Computational-assisted protein engineering to develop thermostable and highly active catalase for industrial and biocatalytic applications

Shuai Xu [a], Ya-min Chen [a] , Xiang-yu Meng [a], Ru Pan [a], Ao-xuan Yan [a] , Zhi-min Li [a,b,*],
Zong-lin Li [a,*]

[a] *State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, 200237 Shanghai, China*
[b] *Shanghai Collaborative Innovation Center for Biomanufacturing Technology, 200237 Shanghai, China*

HIGHLIGHTS

- A high-activity catalase was identified via deep learning models.
- ProteinMPNN-ΔΔG design enhanced catalase thermostability and activity.
- The D78P/K201R/E384Y/T435A variant showed 1.9 × activity and 4.9 × half-life.
- The engineered catalase removed H_2O_2 from three simulated industrial matrices.
- Coupling with alditol oxidase enabled glycerol-to-glycerate conversion.

GRAPHICAL ABSTRACT
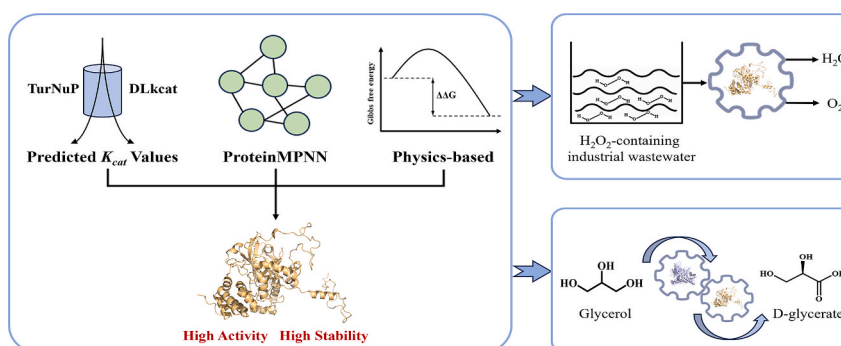
ARTICLE INFO

ABSTRACT

Catalases are ideal biocatalysts for hydrogen peroxide removal in industrial wastewater and enzymatic processes, yet the practical application is hindered by poor thermostability. In this study, a highly active catalase was identified from *Fictibacillus enclensis* using a computer-assisted screening strategy. To overcome thermal instability, a computational design framework integrating ProteinMPNN-based sequence optimization with physics-based energy calculations was developed. The engineered variant, D78P/K201R/E384Y/T435A, exhibited a 1.9-fold increase in catalytic efficiency and a 4.9-fold extension of half-life at 40°C. Molecular dynamics simulations and structural analyses revealed that the mutations conferred enhanced global rigidity through stabilized hydrogen-bond networks. Moreover, the variant was employed for the treatment of industrial effluents containing hydrogen peroxide residues and the biocatalytic upgrading of glycerol. This study not only paved the way for industrial applications of catalase but also established a strategic framework for enhancing both stability and catalytic activity in enzyme engineering.

1. Introduction

Hydrogen peroxide (H_2O_2) functions as a broad-spectrum biocide and bleaching agent, with pervasive applications spanning diverse industries, including food processing, textile bleaching, and medical sterilization (Ciriminna et al., 2016; Fukuzumi et al., 2021). However, the discharge of H_2O_2-laden wastewater poses significant environmental risks, and residual H_2O_2 in food products raises concerns regarding health implications upon ingestion (Arslan-Alaton et al., 2012; Bopitiya et al., 2021). The removal of residual H_2O_2 has emerged as a critical priority in textile production, food safety protocols, and environmental pollution mitigation strategies (Tian et al., 2012). Conventional treatment methodologies for H_2O_2 effluents involve water dilution, thermal degradation, and chemical reduction approaches, which are associated with unsustainable energy demands and risks of secondary contamination (Yu et al., 2016).

Catalase has been demonstrated to exhibit exceptional catalytic efficiency in rapidly decomposing H_2O_2 into water and molecular oxygen (Tehrani and Moosavi-Movahedi, 2018). This unique property renders it an ideal candidate for the elimination of residual H_2O_2 in diverse applications (Lončar and Fraaije, 2015). The direct introduction of catalase into H_2O_2-containing wastewater not only ensures rapid detoxification but also minimises secondary environmental impacts (Vasudhevan et al., 2025; Fruhwirth et al., 2002). Beyond industrial applications, coupling catalase with oxidases is critical for enzymatic cascade reactions, as it prevents enzyme inactivation caused by H_2O_2 accumulation and facilitates oxygen regeneration (Wang et al., 2024; Chen et al., 2022). However, the commercially catalases are extracted from animal liver tissues, involving complex processing steps and high production costs (Meyer et al., 1997). Although microbial fermentation offers an alternative, it often encounters a trade-off between activity and stability. Catalases derived from sources such as *Aspergillus niger* and *Bacillus subtilis* typically exhibit high catalytic activity but limited stability, whereas thermophilic-origin catalases display enhanced thermostability at the expense of reduced catalytic activity (Eberhardt et al., 2004; Kocabas et al., 2008; Wang et al., 1998). The trade-off between activity and stability imposes limitations on industrial feasibility, underscoring the importance of engineering catalases with both high catalytic efficiency and thermal stability.

The integration of artificial intelligence and protein engineering has introduced unprecedented opportunities for enzyme discovery and enhancement of protein stability (Bian et al., 2024; Silverstein et al., 2025). Accurate prediction of catalytic constants is key for the identification of enzymes with high catalytic efficiency. TurNuP introduced a novel approach for identifying highly active enzymes, employing reaction fingerprints to comprehensively consider substrate and product profiles (Kroll et al., 2023). The DLKcat model implemented a hybrid architecture that demonstrates superior predictive capability in k_{cat} determination (Li et al., 2022). In parallel, AI-driven strategies have been leveraged to improve protein stability. The deep learning-based protein sequence design model ProteinMPNN has demonstrated remarkable capabilities in guiding stability optimization (Justas et al., 2022; Sumida et al., 2024). Nevertheless, a critical limitation arises from the inherent design paradigm of ProteinMPNN, prioritising native sequence recovery over explicit identification of physical stability determinants (Dieckhaus et al., 2024). Conventional physics-based protein stabilization strategies utilize computational predictions of Gibbs free energy changes to evaluate mutational effects, yet frequently neglect the accompanying alterations in enzyme activity induced by such mutations (Weinstein et al., 2021; Musil et al., 2017). Therefore, strategic integration of the sequence design capabilities of ProteinMPNN with physics-based energy modeling enables precise identification of mutations that enhance stability while maintaining catalytic activity, and this strategy has not yet been applied to catalase engineering.

In this study, the high-activity catalase A0A0V8J8H6 was identified through the utilisation of learning models. To address industrial application requirements, we developed a computational workflow integrating ProteinMPNN-guided sequence design with ΔΔG profiling to construct a high-confidence variant library. Guided by this strategy, the optimal variant D78P/K201R/E384Y/T435A exhibited a 4.9-fold increase in half-life at 40 °C and a 1.9-fold improvement in catalytic efficiency. In addition, the optimal variant exhibited robust efficacy in the treatment of H_2O_2-containing wastewater. Coupling the optimal variant with alditol oxidase enabled a cascade system for the value-added conversion of glycerol. Overall, this work demonstrated the potential of artificial intelligence in the computational-aided identify and engineering of catalase with enhanced catalytic activity and thermal stability.

2. Materials and methods

2.1. Database mining for enzymes

To acquire highly active and thermostable catalase sequences, the uniprot database was searched using the catalase from *Ureibacillus thermosphaericus* (uniprot ID: A0A2H4YCD5) as the query sequence, with a sequence identity threshold of < 90 %. Phylogenetic analysis of the candidate proteins was performed using the Maximum Likelihood algorithm implemented in MEGA 11 (Tamura et al., 2021). Computational prediction of thermal stability for the candidate proteins was conducted via DeepSTABp, with the melting temperature serving as a quantitative metric for structural robustness (Jung et al., 2023). To effectively screen candidate proteins with high catalytic potential, two deep learning models, DLKcat and the Kcat prediction model, were employed.

2.2. Protein structure modeling and proteinMPNN design

The protein structures of catalase were predicted using AlphaFold3 (https://golgi.sandbox.google.com/), with heme incorporated as the ligand into protein structure (Abramson et al., 2024). All parameters were set to their default values during the prediction process. The model with the highest pTM and ipTM scores was selected from the five predicted structures for subsequent structural analysis. To validate the reliability of the selected model, a comprehensive quality assessment was performed using the SAVES v6.1 suite (Colovos and Yeates, 1993). The tunnel analysis was conducted utilizing the CAVER 3.0 PyMOL plugin with the following computational parameters applied to the protein's three-dimensional structure: a probe radius of 1.4 Å, shell radius of 3 Å, shell depth of 4 Å, max distance of 3 Å, and desired radius of 5 Å (Chovancova et al., 2012).

