

Article

A Novel Strategy for Further Enhancing Superior Properties of Thermophilic Endoglucanase from *Acidomyces richmondensis*

Shengjie Wang^{1,2,†}, Zherui Zhang^{2,†}, Yi Li^{3,†}, Jie Yuan^{1,2}, Haofan Dong^{1,2}, Tongtong Bao², Xin Wu²,
Lingfang Gu⁴, Jian Zhang^{1,*} and Le Gao^{2,*}

- ¹ College of Bioengineering, Tianjin University of Science and Technology, Tianjin 300457, China; wangshj@tib.cas.cn (S.W.); yuanjie@tib.cas.cn (J.Y.); donghf@tib.cas.cn (H.D.)
² Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, National Center of Technology Innovation for Synthetic Biology, Tianjin 300308, China; zhangzhr2022@tib.cas.cn (Z.Z.); baott@tib.cas.cn (T.B.); wuxin@tib.cas.cn (X.W.)
³ College of Mathematics and Computer Science, Dali University, Dali 671000, China; liyigerry@163.com
⁴ Wuhan Sunhy Biology Co., Ltd., Wuhan 430074, China; glf8254@163.com
* Correspondence: zj96sk@tust.edu.cn (J.Z.); gao_l@tib.cas.cn (L.G.); Tel.: +86-22-24828745 (L.G.); Fax: +86-22-84861996 (L.G.)
† These authors contributed equally to this work.

Abstract: Thermophilic β -1,4-endoglucanases (Cel5A) have garnered significant interest due to their potential applications in various industries, particularly in biofuel production and biorefineries. However, despite inherent stability, thermophilic Cel5A still face challenges in terms of further enhancing their catalytic efficiency and thermostability. In this study, a novel B-factor analysis method was used to predict beneficial amino acid substitutions within a 4 Å radius of the catalytic site in the tunnel of thermophilic Cel5A from *Acidomyces richmondensis* (ArCel5A). A combined strategy involving site-saturation mutagenesis and high-throughput screening was employed to identify the variants with the highest endoglucanase activity. Genomic sequencing revealed a mutation at position 299 in the starting strain *T. reesei* A2H, where the nucleotide sequence changed from TAC to TGC, resulting in a corresponding amino acid substitution from Tyrosine(Y) to Cystine(C). The endoglucanase activity of the mutant ArCel5A reached 3251 IU/mL, representing an 85.2% increase compared to wild-type ArCel5A at the fermentation time of 94 h. Significantly, the ArCel5A-Y299C mutant showed superior thermostability, retaining 93.8% of its initial activity after 30 min at 70 °C, and 91.5% after 10 min at 80 °C. Various computational simulation methods confirmed that the Y299C mutation enhanced the stability of the catalytic pocket, thereby improving the overall stability and catalytic efficiency of ArCel5A. This study offers an effective strategy for mining target sites for rational mutagenesis based on highly conserved sequences, which simultaneously improves both the thermostability and catalytic efficiency of thermophilic Cel5A.

Keywords: thermophilic endoglucanase; rational design; novel B-factor analysis method; thermostability; computational simulation



Citation: Wang, S.; Zhang, Z.; Li, Y.; Yuan, J.; Dong, H.; Bao, T.; Wu, X.; Gu, L.; Zhang, J.; Gao, L. A Novel Strategy for Further Enhancing Superior Properties of Thermophilic Endoglucanase from *Acidomyces richmondensis*. *Fermentation* **2024**, *10*, 27. <https://doi.org/10.3390/fermentation10010027>

Academic Editor: Mohammad Taherzadeh

Received: 16 November 2023

Revised: 12 December 2023

Accepted: 15 December 2023

Published: 29 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As a highly valuable and abundant natural resource, the use of lignocellulosic biomass as a renewable feedstock presents a great opportunity for the production of various value-added products [1]. However, the intricate structure of this biomass resource, composed of cellulose (35–50%), hemicellulose (20–30%), and lignin (20–30%), poses a challenge to its conversion into fermentable sugars [2]. This recalcitrant structure is a significant obstacle in the development of a biomass-based biorefinery, which relies on fermentable sugars as intermediates for the conversion of carbohydrates from lignocellulosic biomass into value-added chemicals [3].

Endo- β -1,4-glucanase (EG, EC 3.2.1.4), a crucial cellulase for the deconstruction of lignocellulosic biomass, plays a vital role in cleaving the internal β -1,4-glycosidic bonds of cellulose fibers. By initiating the catalytic attack on biopolymer chains in amorphous regions, EG paves the way for the efficient conversion of biomass into valuable products [4]. Based on sequence and structure similarity, endoglucanases are classified into 13 glycoside hydrolase (GH) families on CAZymes (<http://www.cazy.org>), including GH5–9, 12, 44, 45, 48, 51, 74, 124, and 131 [5]. Among them, endoglucanases represent the largest and most functionally diverse group, with the majority of fungal endoglucanases with endo- β -1,4-glucanase activity belonging to subfamily GH5_5 [6]. Currently, six eukaryotic GH5 endoglucanases from *Piromyces rhizinflata* (PrEglA) [7], *Thermoascus aurantiacus* (TaCel5A) [8], *Hypocrea jecorina* (teleomorph of *Trichoderma reesei*) (TrCel5A) [9], *Ganoderma lucidum* (GICel5A) [10], *Aspergillus niger* (AnCel5A) [11], and *Penicillium verruculosum* (PDB No. 5I6S) have been resolved. The typical catalytic domain of a GH5 cellulase has a canonical $(\beta/\alpha)_8$ TIM-barrel fold, in which eight parallel β -strands and eight α -helices are connected by seven $\beta\alpha$ or $\alpha\beta$ loops [12].

