobacterium* sp. USM2, producing biodegradable plastics

Min Fey Chek¹, Sun-Yong Kim¹, Tomoyuki Mori¹, Hasni Arsad², Mohd. Razip Samian³, Kumar Sudesh³ and Toshio Hakoshima¹

¹ Structural Biology Laboratory, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan

²Advanced Medical and Dental Institute, Universiti Sains Malaysia, Bertam 13200, Kepala Batas, Penang, Malaysia

³ School of Biological Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia

* Corresponding author:

Supplementary Table 1. Crystallographic statistics of the catalytic domain of PhaC_{Cs}-CAT		
Crystallographic Analysis Statistics		
Crystal form	Native	SeMet (SAD) ^a
Space group	<i>P6₁</i>	<i>P6₂</i>
Unit cell		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	117.33, 117.33, 105.91	117.15, 117.15, 52.69
α , β , γ (°)	90, 90, 120	90, 90, 120
Wavelength (Å)	0.90000	0.97906 (Se peak)
Resolution range ^b (Å)	50.00 - 1.48 (1.53 - 1.48)	50.00 - 2.50 (2.54 - 2.50)
Reflections ^c		
Oscillation range (°)	360	360
Measured/Unique	1535,125/137,720	304,374/14,563
Multiplicity	11.2 (11.2)	20.9 (22.5)
Mosaicity (°)	0.17 – 0.34	0.44 - 0.73
<i>I</i> / <i>s</i> (<i>I</i>)	58.6 (6.2)	73.5 (15.5)
<i>R</i> _{merge} (%)	7.5 (64.4)	12.2 (76.4)
Completeness (%)	99.8 (100.0)	100.0 (100.0)
Number of heavy atoms ^d	-	5
FOM (acentric/centric) ^e	-	0.365/0.161
Refinement Statistics		
<i>R</i> _{work} / <i>R</i> _{free} (%) ^f	12.11/15.74	
Number of atoms	6968	
Protein molecules	2 (754 residues)	
Water molecules	1050	
Average B-factors (Å ²)		
Protein	20.1	
Water molecules	33.5	
R.m.s.d. from ideal values		
Bonds (Å)/ angles (°)	0.019/1.603	
Ramachandran plot (%)		
Favored	97.0	
Allowed	2.7	
Outliers	0.4	

^a The SeMet derivative was treated with α -chymotrypsin.

^b Values in parentheses are for the highest-resolution shell.

^c collected at SPring-8 beamline BL44XU with a MX300HE detector (each 1° oscillation) at 100 K.

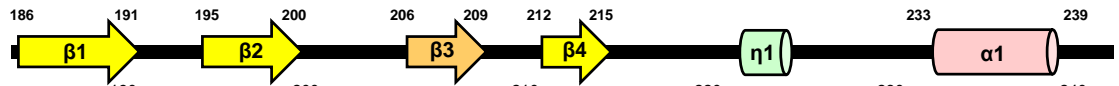
^d Heavy atoms were searched with the program, SHELX C/D.

^e The figure of merit (FOM) was calculated with the program, SHARP/autoSHARP.

^f *R*_{free} was calculated on a random 5 % reflections of the data.

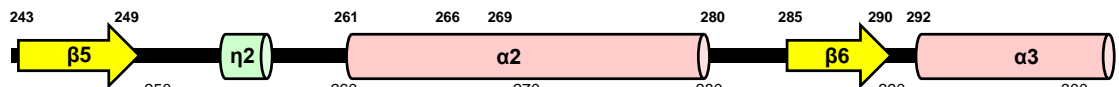
130 140 150 160 170 180

PhaC_{Cs} PSNFMLTNPDDVVKRAIETQGSESLVEGMKNMEDIQKGF - - - HISMSDESKFQIGKNLVVT
PhaC_{Cn} PANFLATNPEAQRLLIESGGESLRAGVRNMEDLTRG - - - KISQTDSESAFEVGRNVAVT
PhaC_{Ap} PSNF L A T N P E L L K L T L E S D G Q N L V R G L A L L A E D L E R S A D Q L N I R L T D E S A F E L G R D L A L T
PhaC1_{P3613} PTNSA - ANPAAVKRFFETGGKSLLDGLTHLAKDLVNNGG - - MPSQVDMGAFEVGKSLGTT
PhaC1_{P3455} PSNSM - ANPAAVKRFFETGGKSLLDGLSHLAKDMVHNGG - - MPSQVNMEAFEVGNLATT
PhaC2_{P3613} PSNTL - LNPLAIKELFNSGGNSLVRGLSHLFDLDMHNNG - - LPSQVTKHAFEIGKTVATT
PhaC2_{P3455} PSNSP - LNPQAVKELFNTGGSSAFKGLRHLDDLLHNDG - - LPSQVSKHAFEVGRNLA CT
PhaC_{Av} - - - - - MFPIDIRPDK - - - LTQEMLDYSRKLGGGMENLLNAAEAI - - - DTGVS
PhaC_{Ta} - - - - - MIPIDIRPDK - - - LAQEMLDYSRKLGGGMENLLNAAEAI - - - DTGVS
PhaC_{Bc} - - - - - MAIPYVQEWKLIKSMPEYKSSARRFKRAYEIMTTEAEP - - - EVGLT
PhaC_{Bm} - - - - - MTTFATEWEKQLELYPEEYRKAYRRVKKRASEILLREPEP - - - QVGLT



