Computational and Structural Biotechnology Journal 23 (2024) 982–989

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research Article

Computational design of α-amylase from *Bacillus licheniformis* to increase its activity and stability at high temperatures

Shuai Fan [a,1], Xudong Lü [a,1], Xiyu Wei [a], Ruijie Lü [b], Cuiyue Feng [b], Yuanyuan Jin [a], Maocai Yan [c,*], Zhaoyong Yang [a,*]

[a] Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100050, China
[b] School of Pharmacy, North China University of Science and Technology, Tangshan 063210, Hebei, China
[c] School of Pharmacy, Jining Medical University, Rizhao 276800, Shandong, China

ARTICLE INFO

Keywords:
Protein engineering
Molecular dynamics
α-amylase
Thermostability

ABSTRACT

The thermostable α-amylase derived from *Bacillus licheniformis* (BLA) has multiple advantages, including enhancing the mass transfer rate and by reducing microbial contamination in starch hydrolysis. Nonetheless, the application of BLA is constrained by the accessibility and stability of enzymes capable of achieving high conversion rates at elevated temperatures. Moreover, the thermotolerance of BLA requires further enhancement. Here, we developed a computational strategy for constructing small and smart mutant libraries to identify variants with enhanced thermostability. Initially, molecular dynamics (MD) simulations were employed to identify the regions with high flexibility. Subsequently, FoldX, a computational design predictor, was used to design mutants by rigidifying highly flexible residues, whereas the simultaneous decrease in folding free energy assisted in improving thermostability. Through the utilization of MD and FoldX, residues K251, T277, N278, K319, and E336, situated at a distance of 5 Å from the catalytic triad, were chosen for mutation. Seventeen mutants were identified and characterized by evaluating enzymatic characteristics and kinetic parameters. The catalytic efficiency of the E271L/N278K mutant reached 184.1 g L^{-1} s^{-1}, which is 1.88-fold larger than the corresponding value determined for the WT. Furthermore, the most thermostable mutant, E336S, exhibited a 1.43-fold improvement in half-life at 95 ℃, compared with that of the WT. This study, by combining computational simulation with experimental verification, establishes that potential sites can be computationally predicted to increase the activity and stability of BLA and thus provide a possible strategy by which to guide protein design.

1. Introduction

As natural biocatalysts, enzymes have been extensively utilized in different industries [1], especially in the paper, food, textile, detergent and pharmaceutical industries. There are many applications in industrial and biotechnological fields that benefit from the use of thermostable enzymes since they have innate advantages over mesophilic enzymes, including the enhancement in mass transfer rate, the reduction in contamination risks and substrate viscosity [2]. However, not all enzymes can tolerate high temperature. To enhance the thermostability of enzymes for widespread industrial applications, directed evolution emerges as a formidable strategy [3,4]. However, this approach is a time-consuming and labor-intensive process. Fortunately, with the evolution of computer technology and deepening of studies on the structure and function of enzymes over the last decade, computer-aided enzyme design holds great promise for generating functional changes in numerous enzymes [5]. Meanwhile, more efficient, smaller and smarter libraries can be created by sequence-based, structure-based or MD-based computational tools, which will facilitate the detection of beneficial mutations and accelerate protein engineering [6–8].

Bacillus licheniformis α-amylase (EC 3.2.1.1, BLA) is one of the most widely used enzymes in high-temperature industrial enzymatic processes, such as starch processing, and paper sectors [9]. Specifically, in the process of starch liquefaction, BLA is able to withstand exposure to steam heating at 105 °C, and then continue to effectively hydrolyze starch at 90 °C for 60–90 min. Due to its industrial and academic

* Corresponding authors.
E-mail addresses: maocaiyan@mail.jnmc.edu.cn (M. Yan), zhaoyongy@imb.pumc.edu.cn (Z. Yang).
[1] These authors equally contributed to this work.

https://doi.org/10.1016/j.csbj.2024.02.005
Received 12 September 2023; Received in revised form 7 February 2024; Accepted 7 February 2024
Available online 13 February 2024

importance, the BLA has been extensively engineered for improved thermostability and activity. To date, rational design, directed evolution, and semirational design have been applied to improve BLA activity [10–12], thermal stability [13], acid stability [14–16], and product specificity [10,17]. Since BLA needs to be utilized in gelatinization and liquefaction at the temperature of 90–110 °C to efficiently hydrolyze the starch molecule, it has drawn the greatest interest to identify thermostable BLA variants. The routine methods to enhance enzyme thermostability involve the introduction of additional disulfide bonds [18], salt bridges [19], or mutations targeting amino acid residues with high B-factors [20]. Previous studies have demonstrated that modifying a long loop of BLA and Ca^{2+}-binding sites can enhance the thermostability and activity of BLA. Li et al. [13] reported that the triple mutant (A269K/S187D/N188T) obtained by modifying the long loop region in BLA through sequence alignment and computational analysis using the PoPMuSiC algorithm showed an 84 % increase in activity at pH 5.5 and 70 °C. Another important consideration when designing thermostable BLA variants is their structure containing flexible regions [8]. The highly flexible regions of enzymes, under high temperatures, can lead to enzyme unfolding and denaturation, resulting in enzyme inactivation. Consequently, targeting these highly flexible regions for design can enhance the thermostability of enzymes, a fact that has been confirmed in multiple enzymes such as pullulanase [21], nitrile hydratase [22], and levansucrase [23,24].

Highly flexible regions can be predicted by flexibility analysis (B-factor [8] or root-mean-square fluctuation per residue, RMSF) using MD simulation, which can be performed at high temperatures to accelerate protein unfolding without significantly changing the unfolding pathway [24]. Here, we combined MD simulation and FoldX to obtain and redesign highly flexible regions of BLA, which in turn increase BLA thermostability and catalytic efficiency, enabling more effective enzymatic hydrolysis of starch at high temperature. Therefore, seventeen mutants carrying the mutations K251E, E271L, E271R, E271M, E271Q, E271T, E271V, E271I, E277C, E277I, T277V, N278K, K319W, E336S, E271L/K319W, E271L/N278K and E271L/K251E were generated and characterized. We show results of thermostability and additional characterizations of the mutant of BLA designed, and then the influence of all mutations is explained.

2. Materials and methods

2.1. Molecular dynamics simulations

2.1.1. Structure preparation

The X-ray crystal structure of BLA was obtained from the PDB library under PDB entry code: 1BLI [25]. Water molecules and heteroatoms were removed, and the protonation state of the amino acid residues was assigned using the H++ server [26] at a pH value of 6.5, which is the optimum pH value for BLA catalysis [13].

