Computational redesign of PETase for plastic biodegradation by GRAPE strategy

Yinglu Cui[1#], Yanchun Chen[1,2#], Xinyue Liu[1,4], Saijun Dong[1], Yu'e Tian[1], Yuxin Qiao[1], Jing Han[1], Chunli Li[1], Xu Han[3], Weidong Liu[3], Quan Chen[4], Wenbin Du[1], Shuangyan Tang[1], Hua Xiang[1], Haiyan Liu[4], Bian Wu[1*]

[1] CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China.

[2] University of Chinese Academy of Sciences, Beijing, China.

[3] Industrial Enzymes National Engineering Laboratory, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China.

[4] School of Life Sciences, Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei, China.

[#] Equal contribution

[*] Email: wub@im.ac.cn

Abstract

The excessive use of plastics has been accompanied by severe ecologically-damaging effects. The recent discovery of PETase from *Ideonalla sakaiensis* that decomposes poly(ethylene terephthalate) (PET) under mild conditions provides an attractive avenue for biodegradation of plastics. However, the inherent instability of the enzyme limits its practical utilization. Here, we devised a novel computational strategy (Greedy Accumulated strategy for Protein Engineering, GRAPE). A systematic clustering analysis combined with greedy accumulation of the beneficial mutations in a computationally derived library enabled a redesigned variant, DuraPETase, with drastically elevated apparent melting temperature by 31 °C and strikingly enhanced degradation activity toward high-crystallinity PET film (49.2%) at moderate temperatures (product concentration raised from μM range to mM levels). The underlying mechanism of the robust promotion of enzyme performance has been demonstrated via crystal structure and molecular dynamics simulations. Collectively, this work promotes the methodological capabilities of computational enzyme design to circumvent antagonistic epistatic effects and provides a valuable tool for further understanding and advancing polyester hydrolysis in the natural environment.

Introduction

Earth's ecosystem is approaching a planetary-scale transformation as a result of human influence[1]. Nowadays, a wide variety of petroleum-based synthetic polymers are produced worldwide to the extent of 20-fold increase over the five decades since 1960s and a slowdown of this trend is not expected[2]. Accumulated marine and terrestrial pollution problems, as well as the negative impacts of microplastic exposure on human health, have caused escalating public and governmental concerns[3, 4]. With the excessive use of plastics and increasing pressure on the capacity for plastic waste disposal, the demand for eco-friendly waste management has assumed dramatically importance in the last few years[5, 6].

Poly(ethylene terephthalate) (PET) is one of the most widely used man-made synthetic plastics worldwide with an annual manufacturing capacity of over 30 million tons[7]. Its excellent durability, however, has now become an environmental villain[8]. The aromatic building block terephthalate units in crystalline PET largely reduces the chain mobility and leads to extremely

limited accessibility of the ester linkages[9]. Therefore, the degradation activities of most PET hydrolases have been limited by extremely low turnover rates towards high-crystallinity PET[10]. Despite the identified PET hydrolases from thermophilic organisms requiring operational temperature near PET's glass transition temperature (~76ºC), Yoshida et al. recently reported a new PETase from *Ideonella sakaiensis* 201-F6 (*Is*PETase) with the highest activity among PET-degrading enzymes at ambient conditions[11], which indicates a promising application for *in situ* disposals of hard-to-collect plastic disintegrates with low energy consumption. Although *Is*PETase has caught much attention for its intriguing capability, the enzyme exhibits poor durability that most of its activity would be lost within 12 h of incubation at 37 °C[12]. Apart from this, *Is*PETase only exhibits reasonable activity toward low-crystallinity PET (1.9%), whereas most of PET applications, including commercial containers, employ high-crystallinity PET with 30%-50% crystallinity[13].

Not content with the finite repertoire of naturally occurring enzymes, numerous research groups have concerned on the engineering of *Is*PETase, which has been summarized by Taniguchi recently[14]. Recently crystal structures of PET degrading enzymes allow for expedients that exploit rational design to improve the PET degradation activity[15-21]. The successful single point mutations afford 1.2 to 3.1 folds higher affinity to PET[18-22], and very recently, Kim's group has succeeded in creating *Is*PETase$^{S121E/D186H/R280A}$ variant with enhanced thermal stability by 8.81 °C and 14-fold higher PET degradation activity at 40 °C[12].

Despite these attempts, there remains a key challenge whether protein engineering can achieve improved enzymatic stability and effective catalytic properties to high-crystallinity PET by allowing large jumps along the sequence fitness landscape. There is usually a pathway whereby some new functions could be acquired by individually beneficial mutations, however, when the desired function is beyond what a single mutation or double mutations can accomplish, possible paths grow exponentially as the mutations accumulate and most paths result in downhill or even unfolded proteins[23]. Since a majority of protein engineering studies involves simple uphill walks, the main demand lies in identifying an efficient path of accumulated mutations to achieve the desired protein performance[24].

The last few years have witnessed impressive progress in dealing with multi-dimensional space by a number of metaheuristic methods[25]. Inspired by the widespread greedy algorithm

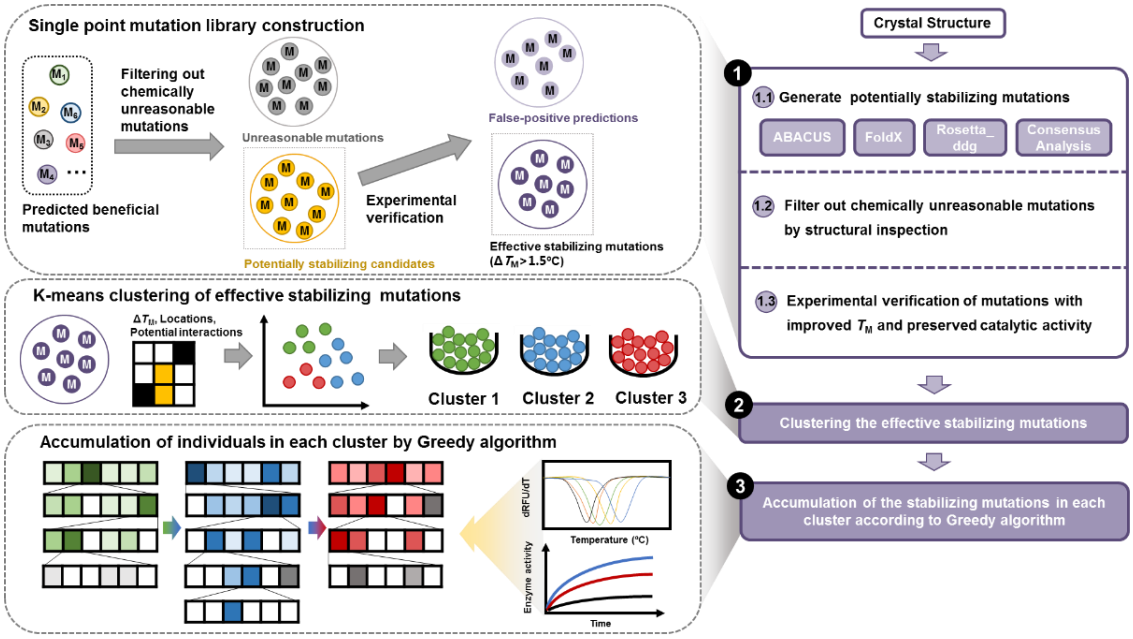
applications in artificial intelligence, we introduce a novel strategy, termed GRAPE, to tackle the evolutionary hard problem effectively by computational design that enhances the probability of discovering the adaptive routes to higher fitness. The redesigned *Is*PETase (DuraPETase) derived from this redesign campaign exhibited an increased apparent melting temperature by 31 °C with good performance toward high-crystallinity PET and vastly improved long-term survival under mild condition. The PET degradation activity was enhanced by 400-fold at 40 °C for 10 days. Ameliorated biodegradation of other plastics, such as PBT and PEN, suggested DuraPETase a broader ability to degrade semiaromatic polyesters. To evaluate the underlying molecular mechanism for enhanced performance, the three-dimensional crystal structure of the variant was determined, highlighting a fine-tuned 'aromatic tunnel' flanker by synergistic hydrophobic interactions.