Endo- β -1,4-glucanases have numerous applications in various industries, including brewing as well as the production of feed, paper, and biofuels. For instance, the inclusion of endo- β -1,4-glucanase can alleviate the adverse effects of barley β -glucan during the mashing process in the brewing industry, which is typically carried out at temperatures ranging from 50 to 60 °C. Moreover, it can also improve the digestibility of β -glucan in poultry feedstuffs [13]. In other industries, it is used to hydrolyze cellulosic materials to release sugars, reduce mash viscosity and turbidity, bioremediate pulp waste, and increase β -glucan digestibility to improve feed conversion efficiency at 50 °C [3]. Thermophilic cellulases, which exhibit high activity and stability at high temperatures, are particularly attractive for industrial applications [14,15]. Typically, low thermostability poses a significant obstacle to the industrial utilization of enzymes. In the context of industrial production, enhancing the heat resistance and stability of EGs can potentially lower production costs and reduce reaction time [16]. Moreover, the industrial production of granulation processes demands a high degree of heat resistance in endo- β -1,4-glucanase. Industrial enzymes generally must be robust and tolerate various harsh processing conditions as well as long-term usage and storage. To develop more robust cellulases, researchers have turned their attention to identifying and engineering enzymes from microorganisms that can survive in extreme environments, such as hot springs and acidic soils. By analyzing omics data, researchers have screened thermostable endoglucanases that offer both high enzymatic activity and thermostability, which are key factors for achieving superior hydrolytic performance. These enzymes have great potential for use in second-generation biorefineries [17,18].

Directed evolution has been the primary strategy for improving endoglucanases, utilizing various approaches such as the hybridization strategy [19], engineering of conserved non-catalytic residues and *N*-glycosylation sites [16], modification of disulfide bridges [20], and optimization of hydrophobic interactions [21]. However, while significant improvements of enzyme thermostability have been achieved, the resulting variants mostly suffer a loss of activity to varying degrees. Although overcoming the challenge of improving stability without compromising activity remains a significant hurdle [22], studies suggest that selecting mutational target sites in regions distant from the catalytic center may be a viable approach for designing enzymes with enhanced thermostability and catalytic efficiency. For example, optimization of the loop regions of endoglucanases resulted in combined improvements of both enzymatic activity (17.2–18.0%) and thermostability (49.9–62.9%) compared to wild-type endoglucanases [23].

In this study, a novel B-factor strategy was employed to accurately predict mutant substitutions of highly conserved amino acids located within the active-site tunnel of the thermophilic β -1,4-endoglucanases from *Acidomyces richmondensis* (ArCel5A). A combined strategy involving site-saturation mutagenesis and high-throughput screening was employed to identify ArCel5A variants with high enzymatic activity. The mechanism for overcoming the activity–stability trade-off in the best variant was explored through com-

putational design and experimental validation. This strategy may provide new insights for further mining potential rational modification sites among highly conserved sequences surrounding the catalytic center of ArCel5A.

2. Materials and Methods

2.1. Strains, Media, and Cultivation Conditions

T. reesei A2H, a cellulase-hyperproducing mutant strain, was obtained via ARTP mutagenesis in the laboratory. The fungal strain was preserved in the China General Microbiological Culture Collection Center (CGMCC21470) and used as control reference strains in the study. For cellulase production, the medium contained 3.3% microcrystalline cellulose, 1.7% corn steep liquor, 0.5% $(\text{NH}_4)_2\text{SO}_4$, 0.6% KH_2PO_4 , 0.1% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.25% CaCO_3 , and 0.2% Tween-80. The initial pH of the fermentation medium was adjusted to 5.0 using 10 M KOH. The spore suspension (10^6 spores/mL) was then used to inoculate 50 mL of the liquid medium in a 250 mL conical flask and was grown at 28 °C with agitation at 180 rpm for 5 days [23].

For efficient screening of medium components, 2% lactose, 0.5% $(\text{NH}_4)_2\text{SO}_4$, 1.5% KH_2PO_4 , 0.06% MgSO_4 , and 0.06% CaCl_2 were used. Transformants were placed in 96-well plates and grown at 28 °C with agitation at 180 rpm for 5 days. For high-efficiency screening, a reagent containing 2% AZO-CMC and 2 M sodium acetate (pH 4.5) was used. The buffer solution contained 10.5% glacial acetic acid and 20% sodium hydroxide solution (pH 4.6), while the precipitant solution contained 20% sodium acetate trihydrate, 2% zinc acetate, and 5 M HCl (pH 5.0). This solution was added to 800 mL of industrial methylated spirits (IMS, 95%) or ethanol (95%), after which 200 mL of the solution was added.

2.2. Rational Mining of Candidate Modification Sites

The cel5A from the acidophilic fungus *Acidomyces richmondensis* belongs to glycoside hydrolase family 5, and the amino acid similarity of *T. reesei* cel5A reaches 63%. The PDB template 3QR3 from the PDB database (<https://www.rcsb.org/>) was used for homology modeling in SwissModel (<https://swissmodel.expasy.org/>).

2.3. Mutant Library and High-Throughput Screening

Overlap-extension PCR was performed to substitute a specific site of Arcel5A with other amino acid residues. The mutagenic PCR fragments, the 22 amino-acid modified FMDV 2A peptide sequence (VKQTLNFDLLKLAGDVESNPGP) and the GFP gene were amplified and cloned into pAN52-Ptef1-TtrpC-bar using the NEB Gibson assembly kit to generate plasmids Ptef1-Arcel5A-2A-GFP-TtrpC. The transformants with resistance to phosphinothricin were confirmed using GFP fluorescence detection [24]. The GFP fluorescence signal of each colony (1×10^9 conidia per sample) was measured with excitation at 480 nm and emission at 520 nm using a SpectraMax M2e microplate reader (Molecular Devices, LLC., San Jose, CA, USA). The transformants identified by the fluorescence screening system were further cultured through high-throughput screening in 96 deep-well plates.

2.4. High-Throughput Screening for Activity

To introduce the expression cassette obtained using PCR purification in *T. reesei*, protoplasts were transformed in accordance with a previously described procedure [25]. For efficient screening of cultivated transformants, 100 μL of enzyme solution (pre-equilibrated at 40 °C) was added to 100 μL of 1% (*w/v*) Azo-CMC and 50 mM sodium citrate buffer (pH 4.80) in 96-well plates. After incubation at 40 °C for 10 min, the reaction was terminated with the addition of 500 μL of precipitant solution with vigorous stirring for 10 s on a vortex mixer. After cooling to room temperature, the assay plates were spun down at 1000 rpm and absorbance measurements were recorded at 540 nm.

