J. Microbiol. Biotechnol. (2016), 26(7), 1163–1172
http://dx.doi.org/10.4014/jmb.1512.12070

jmb

A Novel Strategy for Thermostability Improvement of Trypsin Based on *N*-Glycosylation within the Ω-Loop Region

Chao Guo[1,2,3†], Ye Liu[1,2,3†], Haoran Yu[1,2,3], Kun Du[1,2,3], Yiru Gan[1,2,3], and He Huang[1,2,3*]

[1]*Department of Biochemical Engineering, School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, P.R. China*
[2]*Key Laboratory of System Bioengineering, Ministry of Education, Tianjin University, Tianjin 300072, P.R. China*
[3]*Collaborative Innovation Center of Chemical Science and Engineering, Tianjin 300072, P.R. China*

Received: December 28, 2015
Revised: March 20, 2016
Accepted: March 23, 2016

First published online
March 24, 2016

*Corresponding author
Phone: +86-22-2740-9598;
Fax: +86-22-27409598;
E-mail: huang@tju.edu.cn

†These authors contributed equally to this work.

pISSN 1017-7825, eISSN 1738-8872

The Ω-loop is a nonregular and flexible structure that plays an important role in molecular recognition, protein folding, and thermostability. In the present study, molecular dynamics simulation was carried out to assess the molecular stability and flexibility profile of the porcine trypsin structures. Two Ω-Loops (fragment 57–67 and fragment 78–91) were confirmed to represent the flexible region. Subsequently, glycosylation site-directed mutations (A73S, N84S, and R104S) were introduced within the Ω-loop region and its wing chain based on its potential *N*-glycosylation sites (Asn-Xaa-Ser/Thr consensus sequences) and structure information to improve the thermostability of trypsin. The result demonstrated that the half-life of the N84S mutant at 50°C increased by 177.89 min when compared with that of the wild-type enzyme. Furthermore, the significant increase in the thermal stability of the N84S mutant has also been proven by an increase in the T_m values determined by circular dichroism. Additionally, the optimum temperatures of the wild-type enzyme and the N84S mutant were 75°C and 80°C, respectively. In conclusion, we obtained the thermostability-improved enzyme N84S mutant, and the strategy used to design this mutant based on its structural information and *N*-linked glycosylation modification could be applied to engineer other enzymes to meet the needs of the biotechnological industry.

Keywords: Ω-Loop, *N*-glycosylation, trypsin, thermostability

Introduction

The ability to engineer proteins for improved themostability is an exciting and challenging field since it is critical for broadening the industrial applications of recombinant proteins. The approach of rational design has been proven to be a powerful tool to improve the thermostability of proteins [6, 7, 9]. Compared with directed evolution, the rational design methods are faster and universal, and have the potential to be developed into algorithms that can quantitatively predict the stabilities of the designed sequence [27]. Using the rational design method of biochemical dating, the thermostability of proteins has been thought to originate from the simultaneous effect of different interactions, such as hydrophobic interactions, disulfide bonds, salt bridges, and hydrogen bonds [36]. The critical point of designing thermostable proteins was to pinpoint the weak spots or the "unfolding nucleus." Once the weak spots were identified, further thermostability could be achieved by optimizing the weak spot regions.

Ω-Loops are found almost exclusively at the protein surface and they were shown to be involved in protein function and molecular recognition [10]. Moreover, these loops may also play significant roles in protein stability or folding. Single and multiple mutations in some Ω-loops could have drastic effects on overall protein stability [15, 24]. Hydrogen-exchange experiments indicated that the Ω-loop was the most susceptible element in cytochrome *c* denaturation [2]. By minimizing conformational flexibility through the hydrogen bonds and the existence of the Pro residue, the extra helix and adjacent loops in cytochrome c_{555} could help protect against protein denaturation [23].

Many methods have been applied to stabilize the thermostability of proteins. Among them, glycation is gaining increasing attention owing to the fact that it is a nontoxic method and can be easily performed. There is substantial evidence indicating that *N*-linked oligosaccharides are important for protein folding and stability [14, 22, 35]. With polysaccharides attached enzymatically to the asparagine, serine, and threonine side chains, naturally glycosylated proteins have increased structural stability. It would clearly be advantageous if improved structural stability could be achieved for nonglycosylated proteins by the covalent attachment of sugars or polysaccharides.

As one of the members in the serine protease family, trypsin (E.C. 3.4.21.4) is commonly used as a tool for enzyme protein cleavage and digestion in the field of biotechnology. For instance, it is used for the activation of some zymogens and as a digester of proteins prior to the analysis of peptide fingerprints during mass spectrometry [11, 18]. When used for these purposes, the enzyme often suffers from autolysis and denaturation, which would be ideally solved by improved stability [11]. Moreover, the apparent increase in the thermostability of the trypsin could be used to digest native ribonuclease without the need for prior denaturation [25]. In this work, *N*-linked oligosaccharides have been introduced within the Ω-loop regions to enhance the thermostability of porcine trypsin. As a result, a more thermostable mutant protein was expressed in *Pichia pastoris* GS115, and its enzymatic properties, including its activity, secretion, thermostability, and temperature optima, were characterized and compared with the wild-type enzyme.

Materials and Methods

Microorganisms, Plasmids, and Reagents

Plasmid pPIC9K and the *P. pastoris* host strain GS115 were obtained from Invitrogen (Thermo-Fisher Scientific, USA). The *Escherichia coli* Top 10 host strain was from our laboratory collection. Phusion High-Fidelity DNA Polymerases were obtained from Thermo-Fisher Scientific. The restriction enzyme SalI was purchased from Takara (China). The BCA protein assay kit was purchased from Dingguo Changsheng Biotechnology Co., Ltd. (China). All other chemicals used in the experiments were of analytical grade.

