International Journal of Biological Macromolecules 126 (2019) 229–237

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: http://www.elsevier.com/locate/ijbiomac

Engineering of serine protease for improved thermostability and catalytic activity using rational design

Naeem Mahmood Ashraf [a,1], Akshaya Krishnagopal [b,1], Aadil Hussain [a], David Kastner [c], Ahmed Mahmoud Mohammed Sayed [b], Yu-Keung Mok [b], Kunchithapadam Swaminathan [b,*], Nadia Zeeshan [a,*]

[a] Department of Biochemistry and Biotechnology, University of Gujrat, Hafiz Hayat Campus, Gujrat, Punjab 50700, Pakistan
[b] Department of Biological Sciences, National University of Science, 117543, Singapore
[c] Department of Biophysics, Brigham Young University, Provo, UT 84602, USA

ARTICLE INFO

Article history:
Received 19 October 2018
Received in revised form 15 December 2018
Accepted 22 December 2018
Available online 24 December 2018

Keywords:
Pseudomonas aeruginosa
Serine protease
Protein engineering
Site-directed mutagenesis
Thermostability
Catalytic efficiency

ABSTRACT

The study involves the isolation and characterization of a serine peptidase, named SP, from *Pseudomonas aeruginosa*. In addition to basic characterization, the protein was engineered, by site-directed mutagenesis of selected non-catalytic residues, to increase its thermal stability and catalytic activity. Among the eight-point mutations, predicted by FireProt, two mutants, A29G and V336I, yielded a positive impact. The T_m of A29G and V336I showed an increase by 5 °C and also a substantial increase in residual activity of the enzyme at elevated temperature. Moreover, the catalytic activity of A29G and V336I also showed an increase of 1.4-fold activity, compared to the wild-type (WT). Moreover, molecular docking simulations also predicted better substrate affinity of the mutants. We have also performed molecular dynamics (MD) simulations at 315 and 345 K, and the MD data at 345 K demonstrates improved thermostability for the mutants, compared to the WT. Our findings not only contribute to a better understanding of the structure-stability-activity relationship of SP but also highlights, that modification of non-catalytic residues could also promote favourable catalytic behaviour.

1. Introduction

Proteases represent one of the principal groups of industrial enzymes [1]. Among them, serine proteases of microbial origin are well studied and possess substantial industrial potential due to their biochemical diversity and extensive applications in the silk, tannery and food industries, medicinal formulations, silver recovery, detergents and waste treatment [2]. Serine proteases use the classical Ser/His/Asp catalytic triad to break a peptide bond, where the serine is the nucleophile, histidine being the general base and acid, and aspartate that helps orient the histidine residue and neutralizes the charge taken by the histidine during transition states [3].

High thermostability and catalytic efficiency are highly required hallmark properties of proteases. Use of engineered proteases for therapeutic applications has been an ongoing objective, and several protease therapies are currently under clinical investigation [4]. Several wild-type (WT) enzymes are not readily usable in the industrial process due to their limitations under harsh industrial conditions, particularly temperature [5], an important factor to determine an enzyme's catalytic efficiency [6]. Moreover, thermostable enzymes also withstand extreme conditions [7] and can facilitate cost-effective strategies to large-scale applications in the industry [8,9]. Protein engineering has provided effective ways to address these problems [10] and modulate protease stability to achieve desired requirements. There are three main protein engineering approaches, namely, directed evolution, rational design, and de novo design. The use of rational design has been the preferred choice for enzyme engineering due to accelerated process time and smaller mutant libraries. Furthermore, studies have reiterated the efficiency of this approach for augmenting desirable characteristics through in-depth analysis of structures, aiming at improved catalytic efficiency and thermostability of enzymes [11,12]. Correspondingly, this approach is proven to produce cheaper and more effective enzymes for large-scale industrial use [13–15]. Serine alkaline protease of *Bacillus pumilus* (SAPB) was rationally modified for its increased thermal stability and catalytic activity by constructing a single mutant N99Y [16]. Similarly, it was observed that uncharged amino acid substitution (L31 and T33) resulted in increased substrate specificity and catalytic activity of SABP. Notably, these mutants were also recommended as potential candidates for industrial use in leather treatment industry [17]. Identical findings to the above were also reported in a study by Takagi, H et al. [18] on an alkaline serine protease, *Bacillus subtilis*, in which 2–6 fold

* Corresponding authors.
E-mail addresses: dbsks@nus.edu.sg (K. Swaminathan), nadia.zeeshan@uog.edu.pk (N. Zeeshan).
[1] Contributed equally.

https://doi.org/10.1016/j.ijbiomac.2018.12.218
0141-8130/

catalytic efficiency was achieved. Another study also reported a modification of *Bacillus pumilus* bacterial strain for efficient production of alkaline proteases using a similar screening and mutagenesis method. [19]

We have recently isolated a serine endoprotease (SP) from *Pseudomonas aeruginosa* strain BMB1. In addition to preliminary characterization of this enzyme, we investigated its thermostability and catalytic efficiency using a protein engineering approach by employing site-directed mutagenesis of selective non-catalytic residues. We have also performed molecular dynamics (MD) simulations on the WT and mutants to grasp insights into the molecular basis of its thermostability.