2.5. Nucleotide Sequencing and Sequence Analysis

The plasmids were amplified and extracted, and the nucleotide sequences of cel5a_Tma mutants were determined using DNA sequencing. Vector NTI Advance was used to align and analyze the nucleotide and amino acid sequences.

2.6. Determination of Endoglucanase Activity

Endoglucanase activity was determined using 0.8% (*w/v*) carboxymethyl cellulose. Reactions were performed in 0.1 M sodium acetate buffer (pH 5.5) at 37 °C for 30 min [23]. The total amount of reduced sugars in the supernatants was determined using the dinitrosalicylic acid method.

2.7. Thermostability Assay of EG

Temperature stability was investigated by incubating the enzyme without substrate in a water bath at 70 °C and 80 °C, respectively. Samples were taken at different time intervals, and the residual activity was assayed. The control used was the inactive sample taken at different time intervals. The sample was inactivated through boiling at 100 °C for 5 min [23].

2.8. SDS-PAGE Analysis

SDS-PAGE analysis was performed to detect extracellular proteins. The purified proteins were electrophoresed using a 12% separating gel with 0.1% SDS and a 5% stacking gel. Protein bands were stained with Coomassie Brilliant Blue R-250.

2.9. B Factor Values Analysis

B factor, available from the Protein Data Bank, can be used to identify the flexibility of atoms, side chains, or even whole regions for predicting the mutation sites. Based on the first order sequence information, a deep learning method was used to select the mutation sites. The B factor value was obtained from the X-ray diffraction crystal structure of the protein, which reflects the degree of dispersion of the electron shell of the atoms. The B factor value of the amino acid reflects the flexibility of the amino acid in the current position. On average, the B factor in the flexible position is higher than in the rigid position [26]. Therefore, we used a deep learning model combining a convolutional neural network (CNN) and gated cycle unit (GRU) to predict the B factor values. The model learns potential associations between amino acid sequences and B factors from protein databases, ranking predicted B factor values to select the range of potential mutation sites.

2.10. Molecular Dynamics (MD) Simulations

Molecular dynamics simulations of the wild-type enzyme and the indicated mutants were conducted using AMBER22 software. The protein was modeled using the Amber ff19SB force field, while water molecules were based on the TIP3P model. The protein was placed in a 10 Å octahedral TIP3P water solvent box using the tleap algorithm. Na⁺ or Cl⁻ ions were added to balance the charge, after which initial coordinates and top files were generated. To prepare the system for simulations, an energy minimization step was performed on the entire system using 5000 steps of steepest descent followed by 5000 steps of conjugate gradient minimization. After energy minimization, the system was gradually heated from 0 to 300 K over a period of 100 ps [23]. The temperature was ramped up slowly during the initial 90 ps, and then maintained at 300 K for the final 10 ps to allow for equilibration. Subsequently, a 5 ns simulation was performed to bring the system to an equilibrium state. This was followed by 100 ns of NVT production MD simulations at a target temperature of 300 K to generate trajectories. During the MD simulations, periodic boundary conditions were applied, and long-range interactions were handled using the particle mesh Ewald (PME) method. Short-range interactions were truncated at a cutoff radius of 8 Å. Additionally, the SHAKE algorithm was used to constrain the high-frequency stretching vibrations of hydrogen-containing bonds.

3. Results and Discussion

3.1. Three-Dimensional Model of Cel5A from *Acidomyces Richmondensis*

Cel5A from *Acidomyces richmondensis* belongs to the glycoside hydrolase family 5 (GH5). The modeled structure of ArCel5A exhibits a $(\beta/\alpha)_8$ -barrel typical of GH5 members with an additional one-turn helix. After conducting a homologous alignment of the amino acid sequence, it was discovered that the -2 to $+1$ subunits are highly conserved and interact with cellulose in multiple ways. These subunits play a critical role in substrate recognition and binding [27]. The most crucial seven sites within Cel5A’s catalytic domain have been identified, including the five conserved sites from -2 to $+1$ (>90% sequence conservation). The corresponding residues of 3QR3 are H96, N139, E140, Y211, and E249, while those of Arcel5A are 182D, N227, E228, Y299, and E337, respectively. The -1 and $+1$ sub-sites are the catalytic bond-breaking sites that directly interact with the substrate, and the amino acids present at the -2 sub-site lead to further distortions of the substrate in the catalytic process. From the three-dimensional structure, it can be seen that 299Y is located at the -1 position in the catalytic domain, which is involved in the bond-breaking reaction. Additionally, W367 and Y379 are two key amino acids that also interact with the substrate.

3.2. Selection of Mutational Target Sites through B-Factor Analysis

As noted by Guo et al. [28], improving the stability of the local structure surrounding the catalytic site is an effective approach to increase enzymatic activity under harsh conditions. Upon analyzing the structure of ArCel5A, it was determined that 27 amino acid residues within a 4 Å radius of the catalytic site in the tunnel, including K91, Y92, G93, N96, I97, A139, M178, I180, D183, I183, H184, T222, E228, K257, L259, I293, DD295, H297, Y299, M334, L335, T336, G366, W367, and T368, which are critical for stability (Figure 1a).