Designing catalase using ProteinMPNN. The protein structure of catalase from *Fictibacillus enclensis* (UniProt ID: A0A0V8J8H6) was employed as input to ProteinMPNN. To maintain the integrity of the catalytic function, all residues within a 6 Å radius of the heme cofactor were excluded from the designable positions. Sequence generation was conducted using a ProteinMPNN model that was trained with the application of 0.2 Å backbone noise to enhance its robustness. Three different temperatures (0.1, 0.2, and 0.3) were sampled, with 20 sequences generated at each temperature to ensure a comprehensive exploration of potential sequences. The structures of sequences generated by ProteinMPNN were predicted using AlphaFold3. The sequences generated by ProteinMPNN were filtered based on the predicted local distance difference test (plDDT) scores and Cα root-mean-square deviation (RMSD) values, with the selection criteria set at Cα RMSD < 1.2 Å and plDDT > 87.0 to ensure high accuracy and reliability.

2.3. Gibbs free energy change calculations

Utilizing the protein structure, Gibbs free energy changes (ΔΔG) were calculated using both the FoldX and Rosetta ddg_monomer methodologies. FoldX was employed using standard parameters, with each

calculation repeated for five replicates to determine the ΔΔG. Based on the reported FoldX error margin (0.46 kcal/mol), mutations yielding ΔΔG values exceeding + 0.92 kcal/mol were classified as destabilizing (Studer et al., 2014). For Rosetta ddg_monomer calculations, we employed the settings described by Kellogg *et al.* (options: −ddg::local_opt_only true-ddg::opt_radius 8.0-ddg::weight_file soft_rep_design −ddg::iterations 50-ddg::min_cst false-ddg::mean true-ddg::min false-ddg::sc_min_only false-ddg::ramp_repulsive false). Within this protocol, destabilizing mutations are defined as those exhibiting ΔΔG ≥ + 3 kcal/mol (Kellogg et al., 2011).

2.4. Protein expression and purification

For recombinant protein expression, the *E. coli* BL21 (DE3) expression system was employed in combination with the pET-28a (+) expression vector. The cells harboring the recombinant plasmid were cultured in 100 mL of LB medium supplemented with 50 mg L^{-1} kanamycin at 37 °C with shaking at 220 rpm. Protein expression was induced by adding 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when the OD_{600} reached 0.6–0.8. Following induction at 18 °C for 16 h, the cells were harvested by centrifugation and resuspended in 50 mL of KPi buffer (pH 8.0) containing 10 mM imidazole and 500 mM NaCl. Cell lysis was then performed using high-pressure homogenization, and the lysate was centrifuged at 7000 rpm for 20 min at 4 °C. The His-tagged target protein was purified from the supernatant using nickel-chelating affinity chromatography. The purification process involves washing with 60 mM imidazole to remove impurity proteins and elution of the target protein with 200 mM imidazole. The enzyme purity was assessed via SDS-PAGE and quantified utilizing the BCA Protein Assay Kit from Tiangen (Beijing, China).

2.5. Construction of variants

Site-directed mutations were introduced by PCR. Primers for site-directed mutagenesis were designed using the wildtype catalase gene as a template. Whole plasmid PCR amplification was performed to generate the mutants, which were subsequently transformed into *E. coli* DH5α for cloning amplification and DNA sequencing verification. The mutant plasmids were transformed into *E. coli* BL21 (DE3) cells for recombinant protein expression.

2.6. Enzyme activity assays and kinetic parameters analysis

The activity of catalase was assayed using the reaction mixture containing 50 mM H_2O_2 and 0.0001 g L^{-1} catalase in 100 mM KPi buffer. The reaction solution was incubated in a thermostatic metal bath for 1 min. After the reaction, 200 μL of the reaction mixture was added to 100 μL of 20 % trichloroacetic acid (TCA) solution to terminate the reaction. Subsequently, 100 μL of the terminated reaction solution was mixed with 100 μL of 50 mM ammonium molybdate solution. The color of the mixture changed from colorless to yellow, and the absorbance at 405 nm was measured as an indicator of catalase activity. One unit of catalase activity was defined as the amount of enzyme that converted 1 μmol of substrate H_2O_2 per minute. In addition, to estimate the optimal reaction temperature and pH, catalase activities were determined to be in the range of 20–70 °C and pH 5.0–9.0, respectively.

Under optimal conditions, the initial reaction rates were determined at varying concentrations of H_2O_2 to calculate the kinetic parameters. K_m, V_{max}, and k_{cat} were derived by fitting the experimental data to the Michaelis-Menten equation through nonlinear regression analysis.

2.7. Thermal stability calculation

The thermal stability of the purified enzyme was assessed by measuring its residual activity after incubation at 40 °C or 50 °C for varying time intervals. The thermal inactivation kinetics of the enzyme were analyzed using a first-order kinetic model (Eq. (1)), where A represents the residual enzyme activity at time t (in hours), A_0 corresponds to the initial enzyme activity, and *k* denotes the first-order inactivation rate constant (h^{-1}). Based on this kinetic analysis, the thermal stability of the enzyme was quantitatively expressed through its half-life ($t_{1/2}$), which was defined as the time required for the residual activity to decrease to 50 % of the initial activity.

$$\ln A = \ln A_0 - kt \tag{1}$$

The melting temperature (Tm) of the enzyme was determined using circular dichroism (CD) spectroscopy. Purified enzyme was diluted to a final concentration of 0.2 g L^{-1} in phosphate buffer (100 mM, pH 8.0). The temperature was ramped from 20-90 °C at a rate of 1 °C/min, and the detection wavelength range was 180–260 nm. The Tm value was calculated by analyzing data using Global 3 software.

2.8. MD simulation

Molecular dynamics (MD) simulations of the wildtype and the optimal variant were performed using the GROMACS 2020.3 software package with the AMBER14SB force field (Spoel et al., 2005). The simulation system was constructed using a rectangular box filled with TIP3P water molecules, with a minimum distance of 10 Å between the protein and the box boundaries in all directions. The system was first subjected to energy minimization using the steepest descent algorithm for 50,000 steps, followed by the conjugate gradient algorithm, with a force constant of 100 kJ/mol/nm^2 applied to the protein backbone atoms to ensure the reasonability of overall protein conformation. Afterward, the system was gradually heated to 323 K or 303 K over 100 ps under canonical ensemble (NVT) using the Nosé-Hoover thermostat, followed by pressure equilibration at 1 bar under the isothermal-isobaric (NPT) ensemble using the Parrinello-Rahman barostat for 1 ns. After equilibrium, a 200 ns position-free MD simulation was performed, with snapshots taken every 10 ps for subsequent analysis.

2.9. H_2O_2-containing wastewater treatment

To evaluate the industrial application potential of catalase, three distinct systems were investigated. The medical wastewater system contained 3 % H_2O_2 supplemented with 100 mg L^{-1} glucose, 700 mg L^{-1} NaCl, 10 mg L^{-1} bovine serum albumin (BSA), 10 mg L^{-1} FeCl$_3$, and 0.002 g L^{-1} enzyme. The textile bleaching effluent system comprised 1 % H_2O_2 with 200 mg L^{-1} lignin, 0.25 % (w/v) EDTA, 1 % (w/v) NaOH, 100 mg L^{-1} MgSO$_4$, 100 mg L^{-1} CaCl$_2$, and 0.001 g L^{-1} enzyme. The dairy system consisted of skim milk containing 14.7 mM hydrogen peroxide combined with 0.0001 g L^{-1} enzyme. In medical wastewater and textile bleaching treatments, enzymatic supplementation (0.001 g L^{-1} and 0.0005 g L^{-1}) was administered at 10-minute intervals when the residual H_2O_2 concentration exceeded 100 mM.

2.10. Cascade reaction

D-glycerate was synthesized in 10 mL of 100 mM KPi buffer (pH 7.0) containing glycerol, 2.0 g L^{-1} *Tf*AldO$^{V256L/P257I}$, and 0.0005 g L^{-1} catalase. The reaction was conducted at 40 °C, and samples were collected at various time points to monitor the reaction.