Results

Computational redesign of a robust DuraPETase using GRAPE strategy

GRAPE strategy introduced here uses greedy strategy for global optimization of mutations in each cluster to create functional variations and select the fittest variants to direct the search to higher elevations on the fitness landscape. The initial step in GRAPE strategy (Scheme 1) consists of computational predictions of potentially stabilizing mutations along the whole protein sequence. We previously developed ABACUS[26] algorithm for *de novo* protein design based on statistical energy function. Here, this algorithm was applied to improve the protein stability, together with three complementary algorithms, including FoldX (force field-based energy function)[27], Rosetta_ddg (force field-based energy function)[28] and Consensus Analysis (phylogeny-based method)[29]. Mutations with a free energy of ABACUS energy < -3 A.e.u, folding ($\Delta\Delta G_{fold}$) < -5 kJ/mol or consensus score > 2.0 were selected (Table S1 and Figure S1). A total of 253 unique predicted mutations were obtained as potentially stabilized candidates after filtering, among which ABACUS, FoldX, Rosetta_ddg, and consensus analysis algorithms provided 85, 63, 65, and 54 mutations, respectively, with some overlap. In the second step, the potentially stabilized candidates were inspected for biophysical pitfalls, such as the introduction of internal cavities, loss of hydrogen-bonding interactions, or exposure of hydrophobic residues at the surface of the enzyme, providing a sub-library of 85 candidates. After experimental validation, 21 well-expressed mutants displayed higher stability (> 1.5 °C increase in the

apparent T_m) (Figures S2 and S3). However, owing to the ubiquity of epistatic effects[30-32], these positive mutations may not cooperate to reach the target function. Indeed, the *Is*PETase variant containing all 21 beneficial mutations is totally inactive owing to antagonistic epistatic effects.



Scheme 1 Schematic representation of the GRAPE strategy. In step 1, stabilizing mutations are generated with multiple algorithms. The computational designs with typical known pitfalls are eliminated. Then the remaining designs are selected for experimental validation. Step 2 characterizes the variants according to their positions, efficacies and presumable effects. Accumulation of the mutations in each cluster according to greedy algorithm is performed in Step 3. Details regarding Step 3 are demonstrated in Figure 1(A).

Hence, crucial for GRAPE strategy is the next step that systematically accumulating the beneficial variants in the well-defined library. Searching in the high-dimensional space contains an incomprehensibly large number of possible pathways, in which only an infinitesimal fraction can escape from the local optima traps to achieve the desired goal. The proposed GRAPE strategy mixes advantage of greedy and clustering algorithms to provide a viable solution to minimize the experimental efforts but maximize the exploration of epistatic effects in terms of additivity and/or synergism between sets of mutations at each branching point. In GRAPE strategy, the K-means algorithm[33], which has been proven to be a very powerful tool for data mining problems and was adopted to perform knowledge discovery in bioinformatics research[34],

was applied to cluster the stabilizing mutations into several groups. Variants were characterized based on their positions, efficacies, and presumable effects. Subsequently, the best individual of each population served as the parent to attract offspring to its region of the fitness landscape. Each individual in the cluster was crossed with the current global best one. Greedy algorithm accepts the newly generated individual only when its fitness is better than that of the parent. Figure 1A involves the stipulation that each individual stays in its historically best position and moves forward toward the global best position.