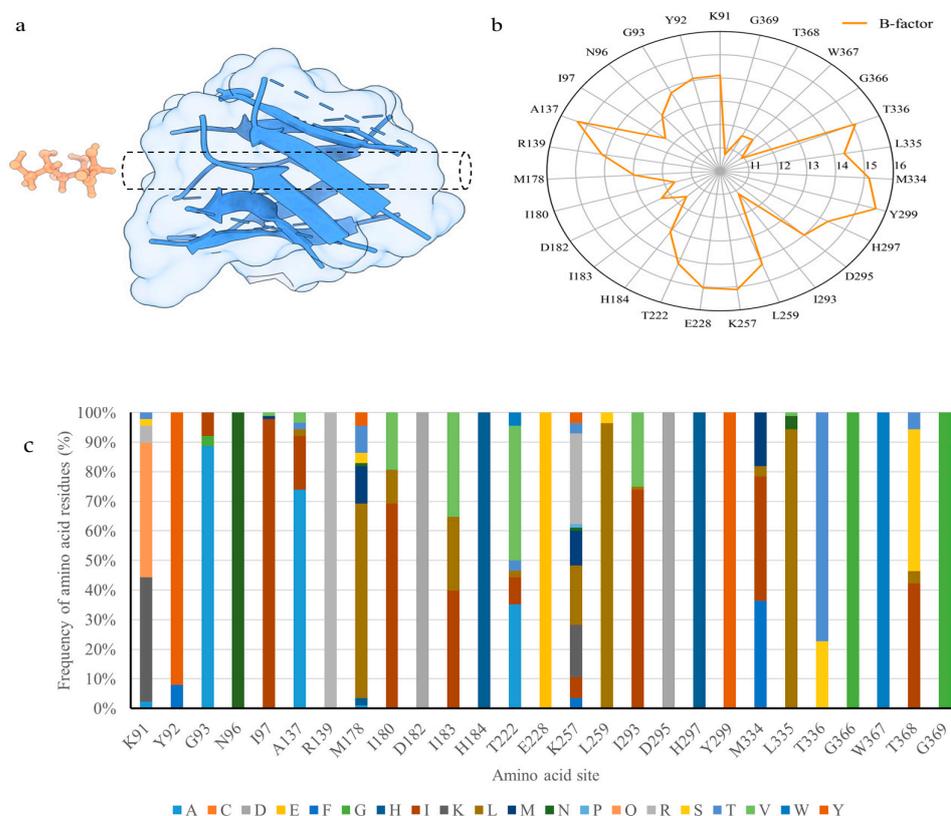


Figure 1. Selection of mutational target sites for thermostability improvement of ArCel5A. (a) Amino acid residues within a 4 Å radius of the catalytic site in the tunnel of ArCel5A; (b) comparison of predicted B-factor values of amino acid residues in ArCel5A; and (c) frequency of the aforementioned amino acids in the aligned sequences of multiple Cel5A orthologs.

The B-factor is an index of atomic displacement parameters obtained from X-ray data that reflects the blurring of atomic electron densities due to thermal motion and positional disorder [29]. Thus, residues with higher B-factors are more flexible and thermally unstable. It is hypothesized that replacing high-B-factor residues with ones with a lower B-factor may decrease flexibility and increase rigidity, resulting in enhanced thermostability of the protein [30]. Zhang et al. used iterative saturation mutagenesis (ISM) to mutate the highest B-factor residues of *B. subtilis* lipase and obtained two more thermostable recombinant lipases with T_m values of 41 and 45 °C [31]. To reduce the variant library size, the predicted B-factor values of the aforementioned amino acids were compared. Notably, after aligning more than 1000 sets of Cel5A orthologs from different species, it was discovered that N96, R139, D182, H184, E228, D295, H297, Y299, T336, G369 were completely conserved (Figure 1c), indicating that these are crucial residues for the function of Cel5A. Except for Y299, the B-factor values of the above-mentioned conserved residues were relatively low. Interestingly, Y299 had the highest B-factor value of 15.63, indicating that it is the most flexible and thermally unstable of the identified residues (Figure 1b). Thus, Y299 was considered as a possible mutational target site, as increasing its rigidity may lead to further improvement in the thermostability of ArCel5A.

3.3. High-Throughput Screening of Transformants with High Enzymatic Activity

To verify the functional roles of Y299 in the catalytic activity of ArCel5A, site-saturation mutagenesis was performed and the variants were expressed in *T. reesei* A2H. The 2A peptide system with high cleavage efficiency was proven to work well in filamentous fungi and led to the generation of two complete gene products, the mutated Cel5A and GFP, with a linear relationship [32]. In this study, overlap-extension PCR was used to conduct saturation mutagenesis on ArCel5A. We next examined their fluorescence intensities using a SpectraMax M2e microplate reader and microscope system. As shown in Figure 2b, the transformants overexpressing ArCel5A displayed varied levels of GFP fluorescence signals.

In the first round, we picked transformants with high levels of GFP fluorescence signals. Subsequently, the transformants were individually cultivated and screened for endoglucanase activity in 48 deep-well plates. After re-screening using the enzyme activity assay, the best transformant showed 83.51% higher endoglucanase activity than the wild-type strain. This positive transformant was further confirmed with gene sequence and SDS-PAGE profiling of the extracellular secretome. Further genomic sequencing confirmed that the gene sequence at position 299 of *T. reesei* A2H was mutated from TAC to TGC, and the corresponding amino acid was mutated from Y to C. Secreted protein production exhibited little change, suggesting that ArCel5A-Y299C was able to undergo proper folding and retained a conformation highly similar to wild-type ArCel5A.

3.4. Improvement of Enzyme Activity and Thermostability in the Mutant EG

ArCel5A-Y299C was successfully heterologously expressed in the culture supernatant of *T. reesei* A2H through fed-batch fermentation (Figure 2). The SDS-PAGE protein profile exhibited significant enhancement of specific bands in ArCel5A-Y299C as the fermentation progressed. The endoglucanase activity of ArCel5A-Y299C reached 3251 IU/mL, representing an 85.2% increase compared to wild-type ArCel5A after 94 h of fermentation (Figure 2c). SDS-PAGE analysis revealed a distinct band at approximately 50 kDa corresponding to ArCel5A-Y299C (Figure 2d). This apparent band indicated that the integration of high-throughput screening technology, which combines GFP fluorescence with enzyme activity analysis, can serve as a rapid and efficient method for protein engineering.

