

Supplementary information

Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation

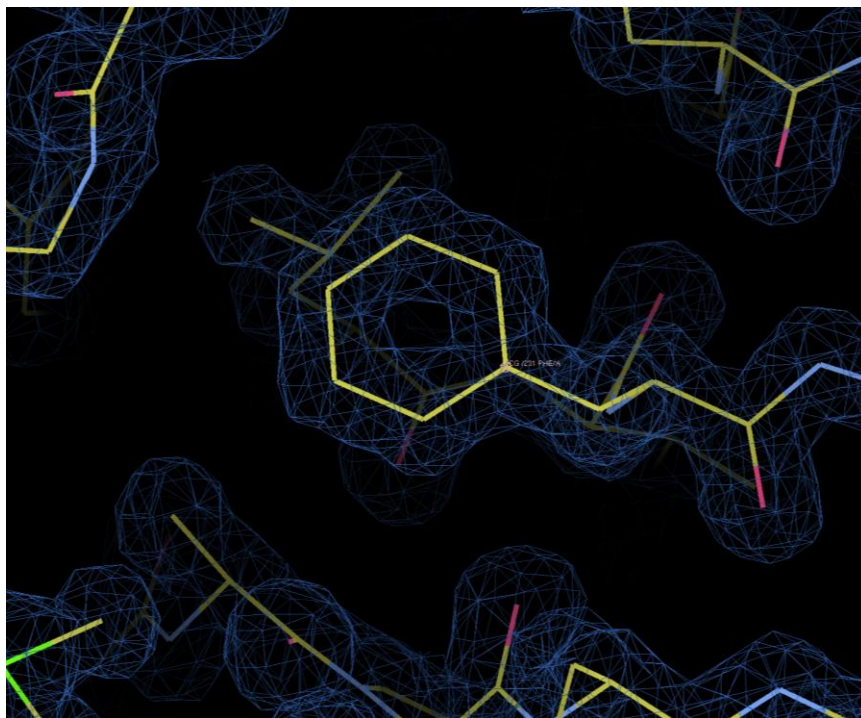
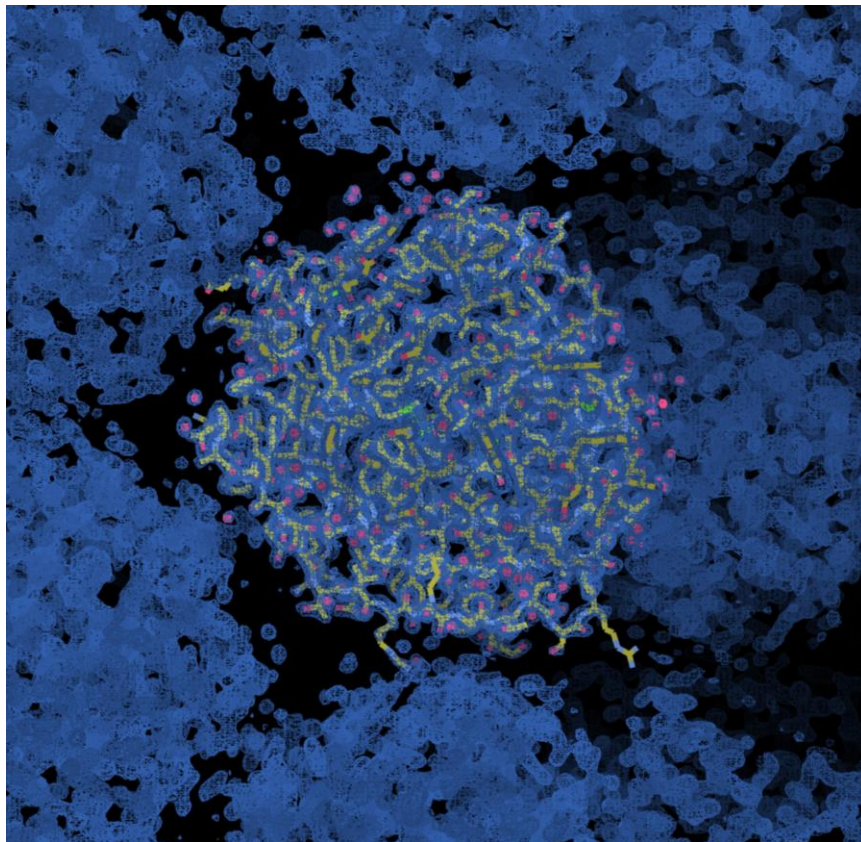
Seongjoon Joo, In Jin Cho, Hogyun Seo, Hyeoncheol Francis Son, Hye-Young Sagong, Tae Joo Shin, So Young Choi, Sang Yup Lee* & Kyung-Jin Kim*