2.1.2. Molecular dynamics simulation

MD Simulations were performed using the Gromacs 2019.6 package [27]. The CHARMM36 force field was applied to the BLA [28]. The system was solvated with TIP3P water [29] molecules in a dodecahedron box with a 10 Å buffer distance in each direction. Three chloride anions were added to neutralize the system to yield a final system with 31081 atoms. During the simulations, the lengths of all bonds involving hydrogen atoms were restricted using the SHAKE algorithm. The time step for the simulations was set to 2 fs. The particle mesh Ewald (PME) method with a cutoff distance of 10 Å was used to calculate the Coulomb interactions. After energy minimization, the system was heated to 358 K (the optimum temperature for BLA) [13] using the v-rescale temperature coupling scheme in a 1000-ps NVT simulation, followed by 1000-ps NPT simulation using the Parrinello Rahman pressure coupling scheme. After equilibration, three 100 ns MD simulations were performed at 358 K and 500 K(NPT ensemble), respectively. To analyze the stability of the system, RMSD values of the simulated structures were calculated against the minimized structure. Three-dimensional models of BLA mutants were constructed by AlphaFold2 [30], and subsequently, 100 ns MD simulations of K251E, E271L, T277I, N278K, K319W and E336S were carried out using the aforementioned method to optimize the structure for analysis. The computation of entropy in BLA and its mutants was carried out by PACKMAN [31]. The electrostatic surface potential was analyzed using the structure of BLA and its mutants by PyMOL.

2.1.3. Identification of highly flexible residues

Cluster analysis was used to generate a stable and representative structure of BLA at 358 K and 500 K, respectively. Then, these structures were subjected to RMSF analysis. This study characterizes its flexibility by calculating the RMSF value of each residue during the simulation. Highly flexible residues exhibit RMSF values markedly exceeding the average. A cluster of these flexible residues, in close spatial arrangement, constitutes a region of significant flexibility. The highly flexible surface region of BLA, located at least 5 Å away from the catalytic residues, was chosen for rational design of BLA residues.

2.2. Rational design of BLA

The stable and representative structure of BLA from the 40 − 100 ns trajectory at 358 K was used as the design template for FoldX 4.0 algorithm calculations of the change in free energy of unfolding. Residues with high RMSF were selected for further investigation. The ΔΔG values, which were computed from the difference between the free energy of the designed mutant (ΔG_{mutant}) and that of WT (ΔG_{WT}) ($\Delta\Delta G = \Delta G_{mutant} - \Delta G_{WT}$), were applied to evaluate the thermal stability of the BLA mutants. The ΔG represents the free energy difference between the folded and unfolded structures. The mutations were generated and analyzed using FoldX 4.0. The Repair PDB command was used to optimize the total energy of the protein based on FoldX's force field before the mutations were introduced. The PositionScan command of FoldX [32] was used to calculate the ΔΔG value between WT and mutant residues in flexible regions due to single site saturation mutagenesis. Other options were set to default. The experiment was conducted three times, and the average value was computed.

2.3. Cloning and site-directed mutagenesis of BLA

The reconstructed sequence of BLA (NCBI: WP_061576039.1) was codon-optimized and synthesized for expression in *E. coli* by Beijing Genomics Institute (Beijing, China). The plasmid pET-21a (+) (Novagen, USA) and strain *E. coli* BL21 (DE3) (Novagen, USA) were used for plasmid construction and protein expression, respectively. Mutations in BLA were introduced using a ClonExpress Ultra One Step Cloning Kit. Each point mutation was checked and confirmed by sequencing; thus, these recombinant plasmids containing point mutations in the BLA gene were transformed into BL21 (DE3) for expression. All the nucleic acid sequences of the primers are listed in Supplementary Table S1.

2.4. Enzyme production and purification

The *E. coli* BL21(DE3) strains harboring recombinant expression plasmids were incubated at 37 °C for 2–3 h at 200 rpm. Following this, isopropyl-β-D-thiogalactopyranoside (IPTG, 0.15 mM) was added to induce protein expression at 18 °C for 20 h. The cell pellet was harvested by centrifugation at 6000g for 15 min at 4 °C and resuspended in 20 mM phosphate buffer (pH 7.5) with 10 mM imidazole and 300 mM NaCl (lysis buffer), with 1 mM phenylmethylsulfonyl fluoride (PMSF), to a total volume of 50 mL. Cells were disrupted by a high-pressure homogenizer. The supernatant was isolated by centrifugation (60000 g at 4 °C for 40 min) and then loaded onto nickel-affinity chromatography in AKTA Purifier 10 (GE Healthcare, USA) with buffer A (20 mM phosphate buffer, 300 mM NaCl, pH 7.4) and buffer B (20 mM phosphate buffer, 50

mM NaCl, 500 mM imidazole, pH 7.4). The purified protein was obtained by linear gradient elution and confirmed by SDS–PAGE. The product has a molecular mass of approximately 55 kDa, in agreement with the expected BLA size.

2.5. Characterization and kinetic properties of WT and mutant BLAs

The activity of WT and mutant BLAs was measured by the dinitrosalicylic acid (DNS) method [14]. The optimum temperature of WT and mutant BLAs for activity was determined by measuring enzyme activity at optimum pH (50 mM citrate-phosphate buffer, pH 6.5) over various temperatures ranging from 80 °C to 100 °C at intervals of 5 degrees. The kinetic parameters (K_M and k_{cat}) of WT and mutant BLAs were determined by adding soluble starch at different concentrations (1, 2, 4, 6, 8,10, 15, and 20 g·L^{-1}) in 50 mM citrate buffer (pH 6.5) at 70 °C for 5 min. Kinetic data were fitted to the Michaelis–Menten equation by using GraphPad Prism software (GraphPad Software Inc., USA). The WT and mutant BLAs were incubated at 90 °C, and aliquots of enzyme solution were sampled at intervals and immediately cooled on ice for 10 min. The residual activities were then measured under standard conditions. Subsequently, the half-life ($t_{1/2}$) of the WT and mutant BLAs was calculated from the slope of ln (residual activities) versus time plots.

2.6. Differential scanning calorimetry measurements

The melting temperature (T_m) of the WT and mutant BLAs was evaluated by differential scanning calorimetry (DSC, Microcal PEAQ-DSC, Malvern). The melting curve analyses were performed by increasing the temperature from 45 °C to 110 °C at a scan rate of 200 °C/h. The experiment was repeated three times, with background corrected by the subtraction of the respective buffer blanks. The data were fitted using Microcal PEAQ-DSC software with a nontwo state model.