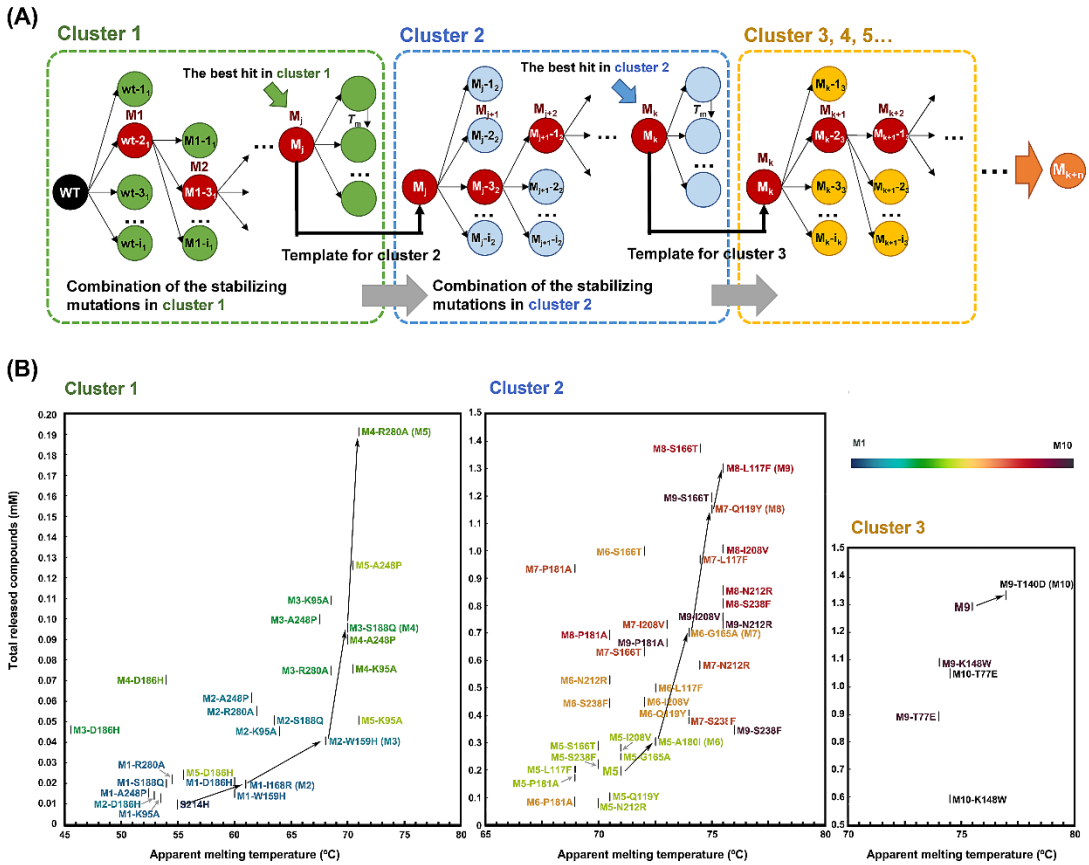


Figure 1 Flowchart of the greedy accumulation. (A) Beneficial mutations in each cluster are crossed with the best hit of the current population until the remaining mutations in the cluster have been traversed or the T_m values of the combined variants decrease. If the variant showed high thermostability but reduced > 50% enzyme activity, the combined variant would not be adopted. For example, the best hit in cluster 1 (M_j) serves as the template for further cycles of accumulation in cluster 2, whereas the best hit in cluster 2 (M_k) is regarded as the parent for further cycles in the next cluster. The best hit in each stage is sequentially referred as M_1, M_2, …, M_j, M_{j+1}, …, M_k, M_{k+1}, …, M_{k+n}. The red cycles represent the mutant chosen from the respective combination stage used as the template for the further accumulation (black arrows). (B) Enhancing the thermostability and activity of IsPETase by greedy accumulation. The accumulating mutations in each cluster are listed in Table 1.

Accordingly, the 21 identified stabilizing mutations were clustered into three groups (Table 1). S214H in the 1st cluster was considered as the starting point for the initial round of accumulation. The exploitation process was then continued in a multi-hierarchical manner until all variants in the cluster have been traversed or no further improvements are found (Figure 1B). Whether a combination stage was positive or deleterious can be conditional on two parameters, i.e., thermostability and enzyme activity. If the combined variant showed high thermostability but seriously reduced activity, it was advisable to make a compromise. After five cycles of the accumulating exploration, the best hit (S214H-I168R-W159H-S188Q-R280A, M5) in the 1st cluster was obtained, with a T_m value of 71 ºC. Further combination of the remaining mutations in the 1st cluster to M5 failed to increase the T_m value. It is worth noting that the stability of S214H-D186H-I168R mutant dramatically decreased, which inferred an antagonistic effect between H186 and R168 residues. Subsequently, the M5 variant was used as the template for crossover of mutations in the 2nd cluster, leading to the best hit variant S214H-I168R-W159H-S188Q-R280A-A180I-G165A-Q119Y-L117F (*Is*PETase-M9) with a T_m value of 75.5 ºC. Although the combination of the remaining mutations in the 2nd cluster has slightly increased the T_m values, largely reduced degradation activities were observed compared to that of M9 (Figure S4). Besides, the introduction of S187W mutation to the best hit in each round exhibited extremely lower expression. Therefore, M9 was set as the parent for next accumulation of the respective other sites in the 3rd cluster. The most thermostable variant S214H-I168R-W159H-S188Q-R280A-A180I-G165A-Q119Y-L117F-T140D (*Is*PETase-M10, referred as DuraPETase) originated from the 3rd cluster exhibited dramatically enhanced thermostability (ΔT_m = 31 ºC). Before addressing the above challenge, we also explored a stepwise combination of the most stabilizing variants (ΔT_m > 7 ºC). In addition to the antagonistic effects between H186 and R168, the introduction of P181A mutation largely reduced the enzymatic activity, leading to an uphill dead end on the fitness landscape. Addressing these multisites progressively in the GRAPE scheme with multilayer upward branches, negative trade-offs can be determined and quantified, thus find effective ways to circumvent apparent dead ends.

Table 1. Clustering of the stabilizing point mutations based on their positions, efficacies and presumable effects.

Residue number	ΔT_m	Interaction cluster[a]	Region[b]	Distance[c]	Hbonds[d]	Hydrophobic interactions[d]	Entropy change[d]	Binding Region[e]	Cluster
S214H	9	0	1	9.62	FALSE	FALSE	FALSE	TRUE	1
W159H	8.5	1	1	3.81	TRUE	FALSE	FALSE	TRUE	1
I168R	7.5	2	1	12.55	TRUE	FALSE	FALSE	TRUE	1
D186H	7	2	1	8.00	TRUE	FALSE	FALSE	TRUE	1
P181A	6	0	1	7.98	FALSE	TRUE	TRUE	FALSE	2
Q119Y	4.5	3	1	10.12	FALSE	TRUE	FALSE	TRUE	2
L117F	3	3	1	12.12	FALSE	TRUE	FALSE	TRUE	2
S187W	3	3	1	11.22	FALSE	TRUE	FALSE	TRUE	2
N212R	3	0	1	14.24	FALSE	TRUE	FALSE	FALSE	2
A248P	3	0	2	21.09	FALSE	FALSE	TRUE	TRUE	1
K95A	2.5	0	2	14.79	FALSE	FALSE	FALSE	TRUE	1
T140D	2	0	3	32.49	TRUE	FALSE	FALSE	FALSE	3
K148W	2	0	3	29.81	FALSE	TRUE	FALSE	FALSE	3
S166T	2	4	1	9.65	FALSE	TRUE	FALSE	TRUE	2
T77E	1.5	0	3	30.22	TRUE	FALSE	FALSE	FALSE	3
G165A	1.5	4	1	8.55	FALSE	TRUE	TRUE	TRUE	2
A180I	1.5	0	1	11.79	FALSE	TRUE	FALSE	TRUE	2
S188Q	1.5	2	1	12.98	TRUE	FALSE	FALSE	TRUE	1
I208V	1.5	0	1	8.59	FALSE	TRUE	FALSE	TRUE	2
S238F	1.5	1	1	9.05	FALSE	TRUE	FALSE	TRUE	2
R280A	1.5	0	2	20.62	FALSE	FALSE	FALSE	TRUE	1

[a] Mutations may have interactions between each other in each interaction cluster

[b] Locations of each mutation. Each number represents a separate region

[c] Distance between the Cα atoms of the mutation and Ser160

[d] Possible mechanism of stabilization

[e] Whether the mutation locates on the PET-binding region

Enhanced PET degradation activity toward high-crystallinity PET film

With significantly enhanced performance of DuraPETase, we would be able to explore the anticipated PET digestion toward commercial high-crystallinity PET film (49.2%) (Figure S5). DuraPETase and the wild-type enzyme were incubated at various temperatures ranging from 37 °C to 60 °C at pH 9.0 (Figure 2A). In both cases, the enzymes converted PET to mono(2-hydroxyethyl) terephthalic acid (MHET), with terephthalic acid (TPA) as secondary products. The degradation activity profiles largely reflect the difference in inherent robustness between the redesigned enzyme and the wild type. As expected, enhanced thermostability of the

DuraPETase coincided with remarkable improvement in long-term survival at moderate temperatures and even withstood elevated temperatures up to 60 °C for 3 days of incubation. As a contrast, *Is*PETase lost complete activity within 12 h at 37 °C, and the degradation product concentration reached 12 μmol and did not increase thereafter, which is consistent with that of Kim's report[12]. When the temperature increased, the activity of *Is*PETase dropped sharply within 1 hour due to the thermal-induced denaturation. Pronounced improvement by ca 100-fold in the overall yield of MHET and TPA products were enabled by DuraPETase after incubation for 72 h at 37 °C. The improved stability of DuraPETase has resulted in dramatically enhanced long-term durability with product yield up to 3.4 mM at 40 °C when the incubation time extended to 10 days, which was a 400-fold increase than that of *Is*PETase. (Figure 2B).