*Correspondence should be addressed to: Email: kkim@knu.ac.kr
leesy@kaist.ac.kr

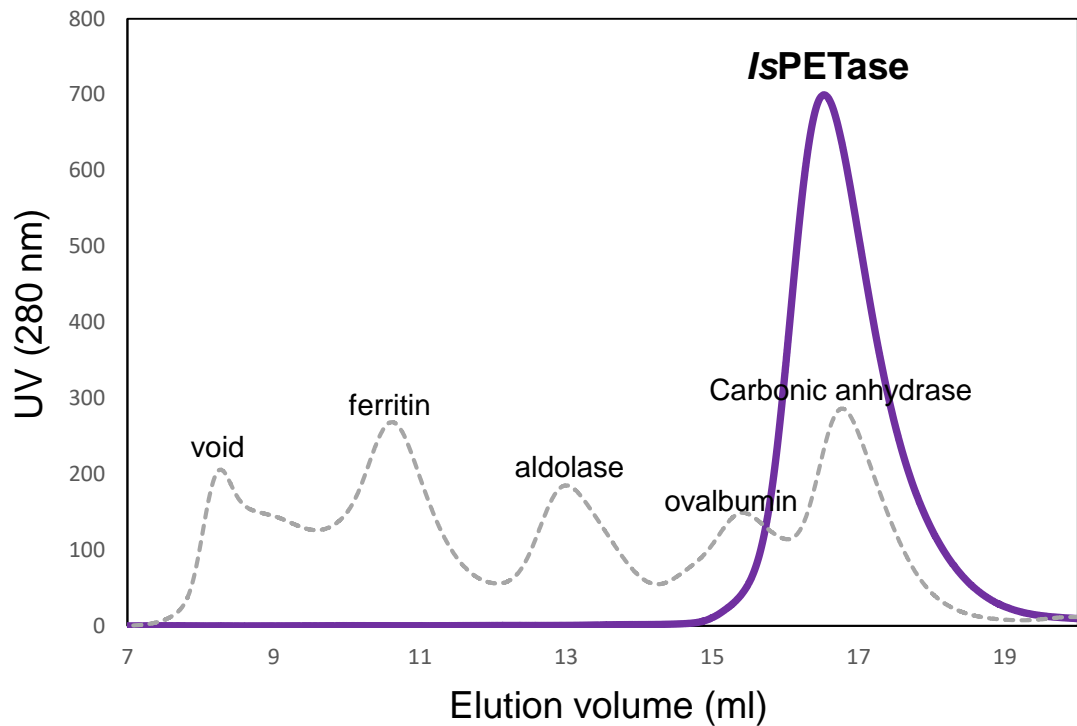
This PDF file includes:

Supplementary Figures 1 to 6.

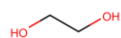
Supplementary Tables 1 and 2



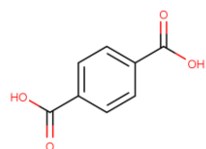
Supplementary Figure 1. The electron density map of *IsPETase*. The figures were generated by *WinCoot*. The $2F_0 - F_c$ map contoured at 1.5σ is shown as blue mesh. The refined atomic model of *IsPETase* are shown as line with different color scheme by atoms; yellow-carbon, red-oxygen, blue-nitrogen, and green-sulfur. No crystallographic symmetry intimates the monomeric structure of *IsPETase* (above). The map represents the high resolution of 1.5 \AA of *IsPETase* (below).



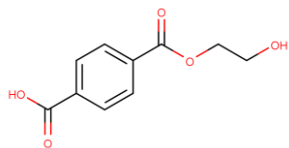
Supplementary Figure 2. Size-exclusion chromatographic analysis of *IsPETase*. A elution pick corresponding to a monomeric state of *IsPETase* in Size-exclusion chromatographic analysis. For precise analysis of the molecular weight, standard samples of ferritin (440kDa), aldolase (158kDa), ovalbumin (44kDa), and carbonic anhydrase (29kDa) are used for calibration and labelled.