2. Materials and methods

2.1. Isolation and cloning of SP

A strain of *Pseudomonas aeruginosa* was isolated from the soil samples of the industrial leather waste and grown in minimal salt medium (MSM). Proteolytic activity was checked in the supernatant by using casein as a substrate to confirm the extra-cellular expression of the protease. The full-length 16S rRNA gene was amplified and sequenced for phylogenetic identification of the bacterial strain. The phylogeny.fr online server [20] was used to construct a phylogenetic tree. The *Pseudomonas aeruginosa* reference strain PA14 serine protease gene was used to design primers for full-length gene sequencing from the highly conserved region using the CEMA suite [21]. The gene SP was amplified by polymerase chain reaction (PCR), using suitable primers (TTTTTT GGATCCATGCATACCCTAAAACGCTG and AAAAAAGCGGCCGCTTATTCG GCCAGCTTGAA). *In-silico* characterization was carried out to gain a better understanding of the protease structure. The physical and chemical parameters of the protein were computed using the ProtParam sever [22] while the presence of the signal peptide at the N-terminal was predicted using SignalP4.0 [23]. InterPro server was employed for identification of domains present in the protease, and lastly, the 3D structure was modelled using I-TASSER, with 1KY9, 3STJ, 3PV2, 2ZLE, and 3MH6 PDB structures as templates [24,25]. The SP gene expressing full length (474 aa) protein, fused with an N-terminal 6× His-tag, was cloned into a modified pET-32a vector, between the BamH1 and Not1 restriction sites. Modified pET-M vector, is pET-32a(+) vector without the Trx•Tag, S•Tag, unlike the original pET-32a. The His-tag was left for Ni2+ -NTA affinity purification and the plasmid conferred ampicillin resistance on transformed BL21 cells. WT SP was cloned in a pET-32a vector, with and without the signal peptide, along with a 6× His tag.

2.2. Protein expression and purification

The confirmed plasmid was transformed into competent *E. coli* BL21 (DE3) cells and grown in Luria-Bertani (LB) medium containing 100 μg/ml ampicillin (Gold Biotechnology) at 37 °C until it reached an OD_{600} of ~0.8, following which protein expression was induced using 500 μM IPTG (Gold Biotechnology) at 37 °C for 16 h. The cells were pelleted at 5000 g for 30 min, washed with 1× phosphate buffered saline (PBS) and stored at −20 °C, until needed. The cell pellet was resuspended in lysis buffer (50 mM Tris, pH 7.5, 300 mM NaCl, 5% glycerol) with 5 mM imidazole (Sigma-Aldrich) and 2 mM DTT (Gold Biotechnology) and sonicated and the lysate was subjected to high-speed centrifugation to separate the soluble supernatant and insoluble cell debris. The recombinant protein was purified using nickel-NTA beads (Merck) and size exclusion chromatography (HiLoad Superdex 200). SDS-PAGE was used to analyze the molecular weight and purity of eluted fractions. Furthermore, the molecular mass of the purified protein was confirmed using matrix-assisted laser desorption ionization–two-stage time of flight (MALDI-TOF/TOF) system (5800 MALDI-TOF/TOF mass spectrometer, ABSciex) and in-house Mascot software for database searching and protein identification.

2.3. Preparation and characterization of stable mutants

Using the rational design approach, we identified stabilizing mutations in SP that would increase its thermostability. In the first step, the protein sequence was submitted to the I-TASSER server to predict a 3D structure of the protein [25]. The most reliable model, on the basis of the C-score, was selected and the predicted model was refined to reduce unwanted side chain clashes using the online ModRef server [26]. The refined model was submitted to the FireProt server to predict stabilizing mutations via evolution-and energy-based approaches [14]. The selected point mutations were incorporated in the WT construct, and the corresponding proteins were over-expressed. Possible changes in the secondary structure of the mutants were verified using far-UV circular dichroism (CD) spectra.

2.4. Circular dichroism (CD) and thermal denaturation analysis

10–15 μM of recombinant WT or mutant SP was prepared in 20 mM Tris (pH 7.5), 5% glycerol and 150 mM NaCl buffer. A Jasco J-1100 spectrophotometer, equipped with a Jasco MCB-100 temperature controller (1 mM quartz cuvette, scanning rate 20 nm/min, 1 nm bandwidth and 1 nm data pitch) was used to obtain far-UV spectra from 190 to 260 nm. The CD spectra were plotted using the online CD data analysis tool CAPITO and curves were made smooth using the Savitzky-Golay filter [27]. Thermal unfolding was monitored at 222 nm (representing the signal from α-helices) in a temperature scan mode starting from 20 to 95 °C with a scanning rate of 1 °C min^{-1}. The recorded denaturation data of the wild-type and mutant enzymes were fitted to a sigmoidal curve using the Boltzmann equation [28]. The T_m values of the proteins were assigned as the midpoint of normalized thermal denaturation data. The CDpal software was used to fit the data sets to calculated ΔH_m values [29].