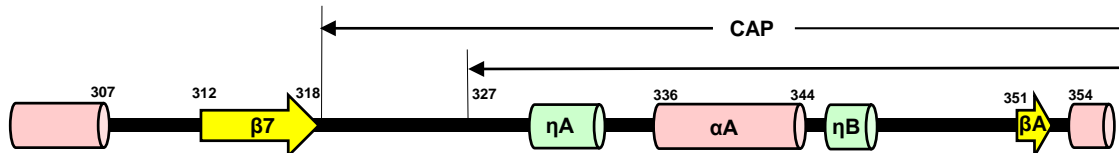
186 191 195 200 206 209 212 215 220 222 230 233 239

PhaC_{Cs} PGEVVF RNELIE LIQYTP TTEK - - VHEK P L L F V P P C I N K Y Y L M D L Q P D N S M V R H F V G Q G Y
PhaC_{Cn} EGAVVFENEYFQLLQYKPLTDK - - VHAR P L L M V P P C I N K Y Y I L D L Q P E S S L V R H V V E Q G H
PhaC_{Ap} PGRVVQRTELYELIQYSPTTET - - VGKT P V L I V P P F I N K Y Y I M D M R P Q N S L V A W L V A Q G Q
PhaC1_{P3613} EGAVVF RNDVLE LIQYRPTTEQ - - VHER P L L V V P P Q I N K F Y V F D L S P D K S L A R F L L R S Q V
PhaC1_{P3455} EGAVVF RNDVLE LIQYKPTITES - - VHER P L L V V P P Q I N K F Y V F D L S P D K S L A R F L L R S Q V
PhaC2_{P3613} AGSVVF RNELLE LMQYKPMSEK - - QYAK P L L I V P P Q I N K Y Y I F D L S P G N S F V Q Y A L K N G L
PhaC2_{P3455} PGAVVF RNELLE LIQYKPMSEK - - QYLR P L L I V P P Q I N K Y Y I F D L S N D K S F V Q Y A L K N G L
PhaC_{Av} PKQAVY SEDKLVLYRYDRPEGAPEAQPV P L L I V Y A L V N R P Y M T D I Q E D R S T I K G L L A T G Q
PhaC_{Ta} PKQP V Y K E D K L V L Y R Y D T P E G V T P - S A V P L L I V Y A L V N R P Y M T D I Q E D R S T I K G L L A T G Q
PhaC_{Bc} PKEVIWKKNKAKLYRYTPVKDN - - LHKT P I L L V Y A L I N K P Y I L D L T P G N S L V E Y L L N R G F
PhaC_{Bm} PKEVIWTKNKTLYRYIPKQEK - - TQRV P I L L I Y A L I N K P Y I M D L T P G N S L V E Y L V D R G F



243 249 250 261 266 269 280 285 290 292 300

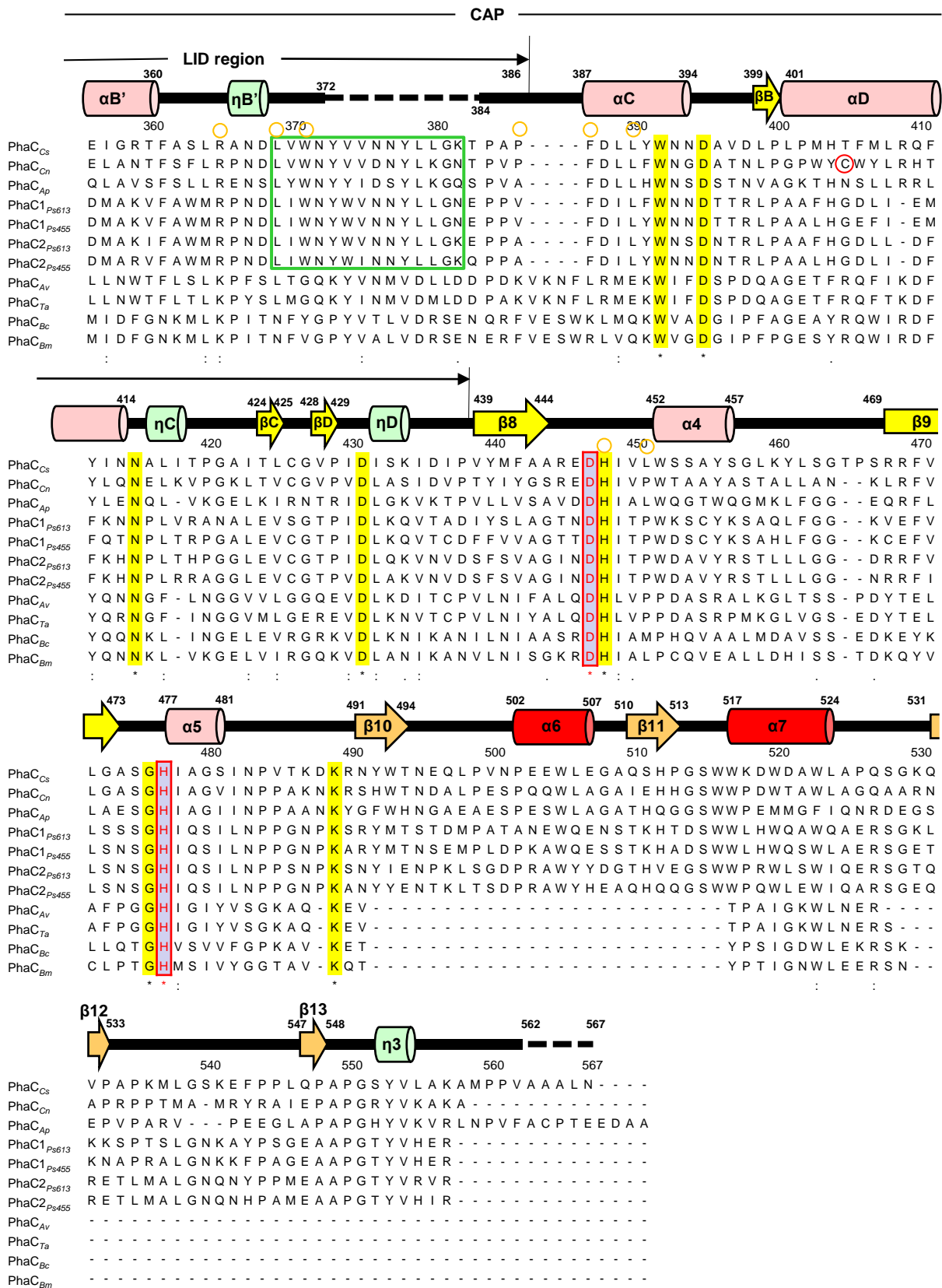
PhaC_{Cs} RVFLVS WRS AVPEMKNFTWETYIEKGVFAAAAEAVQKITKQPTMNALGF C VGG VILTTALC
PhaC_{Cn} TVFLVS WRNPDAS MAGSTWDDYIEHAIRAIEVARDISGQDKINVLGF C VGG TIVSTALA
PhaC_{Ap} TVFMIS WRNPGVAQAQIDLDDYVVDGVIAALDGV EAATGEREVHGI GY C IGG TALSLAMG
PhaC1_{P3613} QTFIVS WRNPTKAQREWGLSTYID - ALKEAVD VVSAITGSKDINMLG A C SGG ITCTALLG
PhaC1_{P3455} QTFVVS WRNPTKAQREWGLSTYIE - ALKEAIDVICAITGSKDINMLG A C SGG LTTASLLG
PhaC2_{P3613} QVFVVS WRNPDVRHREWGLSSYVE - ALEEALNV CRAITGARDVNLMG A C A G G L T I A A L Q G
PhaC2_{P3455} QTFMIS WRNPDPRHREWGLSSYVQ - AVEDAVDACRAIAGSKDINMLG A C A G G L T I A A L Q G
PhaC_{Av} HLQARRQLRKVSSATYMVSLLD S L Q I D S - P A M L F A D E E T L E S A K R R S Y Q - - - Q V L D G R
PhaC_{Ta} DVYLLID WGYPDQADRAINLDDYINGYIDSCVDHLREQLGV D K V N L L G I C Q G G V F S L - - - -
PhaC_{Bc} DVYLLD WGT P G L E D S N M K L D D Y I V D Y I P K A A K V L R T S K S P D L S V L G Y C M G G T M T S - - - -
PhaC_{Bm} DVYMLD WGT F G L E D S H L K F D D F V F D Y I A K A V K K V M R T A K S D E I S L L G Y C M G G T L T S - - - -



