Biochemistry

pubs.acs.org/biochemistry Article

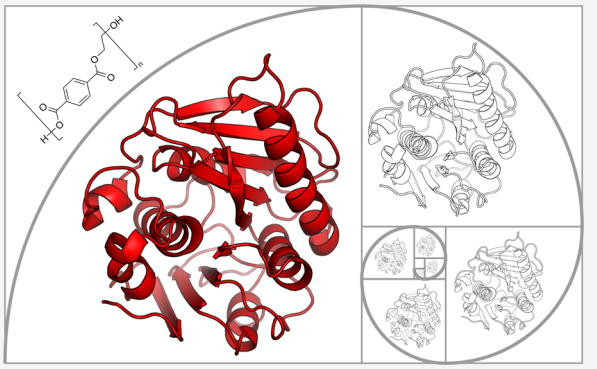
Ancestral Sequence Reconstruction Identifies Structural Changes Underlying the Evolution of *Ideonella sakaiensis* PETase and Variants with Improved Stability and Activity

Yvonne Joho, Vanessa Vongsouthi, Matthew A. Spence, Jennifer Ton, Chloe Gomez, Li Lynn Tan, Joe A. Kaczmarski, Alessandro T. Caputo, Santana Royan, Colin J. Jackson,* and Albert Ardevol*

Cite This: https://doi.org/10.1021/acs.biochem.2c00323 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The improved production, recycling, and removal of plastic waste, such as polyethylene terephthalate (PET), are pressing environmental and economic issues for society. Biocatalytic (enzymatic) PET depolymerization is potentially a sustainable, low-energy solution to PET recycling, especially when compared with current disposal methods such as landfills, incineration, or gasification. IsPETase has been extensively studied for its use in PET depolymerization; however, its evolution from cutinases is not fully understood, and most engineering studies have neglected the majority of the available sequence space remote from the active site. In this study, ancestral protein reconstruction (ASR) has been used to trace the evolutionary trajectory from ancient serine hydrolases to IsPETase, while ASR and the related design approach, protein repair one-stop shop, were used to identify enzyme variants with improved activity and stability. Kinetic and structural characterization of these variants reveals new insights into the evolution of PETase activity and the role of second-shell mutations around the active site. Among the designed and reconstructed variants, we identified several with melting points 20 °C higher than that of IsPETase and two variants with significantly higher catalytic activity.



One of the largest problems facing modern society is the increasing use and mismanagement of plastics. Over the past several decades, the production and use of plastics have dramatically increased to a reported 368 million tons worldwide in 2019.[1,2] The accumulation of plastic waste has impacted marine and terrestrial life, in terms of not only pollution but also health impacts due to microplastics entering the food chain,[3,4] microplastics acting as sinks for toxic chemicals and pollutants,[5] and possibly harmful additives used to modify the polymer properties.[6] Polyethylene terephthalate (PET) is among the most used plastics worldwide, comprising approximately 10% of annual plastic production.[7] PET is a semiaromatic thermoplastic polymer used in industrial settings for many uses, including engineering, packaging, and fibers, with a worldwide annual production of approximately 28 million tons.[8−12] Although the chemical stability and durability of PET allow it to be used ubiquitously, these properties also result in its persistence in landfills and the environment. Deposition in landfills, mechanical and chemical recycling, and incineration have many limitations that make them undesirable as long-term solutions to the problem of PET accumulation.

Due to the detrimental effects of plastics in our environment, and the need to transition to a low-carbon, circular economy, the development of better methods for the breakdown and recycling of plastics is vital. Thermal approaches such as incineration or gasification downcycle plastic waste by either burning (with the consequent release of greenhouse effect gases) or generating hydrocarbons that are subsequently also burned.[13] In contrast, enzymatic depolymerization of plastic is a low-energy, sustainable solution that can be used to upcycle postconsumer plastic waste into high-quality recycled PET (rPET).[14] Using enzymes, PET can be depolymerized into its constituent monomers, terephthalic acid (TPA) and ethylene glycol. One outstanding enzyme candidate in terms of its PET-degrading ability is PETase, isolated from *Ideonella sakaiensis* (IsPETase).[15,16] IsPETase was originally identified in the bacterium *I. sakaiensis*, which could utilize PET as a carbon source.[17] IsPETase degrades PET to a mixture of mono(2-hydroxyethyl) terephthalate (MHET) and TPA. A second enzyme, MHETase, then catalyzes the conversion of MHET to

Special Issue: Protein Engineering

Received: June 7, 2022
Revised: July 27, 2022

TPA and ethylene glycol.[18] IsPETase uses a catalytic triad of Ser160, His237, and Asp206, consistent with the general structure of cutinases, to cleave the ester bonds of PET.[19] Due to the potential biotechnological applications of IsPETase, many studies have recently investigated the structure and catalytic mechanism of the enzyme.[15,16,19] Several years prior to the discovery of IsPETase, several other related serine hydrolases (mostly annotated as cutinases) that also possessed PET-degrading activity had been identified from a range of thermophiles,[20] including a leaf-compost cutinase (LCC).[21] A recent review discussed the benefits of thermophilic and thermostable enzymes, as they are highly active at the glass transition temperature (72−74 °C), where crystalline PET becomes more amorphous and accessible for enzyme degradation.[22] More recently, other mesophilic PETases have been found in plastic-polluted oceans.[23]

For PETases to be applied industrially to aid in the depolymerization and recycling of PET, they typically need to be engineered to increase their kinetic and thermal stability, as well as their catalytic efficiency.[24−26] These approaches have ranged from random mutagenesis to computational design, machine learning, and rational design. Most mutated residues are targeted to modify the substrate binding site, the wobbling tryptophan residue and those in the vicinity of the disulfide bond in the active site.[15,19,27,28] As a result, numerous protein engineering efforts have succeeded in developing more stable and active variants, including DuraPETase, ThermoPETase, HotPETase, and FAST-PETase.[29−32] Recently, Lu et al. demonstrated that FAST-PETase could almost completely degrade untreated, postconsumer PET samples within 1 week and recover the monomers.[31] Previously, Tournier et al. produced a similar result using an engineered LCC variant that showed 90% depolymerization on thermally amorphized postconsumer PET within 10 h,[33] demonstrating how PETase and related enzymes may be implemented in industrial processes. The opportunities and limitations of addressing biotechnological plastic recycling have been discussed extensively elsewhere.[34]

Ancestral sequence reconstruction (ASR) is an approach that can be used to understand the evolutionary history of a particular protein and is increasingly being used in protein engineering[35] owing to its tendency to produce robust proteins with improved properties. Interestingly, ASR tends to produce many mutations in regions of the protein that are remote from the active site and that are currently beyond our ability to rationally engineer in a reliable manner. ASR uses a data set of extant protein sequences to generate a multiple-sequence alignment for constructing a phylogenetic tree. The sequences of ancestral internal nodes from such a phylogeny can be inferred using different reconstruction tools such as CodeML, GRASP, and others.[36,37] Recently, computational design has exploited the tendency for evolutionarily conserved amino acids (such as those that are often introduced through ASR) to stabilize proteins and combined this with force-field-based protein design, leading to the development of the Rosetta-based algorithm protein repair one-stop shop (PROSS),[38] which is an alternative approach for rapidly exploring a large proportion of sequence space for beneficial mutations.

In this study, we have applied ASR to study the evolution of IsPETase from the cutinase family of serine hydrolases. Using stability, activity, and structural analysis (by protein crystallography and homology modeling with AlphaFold),[39,40] we precisely map how substrate specificity has evolved to accommodate the PET polymer. Our results suggest that ancestors of IsPETase likely already possessed promiscuous PET-degrading activity and that this activity was optimized through evolution with a trade-off between thermostability and enhanced activity. Concurrently, we have applied ASR and the related approach of PROSS to generate variants of IsPETase with enhanced industrial characteristics.

MATERIALS AND METHODS

Ancestral Sequence Reconstruction. We performed two independent ancestral sequence reconstructions of *I. sakaiensis* PETase (GenBank accession number A0A0K8P6T7), once with GRASP[37] (*ASR1*) and the second time with PAML[36] (*ASR2*).

ASR1. The online EFI-Enzyme Similarity Tool[41] was used with the IsPETase sequence as input and ran with an *E* value of 10^{-5} set to a maximum number of 1000 retrieved sequences. A total of 914 unique sequences were found. Afterward, CD-Hit[42] was used with a threshold of 65%, reducing the data set to 76 sequences. Several known PET hydrolases were reintroduced, resulting in a final data set of 82 sequences. The sequences were aligned with MAFFT,[43] and signal peptides were identified and manually removed with SignalP4.0.[44] IQ-TREE2[45] and ModelFinder were used to determine the optimal model of sequence evolution, and WAG +F+R6 was found to be the best model. The tree was generated five times, and the best tree was used according to the AU test. Ancestral sequence reconstruction was performed using the online Web server GRASP (WAG model), providing the data set of 82 aligned sequences with the IQTREE2-generated file. Ancestral sequences were retrieved from the marginal reconstruction of GRASP.

ASR2. Sequences were retrieved from NCBI nonredundant protein sequence database bypBLAST using IsPETase as a seed sequence. Default pBLAST parameters were used. Redundancy within the sequence data set was removed to 80% sequence identity using CD-HIT, and signal peptides were identified and manually removed using SignalP4.0.[44] All sequences were aligned using PROMALS3[46] before the alignment was manually refined on the basis of a structural alignment of IsPETase [Protein Data Bank (PDB) entry 5XJH], LCC (PDB entry 4EBO), and TfCut2 (PDB entry 4CG1). Poorly conserved insertions were deleted, and sequences that were consistently poorly aligned were removed from the data set. The final data set had 147 sequences. Phylogenetic reconstruction was performed by maximum likelihood (ML) using IQTREE2. The sequence evolution model was selected by ML using ModelFinder, as implemented in IQTREE2. Tree search parameters were kept as the default, and branch supports were measured by ultrafast bootstrap approximation (ufboot) performed to 1000 replicates. Phylogenetic reconstruction was performed in replicates of 10, and the approximately unbiased (AU) test (conducted to 10000 replicates) was used to differentiate topologies that were significantly more likely than others. The single best tree was used for ASR. Ancestral states were reconstructed by ML using CodeML from the PAML package. The sequence evolution model was WAG+G4 (as CodeML does not support free-rate models), and α parameters were inferred from the data. Only ancestral sequences supported by a minimum mean posterior probability of 0.8 and a topological ufboot of 0.95 were experimentally tested.

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

Protein Repair One-Stop Shop (PROSS). The X-ray crystal structure of the PETase R280A mutant (PDB entry 5YNS)[19] was submitted to the PROSS[38] web server (http://pross.weizmann.ac.il/). The catalytic residues (S160, H237, and D206) and residues comprising the expected binding site (Y87, T88, A89, Q119, W159, M161, W185, S236, S238, N241, and S245) were not allowed to be mutated during the design process. The server produced seven combinatorial designs (PROSS1−7), containing increasing numbers of mutations relative to the input structure. Of these designs, we selected PROSS3−7 for experimental characterization.