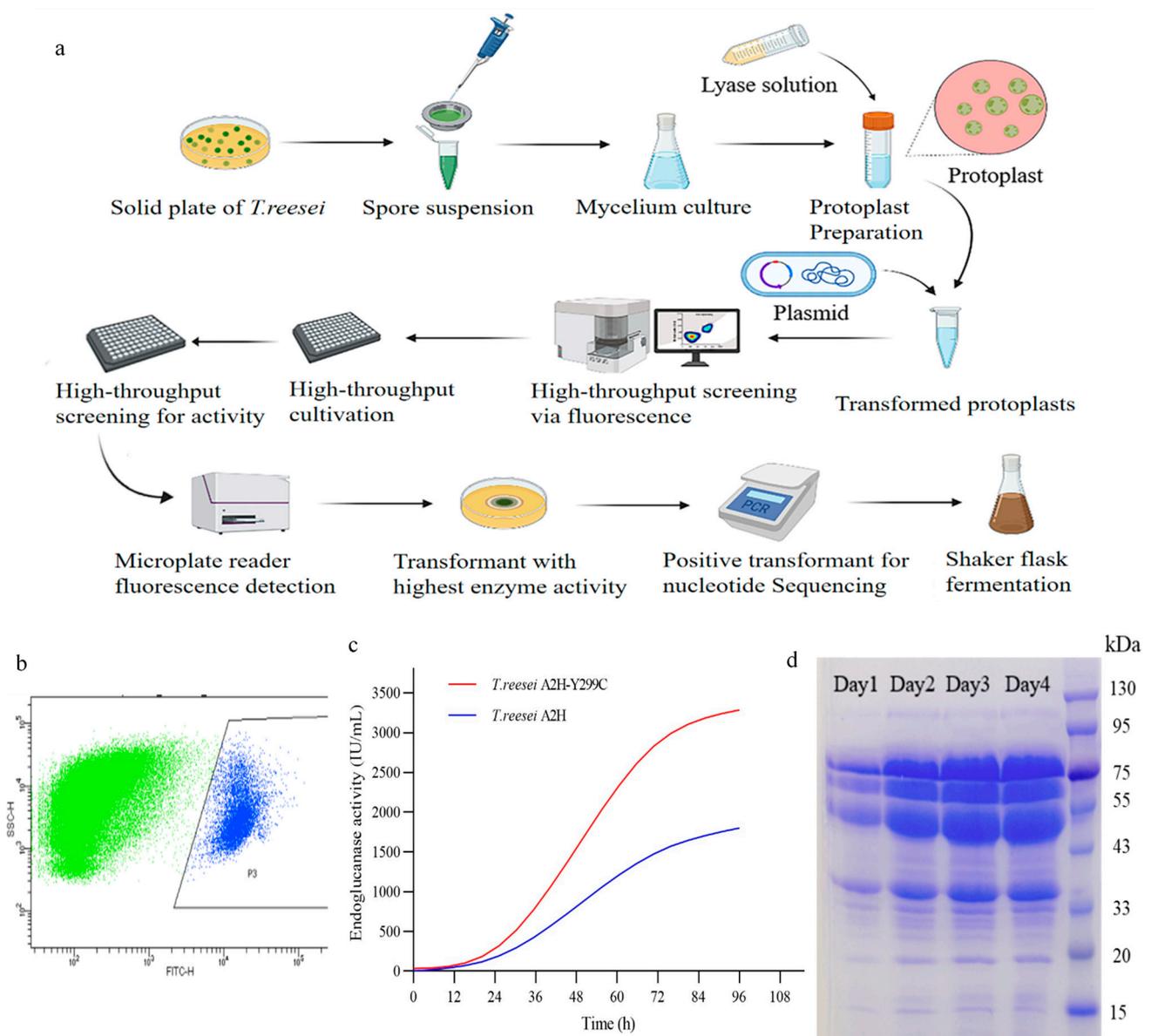


Figure 2. Screening mutants with high endoglucanase activity and comparative analysis of enzyme activity. (a) Schematic of the high-throughput screening approach for obtaining mutants with high endoglucanase activity; (b) high-throughput screening based on the fluorescence intensity of co-translationally expressed GFP; (c) comparison of the endoglucanase activity titer between ArCel5A and ArCel5A-Y299C after fed-batch fermentation; and (d) SDS-PAGE analysis of the fermentation supernatant from ArCel5A-Y299C.

As a thermophilic fungus, *Acidomyces richmondensis* encodes ArCel5A with exceptional thermostability. However, there is a need for further rational improvement of the industrially relevant properties of thermophilic enzymes, which presents a technological challenge. The thermal stability of mutant variants of ArCel5A was evaluated by assessing their enzyme activity after exposure to high temperatures of 70 and 80 °C. The results showed that after incubation at 70 °C for 30 min and 80 °C for 10 min, the residual enzyme activity of wild-type ArCel5A decreased to 76.9% and 71.2%, respectively. By contrast, the mutant ArCel5A-Y299C retained 93.8% and 91.5% of its initial activity under the same conditions. The properties of the endoglucanase were further improved by 17.0–20.3% compared to thermophilic ArCel5A (Table 1). Usually, thermostability improvement of enzymes is accompanied by a decrease of catalytic activity due to the activity–stability

trade-off [33,34]. However, these results indicate that the trade-off between activity and stability has been effectively overcome in the ArCel5A variant.

Table 1. Thermostability comparison between ArCel5A and ArCel5A-Y299C.

Strains	ArCel5A	ArCel5A-Y299C
Conditions	Residual activity	
30 min at 70 °C	76.9 ± 1.1%	93.8 ± 1.4%
10 min at 80 °C	71.2 ± 0.92%	91.5 ± 1.2%

3.5. Mechanisms for Stabilizing EG Mutations

Based on the aforementioned experimental results, it can be observed that the amino acid at position 299 of ArCel5A plays a significant role in enzyme activity and thermostability. To investigate the underlying mechanism, we applied different computational simulation methods to identify the reasons behind the significant increase of enzyme activity in the mutant.

Firstly, we utilized the I-mutant tool (<https://folding.biofold.org/i-mutant/i-mutant2.0.html>, accessed on 1 December 2006) to calculate the change of Gibbs free energy following the mutation of 299Y ($\Delta\Delta G$, Kcal/mol). At pH = 7.0, 50 °C, $\Delta\Delta G > 0$ indicates an increase in stability, while $\Delta\Delta G < 0$ indicates a decrease in stability. Our calculations revealed that when the amino acid at position 299 was mutated from Y to C, the $\Delta\Delta G$ value was the highest, indicating that the Y299C mutation at this position is the most stable (Table 2). These Gibbs free energy calculation results were in agreement with our experimental findings, suggesting that after mutating the amino acid at position 299 from Y to C, the enzyme structure becomes more stable, resulting in higher enzyme activity.

Table 2. Calculation of Gibbs free energy changes following site-saturation mutagenesis of Y299 in ArCel5A.