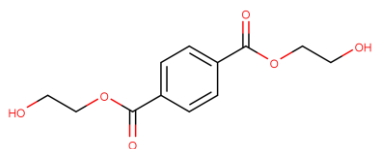
EG



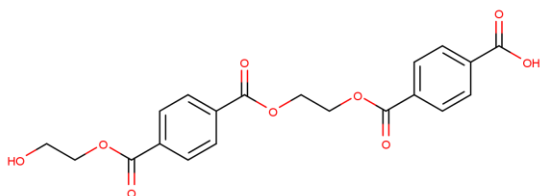
TPA



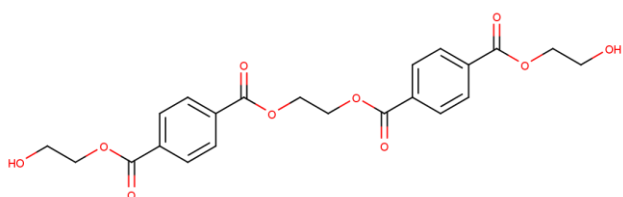
MHET



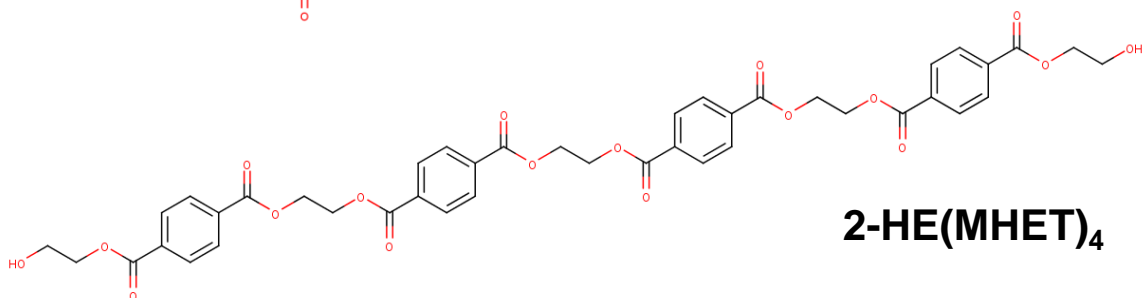
BHET



(MHET)₂

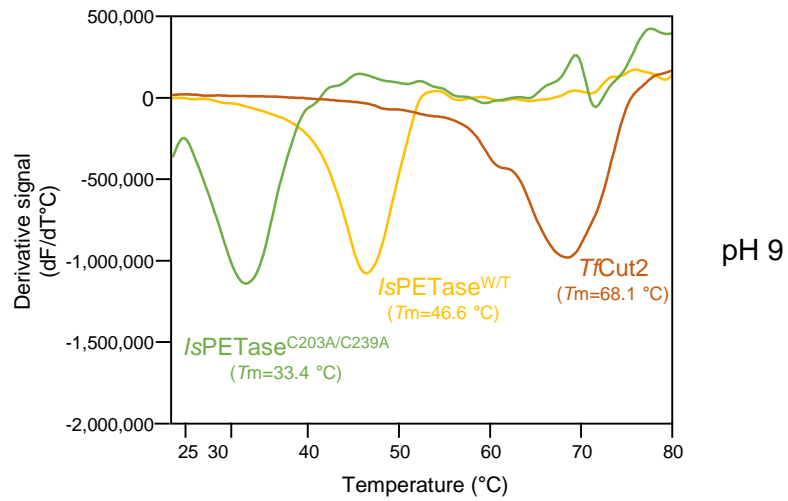
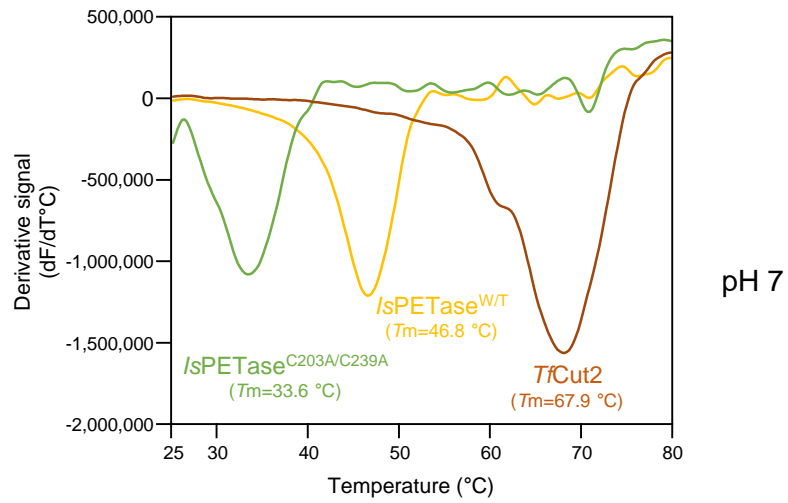


2-HE(MHET)₂

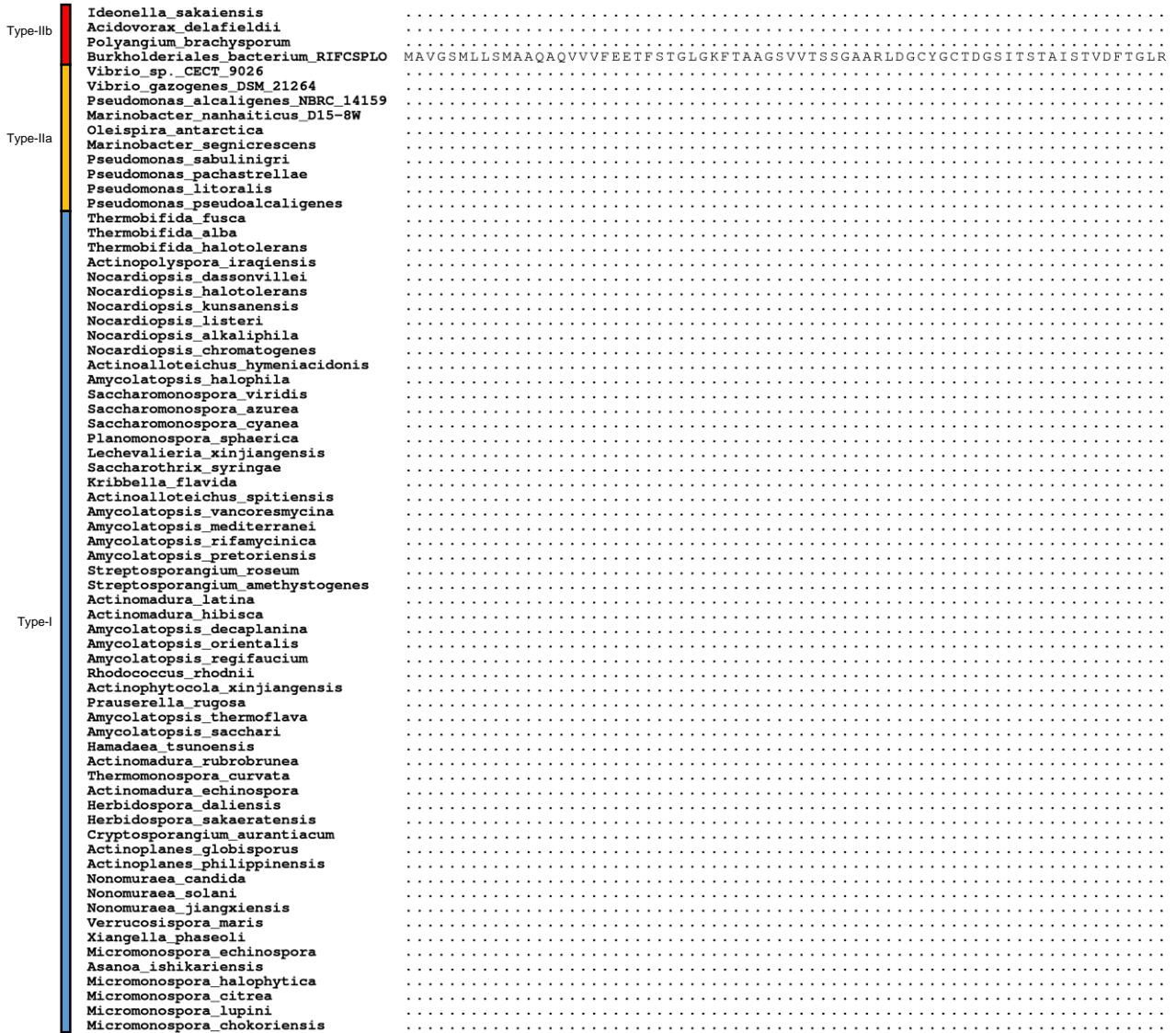


2-HE(MHET)₄

Supplementary Figure 3. Chemical structures of the PET-related molecules used in this study.



Supplementary Figure 4. Thermal stability measurements of *IsPETase*^{W/T}, *IsPETase*^{C203A/C239A}, and *TfCut2*. The T_m values at pH 7 and 9 of *IsPETase*^{W/T}, *IsPETase*^{C203A/C239A}, and *TfCut2* are displayed.