3. Results and discussion

3.1. Identification of highly flexible regions

The redesign of highly flexible regions of BLA is imperative for

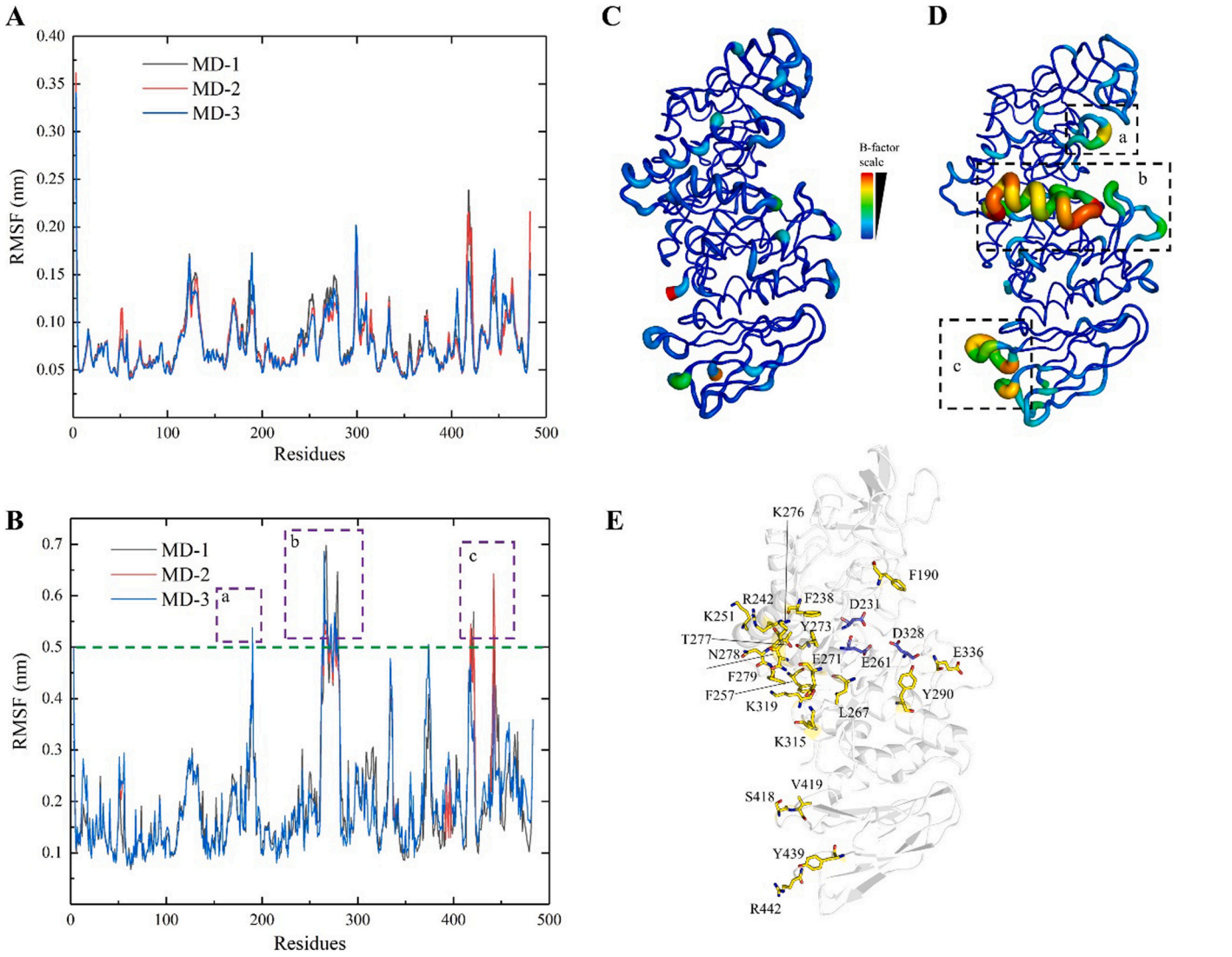

Fig. 1. Computational protein design of the highly flexible region in BLA. The three separate RMSF plots of BLA at 358 K (A) and 500 K (B). The black, red, and blue lines represent the outcomes of three distinct molecular dynamic simulation datasets at 358 K and 500 K. The areas of three highly flexible regions are marked by dotted boxes. The cutoff value for the identification of highly flexible regions is shown by a green dotted line. The B-factor putty visualization of BLA structures, conducted at 358 K (C) and 500 K (D), is depicted using PyMOL (http://www.pymol.org/). The areas with the highest B-factor are highlighted in red, while those with the lowest B-factor are in dark blue, as denoted by the B-factor scale bar. The thickness of the protein backbone correlates with the B factors of Cα atoms. (E) The selected 21 mutants (in yellow) are all situated more than 5 Å away from the catalytic residues (in blue).

enhancing its thermostability and enzymatic activity. To precisely pinpoint these regions, MD simulations were conducted at 358 K and 500 K to calculate RMSD values of BLA's backbone, along with RMSF values for each amino acid residue. The results showed that the RMSD values initially increased and stabilized at both 358 K and 500 K. Specifically, Supplementary Fig. S1 illustrates that the system achieves equilibrium at around 4 ns at 358 K, whereas at 500 K, it takes approximately 40 ns for the system to reach equilibrium. Therefore, the 40 − 100 ns trajectory was selected for subsequent analysis. The average RMSD for BLA at 500 K was observed to be higher than at 358 K, and structurally flexible regions displayed elevated RMSF values compared to more stable areas. Consequently, detailed analysis of the RMSF was also applied to determine the residue flexibility through the simulation period. The RMSF values for BLA at both temperatures exhibit different trends, as illustrated in Fig. 1. Specifically, regions a, b and c, encompassing a total of 21 amino acids, exhibited the highest RMSF values at 500 K, indicating increased disorder compared to their state at 358 K (Figs. 1B and 1C). This suggests that residues within these flexible regions at 500 K undergo more significant structural movements and had higher flexibilities than those at 358 K during simulations. Remarkably, in region b, while remaining stable at 358 K, it heightened flexibility at 500 K. Most amino acids in region b are located in the α-helix, while the amino acids in regions a and c are located in random coils. Notably, none of the 21 sites were located within 5 Å of the catalytic triad (Asp231, Glu261, Asp328). Therefore, we used the RMSF value to identify highly flexible regions of BLA and used this information to design mutants in which these regions may be useful to stabilize protein structure under high temperatures. Finally, twenty-one relatively flexible sites were identified (Fig. 1D).