Protein Expression and Purification. The GRASP ancestral sequences were synthesized at GenScript (containing a C-terminal His tag) in modified expression vector pET-43.1a (+) from Novagen. The CodeML and PROSS sequences were synthesized by IDT and cloned into pETMCSIII[47] by Gibson Assembly.[48] *Escherichia coli* strain BL21(DE3) was used for expression and induced by 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cultures were further incubated for ∼18 h at 16 °C and 180 rpm. Cells were harvested by centrifugation at 5000*g* for 20 min at 4 °C. The cell pellet was resuspended in lysis buffer [10 mM PBS, 350 mM NaCl, 10% (w/v) glycerol, 2 mM $MgCl_2$, and 20 mM imidazole (pH 7.5)] containing the recommended amounts of cOmplete EDTA-free Protease Inhibitor Cocktail (Roche) and Benzonase (Merck-Millipore). The cells were sonicated, and the lysate was filtered (0.22 μm). The ancestral enzymes were purified using a two-step method on an ÄKTAxpress instrument (GE Healthcare) using a prepacked 1 mL HisTrap FF column (GE Healthcare), followed by automated injection of the peak eluate onto a HiLoad Superdex 200 16/60 prep grade column (GE Healthcare). The immobilized metal affinity chromatography (IMAC) column was equilibrated, and samples were washed with IMAC buffer [10 mM PBS, 350 mM NaCl, 10% (w/v) glycerol, and 20 mM imidazole]. The proteins were eluted using an IMAC buffer supplemented with 500 mM imidazole. The Superdex 200 column was equilibrated using SEC buffer [50 mM bicine and 100 mM NaCl (pH 9)]. The corresponding fractions were pooled and concentrated, and the concentration was determined using a NanoDrop instrument at 280 nm according to values calculated using the Expasy, ProtParam tool.[49]

Enzymatic Activity of Ancestral Enzymes. Polyethylene terephthalate microparticles (<5 mm) were generated from commercial PET bottles by being snap-frozen in liquid nitrogen and then ground with a coffee grinder. The enzyme assay reaction was performed in a total volume of 600 μL using SEC buffer [50 mM bicine and 100 mM NaCl (pH 9)] with ≈109 mg of hc-PET microplastics. PET degradation was initiated by adding 500 nM enzyme, and the reaction mixture was incubated for 1 or 6 h at 1000 rpm and 30 °C. Enzymes with increased thermal stability were also tested at higher temperatures between 40 and 70 °C. The hydrolysis reactions were stopped at 90 °C, and the mixtures filtered (0.2 μM) for 2 min via centrifugation at 16110*g*. The PET degradation efficiency was adapted from previous methods.[50,51] Then, 100 μL of the reaction mixture was transferred to a black 96-well plate. The reaction was initiated by adding 25 μL of 5 mM EDTA and 25 μL of 5 mM $FeSO_4$ with a multichannel pipet. Finally, 25 μL of 2% H_2O_2 was added by mixing the reaction mixture. The reaction mixture was incubated for 60 min at room temperature. The fluorescence was then measured with a plate reader from CLARIOstar BMG-Labtech Australia (λ_{ex} = 320 nm, and λ_{em} = 420 nm). At the same time on each plate, three replicates of a calibration curve of terephthalic acid (TPA) between 0 and 400 μM were prepared to interpolate the concentration of TPA in the samples. SEC buffer was used as blank, and its fluorescence was subtracted from the fluorescence of all samples. All activity assays were performed in triplicate and measured in duplicate on the plate.

Thermostability Measurement. The thermostability of the purified PETase variants was determined using differential scanning fluorimetry (DSF) in the CSIRO C3 facilities. The enzymes were diluted to a concentration of 40 μg/mL in SEC buffer [50 mM bicine and 100 mM NaCl (pH 9.0)]. A 20 μL protein sample was pipetted into a 96-well plate, and 0.3 μL of 1× SYPRO orange dye was added using nanodispensing robots (Art Robbins). The thermal shift was observed using a real-time polymerase chain reaction machine (Thermo Fisher Scientific). Melting temperatures were determined from the first-derivative curves using Meltdown.[52] Protein samples were set up in six replicates each, and as an internal control, 0.1 mg/mL lysozyme was used.

Protein Crystallization and Structure Determination. Crystals of GrAnc8 were grown by vapor diffusion setting up drops at a 1:1 ratio with 150 nL of GrAnc8 at 2.2 mg/mL and 150 nL of mother liquor using the Phoenix Crystal (Art Robbins) liquid handling platform. Crystals grew at 293 K after a couple of days in a solution containing 1.2 M trisodium citrate (pH 8.3). Then, 1.6 M trisodium citrate (pH 8.3) was overlaid on the drop for cryoprotection prior to flash cooling in liquid nitrogen. Diffraction experiments were carried out at the Australian Synchrotron MX2 beamline[53] under nitrogen vapor at 100 K at a wavelength of 0.974 Å. Data reduction was carried out with XDS for indexing and integration followed by scaling and merging with Aimless.[54,55] Molecular replacement was carried out with Phaser via a model derived from PDB entry 6EQE.[56] The model was iteratively built and refined with Coot 0.8.9 and Refmac 5.7, respectively, via the CCP4i2 user interface.[57−59] For the rest of the ancestral sequences, we used AlphaFold to predict the three-dimensional protein structure.[39,40] Crystals of PROSS5 PETase were grown at 18 °C using the sitting-drop vapor diffusion method with a reservoir solution containing 0.08 M sodium cacodylate (pH 6.5), 20% (v/v) glycerol, 14.4% (w/v) PEG 8000, and 0.16 M calcium acetate. Drops were set up at a 1:1 reservoir:protein volume ratio. Crystals formed within 2 days and were flash frozen in liquid nitrogen after 14 days. Diffraction data were collected on the macromolecular crystallography beamline (MX2) at the Australian Synchrotron under nitrogen vapor at 100 K at a wavelength of 0.979 Å. Data were processed using XDS[55] and Aimless,[54,59,60] and molecular replacement was performed using Phaser.[56] Iterative cycles of manual model building and refinement were performed using Coot 0.8.9[57,61] and phenix.refine.[62]

Availability of Data. Crystal structures of GrAnc8 and PROSS5 have been deposited as PDB entries 8CRU and 8D1D, respectively.

RESULTS

Phylogenetic Analysis of the PETase/Cutinase Family. Although multiple serine hydrolases that can catalyze the depolymerization of PET have been identified, our study focused predominantly on IsPETase from *I. sakaiensis*. To test the robustness of our hypotheses against evolutionary uncertainty and to generate a range of potential ancestral

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

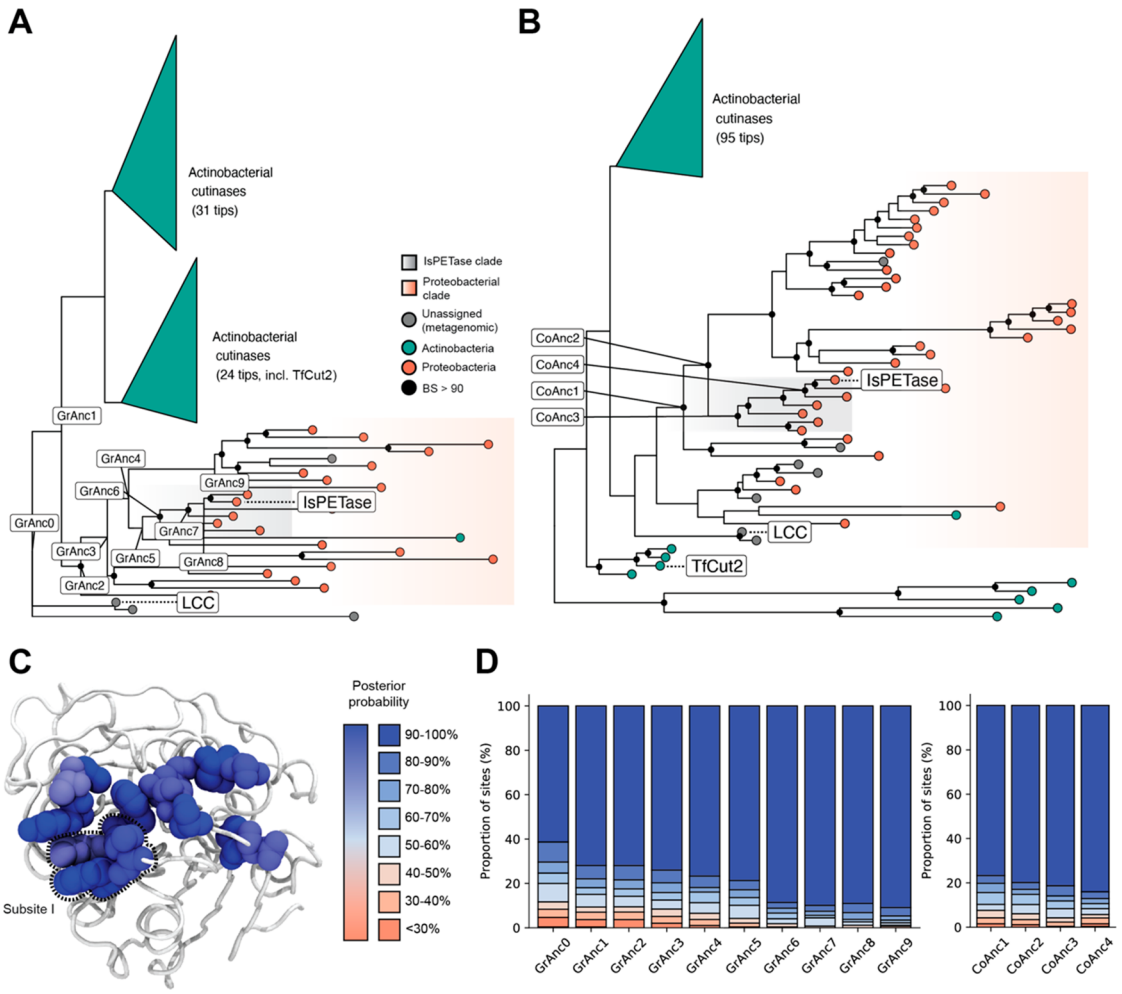


Figure 1. Phylogenetic and ancestral sequence reconstruction. (A) ASR1 data set reconstructed in GRASP (82 sequences). (B) ASR2 data set reconstructed in CodeML (147 sequences). Nodes that were experimentally characterized are labeled in both panels A and B. *I. sakaiensis* PETase is labeled (IsPETase), as well as cutinase-like PETases, leaf compost cutinase (LCC), and *T. fusca* cutinase-2 (TfCut2). The proteobacterial cutinase clade is highlighted, and tips are colored according to phylum. Nodes supported by an ultrafast bootstrap approximation of >95 are colored solid black. IsPETase belongs to a conserved lineage that diverged from ancestral proteobacterial cutinases independently of other PET hydrolytic enzymes. (C) Posterior probabilities for PET binding sites averaged over all reconstructed sequences. All residues belonging to subsites I and II are on average reconstructed with a high posterior probability and are strongly supported. Subsite I, including the catalytic S160, is highlighted. (D) Posterior probability distributions for all reconstructed sequences. GRASP and CodeML ancestors are all supported by high mean posterior probabilities and high per site probabilities at functionally important sites.