Design and Selection of the Proper *N*-Glycosylation Sites

The crystal structure of porcine trypsin (PDB ID: 4AN7) was obtained from the RCSB Protein Data Bank (http://www.rcsb.org). The protein structure information, including its active sites, disulfide bond, β-sheet, and Ω-loop regions, were shown and analyzed using the Visual Molecular Dynamics (VMD) software. Molecular dynamics (MD) simulations were performed using GROMACS ver. 4.5.5 [33], implementing the Gromos96.1 (53A6) force field [28] to examine the flexible regions of the protein. The protein was solvated with the Simple Point Charge water molecule in a cubic box using the GROMACS software package. The minimum distance between any atom in the protein and the box's walls was set to 1.1 nm. Sufficient Cl^- ions were added to neutralize the positive charges in the system. The water box and the whole system were minimized using the steep descent method (1,000 steps) in addition to the conjugate gradient method (3,000 steps). Then, a 20 ps position-restrained MD simulation was performed at 300 K. Finally, an unrestrained MD simulation was performed on the entire system at 300 K for 10 ns. The Particle Mesh Ewald method was used to treat long-range electrostatic interactions and the cut-off value for van der Waals' interactions was set at 1.0 nm. Trajectories were saved at every 1,000 steps (2 ps), and post-processing and analysis were performed using relevant tools in GROMACS. The root mean square fluctuation (RMSF) value of backbone atoms against each residue was calculated for the enzyme using g_rmsf. The proper *N*-glycosylation sites were chosen based on analysis of the crystal structure information and potential *N*-glycosylation sites (Asn-Xaa-Ser/Thr consensus sequences) within the Ω-loop region.

Construction of the Expression Vector and Site-Directed Mutagenesis

The wild-type porcine trypsinogen gene (GenBank Accession No. CS583166.1) was synthesized by GENEWIZ Inc. (China) and inserted into the pPIC9K vector. The recombinant plasmid pPIC9K-Try was linearized by SalI and integrated at the His_4 locus on the competent *Pichia* genome through electroporation using a Bio-Rad Micropulser Electroporater (Bio-Rad, USA). The trypsinogen variants used in this study were constructed using the modified Quik Change method [38]. The mutagenic primers (synthesized by GENEWIZ Inc.) are listed in Table 1. The mutant plasmids were verified by DNA sequencing (GENEWIZ Inc.); they were then transformed into the *P. pastoris* host strain, GS115.

Expression and Purification of Enzymes in *P. pastoris* Recombinants

The selected clones were inoculated into 25 ml of buffered glycerol-complex medium and they were subsequently incubated overnight at 30°C with shaking at 250 rpm. When the OD_{600} reached 2–6, the cells were harvested via centrifugation at 5,000 ×*g* for 5 min. To induce protein expression, the cell pellets were resuspended in 25 ml of buffered methanol-complex medium and transferred into a 100 ml culture flask. The culture was shaken for 96 h at 28°C and supplemented with 100% methanol to a final concentration of 1% (v/v) every 24 h. The fermentation supernatant was centrifuged and collected after being induced with methanol for 96 h. The supernatant was filtered through a 0.22 μm filter, and then was dialyzed through a Millipore 10 kDa cut-off membrane to exclude ions and salts, resuspended in buffer A (20 mM citric acid, pH 3.0), and loaded on a SP Sepharose HP (GE Healthcare, UK), which was pre-equilibrated with at least 5 column volumes

Table 1. Primers for the site mutation.

Mutations	Mutagenic primers[a]	T_m/°C
A73S	5′-AACAATTTATTAATGCTTCTAAAATTATTACTCATCC-3′ (forward)	51.7
	5′-AAGCATTAATAAATTGTTCATTACCTTCCAAAAC-3′ (reverse)	51.3
N84S	5′-TCCAAATTTTAATGGTTCTACTTTGGATAATGATATTAT-3′ (forward)	55.5
	5′-GAACCATTAAAATTTGGATGAGTAATAATTTTAGCAG-3′ (reverse)	54.4
R104S	5′-GCTACTTTAAATTCATCTGTTGCTACTGTTTCTTTG-3′ (forward)	55.7
	5′-AGATGAATTTAAAGTAGCTGGAGAAGACAATTTAATC-3′ (reverse)	56.0

[a]Nucleotides that were changed to encode Ser are underlined.

buffer A until the absorbance reached a steady baseline. The protein was eluted using a 0–500 mM sodium chloride gradient, and the active fraction of peak was pooled, dialyzed against the 50 mM Tris–HCl buffer (pH 7.8). Then, 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was applied to test the result of chromatography.

Deglycosylation of Mutant Trypsinogens

The *N*-glycosylation levels of the trypsinogen mutants (A73S, N84S, and R104S) were confirmed by deglycosylation using Endo H (New England Biolabs, USA) according to the manufacturer's protocol. The enzyme samples (containing 50 μg protein of trypsinogen) were boiled for 10 min in denaturing buffer containing 0.4 M DTT and 0.5% SDS to fully expose the glycosylation sites, and deglycosylation was then performed by treatment with 1,000 NEB units of Endo H in 50 mM sodium citrate (pH 5.5) for 2 h at 37°C. After the reaction, SDS-PAGE was performed to detect the separation of reaction products.

Activity Assay

The trypsin activity of the sample was estimated using Na-benzoyl-L-arginine ethyl ester hydrochloride (BAEE) as a substrate. Since there was an oligopeptide (DDDDK) in the N-terminus of the shortened porcine trypsinogen, enterokinase was required for its activation. The culture was processed by ultrafiltration using a Millipore filter (MW: 10,000 cut-off) and rebuffered by an equivalent volume of 67 mmol/l of sodium phosphate buffer (pH 7.6). Then, 100 μl of the processed sample was mixed with 2 μl of enterokinase and incubated at 25°C for 4 h. At continuous time points, sample mixtures (100 μl) were immediately mixed with 3 ml of assay buffer (67 mmol/l of sodium phosphate buffer at a pH of 7.6 at 25°C, containing 0.25 mmol/l of BAEE) and 100 μl of 1.0 mmol/l HCl. The change in absorbance at 253 nm was then monitored in a Shanghai Youke UV-765 PC spectrophotometer. One BAEE unit (U/ml) will produce a ΔOD_{253} of 0.001 per minute with BAEE as the substrate at a pH of 7.6 at 25°C in a reaction volume of 3.2 ml (1 cm light path). The trypsin activity was calculated according to the following formula:

$$\mathrm{BAEE(U/ml)} = \frac{\Delta \mathrm{OD253/min} \times \mathrm{df}}{0.001 \times 0.1} \quad (1)$$

where df represents the dilution factor.