Supplementary Figure 5. Amino acid sequence alignment of PET-degrading enzymes. All enzymes involved in the tree generation in Fig. 6a were aligned. Key residues involved in the enzyme catalysis and the constitution of subsite I and subsite II are indicated by red-, blue-, and purple-colored triangles, respectively as same in Fig. 1a. At the left side, types of each enzymes are indicated.

	1	10
Type-IIbMNFPPRA.....SR.....LMQA	
MHLPRSRWDIP..FKEE..TTMT....HH.FSVRA	
MP.....VLP.....RR.LAAAA	
	LSFDRVTSGLDSSGEAGIAEFPSTNGSTYTAVESIRTASGRVTFNLP..TSAENQSGLR....LR.FRINA	
M.....NVLTKCKL.....A.LGIIV	
M.....NVLTKCKL.....A.LGIIV	
M.....NKLQKPL.....T.L....	
Type-IIaMS..DHYASNPLR....SV.VAAAS	
MI..NNKQKSTLV....TL.LASSA	
MPFNKKSVL....AL.WGAGA	
MI..NKNLPSSL....SM.LAAGA	
MI..NRTLPSLL....SM.LAAGA	
MAV.M.....TPRR.ERSLLSRALQVTA	
MSV.T.....TPRR.EA.SLLSRVAVAAA	
MRT...RVKSGTA..ARGRFTRAL....GVVTA.AL	
MRT.YPLSPEE....TDGARSGPRS.RVRFARAGM	
MRK.YPLSPSPETP..DSHRARRGPRS.RVSRFAARVGM	
	MRT...TSLPPS....QSHDTLRRS.WTRFASRTAA	
MTE.NP..PSEEG....RTRPRTARRS.WMSGLAAKSAL	
	MRTD.SPPSPIPAE....RTRSGSRARRS.WMSAVAVKSAL	
MSI.ETIP.....HTGGARWA..RAAGTLAATA	
MASRRA....RVRSGRFRPT.RLAGLVLLALTM	
M.....MLFA.TL	
MRI...RRQAGTG..ARASMARAI....GVMTT.AL	
M.....MARTI....GVLTA.AL	
MRI...RSRSLTR..PHGRVSRTI....GVLTA.AL	
M.....NHRA....RTLTKLAL	
MSLRSPLSR.H..ALVFA...	
MSSRSPLSR.R..ALVLV...	
MLGL.V..TTVFA...	
MRS.TLSRPA...P..SRSLRRVLT.S..A..ALTGL...	
MSA.LTSPPTLGLG..EKISRRRPWRT.KAAGVVLA...	
MSA.LTSQPTSSGSS..EKIPRLRGWRA.KAAGVVLA...	
MSA.LTSPPTTSGPS..EKISRRRGWRG.KAAGVVLA...	
MSV.LTSPPTSSGSG..EKISRRRPWRA.KAAGVLA...	
MTLAL	
MTLALAL	
MHLIPR..RAAKTLALAL	
Type-IMPHRTTPSGSP..PPARPGRPR..RGPRRARAVAA	
MCS.....TAN....MRSTTR..ILCGLATVAA	
MRSTTR..ILCGLATVAA	
MRSTTR..ILCGLATVAA	
MRRGVR..ILTLTIAAS	
MRKTLRKGVR..ILGGVALAAT	
MTLLSAAAL	
MRRSPR..NLARLVTALA	
MRRSAK..AWSSLVFAA	
MHTSTKA..GAARVLAALAV	
MRSG.LST....A..PR...RRALR..AVATTALALTL	
MKRT.L.....KRALS....LLPA.AA	
MNTN.L.....ARKLR....TLPALLA	
MSRLAT..TALAF	
MRPHFWSRRLSRIAT..AALAF	
MLTIKPKARRLGLIGK..LTAL	
MTTK...TRR..FLAA..AAA...	
VPTARR..WLVV..AATAI	
MSREYDDVQL...PTPRSGAR..TLTK..LALTV	
MQL...STPHGA...LTK..LTLT	
MQL...STPQGAR..TLTK..LTTL	
MTA	
MRSPTTTRPRSVS....LVAR..LAMAA	
MPSPTTIRPRSLVR..AVAG..LTTAV	
MKTAIR..L...LATA	
MRSPTTTRPRSAVR..TMAR..LAMAA	
MSI.CPDPQGVDPVSSPTTTRPRSAVR..AA.G..LALAA	
MPAR....AVR..LALAA	
MSSPTTTRPRSV....AAR..LGLSA	