2.5. Enzymatic assay

The catalytic activity of the purified recombinant WT enzyme was assayed using casein as a substrate as described in [30], with minor modifications. The optimum temperature was identified by measuring the caseinolytic activity in 100 mM Tris-HCl at 30–80 °C. Similarly, the optimum pH was determined in the pH range 5–10 at the optimum temperature. The residual activities of the wild-type and mutants were identified by incubating 5 μM enzyme at different temperatures for 120 min. The WT and mutant enzymes were assayed for their kinetic parameters, with different concentrations of the substrate (0.1 to 2 mg/ml) for 10 min at the optimum pH and temperature. The kinetic parameters (K_m, V_{max} and K_{cat}) were calculated using the Michaelis-Menten equation.

2.6. Molecular dynamics (MD) simulation

The protein was enclosed in a water box and was neutralized with 0.15 mol/L salt concentration. The system was then prepared with an energy minimization followed by a 1 ns equilibration run using the NAMD software via the QwikMD interface [31,32]. For both the equilibration and production phases, the system was run at either 315 or 345 K. To evaluate the thermostability of the mutants, relative to the WT, we performed MD simulations using the NAMD v2.12 simulator with the CHARMM36 forcefield and the TIP3P water model [33–36]. The production simulations were run for 30 ns each, to give time for adequate sampling of the energy landscape. The non-bonded interactions had a 12 Å cutoff. Long-range electrostatic interactions used the particle mesh Ewald (PME) method and the equations of motion were integrated with 2 fs time step using the multiple time step method r-RESPA [37]. The analysis of the MD simulations was performed for production runs using the VMD software [38]. The frames of the production trajectories were aligned relative to the centre-of-mass coordinates and

the root means square deviation (RMSD) was calculated with the initial equilibrated structures as the reference. The flexibility based on the RMSF was calculated using VMD plugin vmdICE [39].

2.7. Geometry based molecular docking

To predict the binding affinities of the ligand with wild-type and the two mutants, docking algorithm Patchdock, developed by Schneidman-Duhovny et al. was employed [40]. Casein, used during experimental assays was used as a ligand in the docking analysis. To develop the 3D structure of the ligand, SMILES string was obtained from PubChem database and submitted to CORINA to develop a 3D structure for docking [41]. The 3D structures of wild-type, mutants and the ligand were used as input in Patchdock server. The server predicted potential binding confirmations of the enzymes and ligand and ranked them according to their atomic contact energy (ACE) to predict best binding confirmation [42,43].

3. Results

3.1. SP isolation from bacterial strain

A *Pseudomonas aeruginosa* strain was isolated from the soil samples of industrial leather waste. The bacterial strain was selected on the basis of high caseinolytic activity. The phylogenetic tree (Fig. 1A) reveals that the selected bacterial strain is closely related to *Pseudomonas aeruginosa* NBRC and the 16S rRNA sequence has been submitted to the NCBI database under the accession number KY285994.

NCBI Blast-p results show the highest identity with other HtrA-like proteases of *Pseudomonas*. The ProtParam server [22] deduced the molecular weight of SP as 50 kDa and PI as 7.04. Moreover, a 25 amino acid long signal peptide (MHTLKRCMAAMVALLALSLAMTARA) was identified at N-terminal, indicating possible extracellular expression of the enzyme [23]. Three different domains were observed in the enzyme, namely the N-terminal trypsin-like catalytic domain and two of the C-terminal PDZ domains (PDZ1 and PDZ2). Based on the characteristic domain architecture, SP was classified under the S1C family of the PA clan of proteases, according to MEROPS [24,44]. In addition, a 3D structure of the serine protease was also modelled [25] (Fig. 1) with its fold deduced (Fig. S1B) and the predicted 3D model showed the characteristic β-barrel trypsin domain. It was evident from the 3D structure, that there was a low percentage of helices (5.2%) and high content of β-strands (35.1%). This observation was analogous to the previously reported secondary structures of a similar class of proteases [45,46]. Moreover, His89, Asp119 and Ser192 make a relay centre in the enzyme and are involved in catalytic activity. It is known that DegP, DegS and DegQ are the best characterized HtrA serine proteases in bacteria [47]. Notably, our