3.2. Computational protein design

The protein folding free energy (ΔG) is an important characteristic directly related to protein stability. The relative values ΔΔG calculated by FoldX represent the difference in free energy between WT and BLA mutants, which provides a quick way to evaluate the mutational effect on protein stability. The ΔΔG value is universally recognized as a reliable indicator of thermostability, where a negative ΔΔG indicates enhanced stability [33–35]. Subsequently, 399 potential mutants were generated and computed by FoldX through site saturation mutagenesis of the above-described twenty-one relatively flexible sites. Finally, FoldX identified 14 unique mutations across six sites with ΔΔG < -0.85 kJ·mol^{-1} (Fig. 2A, Supplementary Table S2), indicating that these mutations could potentially augment the thermostability of BLA. Interestingly, most studies show that optimizing the loop element could improve the thermotolerance of the enzyme [36–39]. In contrast, in this study, five mutational sites (K251, E271, T277, N278 and K319) show the ability to increase the heat tolerance of BLA, which are located in

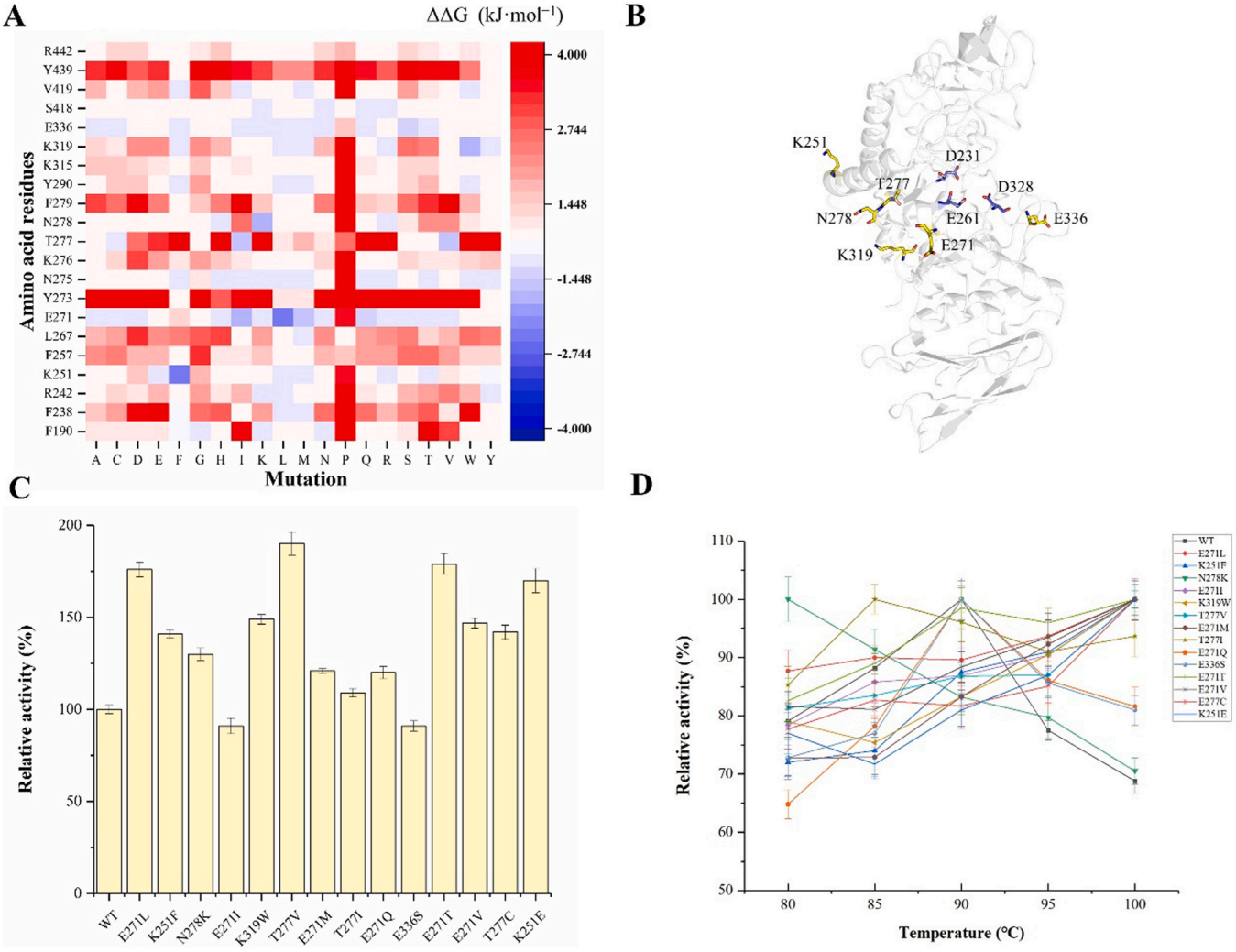

Fig. 2. Computer-aided prediction of mutations and their activity. (A) Heatmap of ΔΔG-value changes in 20 canonical amino acid residues in BLA, (B) The positions of beneficial mutation sites, catalytic residues and mutant residues are marked with blue and yellow, respectively. Relative activity (C) and optimum temperatures (D) of BLA and its mutants.

region b, whose amino acid is located in the α-helix (Fig. 2B).

3.3. Enzymatic assay of single-site mutation

It is widely acknowledged that mutations leading to the rigidification of specific residues, thereby substantially reducing their fluctuations, can increase the thermal stability of the target protein [40,41]. However, proper flexible conformations are beneficial for the occurrence of the catalytic reaction. Hence, improved enzyme thermostability may reduce the activity of the enzyme. Therefore, we first determined the specific activity at elevated temperature and optimum reaction temperature of the WT and BLA mutants. The changes in activity of WT and variants are shown in Fig. 2C. Notably, all variants, with the exception of the E271I and E336S mutants, showed higher residual activity at 95 °C than the WT. For some mutants, the activity was increased by more than 40%. In particular, E271L, T277V, E271T and K251E, increased activity by approximately 76%, 90%, 79% and 70% compared with that of the WT, respectively. In addition, the optimum temperature for catalytic activity was determined the WT and BLA mutants. As shown in Fig. 2D, the optimum temperature for the WT was 90 °C. It was encouraging to note that the optimum temperatures of E271L, K251F, E271I, K319W, T277V, E271M, E271T, E271V and K251E were all higher than that of the WT, reaching 100 °C. E271Q and E336S had no effect on the optimum temperature (90 °C) compared with that of WT, though the relative activity of E271Q increased at 90 °C. However, the optimum temperature for T277I and N278K was reduced by 5 °C and 10 °C, respectively. In comparison to the WT, most of the mutants showed improved activity at 95 °C.