2.11. Analytical methods

The concentration of glyceraldehyde and D-glycerate was determined by HPLC equipped with an Organic Acid Analysis Column (300 × 7.8 mm, Aminex HPX-87H). The column temperature was set to 65 °C, with an injection volume of 20 μL and a detection wavelength of 214 nm. The mobile phase was 5 mM sulfuric acid at a flow rate of 0.6 mL/min.
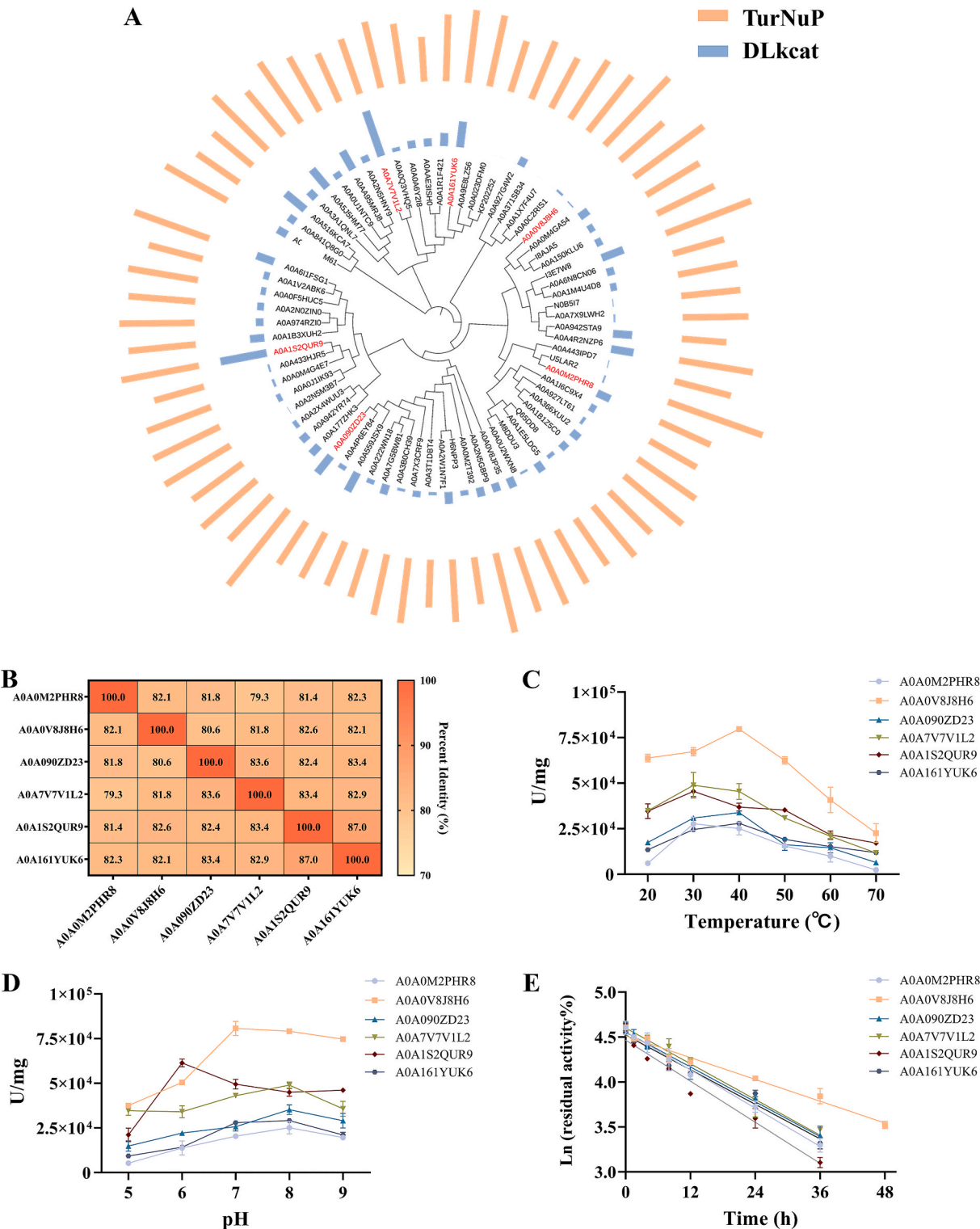
All experiments were conducted in triplicate, and the results are presented as mean ± standard deviation. Statistical analysis was

performed using *t*-test to compare mean values between groups. Differences were considered statistically significant at P < 0.05, P < 0.01, and P < 0.001.

3. Results and discussion

3.1. Computer–Aided discovery of novel catalases

To identify catalases exhibiting both high catalytic activity and

Fig. 1. Deep learning-guided discovery of novel catalases and functional characterization. (A) A phylogenetic tree of 70 catalases with bar charts. The blue and orange bars represent the predicted activity by the DLkcat and TurNup models, respectively. A higher bar indicates a greater predicted k_{cat} value. (B) Sequence similarity of the candidate catalases. (C) Enzyme activity of the candidate catalases across a temperature spanning 20–70 °C. (D) Determination of optimum reaction pH of the candidate catalases. (E) Thermal stability analysis showing the half-life of the candidate catalase measured at 40 °C. Error bars represent standard deviation of at least three replicates. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

thermal stability, a homology search was conducted in the UniProt database. The monofunctional catalase from *Ureibacillus thermosphaericus*, which is characterized by exceptional catalytic efficiency and thermostability, was employed as the query sequence (Jia et al., 2017). 70 catalases derived from different species were retrieved from the database and employed to construct a phylogenetic tree, elucidating their evolutionary relationships and functional divergence (Fig. 1A). The melting temperatures (Tm) of the candidate proteins were predicted using the deep learning model DeepSTABp, revealing narrowly distributed Tm values across all candidates (see supplementary material). Therefore, screening candidate enzymes based on predicted stability became unfeasible, leading to a shift toward selection based on catalytic efficiency. To identify catalases with high catalytic potential, enzyme turnover numbers (k_{cat}) of candidate proteins were predicted using two deep learning models, DLKcat and TurNuP. Integrating predictions from

both models allowed for mutual compensation of individual model limitations, thereby improving the overall reliability of the screening process. Given the discrepancies in k_{cat} predictions between the two models, the top three candidates with the highest predicted k_{cat} values from each model were selected for further analysis (DLKcat: A0A7V7V1L2, A0A1S2QUR9, A0A161YUK6; TurNuP: A0A0M2PHR8, A0A090ZD23, A0A0V8J8H6). These candidates occupy distinct evolutionary branches and share < 90 % sequence similarity (Fig. 1B), ensuring that subsequent experiments would capture the functional diversity of catalases.

Six candidates were successfully heterologously expressed in *Escherichia coli*, and purified enzymes were subsequently employed to assess their catalytic activity, temperature property, and pH optima. The enzymes A0A0M2PHR8, A0A7V7V1L2, and A0A1S2QUR9 exhibited maximal catalytic activity at 30 °C, while A0A1S2QUR9, A0A161YUK6,
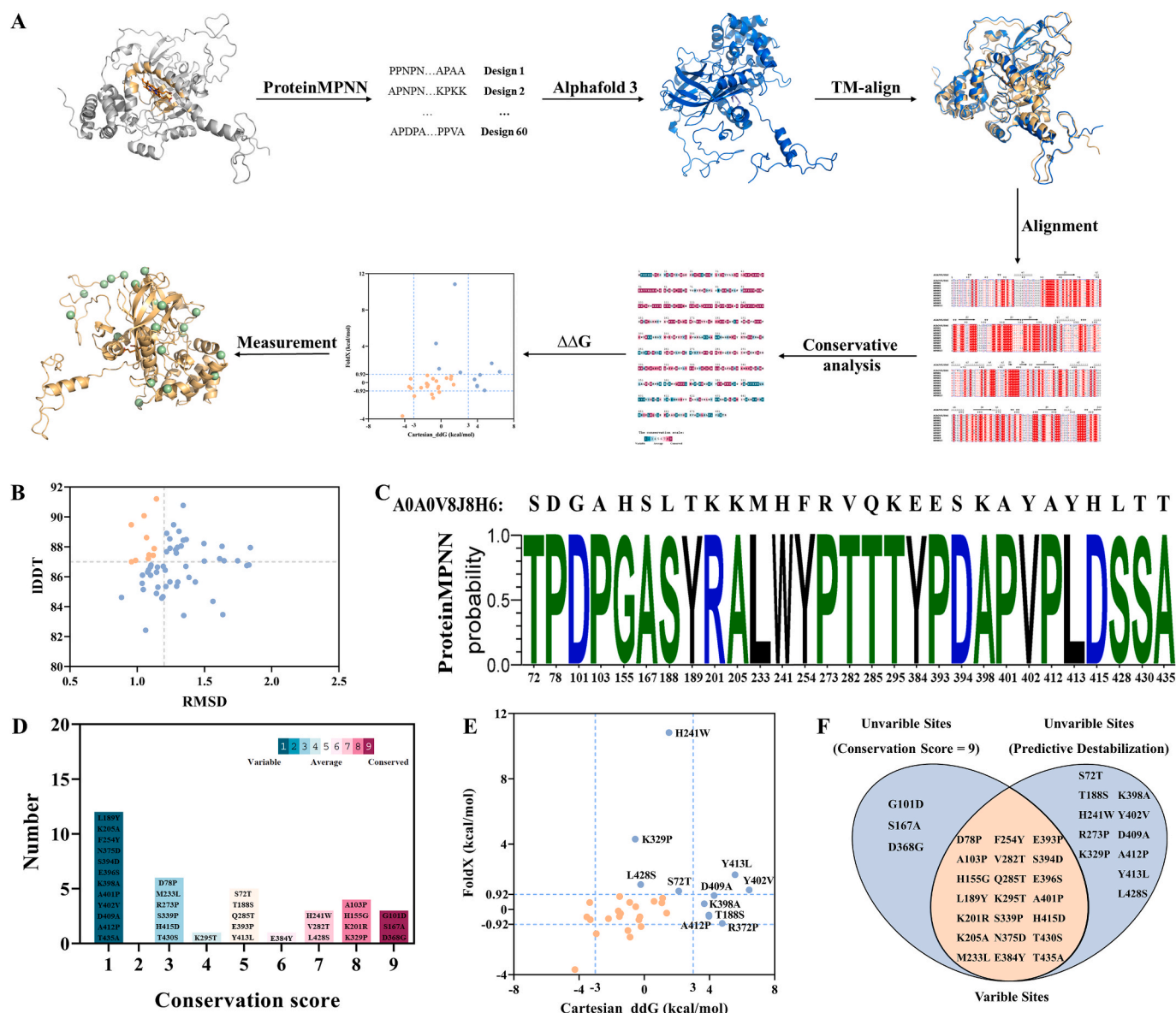
Fig. 2. Deep learning-guided protein engineering workflow for thermostable enzyme **design.** (A) Overview of the computational workflow integrating ProteinMPNN, AlphaFold3, and stability prediction tools (FoldX and Rosetta ddg_monomer) for thermostable enzyme design. (B) Structural alignment of 10 ProteinMPNN-designed variants with the AlphaFold3-predicted structure of the target enzyme A0A0V8J8H6. (C) Identification of 35 candidate mutation sites based on sequence alignment of the refined variant library. (D) Evolutionary conservation analysis of candidate mutation sites using ConSurf, with conservation scores represented by a color gradient (red: conserved, green: variable). (E) ΔΔG predictions of the 32 variable sites using FoldX and Rosetta ddg_monomer. (F) Final selection of 21 high-confidence thermostabilizing mutations retained for experimental validation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and A0A0V8J8H6 indicated optimal reaction temperatures of 40 °C (Fig. 1C). All enzymes retained over 50 % of their activity across a broad pH range of 6.0–9.0 (Fig. 1D). Among the candidates, A0A0V8J8H6 demonstrated the highest catalytic performance under optimal conditions, showing a specific activity of 79727.9 U/mg. Notably, this enzyme retained significant residual activity at elevated temperatures (50 °C) and alkaline pH conditions.