Far UV Circular Dichroism (CD) Spectroscopic Analysis

CD spectra were taken on a Jasco J-810 spectropolarimeter (Jasco, Japan), which was continuously purged with nitrogen. Measurement was performed at 25°C for a protein concentration of 0.5 mg/ml in 50 mM Tris–HCl buffer (pH 7.8) using the cell with 1.0 mm pathlength for far-ultraviolet CD spectra (190–260 nm). The values of the scan rate, response, and band width were 200 nm/min, 1.0 sec, and 1.0 nm, respectively.

Measurement of Thermostability

To study the half-lives of the thermal inactivation of the trypsin variants, the purified trypsins (200 μg/ml) in 67 mM of NaH_2PO_4 (pH 7.6) were incubated at 50°C for different time intervals, ranging from 0 to 180 min; the trypsins were then cooled on ice for 30 sec. The enzyme activities were assayed under room temperature (25°C) and any remaining activity was recorded as a percentage of the original activity. The data were fitted to first-order plots and analyzed, with the first-order rate constant (k_d) measured by the linear regression of ln (remaining activity) versus the incubation time (t). The time required for the residual activity to be reduced by half ($t_{1/2}$) of the trypsin variants at 50°C was calculated by the equation

$$t_{1/2} = \frac{\ln 2}{k_d} \quad (2)$$

The melting temperature (T_m) of the trypsin variants was measured using CD measurements with a J-810 spectrometer, which was equipped with a Julabo temperature control system. The unfolding curves were created by continuously monitoring ellipticity changes at a fixed wavelength of 215 nm over the temperature range of 20°C–95°C in temperature increments of 1°C/min. The measurements were performed in 50 mM of Tris–HCl (pH 7.8) buffer using 1 mm cells and a protein concentration of 0.2 mg/ml. The observed ellipticities were converted to apparent fractions of the denatured protein, using the equation

$$F_d = \frac{\theta_T - \theta_N}{\theta_D - \theta_N}, \quad (3)$$

where θ_T is the sample ellipticity at a specific temperature and θ_N

and θ_D are the corresponding values for the native and denatured states.

Effects of Temperature on Mutant Trypsins

The optimal temperature (T_{opt}) was determined by measuring the activity at a series of temperatures ranging 25°C–95°C at 5°C intervals. The percent age of the remaining activity was determined by considering the activity of the wild-type enzyme incubated at an optimal temperature as a control (100%).

Digestion of Native Ribonuclease A

Native ribonuclease A (Shanghai Sangon Biotech Co., Ltd, China) and wild-type porcine trypsin or glycated porcine trypsin were combined and incubated at 37°C for 3 h. The ratio of the protein to enzyme was 500:1. The digestions were analyzed by SDS-PAGE for each trypsin mutant and wild-type enzyme.

Results

A Three-Dimensional Model of Porcine Trypsin and the Selection of Proper *N*-Glycosylation Sites

The three-dimensional model of porcine trypsin was visualized and analyzed using the VMD software (Fig. 1). The soluble porcine trypsin is mainly composed of two β-barrels (fragments 1–96 and 122–221). The former (fragment 1–96) consists of only one disulfide bond, and fewer than five disulfide bonds are found in the latter (fragment 122–221). As disulfide bonds are covalent linkages, they play important roles in stabilizing the protein structure [1, 31]; hence, we may infer that the fragment 1–96 region may be more flexible.

Furthermore, MD simulations were performed using Gromacs ver. 4.5.5 to determine the most flexible region within the β-barrels. The RMSFs of the backbone atoms against each residue were calculated for the enzyme (Fig. 2). The higher the value of RMSF, the more flexible the protein structure. From the calculation results, the average RMSF value of residual number 57-67 and 78-91 was 0.128 and 0.096, respectively, larger than the rest of the protein. (The average RMSF value of the rest residual number within fragment 1–96 is 0.086.) Therefore, according to the result of the MD simulation, Fragments 57–67 and 78–91 revealed higher RMSF values and showed greater flexibility when compared with the rest of the protein within the fragment 1–96 region. Both of these two fragments contain an Ω-loop, indicating that Ω-loops are flexible regions in the protein.

In addition, we further analyzed the crystal structure information within two Ω-loop regions and their potential *N*-glycosylation sites (Fig. 3). The Ω-loop region (fragment 57–67) contains a Ca^{2+}-binding site, which might feature intrinsic stability and affect overall protein stability [10]. Therefore, we introduced the *N*-glycosylation site (N71 and N102) in its wing chain to avoid affecting the metal ion-binding domain. N80, N82, N84, and N88 are potential *N*-glycosylation sites within fragment 78–91. Considering that N84 and N88 are near the enzyme's active sites, the

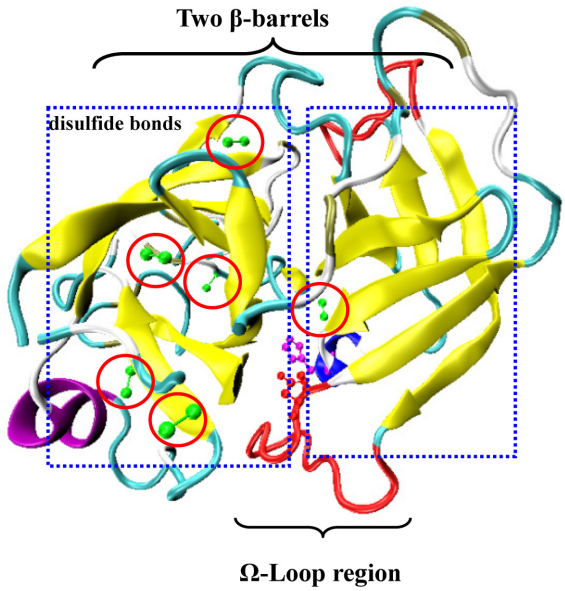

Fig. 1. Crystal structure of porcine trypsin visualized by the Visual Molecular Dynamics software from the Protein Data Bank file of trypsin (PDB code: 4AN7).