3.4. Thermostabilities and kinetic parameters of the designed mutants

To assess the stability of BLA and its mutants, the T_m and half-life ($t_{1/2}$) were evaluated to characterize enzyme kinetic stability (Table 1). As shown in Table 1, the E271T, E271M and E271L/N278K mutants exhibited T_m values similar to that of the WT. However, the T_m values of E271L, E271R, E271Q, E271V, E271I, T277I, N278K, K319W and E336S increased by 4.79 °C, 3.75 °C, 4.45 °C, 3.16 °C, 2.20 °C, 4.53 °C, 3.09 °C, 5.31 °C and 3.98 °C, respectively, and those of K251E and T277C decreased by 3.67 °C and 17.01 °C, respectively. This result suggests that it is possible to decrease the ΔΔG, which demonstrated the higher T_m of the mutant over the WT. The $t_{1/2}$ of the WT was 285 s, but it was prolonged in some mutants. The $t_{1/2}$ values of the mutants were 315 s for E271L, 330 s for E271R, 315 s for E271L, 301 s for N278K, 313 s for K319W and 407 s for E336S at 95 °C, which were 1.11-, 1.16-, 1.11-, 1.06-, 1.10-, 2.07-, and 1.43-fold higher than that of WT, respectively. Furthermore, our findings indicate that T277C significantly reduced enzymatic stability, resulting in a reduced $t_{1/2}$. In contrast, mutants E271L, E271R, N278K and E336S with a high T_m value, exhibit $t_{1/2}$ that surpass that of the WT. However, despite maintaining high T_m values, the $t_{1/2}$ of the three double mutants were significantly decreases.

Kinetic parameters of WT and 14 mutants were determined using soluble starch as the substrate at pH 6.5 and 95 °C, which showed a discrepancy in k_{cat} and K_M for various mutants (Table 1). Specifically, the K_M values of K251E, E271L, E271M, E271Q, E271I and N278K mutants were similar to WT. Excitingly, the E271R and E336S mutants had K_M values approximately 32% and 42% less than that of the WT, respectively. Therefore, the E271R and E336S mutants have more affinity for the substrate soluble starch than the WT. However, the K_M values of K251E, E271M, E271T, E271V, T277I and T277V were increased to varying degrees, which means that these mutants have less affinity for the substrate than WT. Compared with WT, K251E, E271L, E271T, T277C and K319W showed a higher k_{cat} and therefore exhibited an increased k_{cat}/K_M (catalytic efficiency) of 1.72/1.72/1.63/1.72/1.76 times that of WT, respectively. The combined results of $t_{1/2}$ and k_{cat}/K_M showed that theE271L and K319W mutants exhibited higher stability (1.11- and 2.07-fold, respectively) and catalytic efficiency (1.72- and 1.76-fold, respectively) at elevated temperatures than other beneficial mutations. Furthermore, we combined the E271L, K319W, N278K, and K251E mutations into double mutations to enhance activity and thermotolerance. Unfortunately, the $t_{1/2}$ of those double mutations was markedly decreased (Table 1). Thus, the E271L, K319W, N278K, and K251E mutations have an additive effect on catalytic turnover; on the other hand, the combination of these mutants shows negative feedback on heat tolerance.

Table 1
Enzymatic properties of WT and its mutants at pH 6.0 and 95 °C.

	K_M ($g{\cdot}L^{-1}$)	k_{cat} (s^{-1})	k_{cat}/K_M ($g{\cdot}L^{-1}{\cdot}s^{-1}$)	$t_{1/2}$ (s)	T_m (°C)
WT	3.88 ± 0.15	380 ± 10.95	97.9	285 ± 8.26	91.40 ± 0.59
K251E	4.35 ± 0.13	734 ± 15.67	168.7	182 ± 5.79	87.73 ± 0.76
E271L	4.13 ± 0.09	696 ± 12.39	168.5	315 ± 10.36	96.19 ± 0.74
E271R	2.62 ± 0.05	317 ± 9.91	121.0	330 ± 12.45	95.15 ± 0.36
E271M	4.32 ± 0.31	539 ± 11.67	124.8	289 ± 14.32	92.68 ± 0.83
E271Q	3.36 ± 0.28	430 ± 13.27	128.0	257 ± 12.89	95.85 ± 1.24
E271T	5.04 ± 0.19	803 ± 15.86	159.3	231 ± 9.66	90.32 ± 1.28
E271V	4.55 ± 0.20	598 ± 12.56	131.4	228 ± 10.13	94.56 ± 0.57
E271I	3.40 ± 0.17	318 ± 8.72	93.5	315 ± 9.89	93.60 ± 0.94
T277C	2.88 ± 0.08	484 ± 9.26	168.1	92 ± 6.53	74.39 ± 1.04
T277I	4.88 ± 0.11	507 ± 14.16	103.9	267 ± 12.61	95.93 ± 0.85
T277V	5.81 ± 0.26	858 ± 15.50	147.7	161 ± 7.59	90.10 ± 0.69
N278K	3.75 ± 0.11	518 ± 10.29	138.1	301 ± 13.62	94.49 ± 0.91
K319W	2.93 ± 0.09	505 ± 10.09	172.4	323 ± 13.25	96.71 ± 1.18
E336S	2.25 ± 0.08	269 ± 7.23	119.6	407 ± 14.38	95.38 ± 1.37
E271L/K319W	7.88 ± 0.29	1441 ± 32.67	182.9	193 ± 9.37	95.43 ± 1.25
E271L/N278K	8.61 ± 0.32	1585 ± 39.15	184.1	139 ± 7.81	92.86 ± 0.91
E271L/K251E	6.84 ± 0.17	1254 ± 28.94	183.3	210 ± 13.68	95.66 ± 1.04