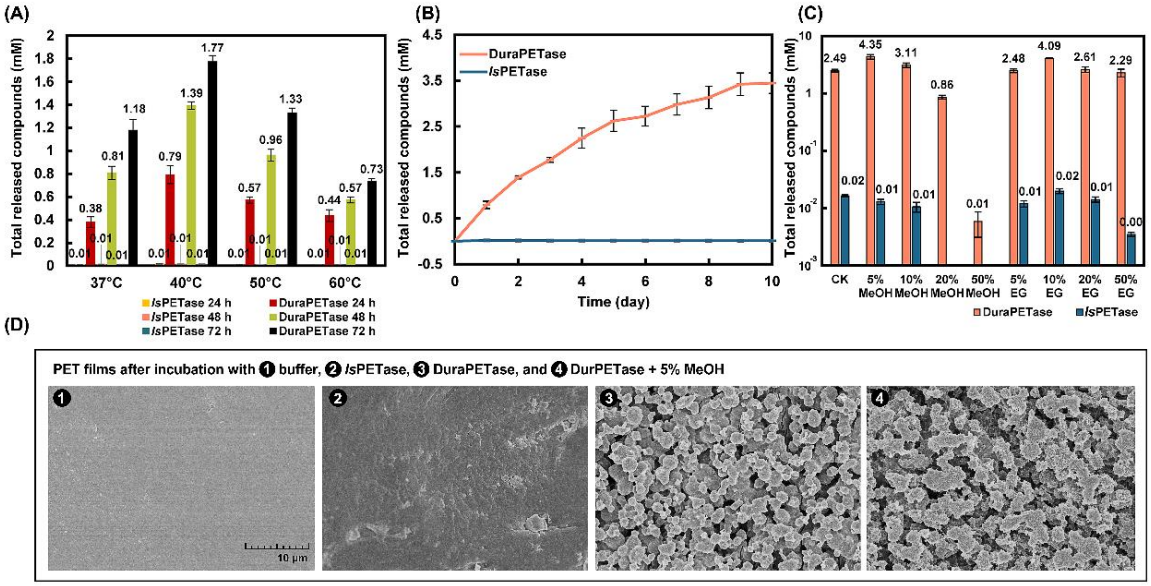


Figure 2 (A) Degrading product release of PET film (49.17% crystallinity) at different temperatures after 3 days incubation with *Is*PETase and DuraPETase; (B) Total product release after incubation with *Is*PETase and DuraPETase for 10 days at 40 °C. (C) Effects of MeOH and EG solvents on the enzymatic activity of DuraPETase. (D) SEM images for buffer-only control of PET film (❶), PET films after incubation with *Is*PETase (❷), DuraPETase (❸), and PET film after incubation with DuraPETase and 5% MeOH (❹). All SEM images were taken after 10 days of incubation at an enzyme loading of 0.01 mg/mL in glycine-NaOH buffer (pH 9.0) or a buffer-only control. Error bars in enzyme assays (A) (B) and (C) represent the s.d. values obtained in triplicate experiments.

Encouraged by the high performance of DuraPETase, it is envisaged that its tolerance to organic solvents would be substantially enhanced, especially for ethylene glycol (EG) solvent,

which can decrease product inhibition during the depolymerizing process. As shown in Figure 2C, DuraPETase enabled a marked increase in tolerance to methanol (MeOH) and EG. The degradation activity was enhanced by 1.7-fold and 1.2-fold at 5% and 10% (v/v) MeOH, respectively, and maintained approximately 35% residual activities at 20% (v/v) MeOH. When incubated with EG solvent, DuraPETase maintained most of its activity at 5%, 20% and 50% (v/v) EG and even increased up to 1.6-fold of degradation activity when incubated at 10% (v/v) EG. This enhanced performance of DuraPETase is considered to be correlated to the increased thermal stability, thus enhancing the potential of this enzyme for auxiliary bio-degradation in glycolysis and alcoholysis application.

The polymer hydrolysis was also confirmed by scanning electron microscopy (SEM) images of the film surfaces following enzymatic hydrolysis for 10 days (Figure 2D). The buffer-only control sample exhibited smooth and uniform surface without characteristic surface defects. The digested PET sample treated by *Is*PETase showed visual modifications with some irregular grooves with different sizes. Remarkably, severe erosions occurred with highly porous foam structures on the film surface after 10 days degradation by DuraPETase, indicating that the biodegradation behavior was not limited to the amorphous area. The hole size after degradation was even larger when 10% MeOH was added (Figure 2D, right panel) and there are many spiny fragments in the corroded area. The film surface became severe nonuniform and partially broke into pieces when touched upon. (Figure S6)

Complete degradation of microplastics by DuraPETase

Pollutions of the marine environment by micro- and even nanoplastics have become ubiquitous. Their persistence continues to increase as they are extremely difficult to be recognized and collected due to the less visibility. The high degradation ability of DuraPETase allows us to explore whether complete degradation of nanoplastics or even microplastics can be achieved. For nanoplastics (⌀ = 50-100 nm), total released product by HPLC indicated complete degradation of nanoplastics by DuraPETase within 1 hour at 37 °C (Figure 3B). While this observation is encouraging, a much slower degradation of PET-microplastic was observed during the incubation period. Figure 3D shows photographs of biodegraded microplastic samples during a period of 20 days. The light transmittance of microplastic solvent exhibited

slowly increase. After 20 days of incubation, the microplastics degraded by DuraPETase almost disappeared with respect to the buffer-only control with microplastics added, which maintained white color over time. Although the *Is*PETase-degraded sample was gradually clarified, its optical transparency was significantly lower than that incubated with DuraPETase. In addition, HPLC analysis of the degrading product concentration reached 3 mM, which also indicated complete degradation of microplastics by DuraPETase.