The total disulfide bonds (green) were shown in the two β-barrel regions (yellow) and Ω-loops (red) in the catalytic domain, containing two active site residues (red and violet), which were visualized.

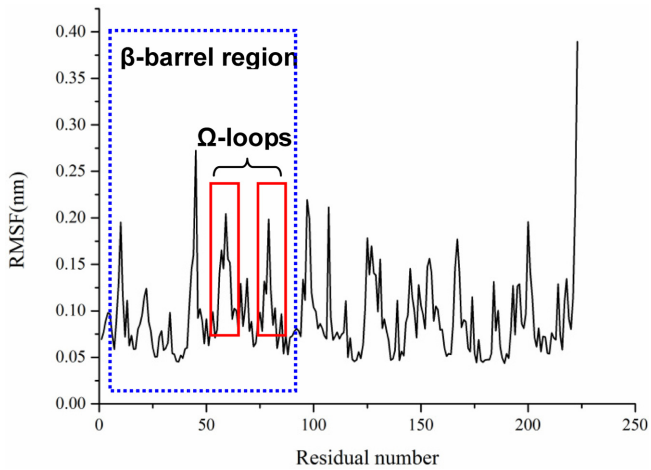

Fig. 2. The residual flexibility of wild-type trypsin during a 10 ns MD simulation.

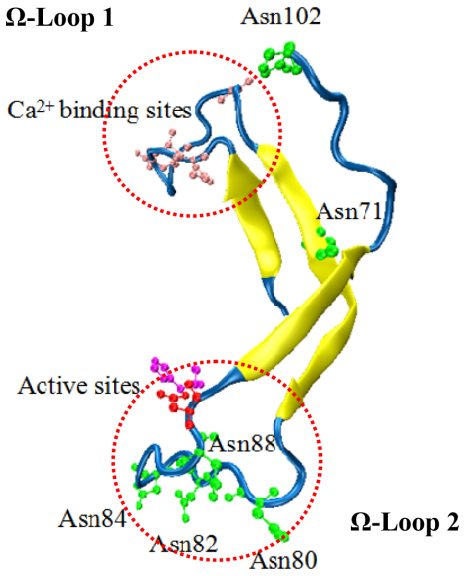

Fig. 3. Ω-loops in the active region and glycosylation of the available sites of trypsin.

The sequence of Ω-loop 2 is HPNFNGNTLDNDIM (fragment 78–91) and the potential *N*-glycosylation sites are shown within the Ω-loop region and its wing chain.

introduction of *N*-linked oligosaccharides could influence enzymatic activity owing to space steric hindrance. Thus, three potential *N*-glycosylation sites (N71, N82, and N102) were ultimately chosen to improve the protein's thermostability.

Identification of *N*-Glycosylation Levels of Trypsinogen Variants

The mutant and wild-type trypsinogens expressed by *P. pastoris* were assessed by SDS-PAGE, as shown in Fig. 4. The *N*-glycosylation levels of the heterologous proteins were determined by the change in their apparent molecular mass following mutation. From the results, the proteins with molecular mass of ~26 kDa were considered non-glycosylated forms (Fig. 4, lane 1), and the proteins with higher molecular mass of ~28 kDa were considered glycosylated forms (Fig. 4, lanes 5 and 7). Furthermore, the *N*-glycosylation levels of the trypsin mutants were confirmed by deglycosylation using the Endo H (NEB), and the Endo H enzyme was also assessed by SDS-PAGE (approximately 28 kDa). From the results, the apparent molecular mass of the wild-type trypsin was not changed by deglycosylation (Fig. 4, lane 2), whereas those of N84S and R104S decreased to approximately 26 kDa after deglycosylation (Fig. 4, lanes 6 and 8). The SDS-PAGE of the A73S mutant did not show any bands (Fig. 4, lane 3),

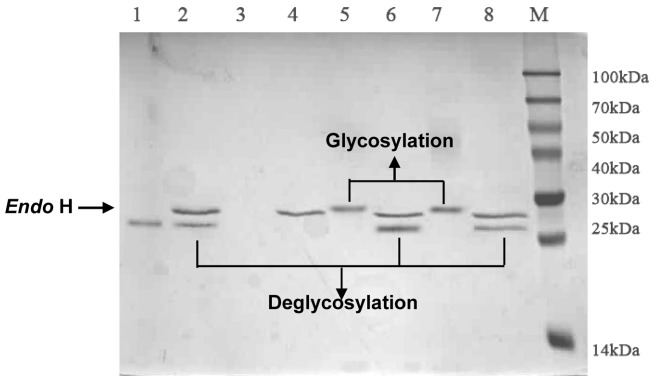

Fig. 4. Expression of mutants and deglycosylation results.

SDS-PAGE was used in the analysis of the wild-type and various mutant trypsins in culture supernatants from various strains. Lanes 1, 3, 5, and 7 show the expression of wild-type and *N*-glycosylation mutants (A73S, N84S, and R104S), respectively; Lanes 2, 4, 6, and 8 show the purified wild-type and mutants (A73S, N84S, and R104S) treated with *Endo* H and the deglycosylation results.

which indicated that the A73S variant was not expressed. After deglycosylation, the figure (Fig. 4, lane 4) just showed the Endo H enzyme. These results demonstrated that the introduction of an *N*-glycosylation site to N71 significantly inhibited trypsin expression.