sequences, we reconstructed ancestral PETase sequences using two independent data sets of 82 and 147 sequences of related extant enzymes. We observed minor differences in the underlying tree topologies and sequence alignment; these were shown previously to have profound effects on both the properties of reconstructed enzymes and the interpretation of evolutionary results.[63,64] By using two independent data sets, we aimed to cross-validate our results across congruent, yet distinct, phylogenetic hypotheses. Maximum likelihood (ML) phylogenetic topologies were built from each respective data set using IQTREE (Figure 1A,C). For the same reason that we used two independent sequence data sets, we also reconstructed ancestral sequences by maximum likelihood (ML) using two separate algorithms, GRASP[37] and CodeML from the PAML suite.[36] To remain computationally efficient on large data sets, GRASP assumes that all sites in an alignment evolve at the same rate and therefore does not support the inclusion of rate heterogeneity parameters in the sequence evolution model. As models that lack rate heterogeneity parameters seldom outperform those that include them,[1,2] reconstructions performed in CodeML typically include a discretized Γ distribution in which evolutionary rates are sampled from during inference of ancestral characters. GRASP is further distinguished from CodeML in its partial order graph framework that allows probabilistic modeling of insertions and deletions, which are not explicitly considered in CodeML.[65,66]

Both sets of trees show that IsPETase diverged from an ancestral lineage that shared a common ancestor with actinobacterial cutinases, independently of other cutinase-like PET hydrolases, such as leaf branch compost cutinase (LCC)[21] and *Thermobifida fusca* cutinase-2 (TfCut2).[67] One clear distinction between the two phylogenetic topologies we investigated (Figure 1A,C) is the placement of the proteobacterial cutinases (including IsPETase) within the broader cutinase family. Indeed, the actinobacterial cutinases are reconstructed as both mono- and polyphyletic across both data sets. Notably, the topology that reconstructs the

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

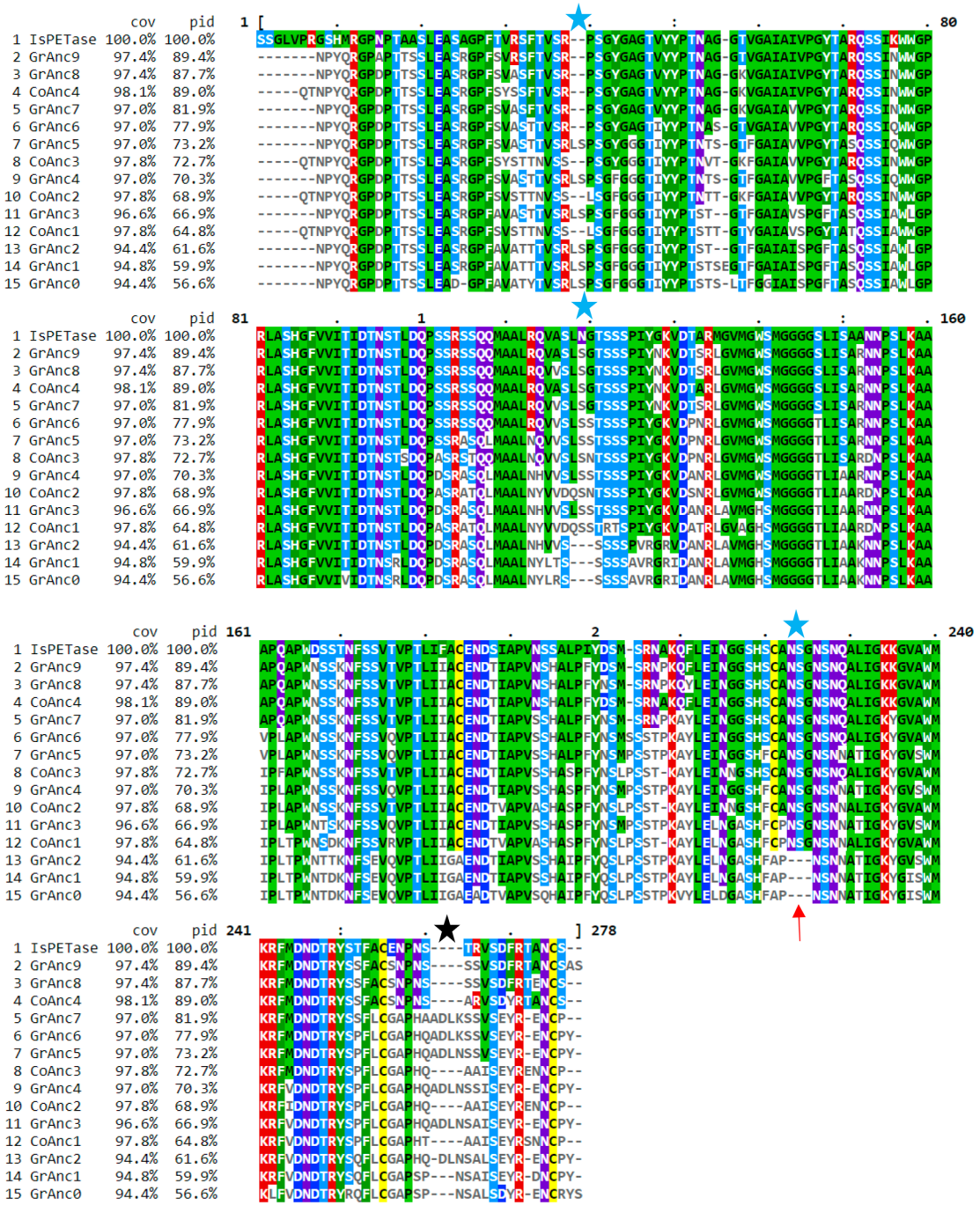


Figure 2. Multiple-sequence alignment of IsPETase with GRASP- and CodeML-generated ancestral sequence predictions. The alignment was created with MAFFT using the default parameters on the EMBL-EBI web server and then colored by MView. A high degree of similarity of the predicted sequences is observed, excluding the C-terminal region with various insertions and deletions (black star). Furthermore, the blue stars denote three other regions with two or three residues inserted, one of those belonging to the IsPETase specific extended loop (red arrow).

proteobacterial cutinase clade as a sister clade to a monophyletic actinobacterial cutinase lineage (the 82-sequence data set) does so with worse branch support [ultrafast bootstrap approximation of <90 for the LCA of the proteobacterial cutinases (Figure 1 and Supplementary Figure 1)] than the alternative 147-sequence data set, indicating less certainty in the placement of these sequences. As our analyses focused on evolution within the proteobacterial lineage, it is unlikely that the placement of the actinobacterial cutinases may significantly alter the reconstruction of ancestral sequences beyond the LCA of the proteobacterial cutinases (GrAnc0); however, it may have ramifications on how the prokaryotic cutinase family has diverged overall.

To investigate the molecular mechanisms of PETase evolution, we reconstructed the sequences of the ancestral nodes spanning from the last common ancestor (LCA) of proteobacterial cutinases to extant IsPETase in both our CodeML (CoAnc1−4) and GRASP (GrAnc0−9) ASR data sets. Because the actinobacterial cutinases are reconstructed as polyphyletic clades in the 147-sequence data set, this means that the LCA of the proteobacterial sequences is closer in evolutionary distance; thus, while 10 sequences were

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

reconstructed in the 82-sequence monophyletic data set (to the shared LCA of actinobacterial and proteobacterial sequences), only four ancestral sequences were reconstructed in the 147-sequence polyphyletic data set. The reconstructed sequences were aligned (Figure 2) and varied by amino acid identity (relative to wild-type IsPETase) between 58.6−91.3% (GrAnc0−9) and 65.3−89.4% (CoAnc1−4). Additionally, all ancestral proteins were reconstructed with relatively high posterior probabilities (Figure 1B,D) both overall and at sites known *a priori* to be involved in PETase activity (Figure 2), and from nodes that were supported by high bootstrap values, suggesting a high degree of statistical confidence in our reconstructions. The alternative reconstructions are strikingly similar at approximately similar positions in the two trees. For example, CoAnc4 differs from GrAnc8 at only 11 positions (Supplementary Table 1), 10 of which are positions with low posterior probability in the CodeML reconstruction. Likewise, CoAnc3 and CoAnc4 are very similar to GrAnc4 and GrAnc5, and CoAnc1 is very similar to GrAnc3. Thus, although these sequences were reconstructed using different trees and different reconstruction algorithms, they can together be viewed as 14 alternative reconstructions that sample closely related sequence space that was likely traversed along this evolutionary trajectory.

Protein Repair One-Stop Shop (PROSS). An alternative approach to ASR for the evolutionarily guided redesign of enzymes is the Rosetta-based algorithm PROSS.[38] In this approach, evolutionarily conserved mutations are incorporated into the gene and the effects of these mutations on stability are calculated using Rosetta-based energy calculations. Here, we used the structure of the R280A IsPETase mutant as our initial model (PDB entry 5YNS), a variant of IsPETase demonstrated to have improved activity compared to that of the wild type in a previous study.[19] Using this structure as the input, and fixing the residues of the active site, we obtained seven designs from the PROSS server (PROSS1−7) that incorporated between 6 and 27 mutations (Supplementary Table 2 and Supplementary Figure 2). Inspection of these designs revealed that two of the mutations common to PROSS1−7 (A152S and S214H) were also present in the reconstructed ancestors. Furthermore, some PROSS designs were mutated at positions common to the reconstructed ancestors, however, to different residues (D186S and T270Q in PROSS vs D186H and T270S in the ancestors). Given that designs PROSS1 and PROSS2 possessed only six and eight mutations, respectively, and that these mutations were also present in PROSS3−7, we synthesized variants PROSS3−7 for experimental characterization.