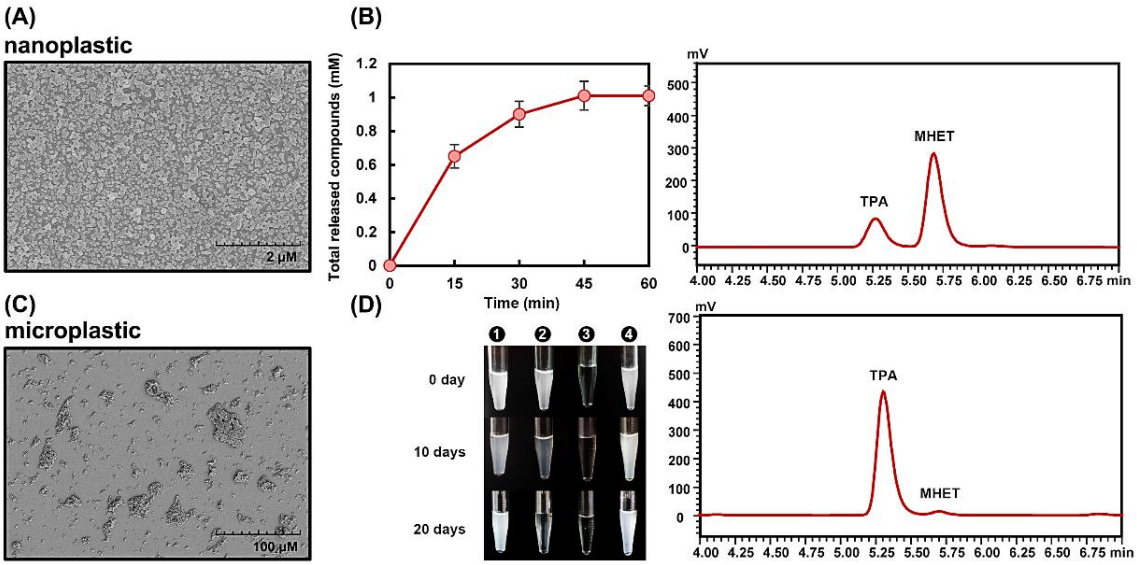


Figure 3 (A) SEM image for PET nanoplastics before incubation. (B) Hydrolytic activities of DuraPETase toward PET nanoplastics. The PET nanoparticles (200 μL) were incubated with 10 μL enzyme (stock concentration 0.1 mg/mL) in 290 μL of 50mM glycine-NaOH buffer (pH 9.0) at 37 ºC for 1 h. The amount of produced MHET and TPA of nanoplastics was monitored by HPLC analysis (right panel). (C) SEM image for PET microplastics before incubation. (D) Time course of PET microplastics degradation. Photographs for transparency changes for microplastics after incubation with *Is*PETase (❶), DuraPETase (❷), buffer-only control (❸) and microplastics incubated in 50mM glycine-NaOH buffer (❹). The PET microparticles (400 μL) were incubated with 10 μL enzyme (stock concentration 0.1 mg/mL) in 290 μL of 50mM glycine-NaOH buffer (pH 9.0) at 37 ºC in ❶ and ❷. The amount of produced MHET and TPA of microplastics was monitored by HPLC analysis (right panel).

Extensions of the potential applications of DuraPETase

To extend the potential applications of DuraPETase, we also examined the use of this enzyme for degradation of other semiaromatic polyesters, including poly(ethylene 2,6-

naphthalenedicarboxylate) (PEN) and poly(tetramethylene terephthalate) (PBT). PEN is similar to PET in chemical structure, but containing the naphthyl rings that produce much stiffer molecular chains than the phenyl rings of PET. Thus, PEN exhibits a high T_g (120 ºC) and good tensile and barrier properties[35]. Poly(butylene terephthalate) (PBT) exhibits higher tensile, flexural and dielectric strengths and faster, more economical moulding characteristics compared to many thermosets. Although the degradation product of PBT is TPA, the longer aliphatic chain of butanediol compared to ethylene glycol endows its excellent resistance to a broad range of chemicals at room temperature ($T_g \approx 50$ ºC)[36].

Biodegradation of PBT was hardly implemented by *Is*PETase due to the notably high specificity of the enzyme toward aromatic groups rather than aliphatic chains (Figure 4). However, a gradual increase of degradation product release was found for the variant at 37 °C, although the degradation activity was substantially reduced with respect to the degradation of PET. SEM analysis also supported this conclusion. In contrast to the slow-degrading experiments with PBT, significant enhancement of PEN degradation by DuraPETase was observed. After 10 days incubation, DuraPETase resulted in maximum product concentration of 48 μM with PEN film, presumably reflecting the high hydrolysis susceptibility of the variant towards semiaromatic polyesters due to the redesigned hydrophobic active cleft, even though the degradation activity was still lower than PET. As the introduction of the naphthalene ring into the main chain stiffens the polymer chains and largely enhanced their dielectric and mechanical properties, PEN holds potential for food packaging, high-performance industrial fibers, and flexible printed circuits applications. Different types of such high-performance plastic would be eventually dispersed into the environment. It therefore inspires further engineering to improve the depolymerization of other semiaromatic polyesters such as PBT.

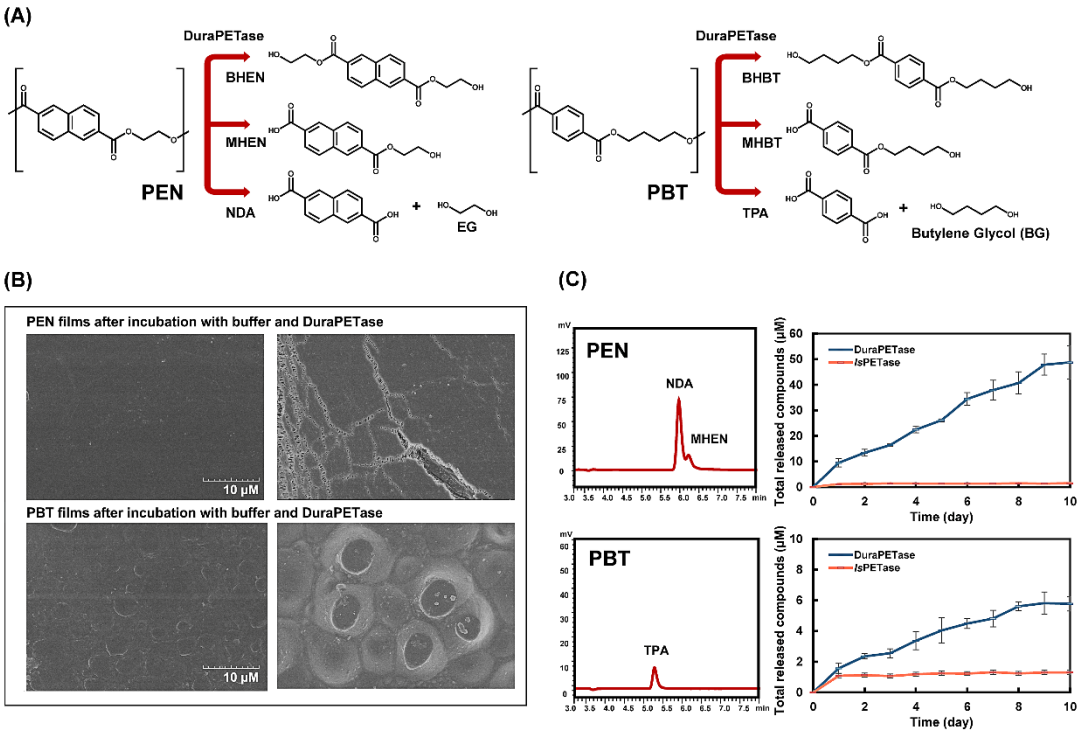


Figure 4 (A) PEN and PBT degradation. DuraPETase catalyzes the hydrolytic cleavage of PEN to produce BHEN, MHEN, and NDA, and converts PBT to BHET, MHET, and TPA. (B) SEM images for PEN and PBT films after incubation with buffer-only and DuraPETase. All SEM images were taken after 10 days of incubation at an enzyme loading of 0.01 mg/mL in glycine-NaOH buffer (pH 9.0) or a buffer-only control at 37 ºC. (C) High-performance liquid chromatography spectrum of the products released from the PEN and PBT films. Error bars represent the s.d. values obtained in triplicate experiments.