307 312 318 327 330 336 344 351 354

PhaC_{Cs} VAQAKGL - KYFDSATFMTSLI D H A E P G - E I S F F I D E A L V S R E A K M A A G - - - - G I S G K
PhaC_{Cn} VLAARGE - HPAASVTLTTL D F A D T G - I L D V F D E G H V Q L R E A T L G G G A G A P C A L L R G L
PhaC_{Ap} WLAARRQKQRVRTATLFTTL D F S Q P G - E L G F I F H E P I I A A L E A Q N E A K - - - - G I M D G R
PhaC1_{P3613} HYAALGE - KKVNALTLVSVL D T T L D S - Q V A L F V D E K T L E A A K R H S Y Q - - - - A G V L E G R
PhaC1_{P3455} HYAALGQ - PKVNALTLVSVL D T Q L D T - Q V A L F A D E K T L E A A K R R S Y Q - - - - A G V L E G S
PhaC2_{P3613} HLQAKRQLRRVSSASYLVSL D S Q I D S - P A T L F A D E Q T L E A A K R H S Y Q - - - - R G V L E G R
PhaC2_{P3455} HLQARRQLRKVSSATYMVSL D S Q I D S - P A M L F A D E E T L E S A K R R S Y Q - - - - Q V L D G R
PhaC_{Av} MYSALHPDK - VRNLVTMVTVP D F K T P D N L L S A W V - - Q N V D I D L A V D T M - - - - G N I P G E
PhaC_{Ta} MYASMHPDK - VNNLVTMVTVP D F K T P D N L L S A W I - - Q N V D V D L A V D T I - - - - G N I P G E
PhaC_{Bc} IFAALNEDLP IKNLIFMTSPF D F S D T G - L Y G A F L D D R Y F N L D K A V D T F - - - - G N I P P E
PhaC_{Bm} IYAALHPHMP I RNLI FMTSPF D F S E T G - L Y G P L L D E K Y F N L D K A V D T F - - - - G N I P P E