Thermal Stability of *N*-Glycosylation

To determine the impact of *N*-glycosylation on trypsin stability, the kinetic stability of trypsin mutants was measured. Purified proteins were incubated at 50°C and the remaining activity was measured at various times. The results are shown in Figs. 5 and 6; the k_d of the trypsin

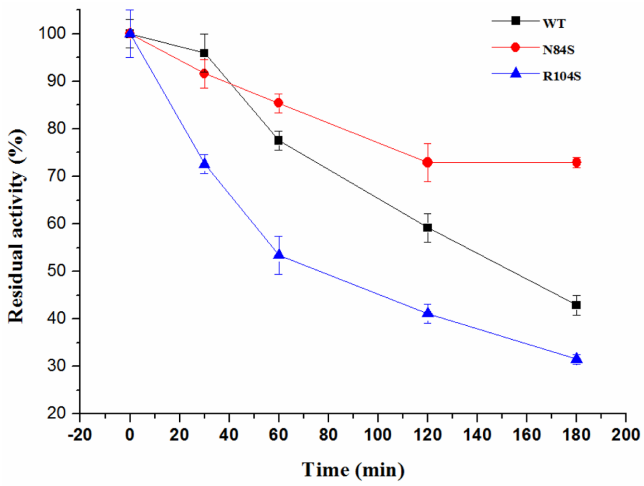

Fig. 5. Comparison of the thermal stability of the wild-type and mutant trypsins.

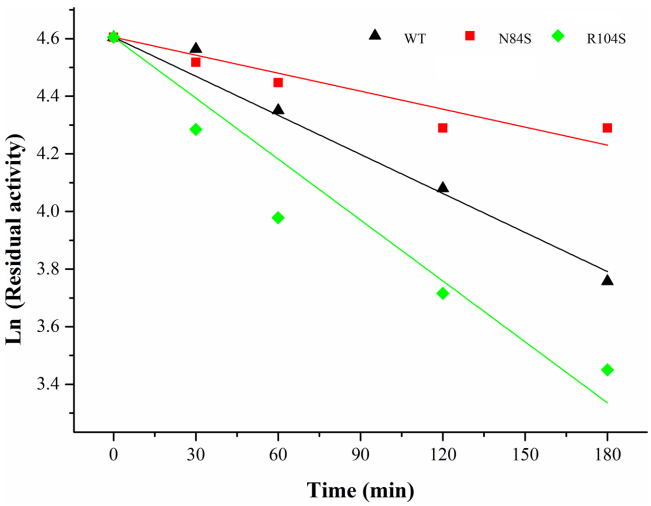

Fig. 6. Investigation of the half-life of the wild-type and mutant trypsins.

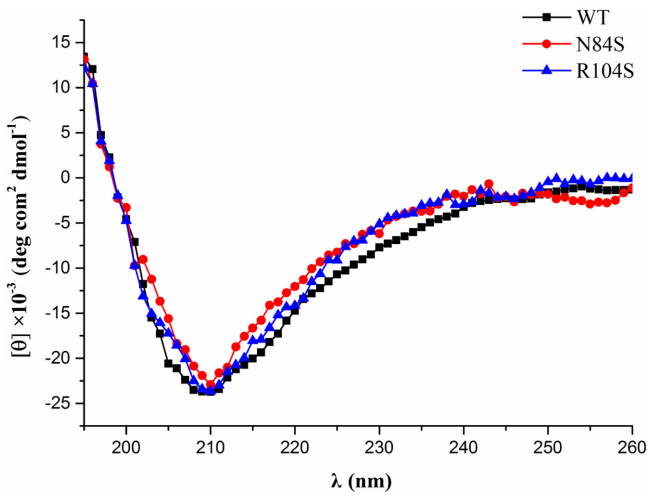

Fig. 7. Circular dichroism spectra of the wild-type and mutant trypsins at 25°C and at a concentration of 0.5 mg/ml.

mutant, N84S, was less than that of the wild type at 50°C, whereas the k_d of R104S was larger than that of the wild type, indicating that the wild-type trypsin had deactivated more rapidly than N84S and more slowly than R104S (Fig. 6). The mutant N84S retained 85% of its original activity after about 60 min at 50°C, whereas the wild-type and R104S retained residual activities of 77% and 53%, respectively (Fig. 5). The half-lives of the wild-type trypsin and the N84S mutant were 155.3 min and 333.2 min, respectively (Table 2). These results indicate that *N*-glycosylation at N82 contributes to the thermal stability of trypsin.

CD Spectroscopy Analysis

To determine the effect of *N*-glycosylation on recombinant porcine trypsin from *P. pastoris*, the secondary structures of recombinant trypsin mutants were analyzed by CD spectroscopy and compared with those of the wild type. As shown in Fig. 7, the spectrum of the wild-type trypsin showed a single minimum near 210 nm, indicating a β-sheet secondary structures, comparable to what has been reported for trypsin [12, 25]. Besides this, the average CD spectra of N84S and R104S were similar to that of the wild-type trypsin, suggesting that the mutations did not disrupt the secondary structure of trypsin and maintained a conformation close to the native protein.

In addition, to confirm the thermal effect on the conformational stability of mutant enzymes, their thermodynamic stability was measured. The proteins' melting temperature (T_m) was assayed using CD spectroscopy (Fig. 8). The thermal denaturation of proteins was detected by monitoring ellipticity changes at 215 nm over the temperature range of 25°C–95°C. The T_m value for the wild-type trypsin was 69.4°C, whereas the N84S and R104S mutants showed T_m values of 78.5°C and 66.0°C,

Table 2. Thermostability of the porcine trypsin mutants.

Trypsin	$t_{1/2}$ (min)	T_m (°C)	T_{opt} (°C)
Wild type	155.35	69.4	75
N84S	333.24	78.5	80
R104S	98.32	66.0	65

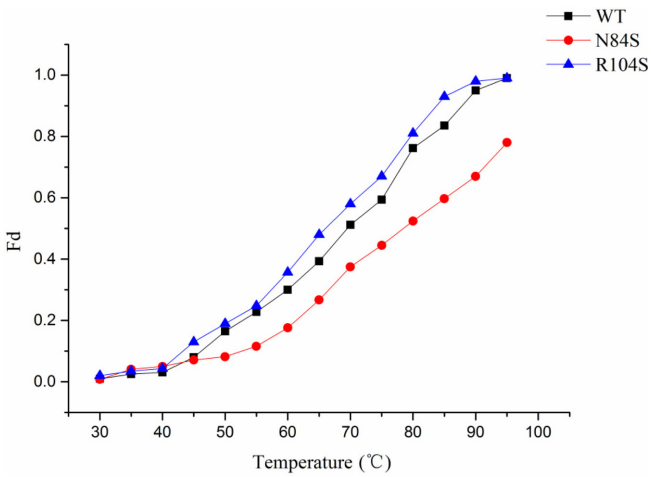

Fig. 8. Temperature-induced unfolding measured using CD spectroscopy for wild-type and mutant trypsins.