3.5. Molecular interaction analysis for improved thermostability

Through MD simulations at various temperatures and calculations of folding free energy, we successfully identified 17 mutations across 6 positions that can enhance either the activity or stability of the enzyme. To probe the molecular basis for increased thermostability, three-dimensional models of BLA mutants were constructed by AlphaFold2 [30]. Subsequently, 100 ns MD simulations for the mutants (K251E, E271L, T277I, N278K, K319W and E336S) were carried out to optimize the structure for further evaluation (Supplemental Fig. S2). Curiously, the RMSF values of the four mutant variants K251E, E271L, T277I and K319W exhibited varying degrees of reduction, aligning with the anticipated outcomes of the designed mutations. However, it is noteworthy that the RMSF values for N278K and E336S showed unexpected increases (Supplemental Table S3). An increase in independence for thermal activity was observed when E271 was mutated to leucine (E271L), arginine (E271R), methionine (E271M), glutamine (E271Q), threonine (E271T), valine (E271V) and isoleucine (E271I). However, not all mutants at the 271th position resulted in improved thermal stability. Only the E271L, E271R and E271I mutants significantly enhanced

heat resistance. E271 lies in a hidden cove on the surface of BLA and forms three H-bonds with K315 and H316, which are localized in a loop motif (Fig. 3C). The E271L mutation disrupts hydrogen bonds, a change typically detrimental to thermal stability. However, the increase in flexibility enhances the entropy change of E271L (Supplemental Table S4), ultimately leading to an improvement in its heat resistance (Table 1). We surmise that similar circumstances may have occurred in other mutant variants (E271I), employing an increase in entropy to counteract the unfavorable enthalpic changes compromising thermal resilience. Therefore, the increased thermotolerance of E271L, E271R and E271I occurs when a favorable folding free energy overcomes an unfavorable breaking intermolecular hydrogen bond. Interestingly, Declerck et al. [42] investigated over 175 mutants and identified key residues for maintaining activity and stability in BLA. Interestingly, some surface residues showed improved stability when mutated to hydrophobic residues [43,44], which is similar to some of the mutations in our study. As lysine is mutated into glutamate (K251E), the charge of this site changes from positive to negative while reducing the size of the amino acid side chain (Fig. 3A). The mutation of N278 to lysine results in these regions being rich in positively charged amino acids (Fig. 3B). In the case of K251E, the mutation led to an increase in entropy (Supplemental Table S4). Furthermore, the results obtained from FoldX calculations also indicate a decrease in folding energy for K251E. Hence, the thermal stability of K251E is indeed improved. Compared to the cationic sidechains of lysine, the K319W mutation not only sustained hydrogen bond formation by the indole NH but also formed new π-π stacking interactions with F279 (Fig. 3D). In general, mutating charged amino acids on the protein surface to hydrophobic amino acids tends to decrease protein stability. However, in studies of BLA [42,43], it has been reported that mutating five surface positions to hydrophobic amino acids (H133V, N190F, A209V, N264S, and Q265Y) can actually enhance protein stability. In this study, despite disrupting the original hydrogen bond and mutating the charged surface residue to a hydrophobic residue, such as E271L and K319W, increased SASA was observed to varying degrees. We speculate that these mutations primarily improve BLA thermostability by increasing local entropy and reducing ΔΔG. However, the accumulation of entropy increase could actually lead to decreased protein stability, as supported by experimental data on a double mutant (E271L/K319W). Further scrutiny was given to the mutant E336S, which disrupts a hydrogen bond with G332 and two salt bridges with R375. The severance of this hydrogen bond is expected to enhance swing movement of the loop region, accompanied a positional conformational shift in the catalytic residue D328, potentially optimizing it for starch hydrolysis. Moreover, E336S reveals significantly improved thermal tolerance compared to WT. Despite the conventional understanding that hydrogen bonds play a pivotal role in enzyme heat resistance, not all hydrogen bonds are crucial for protein stability. Mutants like E271L and E336S disrupt hydrogen bonds or salt bridges, ostensibly reducing protein thermal stability. Unexpectedly, these mutants increase the thermal stability of the protein by reducing the overall free energy of folding. It was also surprising that combinations of these mutations did not appear to have an additive effect as we would expect. Unfortunately, the $t_{1/2}$ of E271L/K319W, E271L/N278K and E271L/K251E were reduced up to 32%, 51% and 26% compared to the WT, respectively. A possible explanation is that individual hydrogen bonds or salt bridges do not substantially contribute to BLA's thermotolerance on their own, but their collective absence leads to decreased stability at elevated temperatures. Notably, E271, K251, and T277 are located on α-helices, which is distinct from previous studies that focused on loop structure or Ca^{2+}-binding sites. By utilizing MD, the number of