Supplementary Figure 1



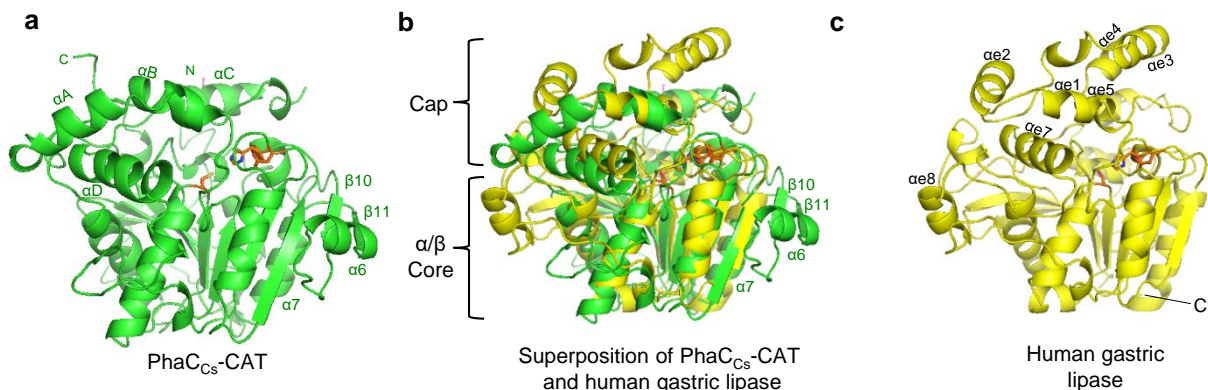
Supplementary Figure 1 (continued)

Supplementary Figure 1 Alignment of amino acid sequences of PhaC catalytic domains.

The secondary structure elements found in the structure of PhaC_{Cs}-CAT are shown at the top of the alignment with α -helices (pink cylinders), β -strands (yellow arrows), loops (bold lines) and missing loops (broken lines). Glu329, Phe332, Phe333, Arg365, His448 and Val450 of PhaC_{Cs}-CAT are marked with orange circles at the top. Two Cys residues (Cys382 and Cys438 of PhaC_{Cs}-CAT) forming a disulfide bond are marked by red circles. Part (Leu402–Asn415 of PhaC_{Cs}-CAT) of the LID region forming $\alpha 4$ helix in the structure is conserved in members of Class I and II synthases. These sequences are marked by circles within a green box.

The abbreviations are as follows:

Class I	PhaC _{Cs} ⁻	<i>Chromobacterium</i> sp. USM2 (ADL70203)
	PhaC _{Cn} ⁻	<i>Cupriavidus necator</i> (AAW65074)
	PhaC _{Ap} ⁻	<i>Aeromonas punctata</i> (BAA21815)
Class II	PhaC1 _{Ps61-3} ⁻	<i>Pseudomonas</i> sp. 61-3 (BAA36200)
	PhaC2 _{Ps61-3} ⁻	<i>Pseudomonas</i> sp. 61-3 (BAA36202)
	PhaC1 _{PsUSM4-55} ⁻	<i>Pseudomonas</i> sp. USM4-55 (ABX64434)
	PhaC2 _{PsUSM4-55} ⁻	<i>Pseudomonas</i> sp. USM4-55 (ABX64435)
Class III	PhaC _{Av} ⁻	<i>Allochromatium vinosum</i> DSM180 (BAE20055)
	PhaC _{Tv} ⁻	<i>Thiocystis violascens</i> DSM198 (AFL75311)
Class IV	PhaC _{Bc} ⁻	<i>Bacillus cereus</i> (BAI68395)
	PhaC _{Bm} ⁻	<i>Bacillus megaterium</i> (AAD05260)

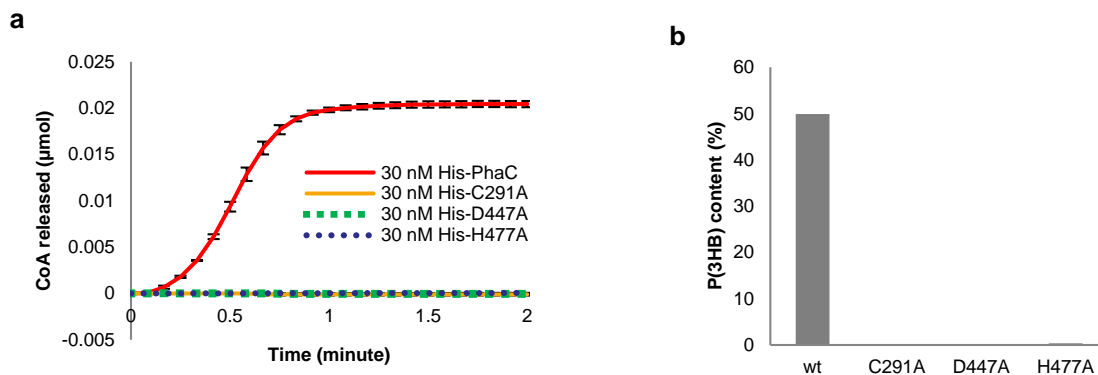


Supplementary Figure 2 Structural comparison of the catalytic domains of PhaC_{Cs} and lipase.

(a) A side-view of PhaC_{Cs}-CAT determined in the current study.

(b) As in a, but overlaid on human gastric lipase. The α/β core subdomains are well overlapped with a small r.m.s. deviation (1.02 Å), whereas the CAP subdomains are poorly overlapped. The α/β core subdomain of PhaC_{Cs}-CAT contains an additional segment comprising $\beta 10$ - $\alpha 6$ - $\beta 11$ - $\alpha 7$.