Underlying mechanism of the improved properties of DuraPETase

To further evaluate the molecular basis for the improved properties of the obtained variant, we determined the crystal structure of DuraPETase at 1.63 Å resolution (PDB ID: 6KY5). Recent progress in the structural characterization of *Is*PETase has a greatly advanced understanding of the reaction mechanism of *Is*PETase. The binding mode of PET polymer with DuraPETase was predicted by molecular docking, which is similar to that in the crystal structure of *Is*PETase bound to 1-(2-hydroxyethyl) 4-methyl terephthalate (HEMT)[17]. The distance between the carbonyl carbon of PET polymer and the Cα atom of Ser160 was 2.75 Å, which was suggested within the relevant reactive distance for nucleophilic attack. The catalytic triad His237 and Asp206 residues form a charge-relay network that activates the nucleophile (Ser160) to attack the carbonyl carbon atom in the scissile ester bond. Oxyanion of the tetrahedral

intermediate is stabilized by an oxyanion hole that consists of backbone NH groups of Tyr87 and Met160.

Key aspects of the molecular mechanism by which the mutations improve the protein robustness were proposed as follows: introduction of new hydrogen-bonding interactions (T140D, W159H, I168R and S188Q), improvement of hydrophobic packing in the protein surface and interior (L117F, Q119Y, A180I, S214H and R280A), and reduction in the conformational entropy of local unfolded protein (G165A).

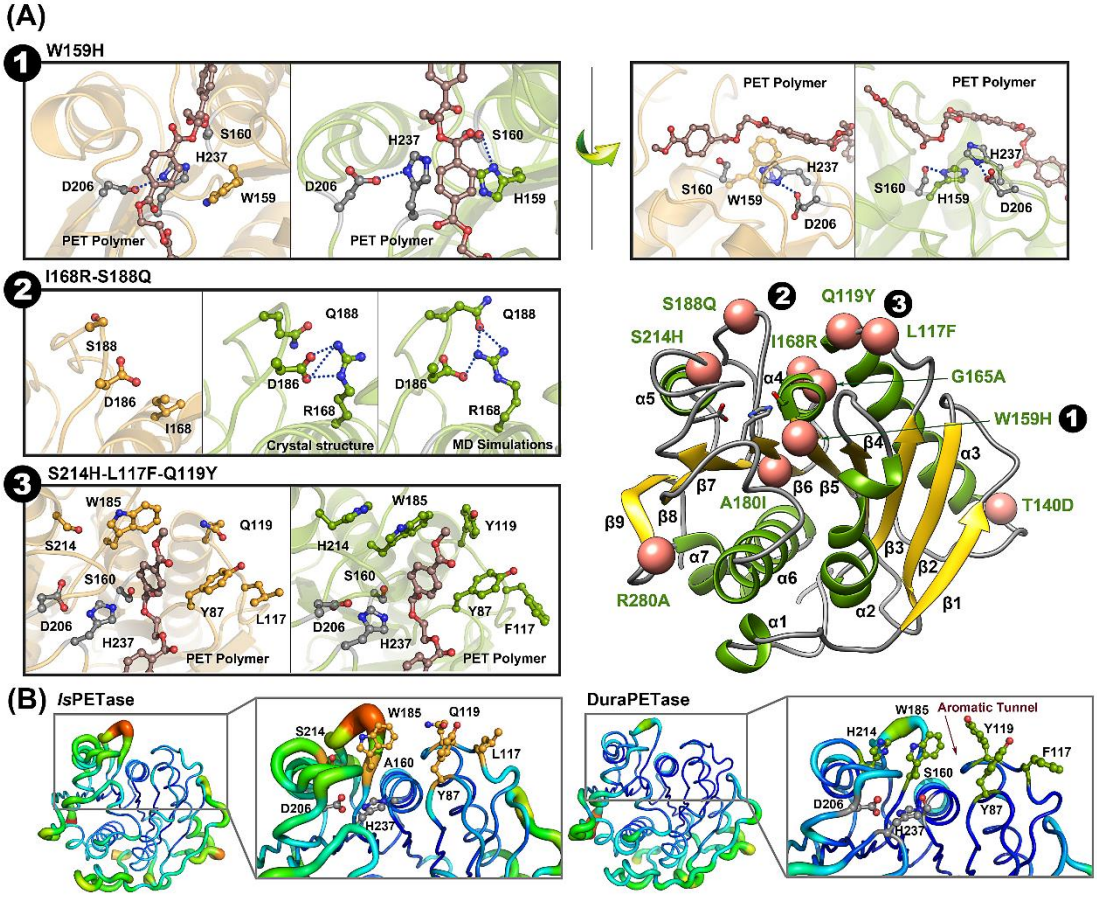


Figure 5 (A) Structural features of DuraPETase. Structural effects of ❶ W159H, ❷ I168R and S188Q, and ❸ S214H, L117F and Q119Y mutations in DuraPETase-PET binding model (green) compared with *Is*PETase (bright orange). Key residues proximal to the stabilizing mutations are shown in ball and sticks representations. Catalytic triad is colored in gray. Right down panel: Location of the stabilizing mutations in the crystal structure of DuraPETase (PDB ID: 6KY5). DuraPETase structure is shown in cartoon while the Cα atoms of the stabilizing mutations are shown as coral atoms. (B) B factors of the X-ray structures of *Is*PETase (PDB ID: 5XH3) and DuraPETase, indicating global changes in protein flexibility, especially for the active site region.

The determined structure of DuraPETase revealed a similar α/β hydrolase fold to the native structure, with a core consisting of seven α helices and nine β sheets. For S140D mutation, the substitution of serine with aspartate conferred a new hydrogen bond with the hydroxyl group of Ser142. The mutating of Trp159 to either alanine or histidine exhibited controversial effects on the enzymatic activity[17-19]. As the assay conditions were similar among these studies, it is possible that differences may be attributed to the degree of PET crystallinity[9]. Although in our crystal structure W159H formed hydrogen bonding with the backbone oxygen atom of H237, molecular dynamics (MD) simulations has revealed a new hydrogen bonding formed between W159H and Ser160, whereas the original catalytic residue His237 may flip "up" out of the catalytic triad to play an aromatic stabilization role with PET (Figure 5A). The reconstruction of the catalytic triad in the active site was also suggested by Austin et al.[15]. For I168R and S188Q mutations, in addition to the new salt-bridge interaction formed between I168R and D186, the guanidine group is also suggested to donate new hydrogen bonds to the amide oxygen and backbone oxygen atoms of S188Q in the MD simulations for 10.36% of the trajectory to stabilize the enzyme structure. Adding D186H to S214H-I168R-W159H (M3) and S214H-I168R-W159H-S188Q (M4) exhibits highly impaired performance, which may be due to the disruption of both the hydrogen-bonding between S188Q and I168R, and the salt-bridge formed by I168R mutation and the native D186 residue. These data confirm the prevailing trend towards negative epistatic interactions when multiple mutations are combined simultaneously.