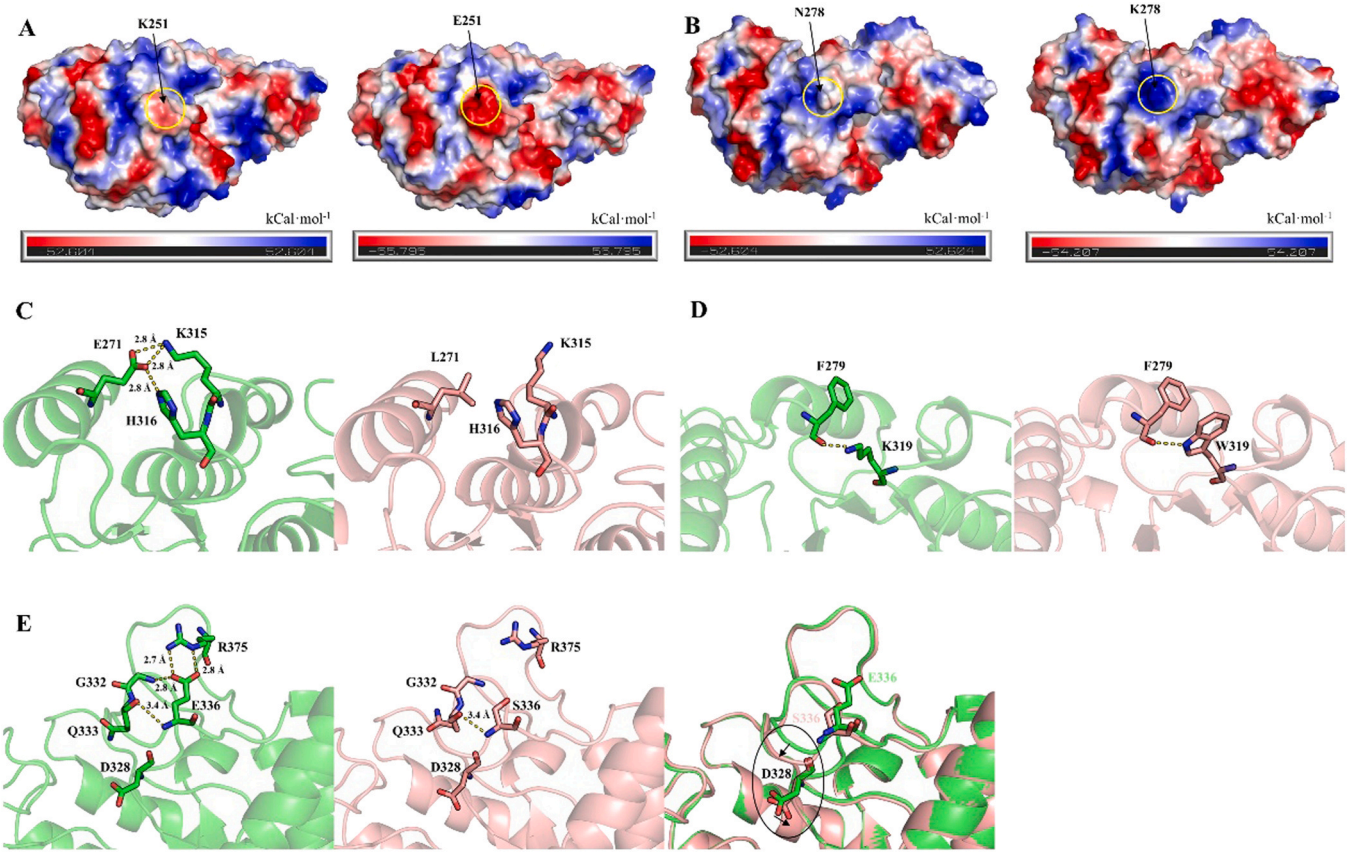

Fig. 3. The interresidue interactions between WT and mutants. Analysis of surface electrostatic interactions in K251E (A) and N278K (B); hydrogen bonds between E271 (C), K319 (D), E336 (E) and their neighboring residues. The WT and mutants are shown as cartoon views in green and pink, respectively. The electrostatic potential was calculated using PyMOL.

candidate mutation sites can be effectively reduced by focusing solely on highly flexible regions. This method not only minimizes the workload needed for experimental validation but also enhances the efficiency of designing proteins in a rational manner. Moreover, in this study, region b demonstrated a relatively low RMSF at 358 K. However, at 500 K, the RMSF increased dramatically. It is worth noting that the final active and thermally stable sites were located in this region. As such, in future studies, we should pay closer attention to regions where the RMSF increases significantly under high-temperature simulations. Rational design in this area may universally enhance the thermal stability and activity of proteins.

4. Conclusions

Overall, computation-guided protein engineering can greatly enhance the productivity of positive mutations. In this study, we combined MD and FoldX to improve thermostability of enzymes by reducing the overall free energy of folding of highly flexible residues using site-directed mutagenesis. Seventeen positive mutants were obtained from 402 potential mutants constructed with computational results by FoldX predictions. Experimentally validated, nine of those seventeen mutants improved thermostability and enzymatic activity at 95 °C compared with that of WT, indicating that these mutations achieved both thermotolerance improvement and activity enhancement. The E271L and K319W are the most effective mutants, showing a 72% and 76% increase in k_{cat}/K_M compared to the WT, respectively. Additionally, their $t_{1/2}$ also improved, with increases to 1.11 and 1.13 times, respectively. This work provides a practical strategy for protein engineering that could be useful for improving the thermostability of other enzymes.

CRediT authorship contribution statement

Z. Y. supervised the study, S. F., X. L. performed the experiments and wrote the manuscript, X. W., R. L., and C. F. participated in the experiments, X. L., M. Y. performed data analyses, S. F., X. W., Y. J. and Z. Y. revised the manuscript. All of the authors have read and approved the final manuscript.

Declaration of Competing Interest

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

Data Availability

Prepared structures and MD simulation input files, parameter files, topology files are available in the Github (https://github.com/shadowlyu/BLA.git).

Acknowledgements

This work was funded by the National Key Research and Development Program of China (No. 2018YFA0901800), National Natural Science Foundation of China (Grants 8187131584), CAMS Innovation Fund for Medical Sciences (CIFMS, 2021-I2M-1-055), Medical and Health Science and Technology Development Plan Project of Shandong Province (202113050691).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2024.02.005.