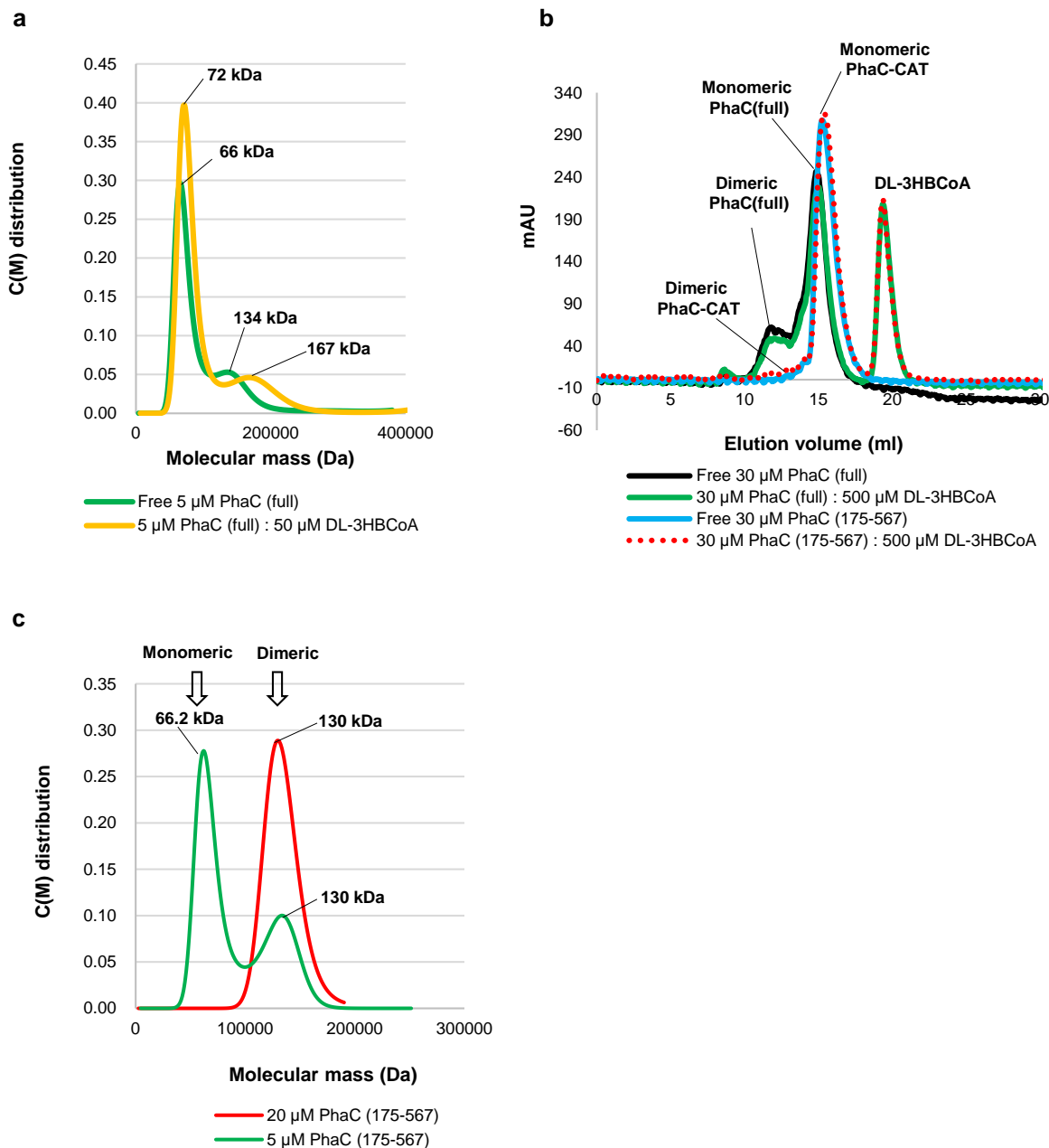
(c) A side-view of human gastric lipase (PDB code 1HLG).



Supplementary Figure 3 Verification of catalytic triad of PhaC_{Cs}

(a) *In vitro* PHA synthase activity of wild-type and mutant PhaC_{Cs} (full-length). Wild-type PhaC_{Cs} showed higher activity compared with the three catalytic mutants – C291A, D447A and H477A. CoA released from the enzymatic reaction by wild-type N-terminal His-tagged fusion PhaC_{Cs} was detected and reached its maximum consumption of substrate within 1 minute, whereas no significant free CoA was detected from the other three individual catalytic triad mutants.

(b) *In vivo* poly(3-hydroxybutyrate) [P(3HB)] production in wild-type and mutant full-length PhaC_{Cs} using 0.5% CPKO as carbon source. Wild-type PhaC_{Cs} (wt) accumulated P(3HB) at 49.9wt% while the accumulation of P(3HB) of mutant PhaC_{Cs} with C291A, D447A, and H477A are 0.1, 0.1, and 0.4wt%, respectively.