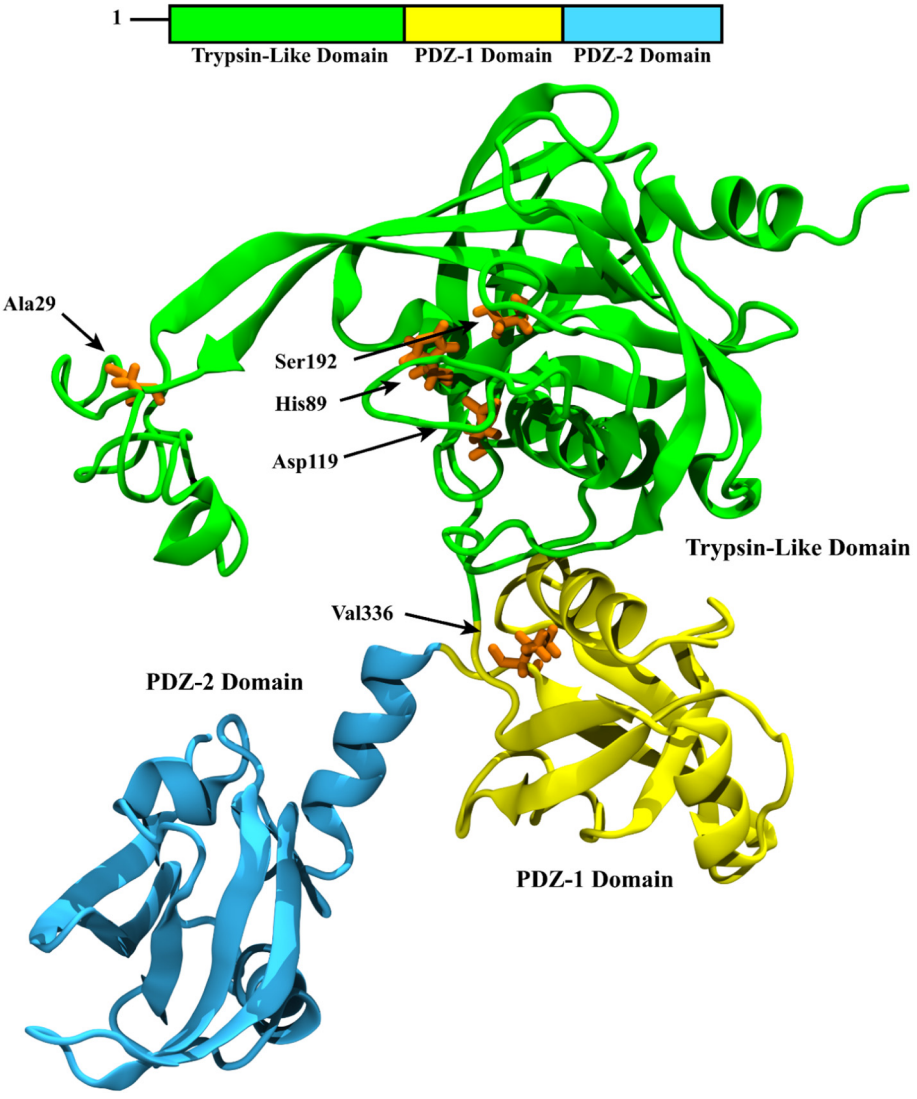

Fig. 1. The 3D structure of SP, modelled by I-TASSER. In the trypsin-like domain (green), the catalytic triad residues His89, Asp119 and Ser192 that form the charge relay centre are shown. The two PDZ domains are also labelled.

protease exhibits high levels of structural similarity in the proteolytic domain with the proteases mentioned above (β-barrel trypsin domain) Additionally, SP appears to be more associated to DegP in terms, of the presence of two C-terminal PDZ domains. Based on these structural resemblances, we could possibly speculate that our protease forms functional oligomers as it is similar to DegP [48].

3.2. Experimental characterization of SP

The construct with the signal peptide resulted in aggregates while the construct lacking the signal peptide resulted in soluble enzyme expression. The experimental molecular weight, western blot and mass spectrometry data also confirmed the presence of SP in the purified fractions (Fig. S1).

Our enzyme shows maximum activity at a temperature around 60 °C and pH around 7.5 by using casein as substrate (Fig. 2). We also pursued circular dichroism (CD) spectroscopy at the far-UV (190 to 260 nm) region to verify correlation coefficients for the respective secondary structure elements. As predicted, our results (Fig. 3A) show a relatively higher proportion of β-strands, a characteristic output of the far-UV CD spectrum similar to all previously reported the similar type of serine proteases [45,46]. The residual activity of the enzyme was also monitored. 5 μM SP was pre-incubated at 60 and 75 °C and the caseinolytic activity was monitored under the optimum condition. The enzyme lost 54% and 85% of the activity after incubation at 60 °C for 10 and 30 min, respectively. Moreover, the enzyme lost 91% of activity after incubation at 75 °C for 10 min (Fig. 4), indicating the low thermal stability and residual activity of wild-type serine protease.

3.3. Construction of more thermostable and more active SP mutants

Rational protein design was employed to identify stabilizing mutations in SP that would confer higher thermostability and catalytic activity. The predicted 3D model was submitted to the FireProt server to predict stabilizing mutations via evolution- and energy-based approaches [14]. A total of eight mutations (by energy-based approach: A29G, A114M, V394Y, V430M and by evolution-based approach: L51F, F211Y, L230M, V336I) were selected (Tables S1 and S2).