Expression and Purification of Reconstructed and Designed PETases. The most likely (by maximum likelihood reconstruction) ancestral enzymes were synthesized, cloned, and heterologously expressed in *E. coli* BL21(DE3) cells. Of the 10 sequences reconstructed using GRASP, all proteins were expressed in soluble form. Notably, for some of the deeper nodes, we observed soluble protein yields >30-fold greater than that of IsPETase (Table 1 and Supplementary Figure 3). In contrast, the nodes closer to IsPETase resulted in lower soluble protein yields. Of the four sequences reconstructed using CodeML, CoAnc4 and CoAnc1 did not yield soluble protein, suggesting that there are issues related to protein folding for these designs, and only CoAnc2 and CoAnc3 were expressed in soluble form. For the PROSS designs, PROSS3 and PROSS6 did not yield soluble protein. Altogether, this produced a data set of 15 purified PETase designs containing from 9 to >50 mutations relative to IsPETase, which is a significant exploration of sequence space.

Table 1. Summary of Thermostabilities, Activities, Theoretical pI Values, and Soluble Yields for All PETase Variants[a]

variant	T_m (°C)	activity (μM TPA/6 h)	pI (calculated)	soluble yield (mg/L)
IsPETase	43.16 ± 0.23	119.35 ± 23.85	9.25	0.5
GrAnc9	61.52 ± 0.25	22.66 ± 2.80	9.76	0.1
GrAnc8	63.18 ± 0.08	27.46 ± 0.68	9.59	1.4
GrAnc7	46.14 ± 0.11	16.83 ± 0.95	9.43	1.4
GrAnc6	50.94 ± 0.09	16.48 ± 0.51	9.19	13.4
GrAnc5	45.55 ± 1.24	10.28 ± 0.16	8.67	1.0
GrAnc4	52.82 ± 0.41	16.20 ± 1.03	8.37	8.0
GrAnc3	50.28 ± 0.74	10.88 ± 0.93	8.38	9.6
GrAnc2	50.86 ± 0.17	17.64 ± 4.40	8.57	9.4
GrAnc1	55.76 ± 0.18	2.79 ± 0.41	8.56	0.4
GrAnc0	48.0 ± 0.09	2.30 ± 0.35	8.54	16.8
CoAnc3	64.38 ± 0.05	–	8.87	0.4
CoAnc2	47.60 ± 0.23	–	8.67	2.7
PROSS4	62.85 ± 0.11	5.54 ± 0.31	7.70	1.6
PROSS5	65.32 ± 0.08	9.48 ± 3.82	8.35	1.3
PROSS7	45.1 ± 0.88	–	6.50	3.8

[a]The melting temperature was determined by differential scanning fluorimetry (DSF). To determine the PET-degrading activity, the concentration of TPA (micromolar) was measured after the enzyme had been incubated with PET for 6 h at 30 °C and 1000 rpm. A dash indicates no measurable TPA was produced by the variant. The pI was calculated with the Expasy ProtParam tool.[49]

Thermostability of Reconstructed and Designed PETases. Thermostability measurements of all 15 synthetic enzymes were obtained from the unfolding midpoint temperature (T_m) using differential scanning fluorimetry (Table 1 and Supplementary Figure 4). The T_m value obtained for wild-type IsPETase was 43 °C, consistent with other reports.[32,68] Several ancestral sequences, reconstructed using CodeML and GRASP, as well as PROSS designs, exhibited T_m values approximately 20 °C higher than that of the wild-type sequence (Table 1 and Figure 3A). Although previous studies have suggested that deeper nodes that are reconstructed using ASR often have higher thermostabilities,[69] in this experiment the most stable variants were relatively close to the wild-type sequence, specifically GrAnc8, GrAnc9, and CoAnc3. The dramatic decrease in thermostability between GrAnc8, GrAnc9, and CoAnc3 and closely related enzymes, such as CoAnc2 and IsPETase, suggests that these differences in thermostability could be due to a small number of mutations that were introduced recently in evolutionary history. Indeed, the same trend is observed with the PROSS variants. While PROSS4 and PROSS5 display T_m values of 63 and 65 °C, respectively, PROSS7 and IsPETase have T_m values of 43−45 °C. These results also suggest that IsPETase is unlikely to have evolved from a recent thermophilic ancestor, which is notable given that most of the other PETases from this family are thermophilic.[20,21] Comparatively, IsPETase is a typical mesophilic enzyme with lower thermostability, although it is considerably more active at low temperatures than other thermophilic enzymes (LCC, TfH, and FsC), which have little to no activity at mesophilic temperatures.[17]

Kinetic Characterization of Ancestral and Designed PETase Activity. One of the most intriguing questions when

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

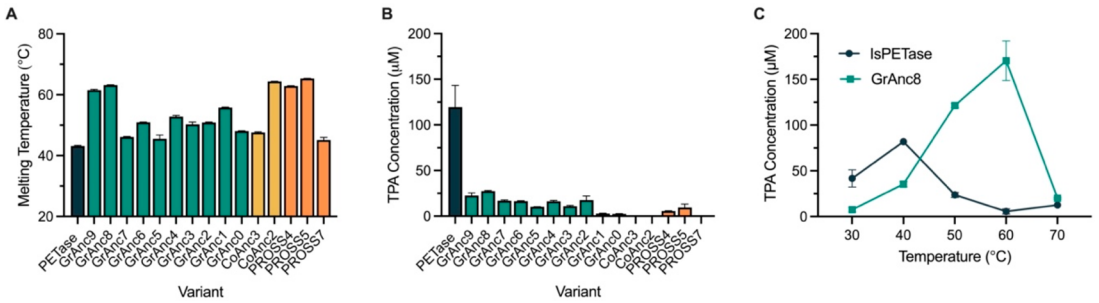


Figure 3. (A) Melting temperatures of GRASP and CodeML variants determined by differential scanning fluorimetry. (B) PET-degrading activity of the GRASP and CodeML enzymes. The terephthalic acid (TPA) concentration was measured after incubation for 6 h with PET at 30 °C and 1000 rpm. (C) Kinetic stability of IsPETase vs GrAnc8. The TPA concentration was measured after incubation for 1 h with PET at 30−70 °C and 1000 rpm. Data are means ± the standard deviation (n = 3).

it comes to the evolution of IsPETase is whether the new PET-degrading activity is a novel gain of function or whether it represents an increase in a preexisting promiscuous activity. As PET has been in the environment for only ∼80 years and all of the reconstructed ancestors are representatives of enzymes that likely existed well before this time, ASR represents a useful method for testing this hypothesis. Accordingly, we measured the catalytic activity of the ancestral enzymes compared to that of the wild type after 6 h at 30 °C (Figure 3B and Table 1). Interestingly, the ancestors from the deepest nodes, GrAnc0 and -1, showed almost no activity against the PET substrate, whereas GrAnc2−9 exhibited promiscuous PETase activity (because PET would not have been present in the environment). However, all ancestors were significantly less active than IsPETase, and there was no consistent increase along the trajectory, in accordance with the expectation that there would be no selective pressure to evolve efficient PETases prior to the synthesis of PET plastic in the 20th century. The ancestors reconstructed using CodeML, CoAnc2 and CoAnc3, did not display any PET hydrolase activity at 30 °C. Given the issues related to expressing other ancestors reconstructed with this phylogeny and/or algorithm, it is possible that this reflects some issues related to the reconstruction (see Discussion).

The PROSS variants revealed some interesting results; despite having far fewer mutations than the ASR variants and despite "fixing" the active site and binding site to the wild-type IsPETase sequence, these variants exhibited considerable (>10-fold) decreases in catalytic activity. This is not entirely due to stability−activity trade-offs, because the PROSS7 variant, which had wild-type-like thermostability, also has no detectable PETase activity.

Altogether, these results highlight a number of interesting observations. (i) Mutations were introduced between GrAnc1 and GrAnc2−9 that conferred upon these enzymes promiscuous PETase activity (this coincides with the separation of the proteobacterial and actinobacterial clades). (ii) Between GrAnc9 and IsPETase, mutations are introduced that result in an ∼10-fold increase in activity and an ∼20 °C decrease in stability. (iii) Despite being extremely close in sequence space (only 10 mutations), CoAnc4 has no detectable activity while GrAnc8 and -9 have significant promiscuous activity. (iv) PROSS4, PROSS5, and PROSS7 have lost significant activity relative to IsPETase despite comparatively few mutations.

Having investigated the activity at 30 °C, we questioned the activity of the most stable and well-expressed ancestor (GrAnc8) at an increased temperature to assess its potential industrial applications. GrAnc8 showed significantly increased turnover rates (as expected, given the increased temperature of the reaction), and we determined the optimal temperature to be 60 °C (Figure 3C). In contrast, IsPETase lost activity beyond 40 °C. As IsPETase could not retain activity at increased temperatures, GrAnc8 had more than twice the maximum turnover rate of IsPETase and represents a significantly enhanced variant for potential industrial use, or an interesting starting point for further engineering.

Molecular Basis for the Functional Transition from Cutinase to PETase. To understand the effects of these mutations, we determined crystal structures of key variants: GrAnc8, the best performing ancestor in terms of stability and activity, and PROSS5, which retained some activity but gained considerable thermostability (Table 2). Both structures diffracted to relatively high resolution (∼1.3−1.4 Å), and all

Table 2. Data Collection and Refinement Statistics of the Crystal Structures of GrAnc8 and PROSS5

	Data Collection[a]	
space group	*P*3121	*C*121
cell dimensions		
a, *b*, *c* (Å)	86.01, 86.01, 88.92	136.57, 159.17, 45.04
α, *β*, *γ* (deg)	90, 90, 120	90, 90.07, 90
resolution (Å)	44.46−1.30 (1.32−1.30)	39.2−1.42 (1.47−1.42)
R_{meas}	0.145 (1.810)	0.1548 (1.01)
R_{pim}	0.032 (0.404)	0.06 (0.409)
$I/\sigma I$	11.7 (1.4)	7.1 (1.0)
$CC_{1/2}$	0.999 (0.747)	0.991 (0.511)
completeness (%)	99.8 (95.6)	98.7 (95.9)
multiplicity	20.3 (19.2)	6.6 (5.9)
	Refinement	
resolution (Å)	43.044−1.30	39.2−1.423
no. of unique reflections	88786	177008
R_{work}/R_{free}	0.1105/0.1331	0.1975/0.2341
no. of atoms	2037	6361
B-factor (Å^2)		
protein	13.16	17.63
ligand	29.32	18.35
water	35.96	29.96
rmsd for bond lengths (Å)	0.0134	0.008
rmsd for bond angles (deg)	1.72	0.94
Ramachandran favored/ outliers (%)	98.47/0.0	98.08/0.0
Clashscore	1.23	2.24

[a]Data for the highest-resolution shell are shown in parentheses.