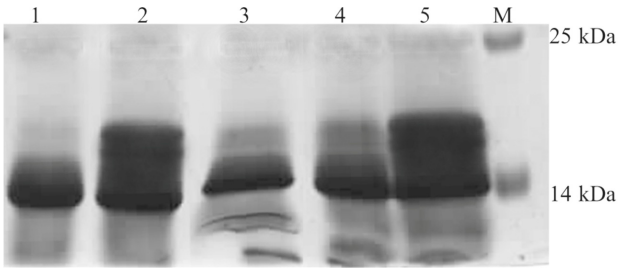

Fig. 9. Digestion specificity of *N*-glycosylation.
Both the trypsin mutants and wild-type enzyme digested the native ribonuclease A at 25°C. Lane 1, RNase A digested by native trypsin; Lanes 2–4, RNaseA digested by the A73S, N84S, and R104S mutant trypsins, respectively; Lane 5, RNase A.

respectively (Table 2). The results demonstrated that the N84S mutant displayed superior thermostability when compared with that of the wild type, whereas R104S did not show improved thermostability.

Digestion Specificity of *N*-Glycosylation

The digestion specificities of the wild-type and *N*-glycosylation mutants were measured using RNase A as the enzyme substrate at 37°C. All enzymes, except for A73S, indeed digested native RNase A (Fig. 9), indicating that introducing *N*-glycosylation within the native trypsin had no effect on enzyme digestion specificity.

Optimum Temperature for *N*-Glycosylation

The optimum temperature for *N*-glycosylation for the wild-type trypsin, R104S, and N84S was 75°C, 65°C, and 80°C, respectively (Table 2). Furthermore, the wild type showed higher activity at low temperatures (50°C–65°C) than N84S, whereas the latter was more active at higher temperatures (70°C–85°C). The N84S mutant enzyme still retained 86% of its activity when compared with the wild-type enzyme (which retained 45% of its activity) after incubation at 90°C (Fig. 10), which could be extensively used to digest native proteins at high temperature without the need for prior denaturation during mass spectrometry.

Discussion

The selection of mutation sites in enzymes or proteins is the most critical step in the rational design method [36]. Inappropriate positions can affect folding and destroy the structure of an active protein [37]. Therefore, two main selection principles have been considered: one was that the chosen mutation sites should be situated away from the active center to guarantee that the overall structure and

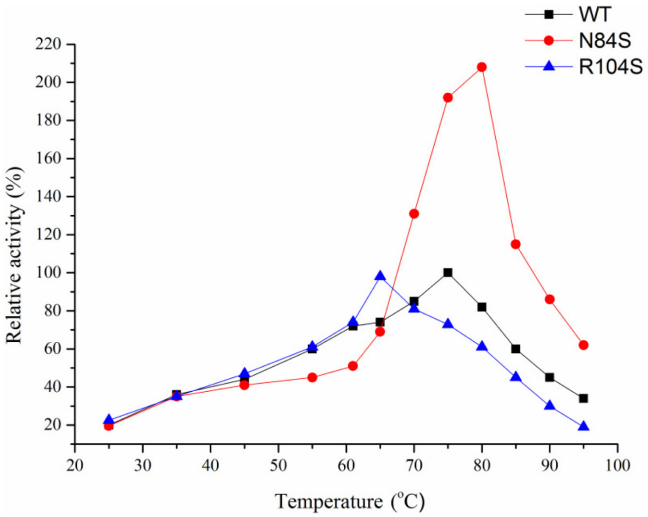

Fig. 10. Effect of temperature on enzyme activity.
The wild-type enzyme and mutants were incubated in 67 mmol/l of sodium phosphate (pH 7.6) at 25°C, 35°C, 45°C, 55°C, 60°C, 65°C, 70°C, 75°C, 80°C, 85°C, 90°C, and 95°C, respectively. Relative activity was shown against the wild-type's highest activity (100%).

catalytic activity of the target enzyme were less affected. The other principle was that the *N*-glycosylation sites would be introduced in a flexible region in the protein. In the present work, MD simulation predicted that two Ω-loops represented the flexible region within the protein structure (Fig. 2). Several research studies have indicated that the Ω-loop was the most susceptible part in protein denaturation [2, 23]; thus, rigidifying the Ω-loop region within trypsin may help to enhance its thermostability.

Glycosylation is one of the most naturally occurring modifications of a protein's covalent structure [34]. The effects of *N*-glycans on enzyme structure and activity have been well reported, due to their importance in biotechnological applications [4, 8, 26]. Several roles have been suggested for the carbohydrate moieties of proteins, one of which is the stabilization of protein conformation. Moreover, the great stabilization effect is achieved when the attached glycan is located at a more flexible region (*e.g.*, a loop) of the protein [29]. Therefore, glycosylation site-directed mutations (A73S, N84S and R104S) were introduced within the Ω-loop region and its wing chain to improve the stability of trypsin. The SDS-PAGE analysis indicated that the N84S and R104S mutants showed obviously increased molecular masses when compared with the wild type (increasing from 26 to 28 kDa). It has been reported that the *N*-glycan chains of the glycoproteins secreted by *P. pastoris* are a high-mannose type ($Man_8GlcNAc_2$ or $Man_9GlcNAc_2$)

Table 3. Protein concentration of trypsin and the activities of glycated trypsin derived from the hydrolysis of BAEE.