Time-dependent activity loss presents a common challenge for industrial enzyme applications. To evaluate the industrial potential of these catalases, their thermal stability was assessed at 40 °C (Fig. 1E). While the remaining five enzymes exhibited comparable half-lives ($t_{1/2}$), A0A0V8J8H6 demonstrated a substantially prolonged $t_{1/2}$ of 29.8 h. However, industrial processes in the food and textile sectors that employ catalase for hydrogen peroxide decomposition are typically conducted under elevated temperatures (50 °C or 60 °C). The $t_{1/2}$ of A0A0V8J8H6 at 50 °C decreased significantly to 2.2 h, highlighting the necessity for further protein engineering to enhance its thermostability for practical industrial deployment.

3.2. ProteinMPNN–Guided selection of mutational residue sites

The deep learning-based protein sequence design method, ProteinMPNN, has been employed to generate variants of native proteins. Sequences generated by ProteinMPNN reliably fold into the native backbone structure with high accuracy. Furthermore, these designed variants exhibit superior performance in enhancing expression levels and thermostability compared to the wildtype counterparts. Therefore, we developed a workflow that integrates ProteinMPNN to guide the rational design of enzyme, specifically targeting enhanced thermostability for industrial biocatalytic applications (Fig. 2A). The backbone structure of the target enzyme was initially input into the ProteinMPNN framework. To preserve catalytic function, residues within the active site were constrained during sequence generation, while temperature sampling parameters were systematically modulated to balance sequence diversity with structural plausibility. The generated sequence was subjected to structural prediction using AlphaFold3, followed by structural alignment with the target protein. Sequence demonstrating reliable folding into the native backbone conformation were retained, thereby constructing a refined library that preserves native-like structural integrity while exploring mutational diversity at non-conserved regions. Subsequently, sequence alignment was performed between the refined variant library and the target protein to identify candidate mutation sites at non-conserved regions. The ΔΔG of these variants were computationally predicted using FoldX and Rosetta ddg_monomer, enabling the elimination of mutation sites that destabilized the protein structure. Finally, the remaining high-confidence mutation sites were experimentally validated through site-directed mutagenesis. Overall, the workflow we have established integrates deep learning-driven sequence diversification with physics-based stability screening, ensuring that variants retain catalytic functionality while acquiring the thermostability required for industrial applications. Notably, the entire computational process can be completed within a few hours.

Based on the aforementioned strategy, the three-dimensional structure of A0A0V8J8H6 was first determined using AlphaFold3. The structural was used as the design template for ProteinMPNN, with residues within 6 Å of the heme cofactor was fixed to preserve heme-binding integrity and catalytic function. Following structural alignment of the variant structures with the A0A0V8J8H6 template, 10 variants meeting thresholds (Cα-RMSD < 1.2 Å and plDDT > 87.0) were retained (Fig. 2B). Sequence alignment identified 35 candidate mutation sites, defined as residues exhibiting amino acid substitutions distinct from A0A0V8J8H6 while maintaining sequence identity > 90 % across the refined library (Fig. 2C). To further safeguard functionally essential residues, the evolutionary conservation profile of the 35 candidate sites was assessed using ConSurf. Analysis revealed that 3 of the 35 residues were evolutionarily invariant (Conservation score = 9; Fig. 2D). The

ΔΔG for the remaining 32 sites were calculated, resulting in the exclusion of 11 mutations predicted to destabilize the protein structure (Fig. 2E). The elimination of these physics-based predictions refined the candidate pool to 21 high-confidence mutations for experimental validation (Fig. 2F).

3.3. Verification of variant activity and thermostability

To evaluate the catalytic activity and thermal stability, 21 variants were overexpressed in *E. coli* and subsequently purified. Specific activity analyses of these purified enzymes revealed that two variants, D78P and E384Y showed significant enhanced enzymatic activity, achieving relative activities of 174 % and 156 %, respectively, in comparison to the wildtype (Fig. 3A). Furthermore, thermal stability assessments were conducted by determining enzyme half-life at 50 °C (Fig. 3B). Seven variants, namely, D78P, K201R, E384Y, E393P, E396S, T430S, and T435A were identified with enhanced thermal stability. However, while the E393P, E396S, and T430S mutants exhibited longer half-life, their catalytic activities were markedly reduced to 64 %, 58 %, and 70 % of the wildtype, respectively. In contrast, the D78P, E384Y, K201R, and T435A variants demonstrated enhanced thermostability while maintaining unaltered or even elevated catalytic activity. Based on these findings, combinatorial mutants were engineered by combining D78P, K201R, E384Y, and T435A mutations to synergistically enhance both catalytic efficiency and thermal stability. This rational design strategy generated six double mutants, four triple mutants, and one quadruple mutant. These 11 combinatorial mutants were subsequently purified and systematically evaluated using the same methodology. As illustrated in Fig. 3C, all 11 combinatorial mutants exhibited further improvements in thermostability compared to single-mutation, with the K201R/E384Y/T435A triple mutant suggesting the most pronounced enhancement. This variant achieved a 4.1-fold increase in half-life relative to the wildtype, reaching 9.1 h at 50 °C. However, the K201R/E384Y/T435A variant retained only 78 % of the catalytic activity relative to the optimal single mutant D78P. Fortunately, the D78P/K201R/E384Y/T435A variant demonstrated a 102 % increase in specific activity compared to the wildtype and a 15 % enhancement over D78P, while simultaneously achieving exceptional thermostability with a half-life of 8.6 h at 50 °C. Under optimal temperature conditions at 40 °C, this quadruple variant exhibited a remarkable 4.9-fold extension in half-life compared to the wildtype, achieving an exceptional stability duration of 145.0 h (see supplementary material). Moreover, the Tm value of the quadruple variant was elevated from 45.9 °C for the wildtype to 54.0 °C. Therefore, the D78P/K201R/E384Y/T435A variant suggested a more favorable trade-off between catalytic activity and thermostability (Fig. 3D).