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

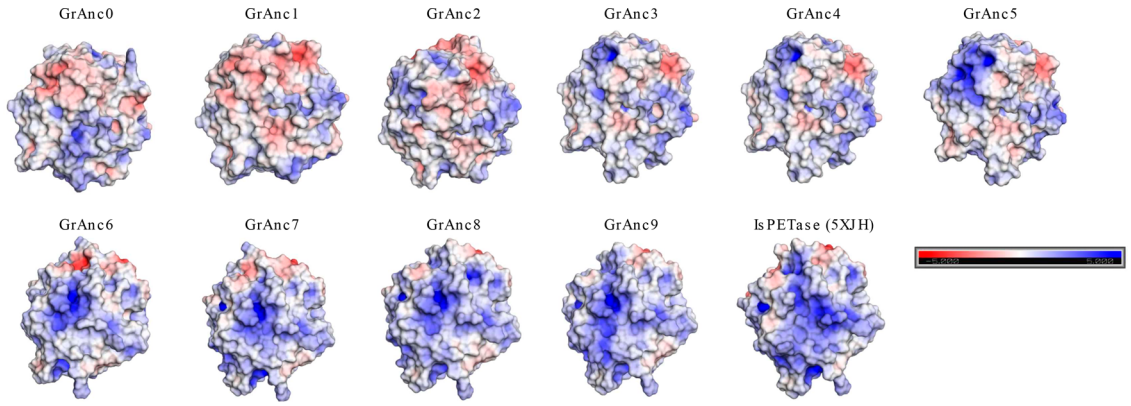


Figure 4. Overview of the surface electrostatic potential changes that occurred during the evolution of IsPETase. The electrostatic surface potential was calculated with the APBS tool[72] in PyMOL[73] showing red for electrostatically negative (−5) and blue for positive (5). The figure shows the surface charge from ancestral enzymes with evenly distributed positive and negative patches to a strongly positive surface electrostatic of more recent ancestors.

residues were resolved in clear density. Alphafold2 was used to obtain models of all of the reconstructed sequences, all with a relatively high degree of confidence given their similarity to the available experimental structures of wild-type IsPETase. The root-mean-square deviation (rmsd) between the crystal structure and the predicted model of GrAnc8, calculated using the Cα atoms, was 0.3 Å.

The various experimental and computational models displayed a high degree of similarity, excluding a small number of regions where short insertions or deletions occurred, as expected from the multiple-sequence alignment (Figure 2). In particular, the regions at the C-terminus seem to be highly variable. We used Consurf[70] to evaluate the evolutionary conservation of different regions of IsPETase (Supplementary Figure 5). The results showed significant variation of surface-exposed residues, especially those lining the active site. This is consistent with previous work that noted that IsPETase has a unique surface profile, having a highly polarized surface charge compared to those of other related enzymes (including cutinases with assumed promiscuous PETase activity).[27] Accordingly, we calculated the surface potential of this evolutionary trajectory to observe how it changed over time. The GRASP data set was most useful for this analysis given its greater completeness and depth. Interestingly, we observed that the electrostatic surface potential of the earlier nodes, which largely lacked any PETase activity, is significantly lower, particularly in the substrate binding groove, which is highly cationic in PETase (Figure 4). From GrAnc2 onward, this channel becomes more cationic due to the accumulation of various substitutions. Moreover, we observe a general negative trend of expression level the closer the ancestor is to IsPETase (Table 1); a trade-off between soluble expression and high pI values has been observed previously in other proteins.[71]

Structural Basis for Thermostability Trade-offs in the Evolution of PETase. We observe a clear difference in the thermostability of some variants produced in this study; these changes often correlate with the gain or loss of InDels or specific point mutations. There are three events in the GRASP trajectory that are particularly noteworthy. First, between GrAnc2 and GrAnc3, a second disulfide bridge (in addition to C273−C289) at C203−C239 is introduced. Although a disulfide such as this is expected to confer significant stability to the folded state,[74] it appears to have a minimal effect on the thermostability of the enzyme, as the two variants have melting temperatures of 50.9 and 50.3 °C, respectively. However, this is predicted to have coincided with the introduction of two insertions, at positions 136−138 and a very close insertion at positions 242−244. The disulfide stabilizes the active site loop on which D206 and H237 of the catalytic triad are located, as has been previously observed.[75] Second, there is a significant increase in stability (from 46 to 63 °C) between GrAnc7 and GrAnc8 and -9 (Table 1). Inspection of the alignment of these sequences suggests this may result from the deletion of a short segment of a surface-exposed loop (ADLK) between S278 and T279. There are otherwise very few sequence differences between GrAnc7 and GrAnc8 that would explain such a large change in thermostability, and it is well-known that truncation of surface-exposed loops can result in increased thermostability.[76] Third, we observe a large decrease in thermostability during the adaptive process from GrAnc9 to IsPETase. In this case, there are no InDels, but there are a moderate number of point mutations, especially around the active site and substrate binding site: A37N, T40A, S41A, R47A, S51T, N95K, S138N, N147G, S152A, L154M, R171A, N186D, K189T, I201F, T207S, H214S, F218I, P226A, S270T, S274E, S279T, and S280R. There is nothing obviously destabilizing in this series of mutations, but the accumulation of small destabilizing effects can be significant.

The comparisons of IsPETase versus PROSS5 reveal some of the effects of mutations at key positions on thermostability. There are several mutations between IsPETase and PROSS5: R90T, S125D, R132D, V134L, G139N, T140D, A152S, Q182M, D186S, S214H, D220N, and T270Q. Notably, positions 152, 186, 214, and 270 were replicated from the ASR. Several of these were then tested as point mutations in IsPETase, of which only D186S was significantly stabilizing (by ∼9 °C) (Supplementary Figure 6). Interestingly, a separate study demonstrated that the mutation of this same position to histidine (D186H) improved the thermostability of IsPETase by stabilizing the β6−β7 connecting loop.[32] Indeed, inspection of the crystal structure for PROSS5 reveals that the side chain of S186 forms a hydrogen bond to the backbone of S188 to stabilize the loop region, providing a structural basis for the increased thermostability.

Structural Basis for the Evolution of PETase Activity. We observe many changes across this evolutionary trajectory; however, as expected given that PET is a recently introduced synthetic molecule, we see a large increase in activity at

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

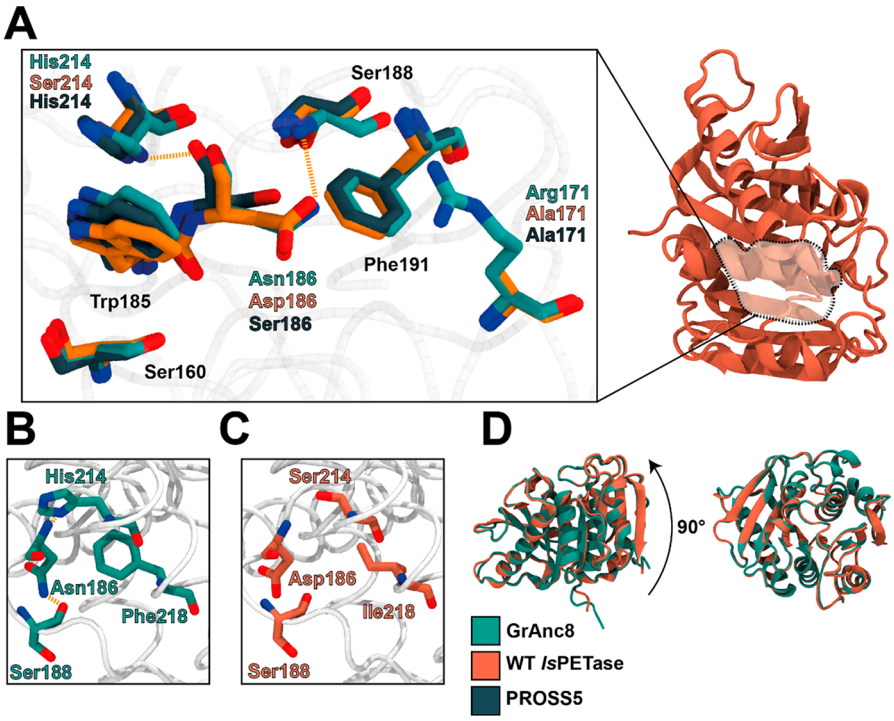


Figure 5. Structural analysis of IsPETase, GrAnc8, and PROSS5 PETase variants. (A) Structural differences among IsPETase, GrAnc8, and PROSS5 near the active site. GrAnc8, IsPETase, and PROSS5 are colored cyan, orange, and green, respectively. Residues important for PETase activity are labeled and colored. The WT IsPETase active site is broader than that of close relatives GrAnc8 and PROSS5 due to the presence of His214 in PROSS5 and GrAnc8 and Arg171 in GrAnc4 (Ser214 and Ala171, respectively, in IsPETase). Mutations Arg171Ala, Asn186Asp, and His214Ser fixed in IsPETase eliminate two hydrogen bonds in the active site (Asn186 with His214 and Asn186 with Ser188) as well as a cation−π interaction (between Arg171 and Phe191), destabilizing the protein to allow a broader and less constrained active site that can accommodate PET. (B and C) Second- and third-shell residues of GrAnc8 and WT IsPETase, respectively. The Phe218Ile mutation in IsPETase broadens the active site for PET binding. (D) Superimposition of the crystal structures of IsPETase (PDB entry 5XJH) and GrAnc8. The overall fold and structure of GrAnc8 are nearly identical to those of IsPETase, consistent with few first- and second-shell mutations driving PETase evolution.

physiologically relevant temperatures only in the extant IsPETase. The structure of GrAnc8 gives us an excellent model of a closely related enzyme that represents a recent ancestor, before the adaptation toward PET. First, there are almost no differences in the position of the Cα residues [0.299 Å (Figure 5)]. The same mutations that result in a decrease in stability from GrAnc8 to IsPETase result in an ∼5-fold increase in PETase activity. It is notable that of the mutations that result in altered charge, contrary to expectation, four result in the loss of cationic residues (R47A, R71A, K189N, and H214S), two result in the gain of anionic residues (N186D and S274E), and only one (N95K) results in the gain of a cationic residue and is remote from the active site. Thus, the altered surface charge, relative to those of other related cutinases, appears to be unrelated to any functional adaptation of IsPETase toward its new function.

Of the sequence differences between GrAnc8 and IsPETase, there is a cluster around the active site that can be rationalized to influence activity (Figure 5). R171A provides additional space behind the loop upon which W185 sits and results in the loss of potential to form a cation-π interaction with F164 and/or a long distance (∼5 Å) electrostatic interaction with D186. The N186D mutation removes a stabilizing H-bond to the main chain carbonyl of S161. H214S appears to be the most obviously beneficial mutation, removing a steric obstruction of W185 and giving it freedom to "wobble" and accommodate the PET chain.[77,78] F218I reduces the extent of hydrophobic packing of the helix neighboring W185. The PROSS5 structure is useful in testing these predictions. The activity of PROSS5 is reduced ∼10-fold relative to that of IsPETase, due to only a handful of mutations. Again, the main chain atoms are essentially invariant (rmsd ∼ 0.196 Å). This loss of activity is likely to be mostly due to the S214H mutation that has previously been shown to reduce catalytic activity while enhancing thermostability.[15,29,79] Of the other mutations in the vicinity, R132D (−8.2 °C) and G139N (−4.4 °C) are marginally destabilizing when introduced individually into IsPETase, while D186S provides higher thermostability (+9.1 °C) and higher hydrolytic activity at 45 °C (Supplementary Figure 6).