Trypsin	WT	A73S	N84S	R104S
Protein concentration (μg/ml)	364.0	–	602.3	777.3
Relative activity	1.00	–	0.98 ± 0.02	1.12 ± 0.04

with a molecular mass close to 2 kDa [5, 14, 20], which may explain why the molecular masses of the wild-type and mutant proteins are displayed at intervals of ~2 kDa.

Furthermore, several research studies have indicated that *N*-glycosylation could significantly influence enzyme activity and secretion [13, 16, 30]. The results reported herein demonstrated that the addition of *N*-glycosylation at the N82 and N102 positions of the porcine trypsin improved the trypsin production level when compared with the wild-type enzyme, but attachment of an *N*-glycan at N71 site would drastically inhibited protein expression (Table 3). *N*-Glycosylation is a common form of post-translational modification in *P. pastoris* and it frequently affects proper protein folding and secretion [3]. Misfolded polypeptides are subject to endoplasmic reticulum-associated degradation, ubiquitinated, and degraded by the 26S proteasome in the cytoplasm [3, 19]. In this paper, the addition of *N*-glycosylation at the N82 and N102 positions of the porcine trypsin increased protein expression levels, implying that there is a positive role of *N*-glycosylation in the proper folding of the enzyme [13]. In contrast, introduction of an *N*-glycosylation site to N71 may influence the proper translation or correct folding of the enzyme, indicating that trypsin was not produced, or that the misfolded protein might have been rapidly degraded. These results are consistent with those of a few studies, which reported that the attachment of the *N*-glycans of glycoproteins exerted negative effects on their secretion [16, 30]. Furthermore, the enzymatic activity and digestion specificity of mutants have been studied. Our data indicated that there were no decreases in activity for the glycated trypsin (Table 3 and Fig. 9). Owing to the fact that the *N*-glycosylation sites added in trypsin are located far away from the substrate-binding pocket, the glycoproteins might have no significant influence on substrate specificity.

In addition, the role of the *N*-glycosylation sites at N82 and N102 on protein stability was investigated by measuring the half-lives of their thermal inactivation ($t_{1/2}$) and melting temperature (T_m). The recombinant glycated trypsin (N84S) exhibited better thermostability (after being incubated at 50°C for 60 min, it still retained 85% of its original activity) than the wild-type enzyme (Fig. 5). Moreover, the CD spectra of the native and glycated enzymes at ambient temperature (Fig. 7) confirmed that *N*-glycosylation site-direced mutations did not disrupt the secondary structure of trypsin, and there were no decreases in activity for the mutant enzymes (Table 3 and Fig. 9). Hence, we may infer that the addition of *N*-glycosylation at the N82 position increases the protein stability, based on the fact that the protein melting temperature (T_m) of glycated protein was significantly increased compared with that of the wide-type trypsin (Fig. 8). Glycosylated forms of a protein are often less susceptible to thermal inactivation than their unglycosylated counterparts, probably because glycosylation renders the protein more resistant to unfolding [21, 32]. Pham *et al*. [25] also claimed that glycation did increase the resistance to unfolding at higher temperatures (as reflected in the increased T_m values), and this effect is magnified owing to the decreased rate of autolysis in the glycated enzymes. As for proteolytic enzymes such as trypsin and chyomotrypsin, protein unfolding at elevated temperatures results in increased exposure of the susceptible sites to autolytic cleavage, which prevents refolding of the denatured protein molecules [25]. The slight decoration of *N*-linked glycosylation modification on the N82 positions of the porcine trypsin would increase the protein's conformational stability, which in turn might result in a decreased rate of autolysis. Moreover, the *N*-linked glycan moiety is a negatively charged molecule with a large branched structure [17]; the thermostability of trypsin might be improved by hydrophilic suger components providing repulsive interactions, which may prevent the formation of large aggregates during heating at high temperatures.

Glycoengineering is a new approach that is currently being applied to improve the thermostability and activity of a given enzyme. In this study, a new strategy to improve the thermostability of trypsin was proposed. MD simulation was used to investigate the molecular stability of trypsin, and oligosaccharides were added within the flexible region of the protein through native glycosylation modified by *P. pastoris* GS115. Based on this strategy, we successfully improved the thermostability of trypsin without altering its activity. This strategy provided a referential method through which to engineer the thermostability of other industrial enzymes; it also has the potential for therapeutic application.

Acknowledgments

The present work was partially supported by the National Natural Science Foundation of China (Grant No. 31470967) and National Science and Technology Major Projects of China (Grant Nos. 2011ZX09201-301-05 and 2014ZX09508006-002-002).