$\Delta\Delta G$	Y299A −2.83	Y299H −0.8	Y299D −0.75	Y299R −0.56	Y299F −0.6
$\Delta\Delta G$	Y299C 0.34	Y299G −3.47	Y299E −0.83	Y299K −1.2	Y299L −0.92
$\Delta\Delta G$	Y299M −0.56	Y299Q −0.99	Y299S −1.68	Y299Y 0	Y299T −0.64
$\Delta\Delta G$	Y299I −1.12	Y299W −0.63	Y299P −2.55	Y299V −0.73	

Secondly, the evaluation of the overall structure after mutation involves the calculation of protein dynamics. This is necessary to gain a better understanding of how ligand binding affects proteins. To achieve this, a simulated kinetic calculation was conducted using the Amber analysis tool. The 100 ns trajectories of the complexes were analyzed, and the RMSD values were calculated for 100 ns of simulation time. The RMSD value (Figure 3a) indicated that the Arcel5a-Y299C mutant exhibited lower deviation from the initial structure compared to the wild type, implying enhanced stability. This indicates that Arcel5a has a more stable overall structure that is conducive to enzymatic catalysis, resulting in better enzyme activity and thermal stability.

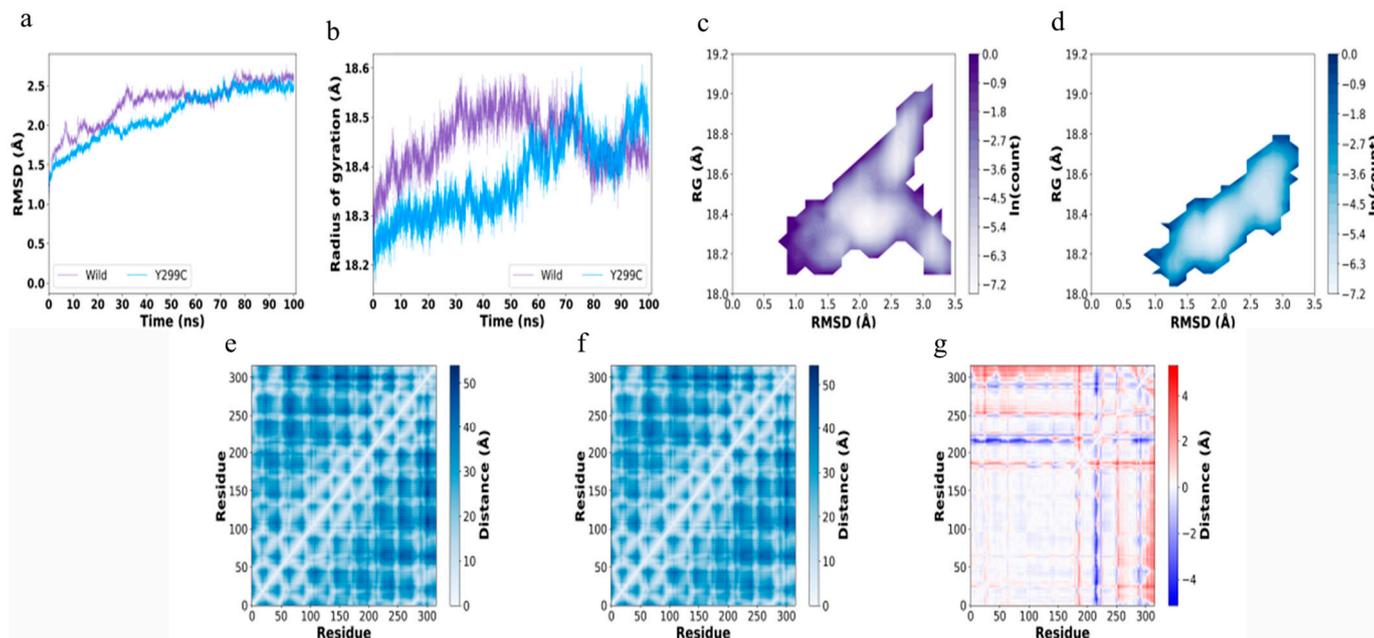


Figure 3. Calculation of the overall structural stability of the enzyme before and after mutation. The time evolution of the backbone root means square deviation (RMSD) (a) and radius of gyration (RG) values (b) of Arcel5a wild-type (Purple) and the indicated mutant (Blue) was calculated from the mean of 3 replicates of molecular dynamics simulations. Conformational distribution of Arcel5a wild-type (c) and mutant (d) in RMSD and RG space (calculated from 3 molecular dynamics replicates). Average C_{α} interatomic distance of Arcel5a wild-type (e) and mutant (f) calculated from the average of 3 molecular dynamics replicates, and the difference between the two (g).

Finally, the RG value of the mutant (Figure 3b) was lower, confirming its improved overall structural stability. The analysis of conformational group distribution (Figure 3c–e) showed that the Arcel5a-Y299C mutant had a narrower range of conformational variations and a more concentrated distribution of conformational groups compared to the wild type. This reduced conformational diversity and indicated higher structural stability in the mutant form. Additionally, the C_{α} distance difference matrix (Figure 3g) indicated shorter distances in several regions of the mutant compared to the wild type, indicating increased compactness and overall stability of the protein. Notably, the region surrounding the mutated residue 299 exhibited a prominent red coloration, indicating that the Arcel5a-Y299C mutation brought the nearby residues closer together. This closer arrangement likely enhanced the stability of the catalytic pocket, thereby improving the overall stability and catalytic efficiency of Arcel5a.

Thus, the trade-off between activity and stability could be overcome by strategically selecting mutation sites that are distant from the catalytic center [22]. Numerous studies have focused on increasing the thermostability of enzymes by modifying residues on the protein surface or away from the catalytic center, including mutations of charged amino acids, increased incorporation of hydrophobic amino acids, alterations of flexible residues on the enzyme surface, and stabilization of the proteins termini [35]. While these modifications close to the catalytic center have improved the thermostability of xylanases, they also had a negative impact on the enzyme's catalytic ability. In comparison to other counterparts, such as EGI from *Stegonsporium opalus* (SoCel5), which lost 80% of the initial activity after being incubated at 70 °C for 10 min [17], and the improved *T. reesei* Cle7B (TrCel7B), which retained only 35% of its initial activity after 15 min at 64 °C [31], the EG mutant obtained in this study exhibited significantly superior thermostability. In this study, a novel B-factor analysis approach has been proven to be an effective method for further thermostability improvement of ArCel5A, which successfully mined potential amino acid substitutions in the catalytic center. Notably, a successful balance was achieved

between activity and stability in the ArCel5A-Y299C variant, resulting in an impressive 85.2% increase of endoglucanase activity and a thermostability improvement of 17.0–20.3% compared to wild-type ArCel5A.