References

[1] Jemli S, Ayadi-Zouari D, Hlima HB, Bejar S. Biocatalysts: application and engineering for industrial purposes. Crit Rev Biotechnol 2016;36(2):246–58.
[2] Lehmann M, Wyss M. Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. Curr Opin Biotechnol 2001;12(4):371–5.
[3] Eijsink VG, Gaseidnes S, Borchert TV, van den Burg B. Directed evolution of enzyme stability. Biomol Eng 2005;22(1-3):21–30.
[4] Guan L, Gao Y, Li J, Wang K, Zhang Z, Yan S, et al. Directed evolution of pseudomonas fluorescens lipase variants with improved thermostability using error-prone PCR. Front Bioeng Biotechnol 2020;8:1034.
[5] Planas-Iglesias J, Marques SM, Pinto GP, Musil M, Stourac J, Damborsky J, et al. Computational design of enzymes for biotechnological applications. Biotechnol Adv 2021;47:107696.
[6] Huang PS, Boyken SE, Baker D. The coming of age of de novo protein design. Nature 2016;537(7620):320–7.
[7] Gupta N, Beliya E, Paul JS, Tiwari S, Kunjam S, Jadhav SK. Molecular strategies to enhance stability and catalysis of extremophile-derived alpha-amylase using computational biology. Extremophiles 2021;25(3):221–33.
[8] Sun Z, Liu Q, Qu G, Feng Y, Reetz MT. Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. Chem Rev 2019;119(3):1626–65.
[9] Jujjavarapu SE, Dhagat S. Evolutionary Trends in Industrial Production of alpha-amylase. Recent Pat Biotechnol 2019;13(1):4–18.
[10] Rivera MH, Lopez-Munguia A, Soberon X, Saab-Rincon G. Alpha-amylase from Bacillus licheniformis mutants near to the catalytic site: effects on hydrolytic and transglycosylation activity. Protein Eng 2003;16(7):505–14.
[11] Ruan Y, Zhang R, Xu Y. Directed evolution of maltogenic amylase from Bacillus licheniformis R-53: enhancing activity and thermostability improves bread quality and extends shelf life. Food Chem 2022;381:132222.
[12] Tran PL, Cha HJ, Lee JS, Park SH, Woo EJ, Park KH. Introducing transglycosylation activity in Bacillus licheniformis alpha-amylase by replacement of His235 with Glu. Biochem Biophys Res Commun 2014;451(4):541–7.
[13] Li Z, Duan X, Chen S, Wu J. Improving the reversibility of thermal denaturation and catalytic efficiency of Bacillus licheniformis alpha-amylase through stabilizing a long loop in domain B. PLoS One 2017;12(3):e0173187.
[14] Liu Y, Fan S, Liu X, Zhang Z, Wang J, Wang Z, et al. A highly active alpha amylase from Bacillus licheniformis: directed evolution, enzyme characterization and structural analysis. J Microbiol Biotechnol 2014;24(7):898–904.
[15] Liu YH, Hu B, Xu YJ, Bo JX, Fan S, Wang JL, et al. Improvement of the acid stability of Bacillus licheniformis alpha amylase by error-prone PCR. J Appl Microbiol 2012;113(3):541–9.
[16] Liu YH, Lu FP, Li Y, Wang JL, Gao C. Acid stabilization of Bacillus licheniformis alpha amylase through introduction of mutations. Appl Microbiol Biotechnol 2008;80(5):795–803.
[17] Fincan SA, Ozdemir S, Karakaya A, Enez B, Mustafov SD, Ulutas MS, et al. Purification and characterization of thermostable alpha-amylase produced from Bacillus licheniformis So-B3 and its potential in hydrolyzing raw starch. Life Sci 2021;264:118639.
[18] Suzuki M, Date M, Kashiwagi T, Suzuki E, Yokoyama K. Rational design of a disulfide bridge increases the thermostability of microbial transglutaminase. Appl Microbiol Biotechnol 2022;106(12):4553–62.
[19] Ban X, Wu J, Kaustubh B, Lahiri P, Dhoble AS, Gu Z, et al. Additional salt bridges improve the thermostability of 1,4-alpha-glucan branching enzyme. Food Chem 2020;316:126348.
[20] Parra-Cruz R, Jager CM, Lau PL, Gomes RL, Pordea A. Rational design of thermostable carbonic anhydrase mutants using molecular dynamics simulations. J Phys Chem B 2018;122(36):8526–36.
[21] Bi J, Chen S, Zhao X, Nie Y, Xu Y. Computation-aided engineering of starch-debranching pullulanase from Bacillus thermoleovorans for enhanced thermostability. Appl Microbiol Biotechnol 2020;104(17):7551–62.
[22] Cheng Z, Lan Y, Guo J, Ma D, Jiang S, Lai Q, et al. Computational design of nitrile hydratase from Pseudonocardia thermophila JCM3095 for improved thermostability. Molecules 2020;25(20).
[23] Klaewkla M, Pichyangkura R, Charoenwongpaiboon T, Wangpaiboon K, Chunsrivirot S. Computational design of oligosaccharide producing levansucrase from Bacillus licheniformis RN-01 to improve its thermostability for production of levan-type fructooligosaccharides from sucrose. Int J Biol Macromol 2020;160:252–63.
[24] Klaewkla M, Pichyangkura R, Chunsrivirot S. Computational design of oligosaccharide-producing levansucrase from Bacillus licheniformis RN-01 to increase its stability at high temperature. J Phys Chem B 2021;125(22):5766–74.
[25] Machius M, Declerck N, Huber R, Wiegand G. Activation of Bacillus licheniformis alpha-amylase through a disorder–>order transition of the substrate-binding site mediated by a calcium-sodium-calcium metal triad. Structure 1998;6(3):281–92.
[26] Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. Nucleic Acids Res 2005;33(Web Server issue):W368–71.
[27] Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. J Comput Chem 2005;26(16):1701–18.
[28] Huang J, MacKerell Jr AD. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem 2013;34(25):2135–45.

[29] Yagasaki T, Matsumoto M, Tanaka H. Lennard-Jones parameters determined to reproduce the solubility of NaCl and KCl in SPC/E, TIP3P, and TIP4P/2005 Water. J Chem Theory Comput 2020;16(4):2460–73.
[30] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596(7873):583–9.
[31] Khade PM, Jernigan RL. Entropies Derived from the Packing Geometries within a Single Protein Structure. ACS Omega 2022;7(24):20719–30.
[32] Delgado J, Radusky LG, Cianferoni D, Serrano L. FoldX 5.0: working with RNA, small molecules and a new graphical interface. Bioinformatics 2019;35(20):4168–9.
[33] Buss O, Muller D, Jager S, Rudat J, Rabe KS. Improvement in the thermostability of a beta-amino acid converting omega-transaminase by using FoldX. Chembiochem 2018;19(4):379–87.
[34] Chen J, Chen D, Chen Q, Xu W, Zhang W, Mu W. Computer-aided targeted mutagenesis of thermoclostridium caenicola d-allulose 3-epimerase for improved thermostability. J Agric Food Chem 2022;70(6):1943–51.
[35] Rahban M, Zolghadri S, Salehi N, Ahmad F, Haertle T, Rezaei-Ghaleh N, et al. Thermal stability enhancement: fundamental concepts of protein engineering strategies to manipulate the flexible structure. Int J Biol Macromol 2022;214:642–54.
[36] Kheirollahi A, Khajeh K, Golestani A. Investigating the role of loop 131-140 in activity and thermal stability of chondroitinase ABC I. Int J Biol Macromol 2018;116:811–6.
[37] Morris P, Rios-Solis L, Garcia-Arrazola R, Lye GJ, Dalby PA. Impact of cofactor-binding loop mutations on thermotolerance and activity of E. coli transketolase. Enzym Micro Technol 2016;89:85–91.
[38] Peng Z, Miao Z, Ji X, Zhang G, Zhang J. Engineering flexible loops to enhance thermal stability of keratinase for efficient keratin degradation. Sci Total Environ 2022;845:157161.
[39] Ruggiero A, Smaldone G, Esposito L, Balasco N, Vitagliano L. Loop size optimization induces a strong thermal stabilization of the thioredoxin fold. FEBS J 2019;286(9):1752–64.
[40] Yu H, Dalby PA. Exploiting correlated molecular-dynamics networks to counteract enzyme activity-stability trade-off. Proc Natl Acad Sci USA 2018;115(52):E12192–200.
[41] Kheirollahi A, Khajeh K, Golestani A. Rigidifying flexible sites: An approach to improve stability of chondroitinase ABC I. Int J Biol Macromol 2017;97:270–8.
[42] Declerck N, Machius M, Wiegand G, Huber R, Gaillardin C. Probing structural determinants specifying high thermostability in Bacillus licheniformis alpha-amylase. J Mol Biol 2000;301(4):1041–57.
[43] Machius M, Declerck N, Huber R, Wiegand G. Kinetic stabilization of Bacillus licheniformis alpha-amylase through introduction of hydrophobic residues at the surface. J Biol Chem 2003;278(13):11546–53.
[44] Declerck N, Machius M, Joyet P, Wiegand G, Huber R, Gaillardin C. Hyperthermostabilization of Bacillus licheniformis alpha-amylase and modulation of its stability over a 50 degrees C temperature range. Protein Eng 2003;16(4):287–93.