We systematically characterized the kinetic parameters of the wildtype and the optimal variants across single-, double-, triple-, and quadruple-mutation. As summarized in Table 1, all engineered variants exhibited significantly enhanced the catalytic efficiency. The D78P variant showed a significantly reduced Michaelis constant (K_m) alongside a 1.5-fold higher turnover number (k_{cat}), compared to the wildtype, collectively driving its superior enzymatic performance. Compared to the D78P variant, the combinatorial mutants D78P/K201R, D78P/K201R/E384Y, and D78P/K201R/E384Y/T435A displayed elevated K_m values. However, introduction of the E384Y and T435A mutations further optimized catalytic activity by demonstrating enhanced K_{cat} values. Ultimately, the D78P/K201R/E384Y/T435A variant displayed the K_m value comparable to the wildtype while achieving a 1.8-fold enhancement in catalytic efficiency (k_{cat}/K_m). The D78P/K201R/E384Y/T435A variant exhibited synergistic improvements in both thermostability and catalytic efficiency, demonstrating that the computational framework not only accelerates the identification of beneficial mutation sites, but also effectively addresses the classical trade-off between stability and activity. This strategy facilitates the rational design of enzymes with dual high activity and enhanced
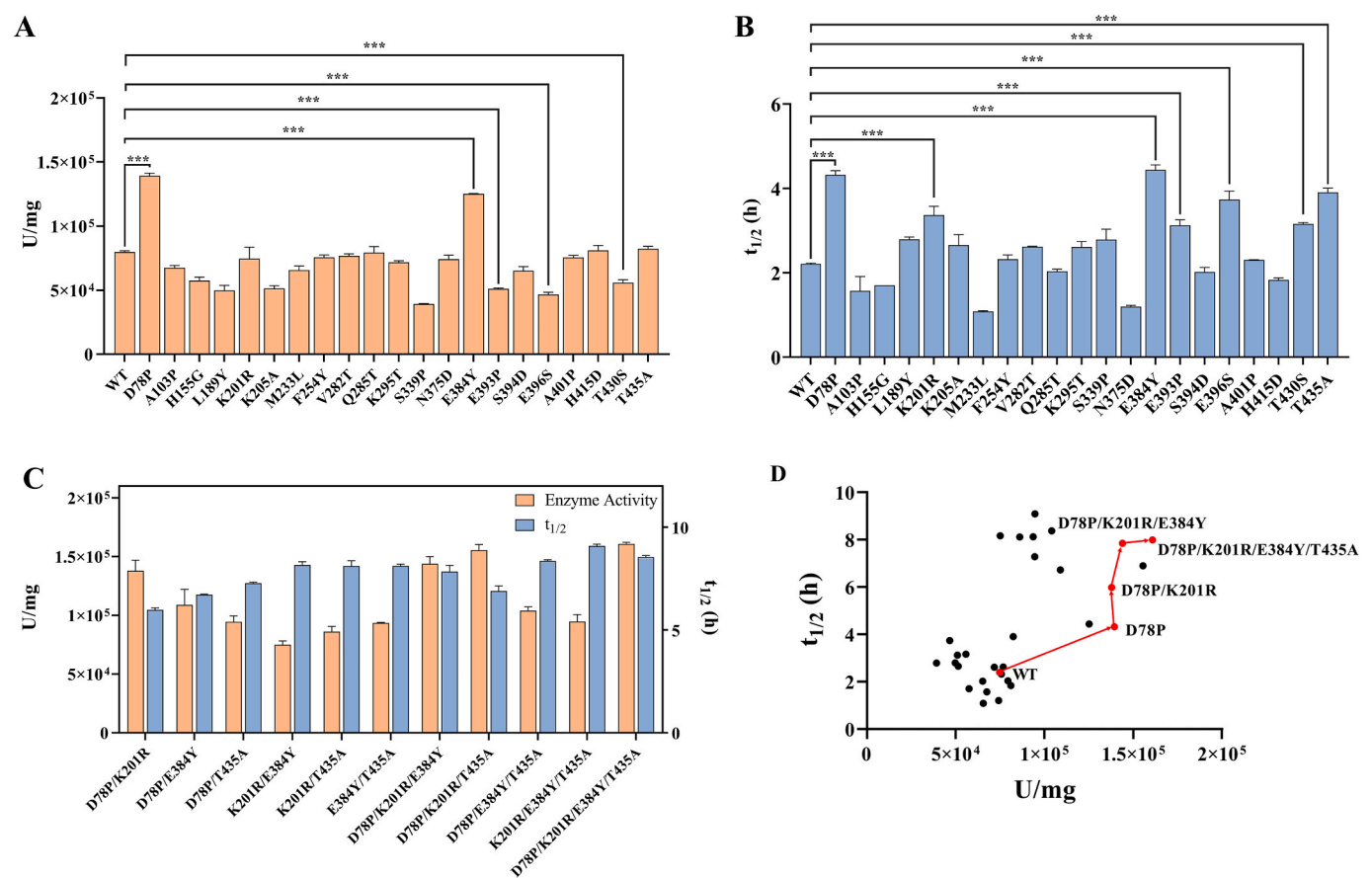
Fig. 3. Characterization of the A0A0V8J8H6 variants activity and thermostability. (A) Enzymatic activity of A0A0V8J8H6 single-point mutants at 40 °C. (B) The half-life of A0A0V8J8H6 single-point mutants at 50 °C. (C) Activity and thermal stability of A0A0V8J8H6 combinatorial mutants. (D) Evolutionary trajectory of activity and thermostability of the A0A0V8J8H6 variants. $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***) indicate the significance levels between wildtype and variants.

Table 1
Kinetic Parameters of A0A0V8J8H6 and Variants.[a]

Enzyme	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})
A0A0V8J8H6	14.81 ± 2.87	107168.43 ± 6470.75	7.24
D78P	12.11 ± 2.27	166767.48 ± 9118.38	13.77
D78P/K201R	15.63 ± 2.78	164790.00 ± 9228.24	10.54
D78P/K201R/E384Y	13.97 ± 1.57	172919.64 ± 6591.60	12.38
D78P/K201R/E384Y/ T435A	14.66 ± 2.00	196759.26 ± 8239.50	13.42

[a] Values represent the mean ± standard deviation (SD) of three independent replicate experiments.

stability, thereby laying a solid foundation for potential industrial applications.
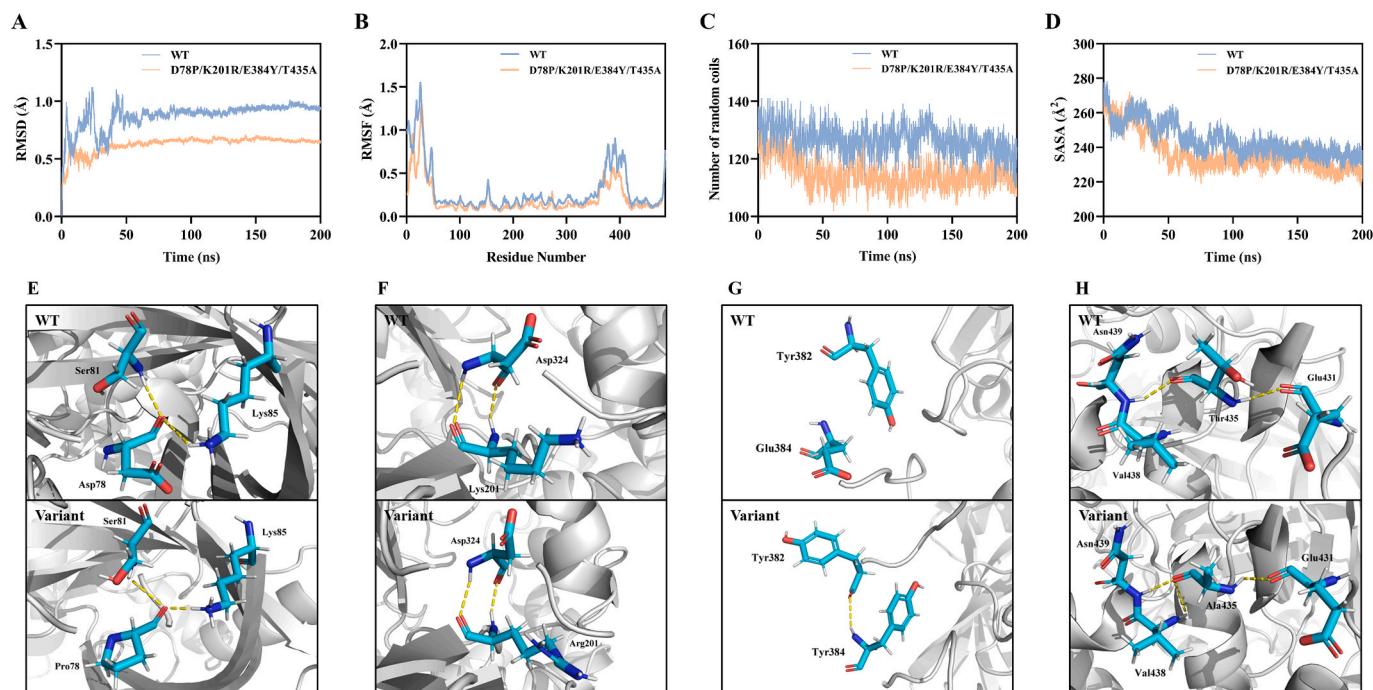
3.4. Molecular dynamics simulation and structural analysis

After evaluating the trade-offs between stability and catalytic activity, the D78P/K201R/E384Y/T435A variant was strategically selected for comprehensive investigation of its thermostability enhancement mechanisms through integrated molecular dynamics (MD) simulation and structural analysis. The variant structure was likewise modeled using AlphaFold3, and 200 ns MD simulations were carried out for both the wildtype and the variant. Under the simulation condition of 323 K, the RMSD of the wildtype gradually stabilized after approximately 60 ns, converging to approximately 0.90 Å (Fig. 4A). While the variant achieved conformational equilibrium within a shorter time (35 ns), demonstrating a notably lower stabilized RMSD of 0.65 Å. This marked reduction in RMSD and accelerated convergence kinetics suggested enhanced global conformational rigidity in the variant. It was also evidenced by the less displacement of residues 370–420 in the variant during the simulation (Fig. 4B). The mutation diminished the unstructured regions in the variant, compared to the wildtype, resulting in a lower number of random coils and a decrease in flexibility (Fig. 4C). SASA serves as a critical parameter in studies of protein folding and structural stability. The average SASA of the variant was lower than that of the wildtype, revealing a more compact and stable conformational state in the variant (Fig. 4D). In addition, MD simulations at the lower temperature of 303 K similarly revealed that the variant exhibited reduced conformational fluctuations, indicative of enhanced overall structural rigidity. Consequently, the attenuation of localized conformational dynamics and reduction in flexible regions collectively enhance global structural rigidity, thereby facilitating improved thermostability of the variant.