References

1. Badieyan S, Bevan DR, Zhang C. 2012. Study and design of stability in GH5 cellulases. *Biotechnol. Bioeng.* **109:** 31-44.
2. Bai Y, Sosnick TR, Mayne L, Englander SW. 1995. Protein folding intermediates: native-state hydrogen exchange. *Science* **269:** 192-197.
3. Çelik E, Çalık P. 2012. Production of recombinant proteins by yeast cells. *Biotechnol. Adv.* **30:** 1108-1118.
4. Chen Y, Chen X, Guo TL, Zhou P. 2015. Improving the thermostability of β-lactoglobulin via glycation: the effect of sugar structures. *Food Res. Int.* **69:** 106-113.
5. Daly R, Hearn MT. 2005. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J. Mol. Recognit.* **18:** 119-138.
6. Deng Z, Yang H, Shin H-D, Li J, Liu L. 2014. Structure-based rational design and introduction of arginines on the surface of an alkaline α-amylase from *Alkalimonas amylolytica* for improved thermostability. *Appl. Microbiol. Biotechnol.* **98:** 8937-8945.
7. Fang Z, Zhou P, Chang F, Yin Q, Fang W, Yuan J, *et al.* 2014. Structure-based rational design to enhance the solubility and thermostability of a bacterial laccase Lac15. *PLoS One* **9:** 1-6.
8. Fatima A, Husain Q. 2007. A role of glycosyl moieties in the stabilization of bitter gourd (*Momordica charantia*) peroxidase. *Int. J. Biol. Macromol.* **41:** 56-63.
9. Fei B, Xu H, Cao Y, Ma S, Guo H, Song T, *et al.* 2013. A multi-factors rational design strategy for enhancing the thermostability of *Escherichia coli* AppA phytase. *J. Ind. Microbiol. Biotechnol.* **40:** 457-464.
10. Fetrow JS. 1995. Omega loops: nonregular secondary structures significant in protein function and stability. *FASEB J.* **9:** 708-717.
11. Finehout EJ, Cantor JR, Lee KH. 2005. Kinetic characterization of sequencing grade modified trypsin. *Proteomics* **5:** 2319-2321.
12. Gizurarson JG, Filippusson H. 2015. Conjugation of D-glucosamine to bovine trypsin increases thermal stability and alters functional properties. *Enzyme Microb. Technol.* **75-76:** 1-9.
13. Han M, Wang W, Jiang G, Wang X, Liu X, Cao H, *et al.* 2014. Enhanced expression of recombinant elastase in *Pichia pastoris* through addition of *N*-glycosylation sites to the propeptide. *Biotechnol. Lett.* **36:** 2467-2471.
14. Han M, Wang X, Yan G, Wang W, Tao Y, Liu X, *et al.* 2014. Modification of recombinant elastase expressed in *Pichia pastoris* by introduction of *N*-glycosylation sites. *J. Biotechnol.* **171:** 3-7.
15. Hoedemaeker FJ, van Eijsden RR, Díaz CL, de Pater BS, Kijne JW. 1993. Destabilization of pea lectin by substitution of a single amino acid in a surface loop. *Plant Mol. Biol.* **22:** 1039-1046.
16. Hoshida H, Fujita T, Cha-aim K, Akada R. 2013. *N*-Glycosylation deficiency enhanced heterologous production of a *Bacillus licheniformis* thermostable α-amylase in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **97:** 5473-5482.
17. Hossler P, Khattak SF, Li ZJ. 2009. Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* **19:** 936-949.
18. Kemmler W, Peterson JD, Steiner DF. 1971. Studies on the conversion of proinsulin to insulin I. Conversion in vitro with trypsin and carboxypeptidase B. *J. Biol. Chem.* **246:** 6786-6791.
19. Kostova Z, Wolf DH. 2003. For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin–proteasome connection. *EMBO J.* **22:** 2309-2317.
20. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* **22:** 249-270.
21. Meldgaard M. 1994. Different effects of *N*-glycosylation on the thermostability of highly homologous bacterial (1, 3-1, 4)-β-glucanases secreted from yeast. *Microbiology* **140:** 159-166.
22. Miyazaki T, Yashiro H, Nishikawa A, Tonozuka T. 2014. The side chain of a glycosylated asparagine residue is important for the stability of isopullulanase. *J. Biochem.* **157:** 225-234.
23. Obuchi M, Kawahara K, Motooka D, Nakamura S, Yamanaka M, Takeda T, *et al.* 2009. Hyperstability and crystal structure of cytochrome c555 from hyperthermophilic *Aquifex aeolicus*. *Acta Crystallogr. D Biol. Crystallogr.* **65:** 804-813.
24. Parge HE, Hallewell RA, Tainer JA. 1992. Atomic structures of wild-type and thermostable mutant recombinant human Cu,Zn superoxide dismutase. *Proc. Natl. Acad. Sci. USA* **89:** 6109-6113.
25. Pham VT, Ewing E, Kaplan H, Choma C, Hefford MA. 2008. Glycation improves the thermostability of trypsin and chymotrypsin. *Biotechnol. Bioeng.* **101:** 452-459.
26. Samoudi M, Tabandeh F, Minuchehr Z, Cohan RA, Inanlou DN, Khodabandeh M, Anvar MS. 2015. Rational design of hyper-glycosylated interferon beta analogs: a computational strategy for glycoengineering. *J. Mol. Graph. Model.* **56:** 31-42.
27. Schweiker KL, Makhatadze GI. 2009. Protein stabilization by the rational design of surface charge–charge interactions. *Methods Mol. Biol.* **490:** 261-283.
28. Scott WR, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, *et al.* 1999. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **103:** 3596-3607.

29. Shental-Bechor D, Levy Y. 2009. Folding of glycoproteins: toward understanding the biophysics of the glycosylation code. *Curr. Opin. Struct. Biol.* **19:** 524-533.
30. Skropeta D. 2009. The effect of individual *N*-glycans on enzyme activity. *Bioorg. Med. Chem.* **17:** 2645-2653.
31. Thangudu RR, Vinayagam A, Pugalenthi G, Manonmani A, Offmann B, Sowdhamini R. 2005. Native and modeled disulfide bonds in proteins: knowledge-based approaches toward structure prediction of disulfide-rich polypeptides. *Proteins* **58:** 866-879.
32. Tull D, Gottschalk TE, Svendsen I, Kramhøft B, Phillipson BA, Bisgård-Frantzen H, *et al.* 2001. Extensive *N*-glycosylation reduces the thermal stability of a recombinant alkalophilic *Bacillus* α-amylase produced in *Pichia pastoris*. *Protein Expr. Purif.* **21:** 13-23.
33. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. 2005. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26:** 1701-1718.
34. Wildt S, Gerngross TU. 2005. The humanization of *N*-glycosylation pathways in yeast. *Nat. Rev. Microbiol.* **3:** 119-128.
35. Xi H, Tian Y, Zhou N, Zhou Z, Shen W. 2015. Characterization of an *N*-glycosylated *Bacillus subtilis* leucine aminopeptidase expressed in *Pichia pastoris*. *J. Basic Microbiol.* **55:** 236-246.
36. Yu H, Huang H. 2014. Engineering proteins for thermostability through rigidifying flexible sites. *Biotechnol. Adv.* **32:** 308-315.
37. Yu H, Zhao Y, Guo C, Gan Y, Huang H. 2015. The role of proline substitutions within flexible regions on thermostability of luciferase. *Biochim. Biophys. Acta* **1854:** 65-72.
38. Zheng L, Baumann U, Reymond J-L. 2004. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32:** e115.