DISCUSSION

This study aimed to understand the evolution of IsPETase, an enzyme discovered within the first bacteria found to be capable of utilizing PET as their main carbon source (*I. sakaiensis*).[17] Indeed, how IsPETase has evolved its PET-degrading activity has been questioned before;[80] however, detailed structural analysis and phylogenetic exploration of ancestral sequence space have not been combined to investigate this. Toward this end, we used ASR to infer and reconstruct 12 plausible ancestral PETases. The results indicate that PET-degrading activity is a function that has recently evolved, although earlier ancestors generated by GRASP (GrAnc2−9) displayed low levels of promiscuous PET-degrading activity, i.e., the activity pre-existed as a promiscuous function, which then evolved to become the primary function, as demonstrated and proposed by Tawfik et al.[81−83] The ancestral sequences generated in this

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

work can be viewed as "likely" possible ancestral sequences; there is a likelihood that evolution has sampled many similar sequences over time. Thus, the finding in this work that some (not all) enzymes that are ancestral to modern PETases could have possessed promiscuous PETase activity is consistent with the hypothesis that the PETase activity in actinobacterial thermophilic cutinases could be a promiscuous, rather than a physiologically adapted, function, given the observation that, unlike *I. sakaiensis* that can assimilate PET via PETase and MHETase,[17] there is less evidence the thermophilic cuntinases utilize PET as a carbon source.

Our structural analysis reveals that the reconstructed PETase ancestors/PROSS variants (represented by GrAnc8 and PROSS5) have almost identical overall structures (Cα rmsd ~ 0.3 Å) but differ from IsPETase by several significant changes near the substrate binding site, such as R171A, N186D, H214S, and F218I mutations (Figure 5). Residues S214 and I218 have been previously identified to broaden the substrate binding site of IsPETase and increase the flexibility of W185.[77] The effects of R171A and N186D, which are "second-shell" remote mutations, are more subtle but entirely consistent with our understanding of the mechanism of IsPETase, and a growing appreciation of the role of second-shell mutations in controlling the conformational sampling of active site residues.[84,85] In IsPETase, these mutations appear to affect the ability of W185 to sample catalytically productive conformations by weakening the structural constraints on the active site (loss of hydrogen bonds and salt bridges) (Figure 5). Thus, the structural basis for the evolution of the new activity of IsPETase appears to conform to a model in which, rather than wholesale changes to an active site, new activity is gained by remote mutations increasing the level of sampling of catalytically productive conformation(s) of active site residues that already existed (and conferred promiscuous activity) and for catalytically nonproductive conformations to be sampled less often.[86]

The activity−stability trade-offs observed here are also notable and relate to another area that Tawfik et al. have studied in depth.[87] Both ASR and PROSS generated a number of variants that were significantly more thermostable (T_m > 60 °C) than IsPETase (T_m ~ 43 °C) (Table 1). These variants all displayed reduced catalytic activity at 30 °C compared to that of IsPETase. This can be partly attributed to the shared S214H substitution that has been previously shown to reduce PETase activity (and increase stability) in IsPETase.[29,78] The thermostable ancestors and PROSS designs share an additional mutation (A152S) and two mutations of the same positions to different residues (D186S and T270Q in PROSS designs vs D186N and T270S in the ancestral reconstructions). The D186S and D186N substitutions are a relatively clear example of how stability and activity can trade off. The D186N substitution (relative to IsPETase) results in the gain of a hydrogen bond to the main chain carbonyl of S188; while increasing stability, this results in less conformational flexibility by constraining the active site loop that accommodates W185, thus affecting the sampling of its productive "wobble" rotamers. The D186S mutation in IsPETase is likewise interesting in this context; it results in a 9 °C increase in thermostability, owing to the hydrogen bond between the Ser side chain and the carbonyl of S188, but reduces activity at 30 °C. However, at 45 °C, a temperature at which this Ser−main chain hydrogen bond is much weaker, we observe a considerable (~3−4-fold) increase in activity, suggesting that many of these mutations might have been specifically accumulated to increase activity at the mesophilic temperature by "loosening" the active site.

In addition to studying the evolution of IsPETase activity using ASR and the related protein design algorithm PROSS, we were interested in assessing the ability of ASR and PROSS to identify PETase variants with improved solubility, thermostability, and/or activity. In comparison to other studies that have screened (experimentally or computationally tens to hundreds of thousands of variants),[29−31] we screened fewer than 20 mutants in this study. With no optimization or fine-tuning, we identified variants with ~16−34-fold improvements in the yield of soluble protein relative to that of IsPETase, thermostable variants with up to 20 °C higher T_m values, and variants with several-fold higher PETase activity (Table 1 and Supplementary Figure 6). Interestingly, many of the mutations identified by ASR and PROSS in this study are common to those that were spontaneously identified in past engineering efforts. For example, PROSS5, the most thermostable variant from our study (T_m = 65.3 °C), shares several mutations with HotPETase[30] (R90T, Q182M, T270Q, and R280A) and DuraPETase[29] (T140D, S214H, and R280A). It is interesting that many of these spontaneous mutations in the high-throughput screens are found at the same positions highlighted by these evolutionary approaches, suggesting that there are regions within the protein fold that are particularly "plastic" and capable of conferring functional change without dramatically compromising the fold integrity.[88]

Finally, this study provided an interesting side-by-side comparison of different ASR approaches and highlights the effect a small number of differences can have on protein stability and activity. The observation that CoAnc4 was not expressed at all in soluble form, whereas GrAnc8 was expressed and was highly thermostable, was notable, given that they were reconstructed from topologically equivalent nodes from their respective data sets, being the LCA of IsPETase and its two closest homologues. Indeed, CoAnc4 and GrAnc8 share 95.1% sequence identity, meaning that at most only 13 mutations (Q1−, T2−, Y25V, S26A, A108V, A125S, D193N, A199P, F202Y, A252S, R253S, Y257F, and A260E, relative to CoAnc4) are responsible for the complete loss of expression in CoAnc4. Interestingly, each of these positions (excluding Y202F and F257Y) is a solvent-accessible, surface mutation. Except for A108 and D193, each is reconstructed in CoAnc4 with low posterior probabilities ranging between 9.9% and 37%, meaning that multiple alternative residues (including those that were introduced into GrAnc8) with similar statistical support are possible at each of these sites (Supplementary Table 2). This observation emphasizes the importance of including multiple and independent data sets in the reconstruction process; i.e., residues reconstructed poorly in one approach (CodeML ASR2 data set) are responsible for the difference between a protein that cannot be expressed in soluble form and one of the most stable and highly active variants generated in this study.

To directly compare GRASP[3] and PAML[4] ASR approaches, we repeated the analysis of the ASR1 data set in CodeML and that of the ASR2 data set in GRASP, i.e., swapped trees and reconstruction methods. This revealed that the reconstruction methods had a weaker effect on the sequences of the reconstructed ancestors, relative to the differences caused through the different phylogenetic tree reconstructions. For example, GrAnc8 and CoAnc4, which are predicted to

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

reconstruct comparable nodes in the two trees, differ by 13 mutations, yet when the sequence of GrAnc8 is reconstructed using CodeML or that of CoAnc4 is reconstructed using GRASP, we observe only five and six positions, respectively, at which the sequences are reconstructed differently (as well as some minor differences in insertions and deletions) (Supplementary Figure 7). This suggests that the major differences between ancestral sequences and their phenotypes are driven more so by the sequence alignment and phylogenetic topology than they are by systematic differences in the sequence evolution model and reconstruction algorithm.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.2c00323.

Additional data and figures mentioned in the text (PDF)

Accession Codes

The structures of GrAnc8 (8CRU) and PROSS5 (8D1D) have been deposited in the Protein Data Bank. The Uniprot accession number of IsPETase is A0A0K8P6T7.

■ AUTHOR INFORMATION

Corresponding Authors

Albert Ardevol − *Manufacturing, Commonwealth Scientific and Industrial Research Organisation, Clayton, Victoria 3168, Australia; CSIRO Synthetic Biology Future Science Platform, Canberra, ACT 2601, Australia*; Email: albert.ardevolgrau@csiro.au

Colin J. Jackson − *Research School of Chemistry, ARC Centre of Excellence for Innovations in Synthetic Biology, Research School of Chemistry, and ARC Centre of Excellence for Innovations in Peptide & Protein Science, Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*; orcid.org/0000-0001-6150-3822; Email: colin.jackson@anu.edu.au

Authors

Yvonne Joho − *Manufacturing, Commonwealth Scientific and Industrial Research Organisation, Clayton, Victoria 3168, Australia; Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*

Vanessa Vongsouthi − *Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*; orcid.org/0000-0001-9645-0003

Matthew A. Spence − *Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*

Jennifer Ton − *Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*

Chloe Gomez − *Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*

Li Lynn Tan − *Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*; orcid.org/0000-0003-0606-4446

Joe A. Kaczmarski − *Research School of Chemistry and ARC Centre of Excellence for Innovations in Synthetic Biology, Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*

Alessandro T. Caputo − *Manufacturing, Commonwealth Scientific and Industrial Research Organisation, Clayton, Victoria 3168, Australia*

Santana Royan − *Manufacturing, Commonwealth Scientific and Industrial Research Organisation, Clayton, Victoria 3168, Australia*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biochem.2c00323

Author Contributions

C.J.J. and A.A. jointly supervised this work.

Funding

This work was supported by the CSIRO SynBio FSP, the ARC Centre of Excellence in Synthetic Biology (CE200100029), and the ARC Centre of Excellence in Peptide and Protein Science (CE200100012).

Notes

The authors declare the following competing financial interest(s): C.J.J., V.V., and M.A.S. hold equity in a plastic-recycling company, SamsaraEco. However, none of the work in this paper is of relevance to that enterprise.

■ ACKNOWLEDGMENTS

All crystallization experiments were carried out by Bevan Marshall and Sophia Newton at the CSIRO Collaborative Crystallization Centre (www.csiro.au/C3), Melbourne, Australia. This research was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector.

■ REFERENCES

(1) PlasticsEurope. Plastics − the Facts 2019: An analysis of European plastics production, demand and waste data. 2019.

(2) PlasticsEurope. Plastics - the Facts 2020: An analysis of European plastics production, demand and waste data. 2020.

(3) Heidbreder, L. M.; Bablok, I.; Drews, S.; Menzel, C. Tackling the plastic problem: A review on perceptions, behaviors, and interventions. *Science of The Total Environment* **2019**, *668*, 1077−1093.