Subsequently, principal component analysis (PCA) was applied to MD simulation trajectories to construct the free energy landscape, from which fully energy-minimized conformations were extracted. The energy-minimized structure was employed to investigate alterations in inter-residue interactions proximal to the mutation site. Comparative analysis of hydrogen bonding patterns revealed that both residues 78 and 201 maintained consistent numbers of hydrogen bonds with neighboring residues before and after mutation (Fig. 4E and F).

Fig. 4. Analysis of the thermal stability mechanism of the D78P/K201R/E384Y/T435A variant by molecular dynamics simulations. (A) Root Mean Square Deviation (RMSD) of the backbone atoms of A0A0V8J8H6 and the D78P/K201R/E384Y/T435A variant during MD simulations conducted at 323 K. (B) Root Mean Square Fluctuation (RMSF) of the residues in A0A0V8J8H6 and the D78P/K201R/E384Y/T435A variant. (C) Number of random coils in A0A0V8J8H6 and the D78P/K201R/E384Y/T435A variant over the course of MD simulations. (D) The alterations in solvent-accessible surface area (SASA) of A0A0V8J8H6 and the D78P/K201R/E384Y/T435A variant through MD simulations. (E-H) Structural comparisons between the wildtype and the D78P/K201R/E384Y/T435A variant near the mutations.

However, alterations were observed in bond lengths following amino acid substitutions. Specifically, the D78P mutation resulted in a significant reduction of the average hydrogen bond distance between P78 and residue K85, decreasing from 4.2 Å–3.3 Å (see supplementary material). The K201R mutation also induced a subtle but consistent shortening of the two hydrogen bonds, with their average length decreasing from 2.0 Å–1.9 Å. Notably, the E483Y and T435A mutations induced the formation of novel hydrogen bonds in the variant, which significantly enhanced structural rigidity (Fig. 4G and Fig. 4H). Analysis of MD simulation trajectories found that when E384 was replaced by tyrosine, Y384 formed a hydrogen bond interaction an average length of 3.0 Å with the Y382 residue in the random coil region. This interaction significantly reduced local dynamics within residues 350–410 of the random coil domain, a finding corroborated by RMSF analysis. The T435A substitution caused an additional hydrogen-bonding interaction to form between residues A435 and V438, with an average bond length of 3.1 Å. These mutations generated novel hydrogen-bond interactions absent in the wildtype structure, thereby stabilizing the variant conformation through reinforced intermolecular forces.

The mutations in the D78P/K201R/E384Y/T435A variant were distant from the catalytic site, exhibiting no direct impact on the conformation of the active center or substrate binding modes. The tunnel connecting the catalytic core to the protein surface, a structural feature known to modulate catalytic efficiency, was hypothesized to mediate the observed activity enhancement. Tunnel analysis based on the static structure revealed that the variant exhibited a shortened tunnel length of 9.4 Å, accompanied by an expansion of the bottleneck diameter from 1.9 Å to 2.1 Å (see supplementary material). Analysis of protein conformations sampled during MD simulations revealed that the wildtype exhibited an average tunnel bottleneck radius of 2.0 Å and an average tunnel length of 13.8 Å. In contrast, the D78P/K201R/E384Y/T435A variant showed a wider average bottleneck radius of 2.4 Å and a shorter tunnel length of 9.7 Å. The presence of a shorter and wider

tunnel in the variant facilitated enhanced substrate accessibility to the catalytic center and promoted product release. Moreover, the variant exhibited enhanced hydrophilicity at the tunnel entrance compared to the wildtype, thereby enabling efficient binding of the polar substrate (H_2O_2). These findings collectively provide a molecular basis for promoting substrate and product translocation through the tunnel in the variant, thereby enhancing enzyme-substrate turnover efficiency.

3.5. Validation of industrial application potential

To assess the practical industrial potential, the H_2O_2 removal capabilities of the wildtype and the engineered variant were systematically evaluated across three distinct industrial matrices, including medical wastewater, textile bleaching effluent, and dairy processing sample. Following sequential additions of enzymes to medical wastewater, the engineered variant demonstrated rapid decomposition efficiency by achieving H_2O_2 elimination within 50 min, compared to 60 min required by the wildtype (Fig. 5A). Notably, the variant's enhanced tolerance to H_2O_2 enabled a reduction in enzyme loading requirements, thereby effectively lowering treatment costs. Similarly, in bleaching wastewater treatment applications, the variant-catalyzed system achieved complete H_2O_2 degradation within 40 min, whereas the wildtype-mediated process required a prolonged duration of 60 min to reach equivalent decomposition levels (Fig. 5B). We then systematically investigated the efficacy of catalase in eliminating residual low-concentration H_2O_2 during dairy processing. The wildtype achieved complete H_2O_2 removal within 30 min, whereas the variant enabled undetectable H_2O_2 levels in dairy matrices within 15 min (Fig. 5C). The variant's superior catalytic efficacy substantially reduced processing duration, thereby not only improving the efficiency of dairy production but also minimizing H_2O_2-induced oxidative degradation of lipid fractions, proteins, and vitamins.

Although real industrial wastewater was not available during this
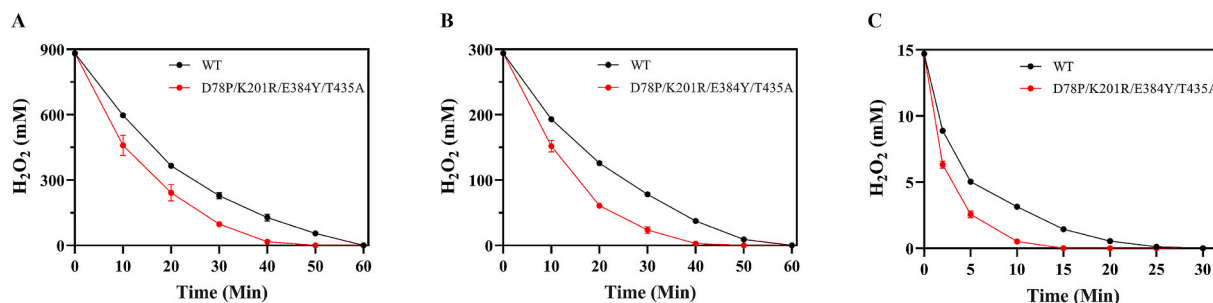
Fig. 5. H_2O_2 degradation performance of the D78P/K201R/E384Y/T435A variant in diverse industrial wastewater matrices. (A) Time-course analysis of H_2O_2 removal in simulated medical wastewater (3% H_2O_2). (B) H_2O_2 decomposition in bleaching effluent (1% H_2O_2) from textile industry. (C) Enzymatic removal of residual low-concentration H_2O_2 in dairy processing.