3.4. A29G and V336I show higher thermal stability

All eight mutants were generated and checked experimentally for their impact on the protein structure. The mutants were characterized for their optimum temperature and pH using casein as substrate. Except for two mutants, A29G and V336I, all other mutants showed negligible proteolytic activity at different temperatures and pHs. CD spectroscopy at the far-UV (190 to 260 nm) was measured to observe the difference between the WT and mutant secondary structures. There were substantial differences noted between the secondary structures of the WT and six mutants L51F, A114M, F211Y, L230M, V394Y, and V430M, indicating the destabilizing impact of the mutation on the protein structure (Fig. S2). This data was consistent with the biochemical characterization of the mutants. Only two non-catalytic residue mutants out of the eight,

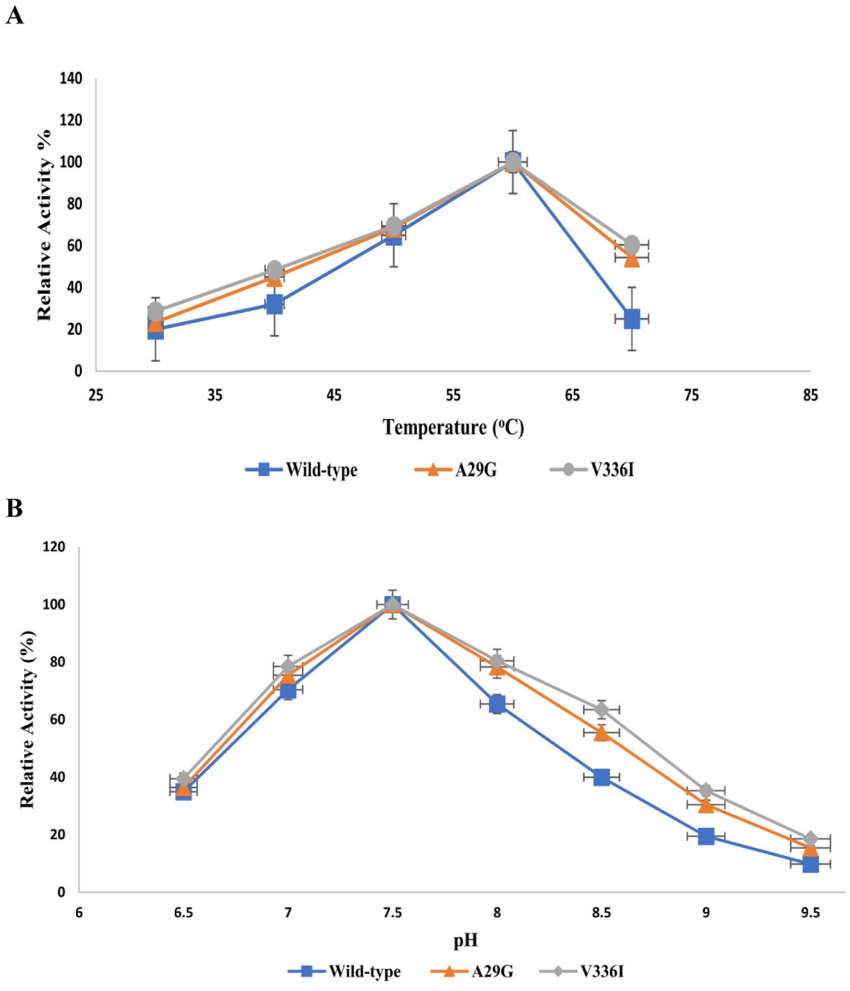

Fig. 2. The optimum (A) temperature and (B) pH of the WT and two mutants of SP (in 100 mM Tris buffer and casein as substrate). The highest activity is taken as 100%. The results are the average of three replicate experiments.