(4) Lindeque, P. K.; Cole, M.; Coppock, R. L.; Lewis, C. N.; Miller, R. Z.; Watts, A. J. R.; Wilson-McNeal, A.; Wright, S. L.; Galloway, T. S. Are we underestimating microplastic abundance in the marine environment? A comparison of microplastic capture with nets of different mesh-size. *Environ. Pollut.* **2020**, *265*, 114721.

(5) Peng, J.; Wang, J.; Cai, L. Current understanding of microplastics in the environment: Occurrence, fate, risks, and what we should do. *Integrated Environmental Assessment and Management* **2017**, *13* (3), 476−482.

(6) Rist, S.; Carney Almroth, B.; Hartmann, N. B.; Karlsson, T. M. A critical perspective on early communications concerning human health aspects of microplastics. *Science of The Total Environment* **2018**, *626*, 720−726.

(7) Geyer, R.; Jambeck, J. R.; Law, K. L. Production, use, and fate of all plastics ever made. *Sci. Adv.* **2017**, *3* (7), e1700782.

(8) PlasticsInsight. Global PET Resin Production Capacity. https://www.plasticsinsight.com/global-pet-resin-production-capacity.

(9) Han, M. Depolymerization of PET bottle via methanolysis and hydrolysis. In *Recycling of Polyethylene Terephthalate Bottles*; Elsevier, 2019; pp 85−108.

(10) Nofar, M. Chapter 6 - PLA binary blends with petroleum-based nondegradable thermoplastics. In *Multiphase Polylactide Blends*; Nofar, M., Ed.; Elsevier, 2021; pp 233−263.

(11) Şimşek, B.; Uygunoğlu, T.; Korucu, H.; Kocakerim, M. Performance of dioctyl terephthalate concrete. In *Use of Recycled Plastics in Eco-efficient Concrete*; Elsevier, 2019; pp 249−267.

(12) Lanaro, M.; Booth, L.; Powell, S. K.; Woodruff, M. A. Electrofluidodynamic technologies for biomaterials and medical

devices: melt electrospinning. In *Electrofluidodynamic technologies (EFDTs) for biomaterials and medical devices*; Elsevier, 2018; pp 37−69.
(13) Schyns, Z. O. G.; Shaver, M. P. Mechanical Recycling of Packaging Plastics: A Review. *Macromol. Rapid Commun.* **2021**, *42* (3), 2000415.
(14) Singh, A.; Rorrer, N. A.; Nicholson, S. R.; Erickson, E.; Desveaux, J. S.; Avelino, A. F. T.; Lamers, P.; Bhatt, A.; Zhang, Y.; Avery, G.; Tao, L.; Pickford, A. R.; Carpenter, A. C.; McGeehan, J. E.; Beckham, G. T. Techno-economic, life-cycle, and socioeconomic impact analysis of enzymatic recycling of poly(ethylene terephthalate). *Joule* **2021**, *5*, 2479.
(15) Liu, B.; He, L.; Wang, L.; Li, T.; Li, C.; Liu, H.; Luo, Y.; Bao, R. Protein crystallography and site-direct mutagenesis analysis of the poly (ethylene terephthalate) hydrolase PETase from Ideonella sakaiensis. *ChemBioChem.* **2018**, *19* (14), 1471−1475.
(16) Chen, C. C.; Han, X.; Ko, T. P.; Liu, W.; Guo, R. T. Structural studies reveal the molecular mechanism of PET ase. *FEBS Journal* **2018**, *285* (20), 3717−3723.
(17) Yoshida, S.; Hiraga, K.; Takehana, T.; Taniguchi, I.; Yamaji, H.; Maeda, Y.; Toyohara, K.; Miyamoto, K.; Kimura, Y.; Oda, K. A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* **2016**, *351* (6278), 1196−1199.
(18) Wallace, N. E.; Adams, M. C.; Chafin, A. C.; Jones, D. D.; Tsui, C. L.; Gruber, T. D. The highly crystalline PET found in plastic water bottles does not support the growth of the PETase -producing bacterium Ideonella sakaiensis. *Environmental Microbiology Reports* **2020**, *12* (5), 578−582.
(19) Joo, S.; Cho, I. J.; Seo, H.; Son, H. F.; Sagong, H.-Y.; Shin, T. J.; Choi, S. Y.; Lee, S. Y.; Kim, K.-J. Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation. *Nat. Commun.* **2018**, *9* (1), 382.
(20) Wei, R.; Oeser, T.; Zimmermann, W. Chapter Seven - Synthetic Polyester-Hydrolyzing Enzymes From Thermophilic Actinomycetes. In *Advances in Applied Microbiology*; Sariaslani, S., Gadd, G. M., Eds.; Academic Press, 2014; Vol. *89*, pp 267−305.
(21) Sulaiman, S.; Yamato, S.; Kanaya, E.; Kim, J. J.; Koga, Y.; Takano, K.; Kanaya, S. Isolation of a novel cutinase homolog with polyethylene terephthalate-degrading activity from leaf-branch compost by using a metagenomic approach. *Appl. Environ. Microbiol.* **2012**, *78* (5), 1556−62.
(22) Wei, R.; von Haugwitz, G.; Pfaff, L.; Mican, J.; Badenhorst, C. P. S.; Liu, W.; Weber, G.; Austin, H. P.; Bednar, D.; Damborsky, J.; Bornscheuer, U. T. Mechanism-Based Design of Efficient PET Hydrolases. *ACS Catal.* **2022**, *12* (6), 3382−3396.
(23) Alam, I.; Aalismail, N.; Martin, C.; Kamau, A.; Guzmán-Vega, F. J.; Jamil, T.; Momin, A. A.; Acinas, S. G.; Gasol, J. M.; Arold, S. T.; Gojobori, T.; Agusti, S.; Duarte, C. M. Rapid Evolution of Plastic-degrading Enzymes Prevalent in the Global Ocean. *bioRxiv* **2020**, DOI: 10.1101/2020.09.07.285692.
(24) Meng, X.; Yang, L.; Liu, H.; Li, Q.; Xu, G.; Zhang, Y.; Guan, F.; Zhang, Y.; Zhang, W.; Wu, N.; Tian, J. Protein engineering of stable IsPETase for PET plastic degradation by Premuse. *Int. J. Biol. Macromol.* **2021**, *180*, 667−676.
(25) Magalhães, R. P.; Fernandes, H. S.; Sousa, S. F. The critical role of Asp206 stabilizing residues on the catalytic mechanism of the Ideonella sakaiensis PETase. *Catal. Sci. Technol.* **2022**, *12*, 3474−3483.
(26) Jia, Y.; Samak, N. A.; Hao, X.; Chen, Z.; Yang, G.; Zhao, X.; Mu, T.; Yang, M.; Xing, J. Nano-immobilization of PETase enzyme for Enhanced Polyethylene Terephthalate Biodegradation. *Biochemical Engineering Journal* **2021**, *176*, 108205.
(27) Austin, H. P.; Allen, M. D.; Donohoe, B. S.; Rorrer, N. A.; Kearns, F. L.; Silveira, R. L.; Pollard, B. C.; Dominick, G.; Duman, R.; El Omari, K.; Mykhaylyk, V.; Wagner, A.; Michener, W. E.; Amore, A.; Skaf, M. S.; Crowley, M. F.; Thorne, A. W.; Johnson, C. W.; Woodcock, H. L.; McGeehan, J. E.; Beckham, G. T. Characterization and engineering of a plastic-degrading aromatic polyesterase. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (19), E4350−E4357.
(28) Nikolaivits, E.; Kanelli, M.; Dimarogona, M.; Topakas, E. A Middle-Aged Enzyme Still in Its Prime: Recent Advances in the Field of Cutinases. *Catalysts* **2018**, *8*, 612.
(29) Cui, Y.; Chen, Y.; Liu, X.; Dong, S.; Tian, Y. e.; Qiao, Y.; Mitra, R.; Han, J.; Li, C.; Han, X.; Liu, W.; Chen, Q.; Wei, W.; Wang, X.; Du, W.; Tang, S.; Xiang, H.; Liu, H.; Liang, Y.; Houk, K. N.; Wu, B. Computational Redesign of a PETase for Plastic Biodegradation under Ambient Condition by the GRAPE Strategy. *ACS Catal.* **2021**, *11* (3), 1340−1350.
(30) Bell, E.; Smithson, R.; Kilbride, S.; Foster, J.; Hardy, F.; Ramachandran, S.; Tedstone, A.; Haigh, S.; Garforth, A.; Day, P. Directed Evolution of an Efficient and Thermostable PET Depolymerase. *chemRxiv* **2021**, DOI: 10.26434/chemrxiv-2021-mcjh6.
(31) Lu, H.; Diaz, D. J.; Czarnecki, N. J.; Zhu, C.; Kim, W.; Shroff, R.; Acosta, D. J.; Alexander, B. R.; Cole, H. O.; Zhang, Y.; Lynd, N. A.; Ellington, A. D.; Alper, H. S. Machine learning-aided engineering of hydrolases for PET depolymerization. *Nature* **2022**, *604* (7907), 662−667.
(32) Son, H. F.; Cho, I. J.; Joo, S.; Seo, H.; Sagong, H.-Y.; Choi, S. Y.; Lee, S. Y.; Kim, K.-J. Rational Protein Engineering of Thermo-Stable PETase from Ideonella sakaiensis for Highly Efficient PET Degradation. *ACS Catal.* **2019**, *9* (4), 3519−3526.
(33) Tournier, V.; Topham, C. M.; Gilles, A.; David, B.; Folgoas, C.; Moya-Leclair, E.; Kamionka, E.; Desrousseaux, M. L.; Texier, H.; Gavalda, S.; Cot, M.; Guémard, E.; Dalibey, M.; Nomme, J.; Cioci, G.; Barbe, S.; Chateau, M.; André, I.; Duquesne, S.; Marty, A. An engineered PET depolymerase to break down and recycle plastic bottles. *Nature* **2020**, *580* (7802), 216−219.
(34) Wei, R.; Tiso, T.; Bertling, J.; O'Connor, K.; Blank, L. M.; Bornscheuer, U. T. Possibilities and limitations of biotechnological plastic degradation and recycling. *Nature Catalysis* **2020**, *3* (11), 867−871.
(35) Spence, M. A.; Kaczmarski, J. A.; Saunders, J. W.; Jackson, C. J. Ancestral sequence reconstruction for protein engineers. *Curr. Opin. Struct. Biol.* **2021**, *69*, 131−141.
(36) Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24* (8), 1586−91.
(37) Foley, G.; Mora, A.; Ross, C. M.; Bottoms, S.; Sützl, L.; Lamprecht, M. L.; Zaugg, J.; Essebier, A.; Balderson, B.; Newell, R.; Thomson, R. E. S.; Kobe, B.; Barnard, R. T.; Guddat, L.; Schenk, G.; Carsten, J.; Gumulya, Y.; Rost, B.; Haltrich, D.; Sieber, V.; Gillam, E. M. J.; Bodén, M. Engineering indel and substitution variants of diverse and ancient enzymes using Graphical Representation of Ancestral Sequence Predictions (GRASP). *bioRxiv* **2019**, DOI: 10.1101/2019.12.30.891457.
(38) Goldenzweig, A.; Goldsmith, M.; Hill, S. E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J.; Lieberman, R. L.; Aharoni, A.; Silman, I.; Sussman, J. L.; Tawfik, D. S.; Fleishman, S. J. Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Mol. Cell* **2016**, *63* (2), 337−346.
(39) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Zidek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50* (D1), D439−D444.
(40) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