In addition to electronic effects, the hydrophobic interactions have a marked influence on enzymatic performance (Figure S7). It is noteworthy that the benzene ring of PET molecular stands at an active site crevice, which consists of five mutually contacting phenylalanine, tyrosine, tryptophan, and histidine side chains. To quantitatively identify π-π interactions, time evolutions of the occupancies during MD simulations were calculated (Figure S8). Residues Tyr87 and Trp185 were observed to form continuous π-π stacking interactions with the aromatic motif of PET polymer throughout the entire trajectory for 54.50% and 31.96% of the simulation time, respectively. Meanwhile, L117F and Q119Y are located in proximity to the active site and promoted "T-shaped" π-π interactions with Tyr87 (the oxyanion hole-forming residue) for 71.42% and 68.62% of the trajectory. The stability of the oxyanion hole may be enhanced and the tetrahedral intermediate can be further stabilized. Another substitution S214H in the active

site, which is consistently observed in the homologous enzymes, is suggested to prevent the wobbling of Trp185 by previous study[20]. In our study, S214H is observed to form offset parallel π-π stacking interaction with Trp185 to finely tune the active-site cleft in the crystal structure and maintained for 60.48% occupancies of the simulation time. Therefore, S214H, Trp185, Tyr87, L117F, and Q119Y are suggested to form a well-organized 'aromatic tunnel' flanker suitable for binding of the aromatic motif of PET. Although the flexibility of the active site region has been largely reduced, the hydrophobic substitutions significantly enhanced the binding affinity for PET polymer (Figure 5B), which endows DuraPETase to hydrolyze ester bonds along high-crystallinity PET polymer chain via targeting neighboring aromatic motif in polymer chain backbone.

Discussion

Durability, the plastic's greatest assets, has now become a lingering curse that stays in our environment for hundreds of years[8]. Even when physically broken, plastics never truly leave the environment but are present as micro- and noanoplastics that are choking marine life and propagating up the food chain. The biodegradation of plastics under ambient conditions, especially for uncollectable microplastics, is highly desirable to boost the change of this scenario. To this end, the seminal discovery of *Is*PETase immediately aroused immense interests from many research groups to investigate its mechanistic basis of the catalytic mechanism and attempt to increase the efficiency and stability of this exciting enzyme. Although random mutagenesis combining with high-throughput screening has proven to be a successful strategy for the modification of enzyme properties, a long-sought alternative to screening-based approaches is reliable in *silico* design of performance-enhancing mutations, especially for degradation of unsolvable solid synthetic polymers. Over the last 20 years, in *silico* design based on energy calculations has taken a long way from fairly simple to more accurate and versatile methods[42-43], particularly with a positive impact in the area of protein stabilization. However, the accuracies based on energy functions are still suboptimal because of several factors, including the insufficient conformational sampling of the static structure, imbalances in the force fields, and the intrinsic problems with existing data sets[24]. Although the drawback can be mitigated by using hybrid methods that incorporate complementary statistical-

based approaches such as ABACUS, most stability strategies are focusing on single-point-mutation or simple stepwise combination process, resulting in higher prediction errors upon application to multiple-point mutants. Whenever epistatic effects are present, the predictions prone to frustrate optimization process.

The GRAPE strategy proposed in this study represents a step forward in the prediction of protein stability because of its capability to reduce the risk of combining mutations with antagonistic effects. Most cases where combination process fails to immediately find efficient pathways occurs when coupled mutations are presumable inert. However, making all possible paths simultaneously creates an exponentially larger library. A systematic clustering and greedy combination strategy, which has been successfully utilized in machine learning, therefore provide a way to tap into the possible beneficial pathways in terms of a well-defined library by reducing the dimensionality of the data. The computationally redesigned DuraPETase described herein substantially improves the operational robustness and the high-crystallinity PET degradation activity, which are important properties for its potential use in industrial applications such as surface modification of PET fibers in textile industry and *in situ* disposal of microplastics. Both crystal structure and MD simulations infer an underlying 'aromatic tunnel' mechanism to explain the promoted enzymatic performance via structural view, indicating a significant role of the synergy effects by various beneficial mutations on stability and activity enhancement.

In summary, this work shows that a collection of subtle variations by *in silico* design provides clues of how to redesign a PET hydrolase to better incorporate high-crystallinity aromatic polyesters. We believe that the proposed GRAPE strategy constitutes a significant advance in enzyme design methodology that is complementary to traditional computational design strategies, which often focus on the exquisitely-tuned energy evaluations of single mutations. The designed DuraPETase by GRAPE strategy here serves as a useful catalyst for efficient PET degradation at moderate temperatures and opens up avenues for research in decreasing environmental micro-plastic accumulation. Furthermore, Zimmermann's group has hypothesized combinatorial exo- and endo-type or endo- type only degradation mechanisms for different PET microstructures when incubated with TfCut2 at 70 ºC[44]. Encouraged by the impressive performance of DuraPETase, investigations of the degradation mechanism of scrap

PET are expected to be achieved based on this robust template under mild conditions. Despite the aforementioned achievements, complete degradation of plastic waste still presents a number of challenges. There is an urgent need to further research dealing with this topic, such as coupling with downstream enzymes such as MHETase, further degradation of TPA product, and scaling up of the combination degradation system for industrial applications.

ACKNOWLEDGMENTS. This work is supported by the National Key R&D Program of China (Grant No. 2018YFA0901600), the National Natural Science Foundation of China (Grant Nos. 31601412, 31822002), Beijing Municipal Natural Science Foundation (8194074), and the Biological Resources Programme (KFJ-BRP-009) from the Chinese Academy of Sciences.

References

1. Selke S, Auras R, Nguyen T A, Aguirre E C, Cheruvathur R, Liu Y. Evaluation of biodegradation-promoting additives for plastics. *Environ. Sci. Technol.* **49(6),** 3769-3777 (2015)
2. Plastics Europe. Plastics-the Facts 2018-An analysis of European plastics production, demand and waste data. Available at https://www.plasticseurope.org/en/resources/publications/619-plastics-facts-2018.
3. Sardon H, Dove A P. Plastics recycling with a difference. *Science* **360(6387),** 380-381 (2018)
4. Worm B, Lotze H K, Jubinville I, Wilcox C, Jambecket J. Plastic as a persistent marine pollutant. *Annu Rev Environ Resour*. **42, 1-26** (2017)
5. Geyer R, Jambeck J R, Law K L. Production, use, and fate of all plastics ever made. *Sci Adv.* **3(7),** e1700782 (2017)
6. Narancic T, Verstichel S, Chaganti S R, Morales-Gamez L, Kenny S T, Wilde B D, Padamati R B. Biodegradable plastic blends create new possibilities for end-of-life management of plastics but they are not a panacea for plastic pollution. *Environ Sci Technol.* **52(18),** 10441-10452 (2018)
7. Polyethylene Terephthalate (PET): Production, Price, Market and its Properties. Available at https://www.plasticsinsight.com/resin-intelligence/resin-prices/polyethylene-terephthalate/