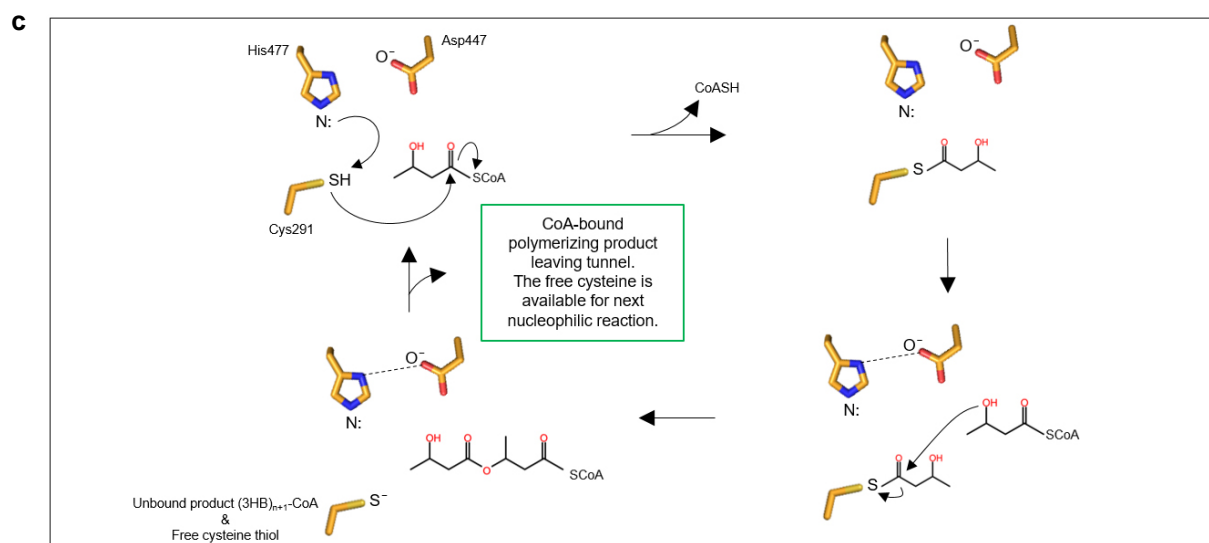
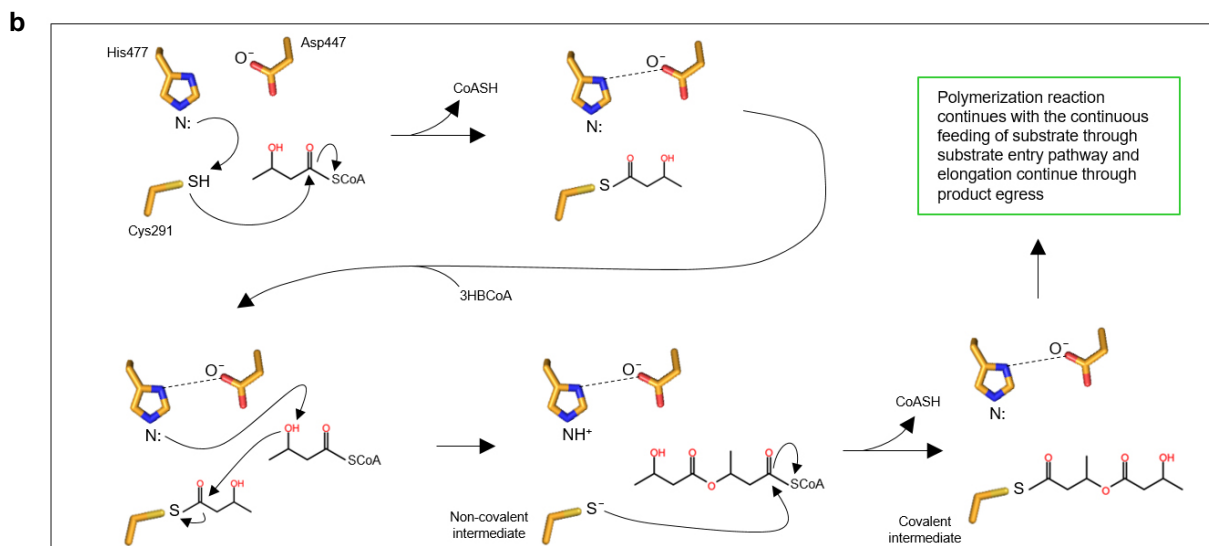
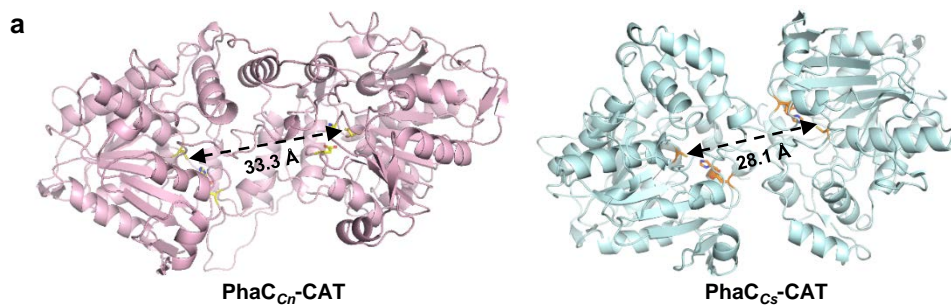
Supplementary Figure 4
Oligomerization equilibrium of PhaC in solution.

(a) Analytical ultracentrifugation (AUC) analyses of PhaC_{Cs} (full-length, 63.4 kDa) using a sedimentation velocity method. The equilibria between monomeric and dimeric forms in solution were observed in the absence or presence of substrate (DL-3HB-CoA). The monomeric form is dominant at 5 μM PhaC_{Cs} in 10 mM Tris-HCl (pH 8.0), 100 mM NaCl and 3 mM β-ME at 20 °C, with a small portion being present in the dimeric form. In the presence of substrate, a higher (presumably tetrameric) form appears.

(b) Size exclusion chromatography (SEC) of PhaC_{Cs} (full-length) and PhaC_{Cs}(175-567), which is PhaC_{Cs}-CAT, in the absence or presence of substrate DL-3HB-CoA. The elution profiles of PhaC_{Cs} and PhaC_{Cs}-CAT were compared.

PhaC_{Cs} (full-length) exists in monomeric and dimeric (dominant) forms. Addition of DL-3HB-CoA induced only a small change with the appearance of a faint peak, presumably that of the tetrameric form. In contrast to these equilibria, PhaC_{Cs}-CAT exists in a monomeric form at a concentration of 30 μM both in the absence or presence of substrate.

(c) AUC analyses of PhaC_{Cs}-CAT using a sedimentation velocity method. The dimeric form of PhaC_{Cs}-CAT was detected at 5 μM (green) and significantly stabilized at a high concentration (20 μM, red).