(41) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry* **2019**, *58* (41), 4169−4182.
(42) Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28* (23), 3150−3152.
(43) Katoh, K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30* (14), 3059−3066.
(44) Petersen, T. N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **2011**, *8* (10), 785−6.
(45) Minh, B. Q.; Schmidt, H. A.; Chernomor, O.; Schrempf, D.; Woodhams, M. D.; Von Haeseler, A.; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37* (5), 1530−1534.
(46) Pei, J.; Grishin, N. V. PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* **2007**, *23* (7), 802−808.
(47) Love, C. A.; Lilley, P. E.; Dixon, N. E. Stable high-copy-number bacteriophage lambda promoter vectors for overproduction of proteins in Escherichia coli. *Gene* **1996**, *176* (1−2), 49−53.
(48) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **2009**, *6* (5), 343−345.
(49) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook* **2005**, 571−607.
(50) Pfaff, L.; Breite, D.; Badenhorst, C. P. S.; Bornscheuer, U. T.; Wei, R. Fluorimetric high-throughput screening method for polyester hydrolase activity using polyethylene terephthalate nanoparticles. *Methods Enzymol* **2021**, *648*, 253−270.
(51) Wei, R.; Oeser, T.; Billig, S.; Zimmermann, W. A high-throughput assay for enzymatic polyester hydrolysis activity by fluorimetric detection. *Biotechnol J.* **2012**, *7* (12), 1517−21.
(52) Rosa, N.; Ristic, M.; Seabrook, S. A.; Lovell, D.; Lucent, D.; Newman, J. Meltdown: A Tool to Help in the Interpretation of Thermal Melt Curves Acquired by Differential Scanning Fluorimetry. *J. Biomol Screen* **2015**, *20* (7), 898−905.
(53) Aragao, D.; Aishima, J.; Cherukuvada, H.; Clarken, R.; Clift, M.; Cowieson, N. P.; Ericsson, D. J.; Gee, C. L.; Macedo, S.; Mudie, N.; Panjikar, S.; Price, J. R.; Riboldi-Tunnicliffe, A.; Rostan, R.; Williamson, R.; Caradoc-Davies, T. T. MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiat.* **2018**, *25*, 885−891.
(54) Evans, P. Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 72−82.
(55) Kabsch, W. Xds. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125−32.
(56) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658−674.
(57) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486−501.
(58) Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355−67.
(59) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography* **2011**, *67* (4), 235−242.
(60) Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography* **2013**, *69* (7), 1204−1214.
(61) Smart, O. S.; Womack, T. O.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallographica Section D: Biological Crystallography* **2012**, *68* (4), 368−380.
(62) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkoczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D: Biological Crystallography* **2019**, *75* (10), 861−877.
(63) Hanson-Smith, V.; Kolaczkowski, B.; Thornton, J. W. Robustness of Ancestral Sequence Reconstruction to Phylogenetic Uncertainty. *Mol. Biol. Evol.* **2010**, *27* (9), 1988−1999.
(64) Park, Y.; Patton, J. E. J.; Hochberg, G. K. A.; Thornton, J. W. Comment on "Ancient origins of allosteric activation in a Ser-Thr kinase. *Science* **2020**, *370* (6519), eabc8301.
(65) Yang, Z. Among-site rate variation and its impact on phylogenetic analyses. *Trends in Ecology & Evolution* **1996**, *11* (9), 367−372.
(66) Jia, F.; Lo, N.; Ho, S. Y. The impact of modelling rate heterogeneity among sites on phylogenetic estimates of intraspecific evolutionary rates and timescales. *PLoS One* **2014**, *9* (5), e95722.
(67) Roth, C.; Wei, R.; Oeser, T.; Then, J.; Föllner, C.; Zimmermann, W.; Sträter, N. Structural and functional studies on a thermostable polyethylene terephthalate degrading hydrolase from Thermobifida fusca. *Appl. Microbiol. Biotechnol.* **2014**, *98* (18), 7815−23.
(68) Brott, S.; Pfaff, L.; Schuricht, J.; Schwarz, J.-N.; Böttcher, D.; Badenhorst, C. P. S.; Wei, R.; Bornscheuer, U. T. Engineering and evaluation of thermostable IsPETase variants for PET degradation. *Engineering in Life Sciences* **2022**, *22* (3−4), 192−203.
(69) Clifton, B. E.; Kaczmarski, J. A.; Carr, P. D.; Gerth, M. L.; Tokuriki, N.; Jackson, C. J. Evolution of cyclohexadienyl dehydratase from an ancestral solute-binding protein. *Nat. Chem. Biol.* **2018**, *14* (6), 542−547.
(70) Glaser, F.; Pupko, T.; Paz, I.; Bell, R. E.; Bechor-Shental, D.; Martz, E.; Ben-Tal, N. ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information. *Bioinformatics* **2003**, *19* (1), 163−164.
(71) Mehlin, C.; Boni, E.; Buckner, F. S.; Engel, L.; Feist, T.; Gelb, M. H.; Haji, L.; Kim, D.; Liu, C.; Mueller, N.; Myler, P. J.; Reddy, J. T.; Sampson, J. N.; Subramanian, E.; Van Voorhis, W. C.; Worthey, E.; Zucker, F.; Hol, W. G. J. Heterologous expression of proteins from Plasmodium falciparum: Results from 1000 genes. *Mol. Biochem. Parasitol.* **2006**, *148* (2), 144−160.
(72) Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L. E.; Brookes, D. H.; Wilson, L.; Chen, J.; Liles, K.; Chun, M.; Li, P.; Gohara, D. W.; Dolinsky, T.; Konecny, R.; Koes, D. R.; Nielsen, J. E.; Head-Gordon, T.; Geng, W.; Krasny, R.; Wei, G.-W.; Holst, M. J.; McCammon, J. A.; Baker, N. A. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **2018**, *27* (1), 112−128.
(73) *PyMOL*; Schrödinger, LLC (http://www.pymol.org/pymol).
(74) Betz, S. F. Disulfide bonds and the stability of globular proteins. *Protein Sci.* **1993**, *2* (10), 1551−1558.
(75) Fecker, T.; Galaz-Davison, P.; Engelberger, F.; Narui, Y.; Sotomayor, M.; Parra, L. P.; Ramírez-Sarmiento, C. A. Active site flexibility as a hallmark for efficient PET degradation by I. sakaiensis PETase. *Biophysical journal* **2018**, *114* (6), 1302−1312.

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX

(76) Kumar, S.; Tsai, C.-J.; Nussinov, R. Factors enhancing protein thermostability. *Protein Engineering, Design and Selection* **2000**, *13* (3), 179−191.
(77) Chen, C.-C.; Han, X.; Li, X.; Jiang, P.; Niu, D.; Ma, L.; Liu, W.; Li, S.; Qu, Y.; Hu, H.; Min, J.; Yang, Y.; Zhang, L.; Zeng, W.; Huang, J.-W.; Dai, L.; Guo, R.-T. General features to enhance enzymatic activity of poly(ethylene terephthalate) hydrolysis. *Nature Catalysis* **2021**, *4* (5), 425−430.
(78) Han, X.; Liu, W.; Huang, J.-W.; Ma, J.; Zheng, Y.; Ko, T.-P.; Xu, L.; Cheng, Y.-S.; Chen, C.-C.; Guo, R.-T. Structural insight into catalytic mechanism of PET hydrolase. *Nat. Commun.* **2017**, *8* (1), 2106.
(79) Clifton, B. E.; Jackson, C. J. Ancestral Protein Reconstruction Yields Insights into Adaptive Evolution of Binding Specificity in Solute-Binding Proteins. *Cell Chem. Biol.* **2016**, *23* (2), 236−245.
(80) Bornscheuer, U. T. Feeding on plastic. *Science* **2016**, *351* (6278), 1154−1155.
(81) Jackson, C.; Toth-Petroczy, A.; Kolodny, R.; Hollfelder, F.; Fuxreiter, M.; Kamerlin, S. C. L.; Tokuriki, N. Adventures on the Routes of Protein Evolution-In Memoriam Dan Salah Tawfik (1955−2021). *J. Mol. Biol.* **2022**, *434* (7), 167462.
(82) Tawfik, D. S. Enzyme promiscuity and evolution in light of cellular metabolism. *Febs Journal* **2020**, *287* (7), 1260−1261.
(83) James, L. C.; Tawfik, D. S. Catalytic and binding poly-reactivities shared by two unrelated proteins: The potential role of promiscuity in enzyme evolution. *Protein Sci.* **2001**, *10* (12), 2600−7.
(84) Campbell, E.; Kaltenbach, M.; Correy, G. J.; Carr, P. D.; Porebski, B. T.; Livingstone, E. K.; Afriat-Jurnou, L.; Buckle, A. M.; Weik, M.; Hollfelder, F.; Tokuriki, N.; Jackson, C. J. The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **2016**, *12* (11), 944−950.
(85) Hong, N. S.; Petrović, D.; Lee, R.; Gryn'ova, G.; Purg, M.; Saunders, J.; Bauer, P.; Carr, P. D.; Lin, C. Y.; Mabbitt, P. D.; Zhang, W.; Altamore, T.; Easton, C.; Coote, M. L.; Kamerlin, S. C. L.; Jackson, C. J. The evolution of multiple active site configurations in a designed enzyme. *Nat. Commun.* **2018**, *9* (1), 3900.
(86) Campbell, E. C.; Correy, G. J.; Mabbitt, P. D.; Buckle, A. M.; Tokuriki, N.; Jackson, C. J. Laboratory evolution of protein conformational dynamics. *Current Opinions in Structural Biology* **2018**, *50*, 49−57.
(87) Tokuriki, N.; Stricher, F.; Serrano, L.; Tawfik, D. S. How protein stability and new functions trade off. *PLOS Computational Biology* **2008**, *4* (2), e1000002.
(88) Dellus-Gur, E.; Toth-Petroczy, A.; Elias, M.; Tawfik, D. S. What Makes a Protein Fold Amenable to Functional Innovation? Fold Polarity and Stability Trade-offs. *J. Mol. Biol.* **2013**, *425* (14), 2609−2621.

https://doi.org/10.1021/acs.biochem.2c00323
Biochemistry XXXX, XXX, XXX−XXX