8. Editorial. The future of plastic. *Nat. Commun.* **9,** 2157 (2018)
9. Wei R, Zimmermann W. Microbial enzymes for the recycling of recalcitrant petroleum-based plastics: how far are we? *Microb Biotechnol.* **10(6),** 1308-1322 (2017)
10. Wei R, Zimmermann W. Biocatalysis as a green route for recycling the recalcitrant plastic polyethylene terephthalate. *Microb Biotechnol.* **10(6),** 1302-1307 (2017)
11. Yoshida S, Hiraga K, Takehana T, Taniguchi I, Yamaji H, Maeda Y, Toyohara K, Mjyamoto K, Kimura Y, Oda K. A bacterium that degrades and assimilates poly (ethylene terephthalate). *Science* **351(6278),** 1196-1199 (2016)
12. Son H F, Cho I J, Joo S, Seo H, Sagong H, Choi S Y, Lee S Y, Kim K. Rational protein engineering of thermo-stable PETase from *Ideonella sakaiensis* for highly efficient PET degradation. *ACS Catal.* **9(4),** 3519-3526 (2019)
13. de Castro A M, Carniel A, Junior J N, da Conceição Gomes A, Valoni E. Screening of commercial enzymes for poly (ethylene terephthalate)(PET) hydrolysis and synergy studies on different substrate sources. *J Ind Microbiol Biotechnol*. **44(6),** 835-844 (2017)
14. Taniguchi I, Yoshida S, Hiraga K, Miyamoto K, Kimura Y, Oda K. Biodegradation of PET: current status and application aspects. *ACS Catal.* **9(5),** 4089-4105 (2019)
15. Palm G J, Reisky L, Böttcher D, Müller H, Michels E A P, Walczak M C, Berndt L, Weiss M S, Bornscheuer U T, Weber G. Structure of the plastic-degrading *Ideonella sakaiensis* MHETase bound to a substrate. *Nat Commun.* **10(1),** 1717 (2019)
16. 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.* **115(19),** E4350-E4357 (2018)
17. 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.* **8(1),** 2106 (2017)
18. 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.* **9(1),** 382 (2018)
19. 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* **19(14),** 1471-1475 (2018)

20. Fecker T, Galaz-Davison P, Engelberger F, Narui Y, Sotomayor M, Parra L P, Ramirez-Sarmiento C A. Active site flexibility as a hallmark for efficient PET degradation by *I. sakaiensis* PETase. *Biophys J.* **114(6),** 1302-1312 (2018)
21. Liu C, Shi C, Zhu S, Wei R, Yin CC. Structural and functional characterization of polyethylene terephthalate hydrolase from *Ideonella sakaiensis*. *Biochem Biophys Res Commun.* **508(1),** 289-294 (2019)
22. Ma Y, Yao M D, Li B Z, Ding M Z, He B, Chen S, Zhou X, Yuan Y J. Enhanced poly (ethylene terephthalate) hydrolase activity by protein engineering. *Engineering* **4(6),** 888-893 (2018)
23. Weinreich D M, Delaney N F, DePristo M A, Hartl D L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312(5770),** 111-114 (2006)
24. Musil M, Konegger H, Hon J, Bednar D, Damborsky J. Computational design of stable and soluble biocatalysts. *ACS Catal.* **9(2),** 1033-1054 (2018)
25. Sammut C, Webb G I. Encyclopedia of machine learning and data mining. Springer Publishing Company, Incorporated, 2017.
26. Xiong P, Wang M, Zhou X Q, Zhang T C, Zhang J H, Chen Q, Liu H Y. Protein design with a comprehensive statistical energy function and boosted by experimental selection for foldability. *Nat Commun.* **5,** 5330 (2014)
27. Guerois R, Nielsen J E, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J Mol Biol.* **320(2),** 369-387 (2002)
28. Kellogg E H, Leaver-Fay A, Baker D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins.* **79(3),** 830-838 (2011)
29. Wheeler T J, Clements J, Finn R D. Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics* **15,** 7 (2014)
30. Romero P A, Arnold F H. Exploring protein fitness landscapes by directed evolution. *Nat Rev Mol Cell Biol.* **10(12),** 866-876 (2009)
31. Bu Y F, Cui Y L, Peng Y, Hu M R, Tian Y E, Tao Y, Wu B. Engineering improved thermostability of the GH11 xylanase from *Neocallimastix patriciarum* via computational

library design. *Appl Microbiol Biotechnol.* **102(8),** 3675-3685 (2018)

32. Mu Q X, Cui Y L, Tian Y E, Hu M R, Tao Y, Wu B. Thermostability improvement of the glucose oxidase from *Aspergillus niger* for efficient gluconic acid production via computational design. *Int J Biol Macromol.* **136,** 1060-1068 (2019)
33. Jain A K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **31(8),** 651-666 (2010)
34. Quackenbush J. Computational analysis of microarray data. *Nat Rev Genet.* **2(6),** 418-427 (2001)
35. Kim J C, Cakmak M, Zhou X. Effect of composition on orientation, optical and mechanical properties of bi-axially drawn PEN and PEN/PEI blend films. *Polymer* **39(18),** 4225-4234 (1998)
36. Androsch R, Rhoades A M, Stolte I, Schick C. Density of heterogeneous and homogeneous crystal nuclei in poly (butylene terephthalate). *Eur Polym J.* 66, 180-189 (2015)
37. Yang K K, Wu Z, Arnold F H. Machine-learning-guided directed evolution for protein engineering. *Nat Methods* **16(8)**, 687-694 (2019)
38. Sun Z, Liu Q, Qu G, Feng Y, Reetz M T. Utility of B-factors in protein science: Interpreting rigidity, flexibility, and internal motion and engineering thermostability. *Chem Rev.* **119(3),** 1626-1665 (2019)
39. Mirts E N, Petrik I D, Hosseinzadeh P, Nilges M J, Lu Y. A designed heme-[4Fe-4S] metalloenzyme catalyzes sulfite reduction like the native enzyme. *Science* **361(6407),** 1098-1101 (2018)
40. Huang P S, Boyken S E, Baker D. The coming of age of de novo protein design. *Nature* **537(7620),** 320-327 (2016)
41. Li R F, Wijma H J, Song L, Cui Y L, Otzen M, Tian Y E, Du J W, Li T, Niu D D, Chen Y C, Feng J, Han J, Chen H, Tao Y, Janssen D B, Wu B. Computational redesign of enzymes for regio- and enantioselective hydroamination. *Nat Chem Biol.* **14(7),** 664-670 (2018)
42. Bednar D, Beerens K, Sebestova E, Bendl J, Khare S, Chaloupkova R, Prokop Z, Brezovsky J, Baker D, Damborsky J. FireProt: energy-and evolution-based computational design of thermostable multiple-point mutants. *PLoS Comput Biol.* **11(11),** e1004556 (2015)
43. Wijma H J, Floor R J, Jekel P A, Baker D, Marrink S J, Janssen D B. Computationally designed

libraries for rapid enzyme stabilization. *Protein Eng Des Sel.* **27(2),** 49-58 (2014)

44. Wei R, Breite D, Song C, Gräsing D, Ploss T, Hille P, Schwerdtfeger R, Matysik J, Schulze A, Zimmermann W. Biocatalytic degradation efficiency of postconsumer polyethylene terephthalate packaging determined by their polymer microstructures. *Adv Sci*. **6(14),** 1900491 (2019)