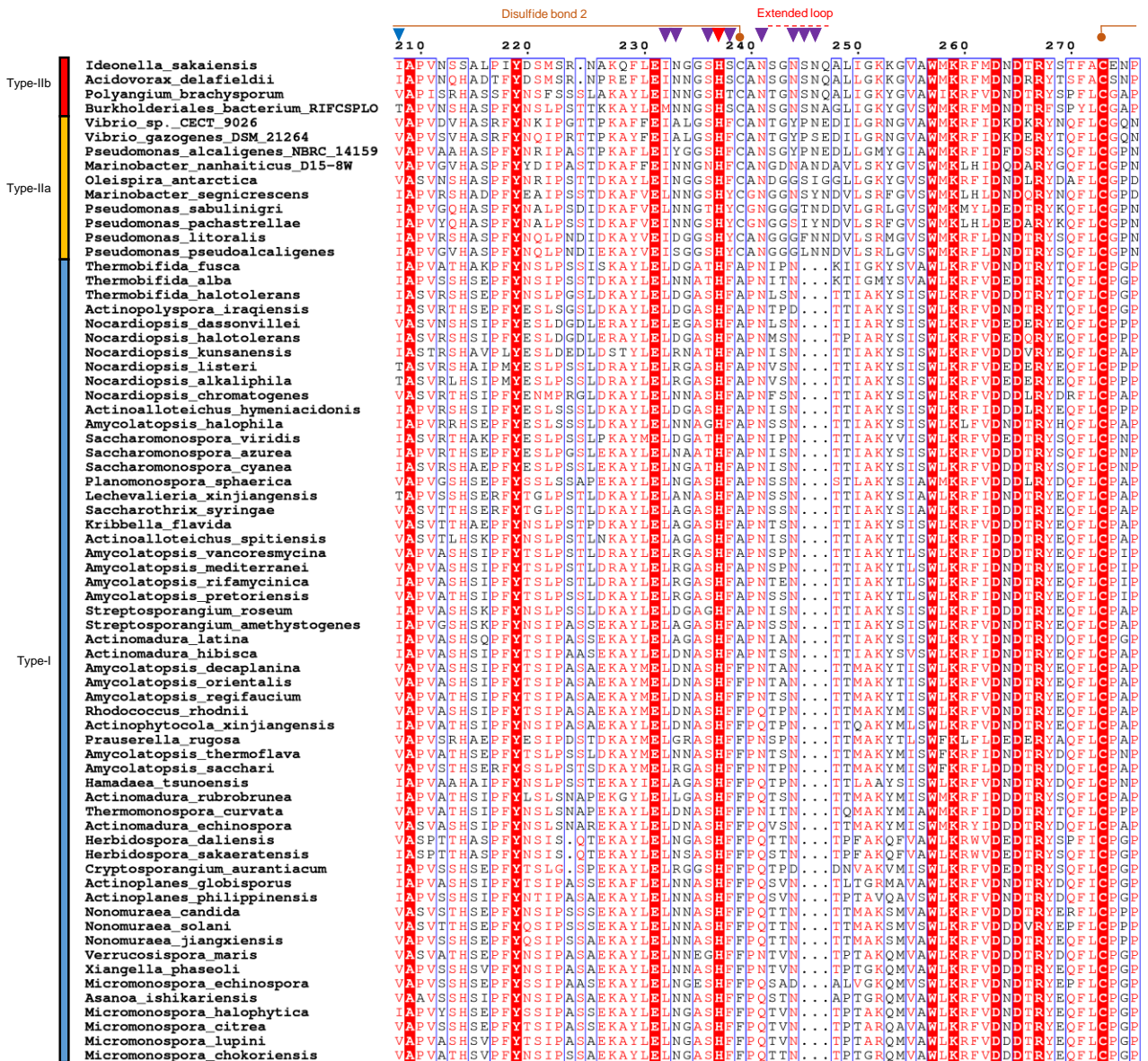
Supplementary Figure 5. Amino acid sequence alignment of PET-degrading enzymes (continued).