study, the simulated systems were carefully designed to reflect realistic industrial conditions. The above results demonstrated that the D78P/K201R/E384Y/T435A variant exhibited exceptional application potential. Its ability to decompose H_2O_2 in complex aqueous matrices is not impeded by the presence of Mg^{2+}, Ca^{2+}, or Fe^{3+} ions, nor by organic constituents such as lignin, glucose, or lactose. Moreover, the variant displayed enhanced tolerance toward H_2O_2, underscoring its unparalleled robustness and catalytic efficiency for industrial biocatalytic applications.
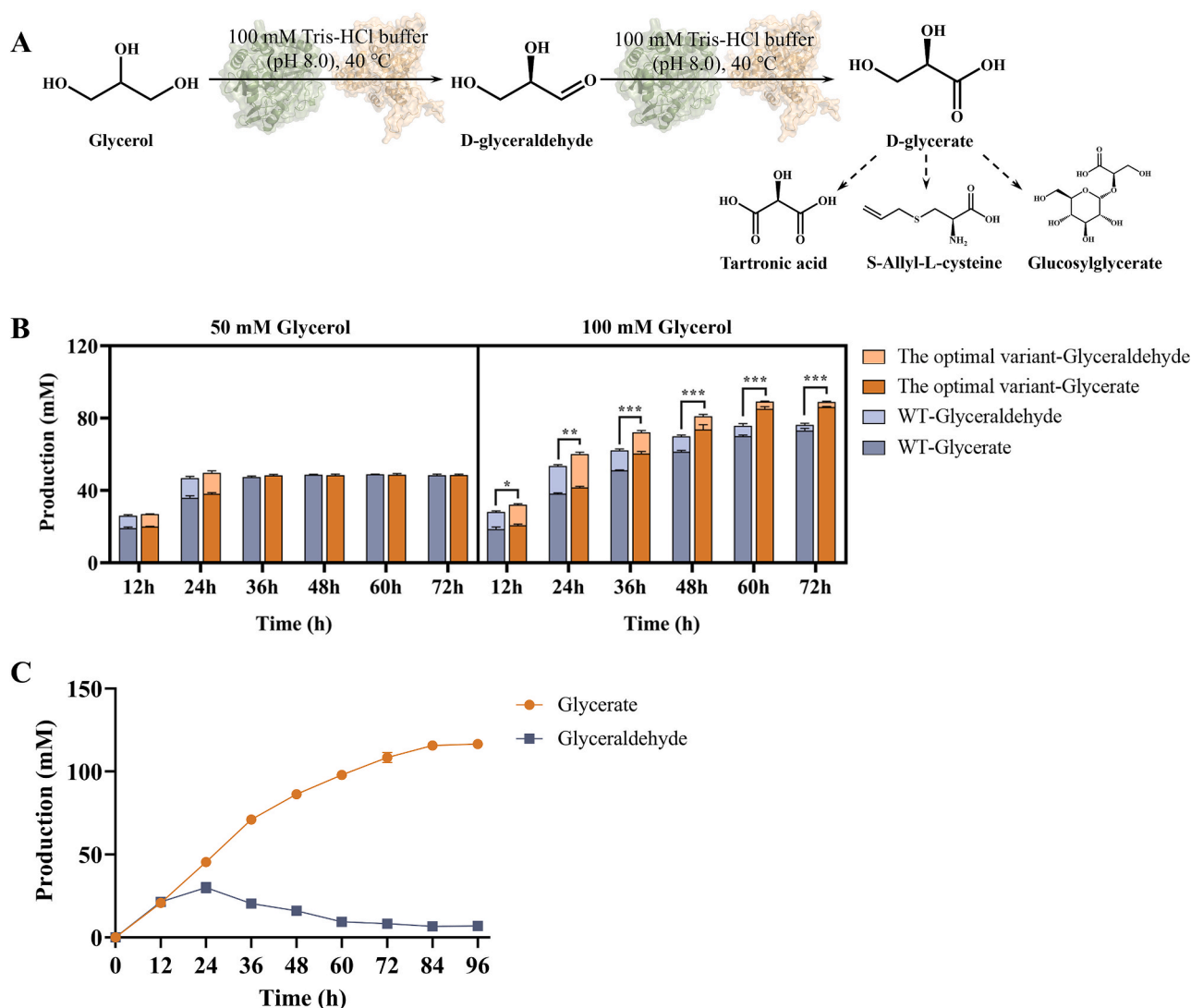
Fig. 6. Enhanced biocatalytic glycerol valorization by thermostable catalase. (A) Schematic representation of the one-pot enzymatic cascade system employed for D-glycerate biosynthesis from glycerol. (B) Time-course analysis of D-glycerate production catalyzed by the wildtype and the D78P/K201R/E384Y/T435A variant under different glycerol concentrations (50 mM and 100 mM). (C) D-glycerate and glyceraldehyde accumulation profiles in high-substrate (200 mM glycerol) systems catalyzed by the D78P/K201R/E384Y/T435A variant. $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***) indicate the significance levels between the wildtype system and the D78P/K201R/E384Y/T435A variant.

3.6. Application of the optimal variant in glycerol value-added conversion

As a major byproduct of biodiesel production, the excessive accumulation of glycerol has been recognized as a critical challenge in the biofuel industry (Pirzadi and Meshkani, 2022). The utilization of glycerol as a key feedstock for biocatalytic production high-value chemical synthesis has emerged as a promising strategy (Fig. 6A) (Kumawat et al., 2024; Liu et al., 2024; Cai et al., 2014). The enzymatic conversion of glycerol to D-glycerate constitutes a rate-limiting step in its biotransformation pathway, while prolonged reaction durations impose significant challenges to the thermostability of catalase. In order to assess the impact of the D78P/K201R/E384Y/T435A variant on D-glycerate biosynthesis, a one-pot enzymatic cascade system coupling the variant with TfAldO$^{V256L/P257I}$ was established and assessed under varying glycerol concentrations (Fig. 6B). In a 50 mM glycerol system, both the variant and the wildtype catalyzed complete conversion of glycerol to D-glycerate within 36 h. Under 100 mM glycerol conditions, the variant-catalyzed reaction exhibited superior conversion efficiency, yielding 85.1 mM D-glycerate at 60 h, whereas the wildtype required 72 h to attain a maximum conversion rate of merely 73 %. The variant consistently evidenced superior conversion rates across all tested systems and time points. The result could be ascribed to its improved catalytic efficiency and thermostability, which facilitates accelerated oxygen regeneration throughout the reaction duration. Notably, the variant exhibited more pronounced advantages under elevated substrate concentrations. When employing 200 mM glycerol as substrate, the reaction system accumulated 116.6 mM D-glycerate and 6.9 mM glyceraldehyde at 84 h, without requiring supplemental catalase addition during the process (Fig. 6C). At higher glycerol concentrations, although conversion efficiency remained constrained by the inherent limitations in TfAldO$^{V256L/P257I}$, the D78P/K201R/E384Y/T435A variant has demonstrated sufficient performance to meet the industrial requirement for high-concentration glycerol valorization.

4. Conclusions

This study presents an integrated biocatalyst development strategy combining deep learning-driven enzyme screening with computational protein engineering. A high-activity catalase was identified using deep learning models, and subsequent ProteinMPNN-guided design with Gibbs free energy calculation yielded variants with enhanced activity and stability. The D78P/K201R/E384Y/T435A variant achieved a specific activity of 160,842 U/mg, alongside a preserved thermal half-life of 145.0 h at 40 °C. The variant was used for implementation in H$_2$O$_2$-containing wastewater treatment and glycerol valorization processes, highlighting its potential for practical industrial deployment. E-supplementary data for this work can be found in e-version of this paper online.

CRediT authorship contribution statement

Shuai Xu: Writing – original draft, Visualization, Software, Methodology, Investigation, Data curation, Conceptualization. **Ya-min Chen:** Validation, Investigation. **Xiang-yu Meng:** Validation, Investigation. **Ru Pan:** Investigation. **Ao-xuan Yan:** Data curation. **Zhi-min Li:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition. **Zong-lin Li:** Writing – review & editing, Software, Methodology, Funding acquisition, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biortech.2025.133081.

Data availability

The data supporting the findings of this study are available within the article and its supplementary materials. Additional data may be obtained from the corresponding author upon reasonable request.

References

Arslan-Alaton, I., Olmez-Hanci, T., Shayin, S., 2012. H$_2$O$_2$/UV-C treatment of textile preparation wastewater: kinetic investigation on alternative combinations of commercial textile preparation auxiliaries. Environ. Technol. 33, 1531–1537. https://doi.org/10.1080/09593330.2012.665489.

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., Jumper, J.M., 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

Bopitiya, D., Christensen, D., Martin, M., Zhang, J., Bennett, L.E., 2021. Production of hydrogen peroxide in formulated beverages is associated with the presence of ascorbic acid combined with selected redox-active functional ingredients. Food Chem. 338, 127947. https://doi.org/10.1016/j.foodchem.2020.127947.

Bian, J., Tan, P., Nie, T., Hong, L., Yang, G., 2024. Optimizing enzyme thermostability by combining multiple mutations using protein language model. mLife 3, 492–504. https://doi.org/10.1002/mlf2.12151.

Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J., Damborsky, J., 2012. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Comput. Biol. 8, e1002708. https://doi.org/10.1371/journal.pcbi.1002708.

Colovos, C., Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2, 1511–1519. https://doi.org/10.1002/pro.5560020916.

Cai, J., Ma, H., Zhang, J., Du, Z., Huang, Y., Gao, J., Xu, J., 2014. Catalytic oxidation of glycerol to tartronic acid over Au/HY catalyst under mild conditions. Chin. J. Catal. 35, 1653–1660. https://doi.org/10.1016/S1872-2067(14)60132-7.

Ciriminna, R., Albanese, L., Meneguzzo, F., Pagliaro, M., 2016. Hydrogen peroxide: a key chemical for today's sustainable development. ChemSusChem 9, 3374–3381. https://doi.org/10.1002/cssc.201600895.

Chen, Z., Fei, K., Hu, Y., Xu, X., Gao, X., Li, Z., 2022. Identification of a novel alditol oxidase from *Thermopolyspora flexuosa* with potential application in D-Glyceric acid production. Mol. Biotechnol. 64, 804–813. https://doi.org/10.1007/s12033-022-00459-3.