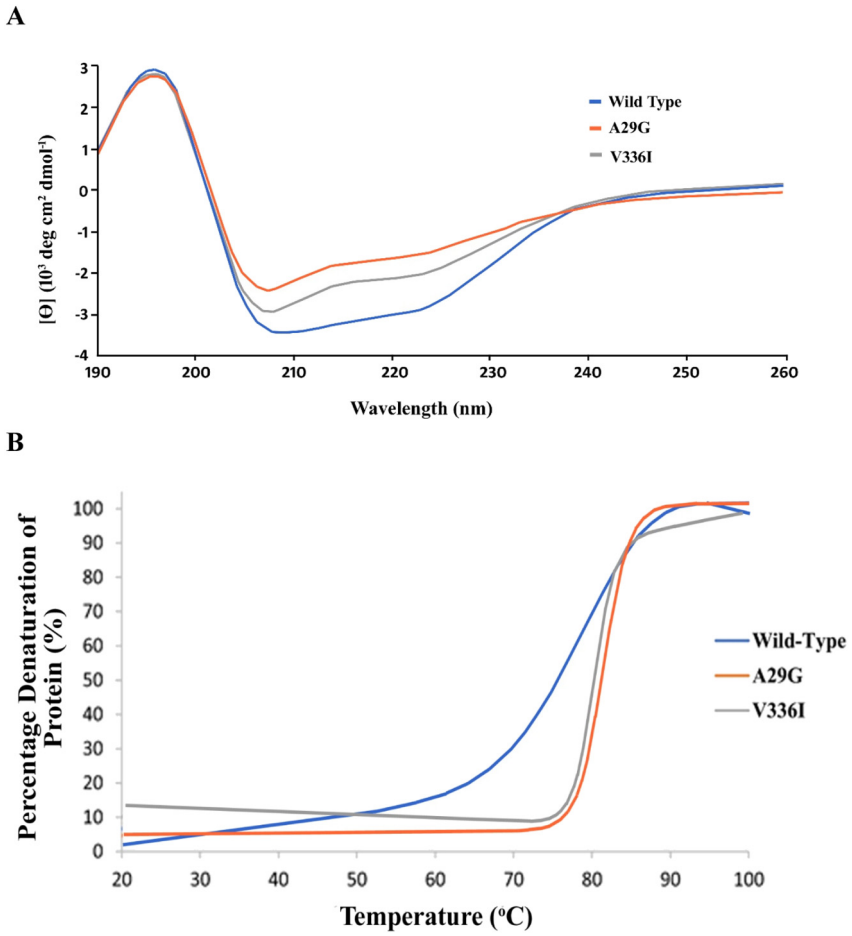

Fig. 3. (A) The far-UV (190–260 nm) CD spectrum of the WT and A29G and V336I mutants of SP, using 10 μM protein at 25 °C. (B) The percentage denaturation of SP with an increase in temperature. A29G and V336I show higher T_m, compared to the WT. The results are the average of three replicate experiments.

A29G and V336I, showed secondary structures comparable to that of the WT (Fig. 3A) and were selected for further downstream thermostability analysis.

Furthermore, the thermostability of the WT and mutants A29G and V336I was assessed using far-UV CD measurements, Fig. 3B. Both mutants, A29G and V336I, showed considerable thermodynamic stability, compared to the WT. The T_m of the WT, A29G and V336I are 71.9, 77.5 and 76.6 °C, respectively. Moreover, values for the enthalpy change (ΔH_m) during transition also gave insights about the thermostability of the mutants. The wild-type illustrated a gentle CD thermal transition curve, compared to the mutants. The gentle transition curve is associated with a small ΔH_m value, indicating that the wild-type enzyme exhibits less thermostability, whereas both mutants indicate large ΔH_m values, suggesting an enhanced thermal resistance (Table 1).

3.5. A29G and V336I show higher activity

The optimum enzyme activity conditions and residual activity were determined for mutants A29G and V336I, Fig. 2. The mutants also show maximum catalytic activity at the same WT optimum conditions (60 °C and pH 7.5). There was a decrease in the activity above and below the optimum temperature and pH.

Residual activity of the mutants was also monitored to determine their thermal resistance. The caseinolytic activities of the mutants were reduced after pre-incubation for 2 h at 60, 70 and 75 °C (Fig. 4). Compared to the WT, both mutants exhibit higher thermostability and this residual activity data is consistent with thermal denaturation CD data where both mutants exhibited better thermal resistance compared to the WT.

3.6. Kinetic parameters of the WT and mutant SP proteins

The Michaelis-Menten equation was used to characterize the kinetics of the WT and mutants. At 5 μM, the WT and mutants were incubated with different concentrations of casein at the optimum temperature and pH. The Microsoft Excel Solver was used to calculate different kinetic parameters (Table 2). The Km values for the two mutants indicate a higher affinity with substrate compared to the WT. Apparently, A29G and V336I had little effect on the turn over number (Kcat) but Kcat/Km values of both mutants showed approximately a 1.4-fold increase.

3.7. Geometry based molecular docking show better affinity of substrate with the mutants

The affinities of the ligand and the enzymes were also predicted using molecular docking simulations. Patchdock server calculated the ACE scores of all possible confirmations of wild-type and the mutants. Both mutants showed lower ACE scores (−652.99 kJ/mol) as compared to the wild-type (−559.37 kJ/mol). These results indicate better binding of the ligand with both mutants, compared to wild-type (Fig. S7). Molecular docking simulation results are consistent with experimentally calculated Km values.