		20	30	40	50	60	
Type-IIb	<i>Ideonella sakaiensis</i>	AVLGG.....	GLMAYVSAATAQIT	NPYARGP	NP	TAAASLEASAGPFTVRS	
	<i>Acidovorax delafieldii</i>	LLAA.....	GALLASAAVSAQIT	NPYERGP	NP	TSSLEASRGPFYSQST	
	<i>Polyangium brachysporum</i>	LLASA.....	TLVPLSAAQIT	NPYQKGP	NP	TRDLEASRGPFYAST	
	<i>Burkholderiales bacterium_RIFCSPL0</i>	SLSSSEYTVVDN	IRLEGTSGSGGGT	NPFEKGP	NP	KTMLEASGPFYKTT	
	<i>Vibrio sp. CECT_9026</i>	IFFSLPSFAV....	P.CSD.....	CNNGFERGP	NP	VRVDQLEASRGPFYKTT	
	<i>Vibrio zarogones DSM_21264</i>	AFSSLPSFAV....	P.CSD.....	CNNGFERGP	NP	VRVDQLEASRGPFYKTT	
	<i>Pseudomonas alcaligenes NBRC_14159</i>	
	<i>Marinobacter nanhaiticus D15-8W</i>	
	<i>Oleispira antarctica</i>	
	<i>Marinobacter segnicrescens</i>	LLFSASVFAAG....	GGGSDGGDDGCTSN	CYGERGP	NP	TESLEASGPFYSVTR	
Type-IIa	<i>Pseudomonas sabulinigri</i>	LLMSSVAMANN....	PPVDPG....	TGEYQKGP	NP	TVSLEASRGPFYSVTR	
	<i>Pseudomonas pachastrellae</i>	LLFMSALANN....	PAPTDPDGG....	GSAVQKGP	NP	SVSLEASRGPFYSVTR	
	<i>Pseudomonas litoralis</i>	LMLSTSAFATN....	PPVDDPSD....	GGAYERGP	NP	TVSLEASGPFYSVTR	
	<i>Pseudomonas pseudocalcigenes</i>	LLLSTSVMATN....	PPVDETPD....	GDSYARGP	NP	VAFLEASGPFYSVTR	
	<i>Thermobifida fusca</i>	AATALVTAVSLAAPAHAA	NPYERGP	NP	TDALLEASGPFYSVEE	
	<i>Thermobifida alba</i>	AATVALAAPAQAAN	NPYERGP	NP	TESMLEASGPFYSVEE	
	<i>Thermobifida halotolerans</i>	MANPYERGP	NP	TNSIEALRGPFYVDE	
	<i>Actinopolyspora iraqiensis</i>	ALTG.....	GVAGAGVASAQGN	NPYERGP	NP	EDSIEDRGPFYVDE	
	<i>Nocardioopsis dassonvillei</i>	TLALT...AGLTGAVVAPAQAAN	NPYERGP	NP	TESVTVAVEGYVDE	
	<i>Nocardioopsis halotolerans</i>	TLTLA...AGLTGGVVAQAAN	NPYERGP	NP	EQSVTALRGYFDTD	
Type-I	<i>Nocardioopsis kunsanensis</i>	VIGAA...AVTTVGLTPVPHAE	NPHERGP	NP	TEESVTAFTGFFDTD	
	<i>Nocardioopsis listeri</i>	VFGLS...LIGAGATVPALAE	NPHERGP	NP	TESVTVARHGFDTD	
	<i>Nocardioopsis alkaliphila</i>	VVGLA...LFTGGATVAPAAAS	NPYERGP	NP	TERSVTARHGFDTD	
	<i>Nocardioopsis chromatogenes</i>	LALGGGTVAASS...A	HAAPLPAAQAAN	NPHERGP	NP	EDSIEALRGHFDVAE
	<i>Actinalloteichus hymeniacidonisi</i>	VA.....	
	<i>Amycolatopsis halophila</i>	LL...VGTPTAATGVATAQD	NPYERGP	NP	TESMIEALRGPFYVAT	
	<i>Saccharomonospora viridis</i>	AVLVG...AV....	GGVACAESTVAQD	NPYERGP	NP	EDSIEALRGPFYVAT	
	<i>Saccharomonospora azurea</i>	AVASG...LV....	SAGVACADVITTEQ	NPYERGP	NP	EDSIEALRGPFYVAT	
	<i>Saccharomonospora cyanea</i>	AVASG...LA....	SAGVACADEVITTAQ	NPYERGP	NP	TEGSEALRGPFYVAT	
	<i>Planomonospora sphaerica</i>	ALALVGGAGVAPAAHAAA	GYERGP	NP	TSSLEASRGPFYVAT	
<i>Lechevalieria xinjiangensis</i>	AV...VALTSSGSIAPAAE	NPYERGP	NP	ATASIEALRGPFYVAT		
<i>Saccharothrix syringae</i>	AA...VALTTPTAVAAQAE	NPYERGP	NP	ATSSIEALRGPFYVAT		
<i>Kribbella flavida</i>	LT...LFTGGATVAPAAAS	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Actinalloteichus spitiensis</i>	LV...AGSLAAPAAQAAN	NPYERGP	NP	TERSIEALRGPFYVAT		
<i>Amycolatopsis vancouveri</i>	AL...LFTGGATVAPAAAS	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Amycolatopsis mediterranea</i>	AL...ALTTGVAAPAPAAA	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Amycolatopsis rifamycinica</i>	AL...AVTTAVAAPAAE	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Amycolatopsis pretoriensis</i>	AL...VVTITLASPAPAA	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Streptosporangium roseum</i>	VLGGALAPAA...HA	ATFTAPAAQVAAN	NPYERGP	NP	TVSIEALRGPFYVAT		
<i>Streptosporangium amethystogenes</i>	LGTGGGLAPAA...Q	AVAAP...AAQAAAAG	PYDRGP	NP	TNASIEALRGPFYVAT		
<i>Actinomadura latina</i>	T...ATALAVPQAAQAAN	NPYERGP	NP	APESERSIEALRGPFYVAT		
<i>Actinomadura hibisca</i>	LSLG.....	AALLTAAPVQAAA	NPYERGP	NP	TEASIEALRGPFYVAT		
<i>Amycolatopsis decaplanina</i>	L...T...VAPTAYAAE	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Amycolatopsis orientalis</i>	L...T...VAPTAYAAE	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Amycolatopsis regifaucium</i>	L...A...IAPTAAHAD	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Rhodococcus rhodnii</i>	VA...AG...VPATALGAE	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Actinophytocola xinjiangensis</i>	V...L...VAG...TATAATAQD	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Prauserella rugosa</i>	A.....	LV...TAVAGPTTALAQDAS	PYQHGP	NP	EDSIEALRGPFYVAT		
<i>Amycolatopsis thermoflava</i>	L...LA...AGAATATASAQAAN	NPYERGP	NP	TESMIEALRGPFYVAT		
<i>Amycolatopsis sacchari</i>	MA...V...TLAANPAEAS	PYQHGP	NP	TESMIEALRGPFYVAT		
<i>Hamadaea tsunoensis</i>	M...LTGVLAAGSPAADV	NPYERGP	NP	ETLASIQAGPFYVAT		
<i>Actinomadura rubrobrunea</i>	TAGTGVLAAPAAAGA	PAAPAAAAAAV	NPYQHGP	NP	EDLIEALRGPFYVAT	
<i>Thermomonospora curvata</i>	LAASA.....	LVAASPAQAAA	NPYQKGP	NP	ETASITAAARGPFYVAT	
<i>Actinomadura echinospora</i>	LAGSA.....	IIVAAPAANAAA	NPYERGP	NP	TESMIEALRGPFYVAT	
<i>Herbidosporea daliensis</i>	ILATGTAVTL....	APSAHAA	GFERGP	NP	TSSIEALRGPFYVAT	
<i>Herbidosporea sakaeratsensis</i>	TLATGTAVTL....	APSAHAA	GFERGP	NP	TSSIEALRGPFYVAT	
<i>Cryptosporangium aurantiacum</i>	ALAGGSIAAAPS....	GVQAAD	GPYWRGP	NP	ETASVATTGPFYVAT	
<i>Actinoplanes globisporus</i>	AAVLTV...PATSAHAAD	NPYQKGP	NP	TLAALDASRGPFYVAT		
<i>Actinoplanes philippinensis</i>	TLVT...GAGVT....	PAAAAAN	NPYERGP	NP	ETLAAQASRGPFYVAT	
<i>Nonomuraea candida</i>	SLIAGALAAATSPAQ....	AAAGPYDRGP	NP	ETDAIEALRGPFYVAT		
<i>Nonomuraea solani</i>	LLIGGTLTPASPAQAAD	PPYDSASLTAAG	NPYDRGP	NP	ETNAIEALRGPFYVAT		
<i>Nonomuraea jiangxiensis</i>	ALIGGALAAAPAAQ....	AAAGPYDRGP	NP	ETNAIEALRGPFYVAT		
<i>Verrucosipora maris</i>	VLAAGGVVATSS....	AAQAAN	NPYERGP	NP	ETNAIEALRGPFYVAT	
<i>Xiangella phaseoli</i>	VLATGGVVATSS....	PAQAAN	NPYERGP	NP	ETNAIEALRGPFYVAT	
<i>Micromonospora echinospora</i>	LLATGGVLAAS....	GAQAAN	NPYERGP	NP	ETNAIEALRGPFYVAT	
<i>Asanoa ishikariensis</i>	AAVLAVGVAGVP....	NALAAD	NPYERGP	NP	ETTSILDAIRGPFYVAT	
<i>Micromonospora halophytica</i>	VLAAGGVVATSS....	GVQAAN	NPYERGP	NP	ETSAIEALRGPFYVAT	
<i>Micromonospora citrea</i>	ALAATAALAAP....	AAASAD	NPYERGP	NP	ETVALLEALRGPFYVAT	
<i>Micromonospora lupini</i>	TLLAAGVLTGPA....	PAQAAN	NPYERGP	NP	ETTAIEALRGPFYVAT	
<i>Micromonospora chokoriensis</i>	ALVAGVVLAGPS....	GAAQAAG	PYERGP	NP	ETTAIEALRGPFYVAT	