Dieckhaus, H., Brocidiacono, M., Randolph, N.Z., Kuhlman, B., 2024. Transfer learning to leverage larger datasets for improved prediction of protein stability changes. PNAS 121, e2314853121. https://doi.org/10.1073/pnas.231485312.

Eberhardt, A.M., Pedroni, V., Volpe, M., Ferreira, M.L., 2004. Immobilization of catalase from *Aspergillus niger* on inorganic and biopolymeric supports for H$_2$O$_2$ decomposition. Appl. Catal. B-Environ. 47, 153–163. https://doi.org/10.1016/j.apcatb.2003.08.007.

Fruhwirth, G., Paar, A., Gudelj, M., Cavaco-Paulo, A., Robra, K.-H., Gübitz, G., 2002. An immobilised catalase peroxidase from the alkalothermophilic bacillus SF for the treatment of textile-bleaching effluents. Appl. Microbiol. Biotechnol. 60, 313–319. https://doi.org/10.1007/s00253-002-1127-0.

Fukuzumi, S., Lee, Y., Nam, W., 2021. Recent progress in production and usage of hydrogen peroxide. Chin. J. Catal. 42, 1241–1252. https://doi.org/10.1016/S1872-2067(20)63767-6.

Jia, X., Lin, X., Tian, Y., Chen, J., You, M., 2017. High production, purification, biochemical characterization and gene analysis of a novel catalase from the thermophilic bacterium *Ureibacillus thermosphaericus* FZSF03. Int. J. Biol. Macromol. 103, 89–98. https://doi.org/10.1016/j.ijbiomac.2017.05.034.

Justas, D., Anishchenko, I., Nathaniel, B., Hua, B., Robert, R., Frederik, M.L., Basile, W., Alexis, C., Robbert, H., Neville, B., Philip, L., Timothy, H., Samuel, P., Doug, T., Chan, F., Brian, K., Hannah, N., Alex, K., Banumathi, S., Asim, B., Neil, K., Baker, D.,

2022. Robust deep learning-based protein sequence design using ProteinMPNN. Science 378, 49–56. https://doi.org/10.1126/science.add2187.

Jung, F., Frey, K., Zimmer, D., Mühlhaus, T., 2023. DeepSTABp: a deep learning approach for the prediction of thermal protein stability. Int. J. Mol. Sci. 24, 7444. https://doi.org/10.3390/ijms24087444.

Kocabas, D.S., Bakir, U., Phillips, S.E.V., McPherson, M.J., Ogel, Z.B., 2008. Purification, characterization, and identification of a novel bifunctional catalase-phenol oxidase from *Scytalidium thermophilum*. Appl. Microbiol. Biotechnol. 79, 407–415. https://doi.org/10.1007/s00253-008-1437-y.

Kellogg, E.H., Leaver-Fay, A., Baker, D., 2011. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. Proteins 79, 830–838. https://doi.org/10.1002/prot.22921.

Kroll, A., Rousset, Y., Hu, X., Liebrand, N.A., Lercher, M.J., 2023. Turnover number predictions for kinetically uncharacterized enzymes using machine and deep learning. Nat. Commun. 14, 4139. https://doi.org/10.1038/s41467-023-39840-4.

Kumawat, S., Singh, S., Bhatt, T., Maurya, A., Vaidyanathan, S., Natte, K., Jagadeesh, R. V., 2024. Valorization of bio-renewable glycerol by catalytic amination reactions. Green Chem. 26, 3021–3038. https://doi.org/10.1039/D3GC02699J.

Lončar, N., Fraaije, M.W., 2015. Catalases as biocatalysts in technical applications: current state and perspectives. Appl. Microbiol. Biotechnol. 99, 3351–3357. https://doi.org/10.1007/s00253-015-6512-6.

Li, F., Yuan, L., Lu, H., Li, G., Chen, Y., Engqvist, M.K.M., Kerkhoven, E.J., Nielsen, J., 2022. Deep learning-based kcat prediction enables improved enzyme-constrained model reconstruction. Nat. Catal. 5, 662–672. https://doi.org/10.1038/s41929-022-00798-z.

Liu, J., Ren, M., Ma, H., Zhang, H., Cui, X., Kang, R., Feng, X., Meng, D., 2024. One-pot sustainable synthesis of glucosylglycerate from starch and glycerol through artificial in vitro enzymatic cascade. Bioresour. Technol. 399, 130611. https://doi.org/10.1016/j.biortech.2024.130611.

Meyer, A.S., Pedersen, L.H., Isaksen, A., 1997. The effect of various food parameters on the activity and stability of catalase from *Aspergillus niger* and catalase from bovine liver. Food Chem. 60, 137–142. https://doi.org/10.1016/S0308-8146(95)00252-9.

Musil, M., Stourac, J., Bendl, J., Brezovsky, J., Prokop, Z., Zendulka, J., Martinek, T., Bednar, D., Damborsky, J., 2017. FireProt: web server for automated design of thermostable proteins. Nucleic Acids Res. 45, W393–W399. https://doi.org/10.1093/nar/gkx285.

Pirzadi, Z., Meshkani, F., 2022. From glycerol production to its value-added uses: a critical review. Fuel 329, 125044. https://doi.org/10.1016/j.fuel.2022.125044.

Spoel, D.V.D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J.C., 2005. GROMACS: fast, flexible, and free. J. Comput. Chem. 26, 1701–1718. https://doi.org/10.1002/jcc.20291.

Studer, R.A., Christin, P., Williams, M.A., Orengo, C.A., 2014. Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. PNAS 111, 2223–2228. https://doi.org/10.1073/pnas.131081111.

Sumida, K.H., Núñez-Franco, R., Kalvet, I., Pellock, S.J., Wicky, B.I.M., Milles, L.F., Dauparas, J., Wang, J., Kipnis, Y., Jameson, N., Kang, A., Cruz, J.D.L., Sankaran, B., Bera, A.K., Jiménez-Osés, G., Baker, D., 2024. Improving protein expression, stability, and function with ProteinMPNN. J. Am. Chem. Soc. 146, 2054–2061. https://doi.org/10.1021/jacs.3c10941.

Silverstein, R.A., Kim, N., Kroell, A., Walton, R.T., Delano, J., Butcher, R.M., Pacesa, M., Smith, B.K., Christie, K.A., Ha, L.L., Meis, R.J., Clark, A.B., Spinner, A.D., Lazzarotto, C.R., Li, Y., Matsubara, A., Urbina, E.O., Dahl, G.A., Correia, B.E., Marks, D.S., Tsai, S.Q., Pinello, L., Ravin, S.S.D., Liu, Q., Kleinstiver, B.P., 2025. Custom CRISPR-Cas9 PAM variants via scalable engineering and machine learning. Nature. https://doi.org/10.1038/s41586-025-09021-y.

Tian, L., Branford-White, C., Wang, W., Nie, H., Zhu, L., 2012. Laccase-mediated system pretreatment to enhance the effect of hydrogen peroxide bleaching of cotton fabric. Int. J. Biol. Macromol. 50, 782–787. https://doi.org/10.1016/j.ijbiomac.2011.11.025.

Tehrani, H.S., Moosavi-Movahedi, A.A., 2018. Catalase and its mysteries. Prog. Biophys. Mol. Bio. 140, 5–12. https://doi.org/10.1016/j.pbiomolbio.2018.03.001.

Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol. Biol. Evol. 38, 3022–3027. https://doi.org/10.1093/molbev/msab120.

Vasudhevan, P., Ruoyu, Z., Ma, H., Singh, S., Varshney, D., Pu, S., 2025. Biocatalytic enzymes in food packaging, biomedical, and biotechnological applications: a comprehensive review. Int. J. Biol. Macromol. 300, 140069. https://doi.org/10.1016/j.ijbiomac.2025.140069.

Wang, H., Tokusige, Y., Shinoyama, H., Fujii, T., Urakami, T., 1998. Purification and characterization of a thermostable catalase from culture broth of *Thermoascus aurantiacus*. J. Ferment. Bioeng. 85, 169–173. https://doi.org/10.1016/S0922-338X(97)86762-1.

Weinstein, J.J., Goldenzweig, A., Hoch, S., Fleishman, S.J., 2021. PROSS 2: a new server for the design of stable and highly expressed protein variants. Bioinformatics 37, 123–125. https://doi.org/10.1093/bioinformatics/btaa1071.

Wang, Y., Cao, X., Jiang, S., Gao, L., Han, X., Qu, J., Jiang, X., Liu, G., Quab, Y., 2024. Engineering the substrate preference of glucose oxidase for the enzymatic oxidation of xylose. Green Chem. 26, 4851–4859. https://doi.org/10.1039/D3GC04981G.

Yu, Z., Zheng, H., Zhao, X., Li, S., Xu, J., Song, H., 2016. High level extracellular production of a recombinant alkaline catalase in *E. coli* BL21 under ethanol stress and its application in hydrogen peroxide removal after cotton fabrics bleaching. Bioresour. Technol. 214, 303–310. https://doi.org/10.1016/j.biortech.2016.04.110.