4. Conclusions

In the present paper, we demonstrate an innovative strategy for rational modification of thermophilic endoglucanases, resulting in superior characteristics that overcome the enzyme activity–stability trade-off. To further understand the mechanisms responsible for enhanced thermostability and catalytic efficiency of the thermophilic EG, we conducted various computational analyses, including Gibbs free energy calculations and MD simulations. This strategy may provide new insights for further mining of potential rational modification sites among highly conserved sequences in the catalytic center of ArCel5A.

Author Contributions: Experimental work, S.W. and Z.Z.; formal analysis, Z.Z. and T.B.; writing original draft preparation, L.G. (Le Gao) and Y.L.; model analysis, H.D. and Y.L.; results analysis, J.Y., J.Z., X.W. and L.G. (Lingfang Gu); manuscript reviewing, L.G. (Le Gao) and X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National key R&D program of China (2021YFD1301002), Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA28030301), National Natural Science Foundation of China (62241602), as well as the Jilin Province and Chinese Academy of Sciences Science and Technology Cooperation High-tech Industrialization Fund Project (2022SYHZ0015, 2022SYHZ0019).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated or analyzed during this study are included in this published article.

Conflicts of Interest: Author Lingfang Gu was employed by the company Wuhan Sunhy Biology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Qaiser, H.; Kaleem, A.; Abdullah, R.; Iqtedar, M.; Hoessli, D.C. Overview of lignocellulolytic enzyme systems with special reference to valorization of lignocellulosic biomass. *Protein. Pept. Lett.* **2021**, *28*, 1349–1364. [[CrossRef](#)] [[PubMed](#)]
2. Marriott, P.E.; Gómez, L.D.; McQueen-Mason, S.J. Unlocking the potential of lignocellulosic biomass through plant science. *New Phytol.* **2016**, *209*, 1366–1381. [[CrossRef](#)] [[PubMed](#)]
3. Liu, G.; Qu, Y. Integrated engineering of enzymes and microorganisms for improving the efficiency of industrial lignocellulose deconstruction. *Eng. Microbiol.* **2021**, *1*, 100005. [[CrossRef](#)]
4. Davies, G.J.; Dodson, G.G.; Hubbard, R.E.; Tolley, S.P.; Dauter, Z.; Wilson, K.S.; Hjort, C.; Mikelsen, J.M.; Rasmussen, G.; Schülein, M. Structure and function of endoglucanase V. *Nature* **1993**, *365*, 362–364. [[CrossRef](#)] [[PubMed](#)]
5. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic. Acids Res.* **2014**, *42*, 490–495. [[CrossRef](#)]
6. Aspeborg, H.; Coutinho, P.M.; Wang, Y.; Brumer, H.; Henrissat, B. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol. Biol.* **2013**, *12*, 186. [[CrossRef](#)]
7. Tseng, C.W.; Ko, T.P.; Guo, R.T.; Huang, J.W.; Wang, H.C.; Huang, C.H.; Cheng, Y.S.; Wang, A.H.J.; Liu, J.R. Substrate binding of a GH5 endoglucanase from the ruminal fungus *Piromyces rhizinflata*. *Acta Crystallogr. Sect. F* **2011**, *67*, 1189–1194. [[CrossRef](#)]
8. Leggio, L.L.; Larsen, S. The 1.62 Å structure of *Thermoascus aurantiacus* endoglucanase: Completing the structural picture of subfamilies in glycol-side hydrolase family 5. *FEBS Lett.* **2002**, *523*, 103–108. [[CrossRef](#)]
9. Lee, T.M.; Farrow, M.F.; Arnold, F.H.; Mayo, S.L. A structural study of *Hypocrea jecorina* Cel5A. *Protein. Sci.* **2011**, *20*, 1935–1940. [[CrossRef](#)]
10. Liu, G.; Li, Q.; Shang, N.; Huang, J.W.; Ko, T.P.; Liu, W.; Zheng, Y.; Han, X.; Chen, Y.; Chen, C.C.; et al. Functional and structural analyses of a 1,4-β-endoglucanase from *Ganoderma lucidum*. *Enzyme. Microb. Technol.* **2016**, *86*, 67–74. [[CrossRef](#)]
11. Yan, J.; Liu, W.; Li, Y.; Lai, H.L.; Zheng, Y.; Huang, J.W.; Chen, C.C.; Chen, Y.; Jin, J.; Li, H.; et al. Functional and structural analysis of *Pichia pastoris*-expressed *Aspergillus niger* 1,4-β-endoglucanase. *Biochem. Biophys. Res. Commun.* **2016**, *475*, 8–12. [[CrossRef](#)]
12. Wierenga, R. The TIM-barrel fold: A versatile framework for efficient enzymes. *FEBS Lett.* **2001**, *492*, 193–198. [[CrossRef](#)] [[PubMed](#)]