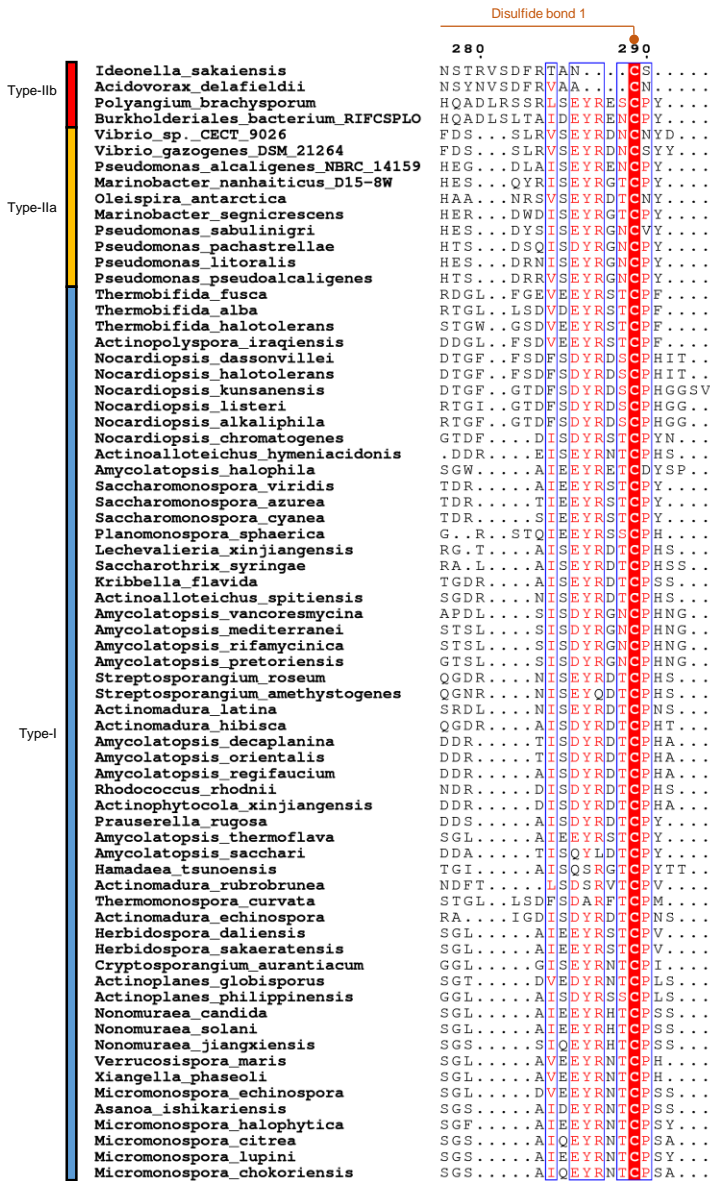
Supplementary Figure 5. Amino acid sequence alignment of PET-degrading enzymes (continued).



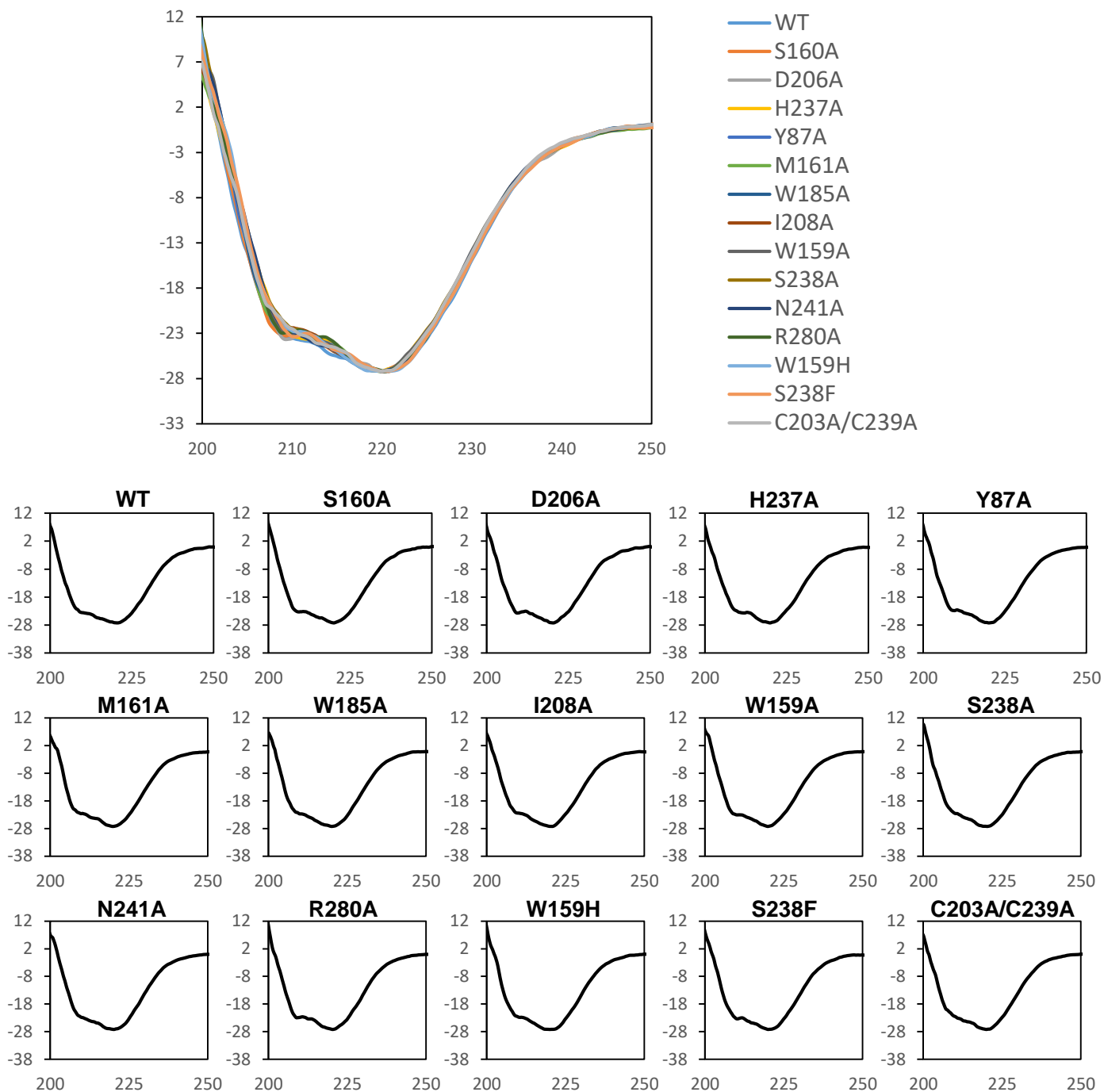
Supplementary Figure 5. Amino acid sequence alignment of PET-degrading enzymes (continued).