3.8. Molecular dynamics simulation confirms higher thermostability for A29G and V336I

We used molecular dynamics (MD) simulations to confirm our experimental observations of thermostability and the conformational

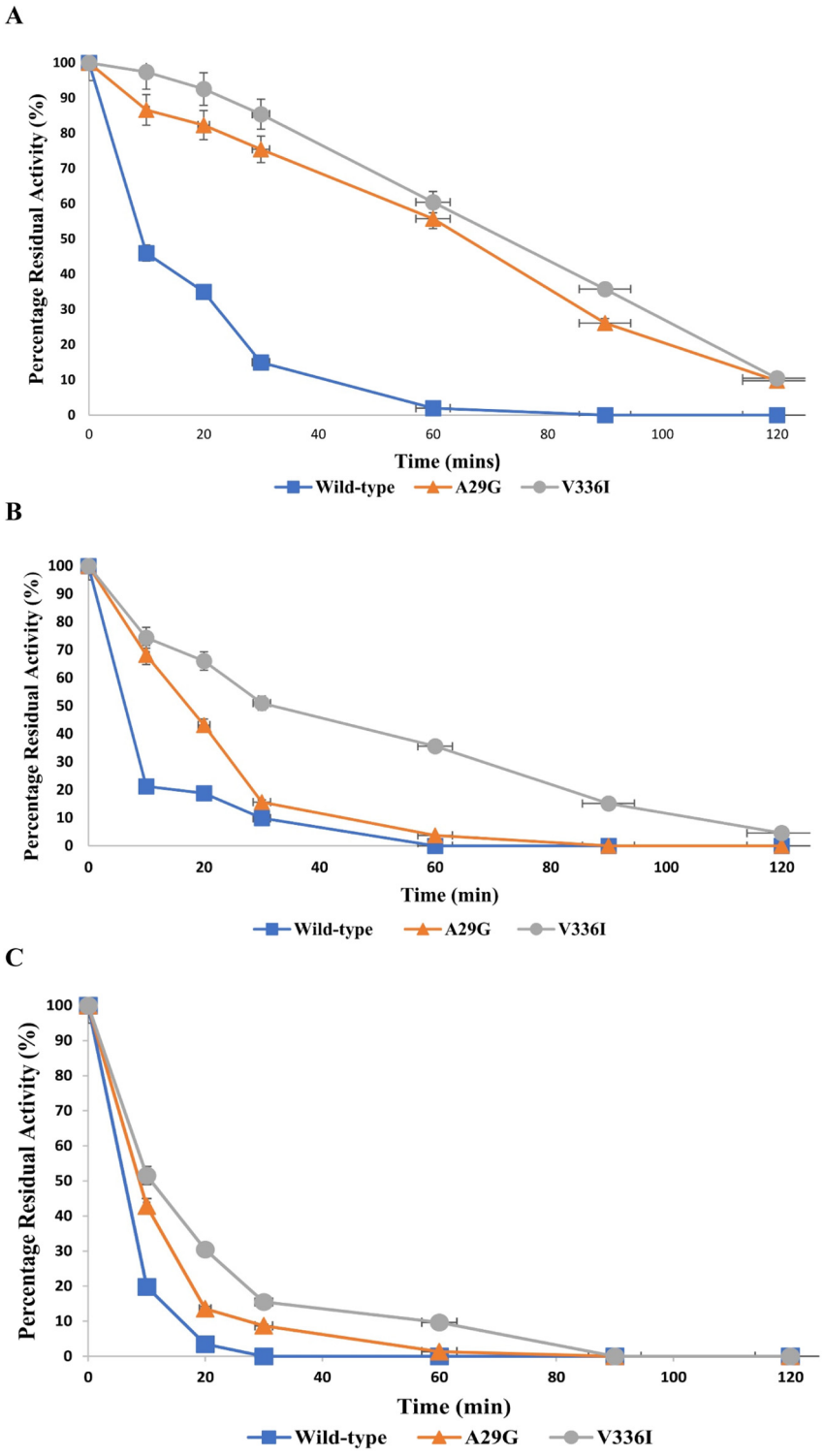

Fig. 4. Percentage residual activity of the WT and two mutants A29G and V336I after incubation for 120 min at (A) 60, (B) 70 and (C) 75 °C. The results are the average of three replicate experiments.

Table 1
Enthalpy changes (ΔH_m) of wild-type, A29G and V336I mutants along with the T_m values.

Enzyme	T_m (°C)	ΔH_m Kj/mol
Wil-type	71.9	198.87
A29G	77.5	798.74
V336I	76.6	657.29

Table 2
Kinetic parameters of the WT and mutants SP, treated at 60 °C and pH 7.5 for 10 min, using casein as substrate (0.1 to 2 mg/ml). The results are the average of three replicate experiments.

Enzyme	K_m (mg/ml)	V_{max} (μM/min ml)	K_{cat} (s^{-1})	K_{cat}/K_m (μL/s/mg)
Wild-type	0.90 ± 0.1	221	0.73	0.81
A29G	0.60 ± 0.1	204	0.67	1.11
V336I	0.56 ± 0.1	198	0.65	1.15

behaviour of the mutants A29G and V336I, compared to that of the WT. The MD simulations were conducted at two different temperatures, 315 and 345 K. The RMSD values of the WT and mutant backbone atoms were compared. At 315 K (Fig. 5A), both the WT and mutants show the same RMSD tendencies for 30 ns, indicating the same level of thermostability. However, at 345 K (Fig. 5B), both mutants demonstrated steadier RMSD values than the WT, signifying improved thermal tolerance in comparison with the WT. The regions of highest flexibility and sensitivity, to changes in temperature of the WT and mutants, are shown in Fig. 5C. These data were consistent with the CD thermal denaturation analysis and enzyme percentage residual activity results.