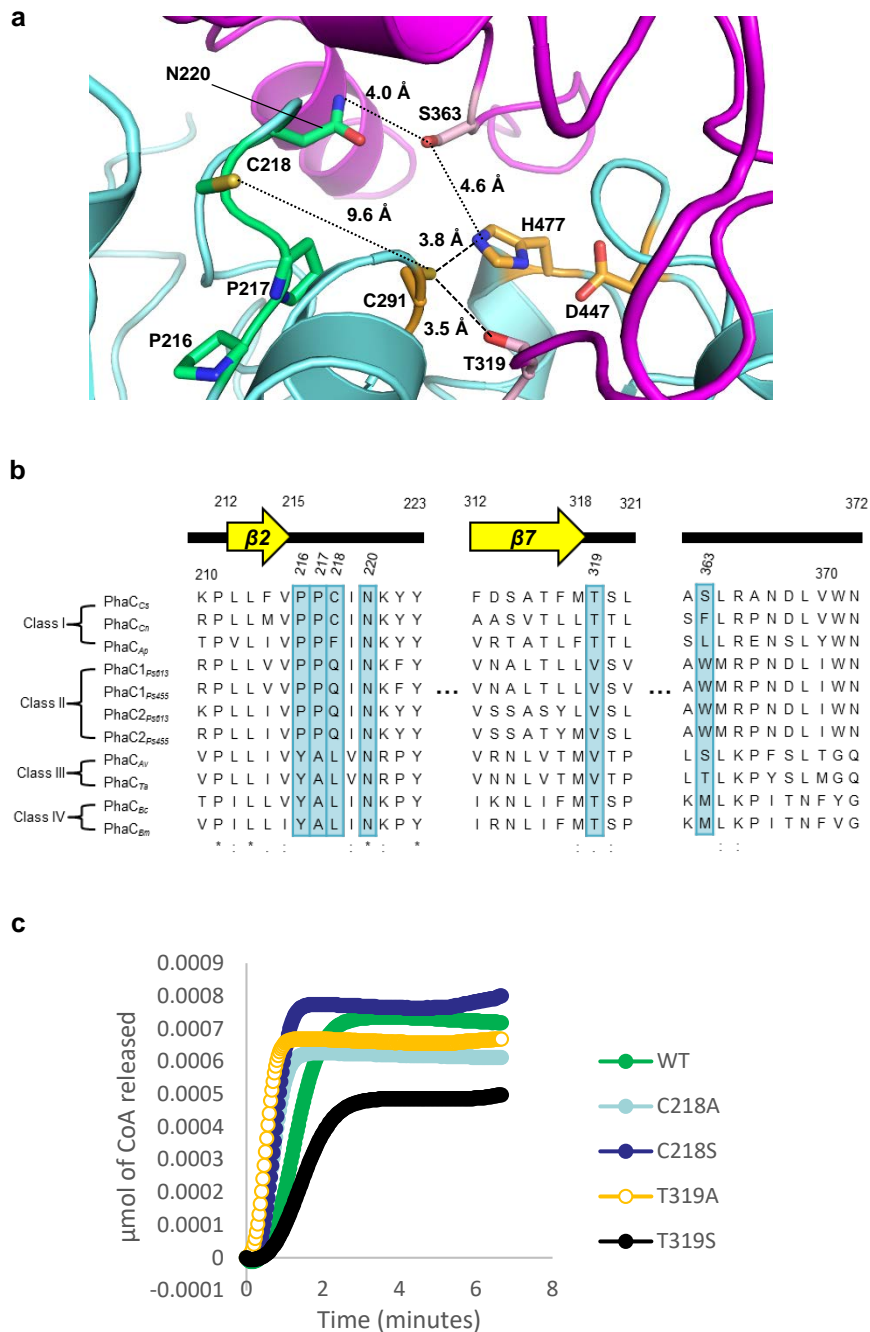


Supplementary Figure 5 Proposed mechanism

(a) The distance of the catalytic cysteines from both protomers. The distance (broken lines) between the S_V from both mol A and mol B is 33.3 Å in dimer PhaC_{Cr}-CAT (5HZ2)²³ (left) and 28.1 Å in dimeric PhaC_{Cs}-CAT (right).

(b) Processive single active site model (In and out tunnels) which is also proposed in reference²². This model requires a single active site for PHA chain elongation and a non-covalent intermediate, in addition to a covalent intermediate bound to the Cys residue at the active center during the catalytic cycle.

(c) Alternative processive single active site model (Single tunnel) which is proposed in reference²³. In this model, two substrates share the same substrate-binding tunnel and the first 3HB-CoA produces 3HB-Cys. The second 3HB-CoA attacks 3HB-Cys to produce (3HB)₂-CoA, which is released from the active site. The cycle is repeated with newly entered 3HB-CoA to produce 3HB-Cys, and then the following (3HB)₂-CoA enters the active site to produce (3HB)₃-CoA, which is again released from the active site.