Supplementary Figure 5. Amino acid sequence alignment of PET-degrading enzymes (continued).



Supplementary Figure 5. Amino acid sequence alignment of PET-degrading enzymes.



Supplementary Figure 6. Far-UV Circular dichroism. Above figure includes superposed CD spectra of the *IsPETase* variants used in this study. Each spectrum of the variants is also shown in below. The spectra indicate that there are no significant difference among the variants and *IsPETase*^{WT}.

	<i>Is</i> PETase	<i>Is</i> PETase ^{R280A}
PDB code	5XJH	5YNS
Data collection		
Wavelength (Å)	0.97934	0.97934
Unit cell (<i>a</i> , <i>b</i> , <i>c</i> ; γ) (Å; °)	43.48, 50.40, 129.49; 90.0	43.61, 50.59, 129.58; 90.0
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Solvent content (%)	53.38	52.66
Protein chains in AU	1	1
Resolution range (Å)	50.00-1.55	50.00-1.36
Highest resolution shell (Å)	1.58-1.55	1.38-1.36
Unique reflections	42939	62434
Redundancy	6.5(5.9)	9.3(5.7)
Completeness (%)	99.9(99.9)	98.7(96.9)
<i>R</i> _{merge} (%)	6.8(29.8)	6.7(30.1)
Average <i>I</i> / σ (<i>I</i>)	31.4(5.2)	40.8(4.6)
B from Wilson plot (Å ²)	14.2	10.8
Refinement		
<i>R</i> (%)	15.9	16.6
<i>R</i> _{free} (%)	19.2	19.2
Mean B value (Å ²)*	16.1	15.0
RMS deviation bond lengths (Å)	0.026	0.025
RMS deviation bond angles (°)	2.286	2.249
Number of amino acid residues	272	264
Number of water molecules	150	294
Ramachandran plot		
Most favored regions (%)	97.3	97.6
Additional allowed Regions (%)	2.7	2.4

*Mean B value is for both protein atoms and the solvent molecules.

Supplementary Table 1. Data collection and refinement statistics of *Is*PETase.

Extra amino acids at N-terminus	MGSSHHHHHSSGLVPRGSHM
Codon optimized DNA sequence used for expression of <i>IsPETase</i>	CGCGGTCCGAATCCGACAGCCGCCAGTTTGGAAAG CGAGCGCTGGTCCATTCACCGTTCGCTCCTTTACC GTGAGTAGACCGAGCGGTTATGGCGCTGGCACCG TTTACTATCCAACAAATGCTGGGGGTACCGTGGGC GCCATAGCCATAGTTCCCGGGTATACGGCACGGCA GTCATCAATTAATGGTGGGGACCGCGTCTGGCAT CCCACGGTTTCGTAGTAATTACAATTGACACAAATT CCACGTTAGACCAGCCATCAAGTCGGAGTTCGCAA CAAATGGCCGCGCTGCGCCAGGTGGCGTCGTTAA ACGGTACAAGTAGCAGCCCGATTTACGGAAAGGTC GATACCGCTCGTATGGGTGTTATGGGGTGGAGTAT GGGAGGTGGAGGCTCCCTGATCTCTGCTGCTAAC AACCTTTCGCTGAAAGCAGCGGCGCCTCAAGCAC CATGGGATTCTTCGACAAATTTTAGTTCTGTAAGTG TGCCACGCTGATCTTCGCATGTGAAAACGATAGT ATAGCCCCGGTCAACTCTTCAGCACTTCCTATCTAT GATTCTATGTCACGCAACGCTAAGCAGTTTCTCGA AATTAATGGTGGCTCACATTCCTGTGCGAATAGCG GCAATTCTAACCAAGCATTAAATCGGAAAAAAGGC GTTGCATGGATGAAACGTTTTATGGACAATGATACT AGGTATTCTACTTTTGCCTGCGAGAACCCGAATAG CACCAGAGTGTCTGATTTTCGTACAGCGAATTGCA GC
Extra amino acids at C-terminus	LEDPAANKARKEAELAAATAEQ

Supplementary Table 2. Codon-optimized DNA sequence used for expression of *IsPETase* and extra amino acids at the N- and the C-termini.