13. You, S.; Tu, T.; Zhang, L.J.; Wang, Y.; Huang, H.Q.; Ma, R.; Shi, P.J.; Bai, Y.G.; Su, X.Y.; Lin, Z.M.; et al. Improvement of the thermostability and catalytic efficiency of a highly active β -glucanase from *Talaromyces leycettanus* JCM 12802 by optimizing residual charge-charge interactions. *Biotechnol Biofuels*. **2016**, *9*, 124. [[CrossRef](#)] [[PubMed](#)]
14. Wilson, D.B. Cellulases and biofuels. *Curr. Opin. Biotechnol.* **2009**, *20*, 295–299. [[CrossRef](#)]
15. Celestino, K.R.S.; Cunha, R.B.; Felix, C.R. Characterization of a β -glucanase produced by *Rhizopus microsporus* var. *microsporus*, and its potential for application in the brewing industry. *BMC Biochem.* **2006**, *7*, 23. [[CrossRef](#)] [[PubMed](#)]
16. Han, C.; Liu, Y.; Liu, M.; Wang, S.; Wang, Q. Improving the thermostability of a thermostable endoglucanase from *Chaetomium thermophilum* by engineering the conserved non-catalytic residue and *N*-glycosylation site. *Int. J. Biol. Macromol.* **2020**, *164*, 3361–3368. [[CrossRef](#)] [[PubMed](#)]
17. Juturu, V.; Wu, J.C. Microbial cellulases: Engineering, production and applications. *Renew. Sustain. Energy. Rev.* **2014**, *33*, 188–203. [[CrossRef](#)]
18. Martínez, A.C.; Balcázar, L.E.; Dantán, G.E.; Folch, M.J.L. Celulasas fúngicas: Aspectos biológicos y aplicaciones en la industria energética. *Rev. Latinoam. Microbiol.* **2008**, *50*, 119–131.
19. Zheng, F.; Vermaas, V.J.; Zheng, J.; Wang, Y.; Tu, T.; Wang, X.; Xie, X.; Yao, B.; Beckham, G.T.; Luo, H. Activity and thermostability of GH5 endoglucanase chimeras from mesophilic and thermophilic parents. *Appl. Environ. Microbiol.* **2019**, *85*, e02079-18. [[CrossRef](#)]
20. Ferrè, F.; Clote, P. DiANNA: A web server for disulfide connectivity prediction. *Nucleic Acids Res.* **2005**, *33*, W230–W232. [[CrossRef](#)]
21. Pace, C.N.; Fu, H.; Fryar, K.L.; Landua, J.; Trevino, S.R.; Shirley, B.A.; Hendricks, M.M.; Limura, S.; Gajiwala, K.; Scholtz, J.M.; et al. Contribution of hydrophobic interactions to protein stability. *J. Mol. Biol.* **2011**, *408*, 514–528. [[CrossRef](#)]
22. Tong, L.; Zheng, J.; Wang, X.; Wang, X.; Huang, H.; Yang, H.; Tu, T.; Wang, Y.; Bai, Y.; Luo, H.; et al. Improvement of thermostability and catalytic efficiency of glucoamylase from *Talaromyces leycettanus* JCM12802 via site-directed mutagenesis to enhance industrial saccharification applications. *Biotechnol. Biofuels* **2021**, *14*, 202. [[CrossRef](#)]
23. Gao, L.; Guo, Q.; Xu, R.; Dong, H.; Zhou, C.; Yu, Z.; Zhang, Z.; Wang, L.; Chen, X.; Wu, X. Loop optimization of *Trichoderma reesei* endoglucanases for balancing the activity–stability trade-off through cross-strategy between machine learning and the B-factor analysis. *GCB Bioenergy* **2022**, *15*, 128–142. [[CrossRef](#)]
24. Li, F.; Xie, Y.; Gao, X.; Shan, M.; Sun, C.; Niu, Y.D.; Shan, A. Screening of cellulose degradation bacteria from Min pigs and optimization of its cellulase production. *Electron. J. Biotechnol.* **2020**, *48*, 29–35. [[CrossRef](#)]
25. Gomi, K.; Iimura, Y.; Hara, S. Integrative transformation of *Aspergillus oryzae* with a plasmid containing the *Aspergillus nidulans* argB gene. *Agr. Biol. Chem.* **1987**, *51*, 2549–2555.
26. Karplus, P.A.; Schulz, G.E. Prediction of chain flexibility in proteins. *Naturwissenschaften* **1985**, *72*, 212–213. [[CrossRef](#)]
27. You, S.; Chen, C.C.; Tu, T.; Wang, X.; Ma, R.; Cai, H.Y.; Guo, R.T.; Luo, H.Y.; Yao, B. Insight into the functional roles of Glu175 in the hyperthermostable xylanase XYL10C- Δ N through structural analysis and site-saturation mutagenesis. *Biotechnol. Biofuels* **2018**, *11*, 159. [[CrossRef](#)]
28. Guo, Y.; Tu, T.; Zheng, J.; Bai, Y.; Huang, H.; Su, X.; Wang, Y.; Wang, Y.; Yao, B.; Luo, H. Improvement of BsAPA Aspartic Protease Thermostability via Autocatalysis-Resistant Mutation. *J. Agric. Food Chem.* **2019**, *67*, 10505–10512. [[CrossRef](#)]
29. Parthasarathy, S.; Murthy, M.R.N. Protein thermal stability: Insights from atomic displacement parameters (B values). *Protein. Eng.* **2000**, *13*, 9–13. [[CrossRef](#)]
30. Reetz, M.T.; Carballeira, J.D.; Vogel, A. Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angew. Chem. Int. Ed.* **2006**, *45*, 7745–7751. [[CrossRef](#)]
31. Zhang, S.; Wang, Y.; Song, X.; Hong, J.; Zhang, Y.; Yao, L. Improving *Trichoderma reesei* Cel7B thermostability by targeting the weak spots. *J. Chem. Inf. Model.* **2014**, *54*, 2826–2833. [[CrossRef](#)]
32. Li, F.; Liu, Q.; Li, X.; Zhang, C.; Li, J.; Sun, W.; Liu, D.; Xiao, D.; Tian, C. Construction of a new thermophilic fungus *Myceliophthora thermophila* platform for enzyme production using a versatile 2A peptide strategy combined with efficient CRISPR-Cas9 system. *Biotechnol. Lett.* **2020**, *42*, 1181–1191. [[CrossRef](#)]
33. Kurahashi, R.; Tanaka, S.; Takano, K. Activity-stability trade-off in random mutant proteins. *J. Biosci. Bioeng.* **2019**, *128*, 405–409. [[CrossRef](#)] [[PubMed](#)]
34. Siddiqui, K.S. Defying the activity–stability trade-off in enzymes: Taking advantage of entropy to enhance activity and thermostability. *Crit. Rev. Biotechnol.* **2017**, *37*, 309–322. [[CrossRef](#)]
35. Li, G.; Zhou, X.; Li, Z.; Liu, Y.; Liu, D.; Miao, Y.; Wan, Q.; Zhang, R. Significantly improving the thermostability of a hyperthermophilic GH10 family xylanase XynAF1 by semi-rational design. *Appl. Microbiol. Biotechnol.* **2021**, *105*, 4561–4576. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.