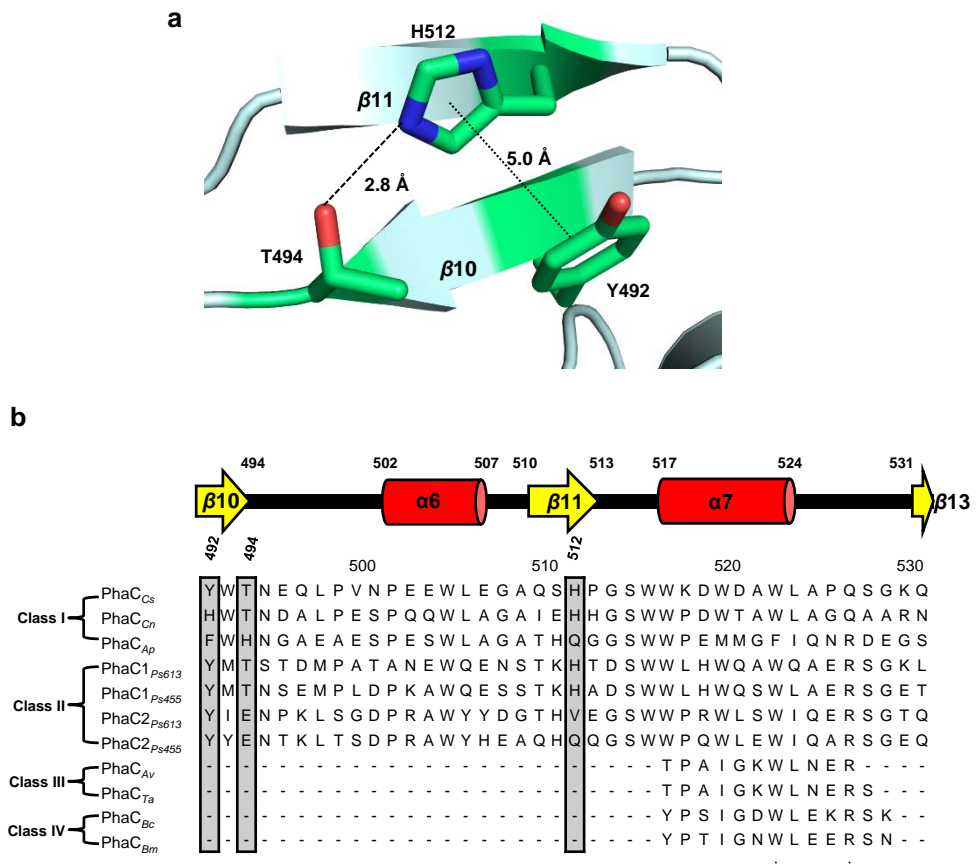
Supplementary Figure 6

The active site of PhaC_{CS}-CAT contains residues at positions where assistance of active center Cys291 or intermediate binding may be possible at the active site.

(a) The active site of PhaC_{CS}-CAT contains two polar residues, and Cys218, which may act to stabilize intermediate binding, and Thr319 may assist with activation of active center Cys291. The active site of the closed form of PhaC_{CS}-CAT also contains Ser363 from the LID region of the CAP subdomain. Hydrogen bonds (broken lines) and interesting distances (dotted lines) are indicated.

(b) Partial alignment of the sequence showing Cys218, Thr319 and Ser363 of PhaC_{CS}-CAT.

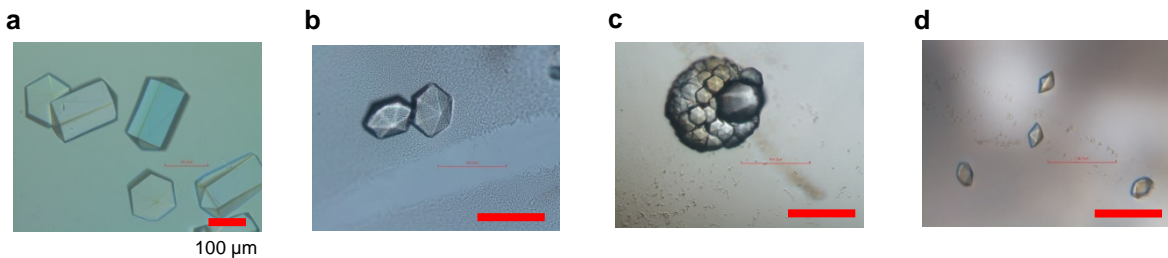
(c) Enzymatic activity of wild-type and mutant PhaC_{CS} with substrate 3HB-CoA. Released CoA was monitored.



Supplementary Figure 7
Conformation of the β 10- β 11 segment of PhaC_{Cs}

(a) A close-up view of the β 9- β 10 antiparallel β -sheet, which is part of the PhaC-specific additional β 9- α 5- β 10 segment of the core subdomain. The segment is located at the edge of the α / β core and projects from the core subdomain without making contacts with the other protomer of the closed form dimer of PhaC_{Cs}-CAT. Hydrogen bonds (broken lines) and the ring-ring distance (dotted line) between His512 and Tyr492 residues are indicated.

(b) Partial alignment of the sequence showing that Tyr492 is replaced with other aromatic residues such as Phe or His in Class I synthases, but is conserved in Class II synthases.



Supplementary Figure 8
Obtained crystals of PhaC_{Cs}-CAT.

- (a) Hexagonal crystals of PhaC_{Cs} (175-567) grown in 50 mM Bis-Tris (pH5.5), 0.06 M Ammonium sulfate, and 5% PEG4000 by seeding.
- (b) Hexagonal crystals of full-length PhaC_{Cs} grown after 8 – 12 months in a sitting drop set up equilibrated against INDEX (75): 0.1 M Bis-Tris (pH6.5), 0.2 M Lithium sulfate, and 25% PEG3350.
- (c) Crystal clusters of α -chymotrypsin-digested PhaC_{Cs} in a hanging drop against 0.1 M Bis-Tris (pH6.5), 0.2 M Lithium sulfate, and 20% PEG4000.
- (d) α -chymotrypsin-digested Se-Met labeled PhaC_{Cs} grown in 0.1 M Bis-Tris (pH6.3), 0.2 M Lithium sulfate, and 15% PEG3350 by seeding.