Highly flexible regions in the wild-type and mutants were compared using RMSF values and regions having fluctuation more than 1 Å were considered as flexible regions (Fig. 5C, D). With respect to WT enzyme thirteen regions were identified as flexible (1–6, 28–62, 170–176, 213–218, 238–241, 255–267, 274–283,324–330, 339–357, 368–379, 385–400, 408–427, 435–449) at 345 K and in the case of A29G, eight regions (1–6,28–63, 255–267, 274–278, 326–330, 349–357, 368–379, 446–449) showed fluctuation higher than 1 Å at 345 K. Less flexible regions noticed in the mutants confer temperature resistance to the protein structure. Interestingly, thirteen flexible regions (1–8, 24–67, 172–175, 213–219, 238–241, 255–267, 326–330, 349–354, 368–379, 384–395, 408–412, 419–428, 435–449) were identified in V336I and the flexible region numbers in V336I resembles wild-type but were shorter as compared to wild-type except for the first two regions. The shortening of the flexible regions could have possibly contributed to the decreased temperature sensitivity in the mutants. The double mutant A29G/V336I was constructed to assess a potential synergetic effect. Unexpectedly, the thermostability of the double mutant was far less than its single mutant counterparts as well as the WT (Supplementary File).

4. Discussion

Despite several effective proteases from different sources being characterized [49], there is still a high demand for more stable and efficient proteases [50–52]. The isolation of a strain of *Pseudomonas aeruginosa* and the subsequent observation of extracellular protease activity led to the characterization of a serine protease based on its biochemical and biophysical features. Our protease resembles the DegP like HtrA bacterial serine protease, having the characteristic N-terminal catalytic domain and a C-terminal substrate binding PDZ domain. Like DegP, there are two C-terminal PDZ domains, namely PDZ1 and PDZ2. The PDZ1 is responsible for binding of substrate and presents it to the catalytic N-terminal domain for proteolytic activity. It is also observed that stabilization of the interface of the catalytic domain and PDZ1 domains play an important role in the overall stability and catalytic activity [45]. In order to upsurge the thermal stability of the protease, we undertook a rational protein design approach that is considered to be effective in improving temperature resistance and activity of enzymes [10,53,54].

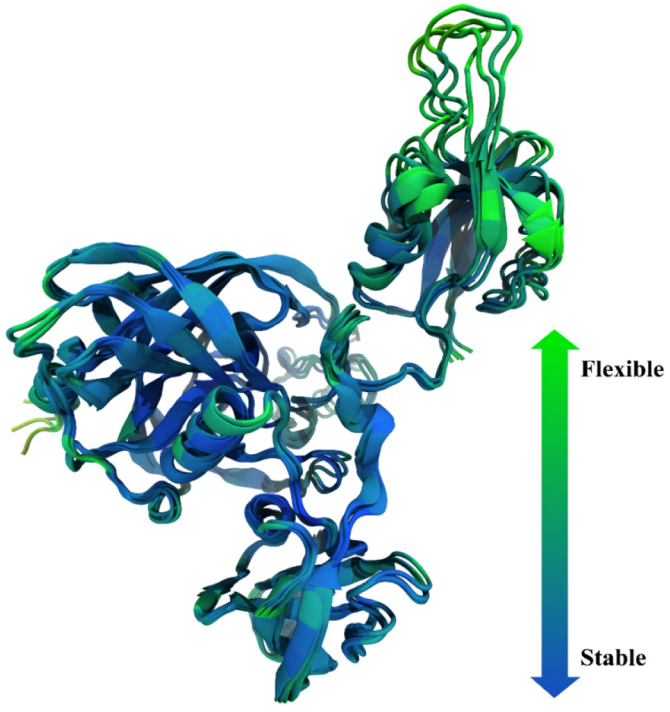

Fig. 5. Regions of high flexibility are indicated in green and more stable regions are shown in blue for the thermally unstable wild type. Periodic structures from the 30 ns simulation are superimposed to show the final stages of unfolding at 345 K.

Eight-point mutations, predicted by FireProt, were selected for screening for increased thermostability. However, among them, only two mutations A29G and V336I exhibited a positive impact on thermostability of the enzyme while rest of mutations either manifested disrupted secondary structures or inactivated the enzymes (Fig. S3). The majority of the previously reported rational designs applied either the consensus-based approach or the energy-based approach to increase the thermal stability of various proteins. Most of the stable mutants predicted by these approaches failed to produce any positive impact on the protein structure and only a small number of predicted substitutions were able to demonstrate their stable impact experimentally and therefore, suggests a need for improvement in computational algorithms to predict stabilizing mutations in protein structures [55,56]. Usually, predicted stabilizing mutations by the currently available algorithm focus on local stability impact of substitution while ignoring their dynamic impact on the protein structure. This may be another reason for the high false positive predictions.

Ala29, located in the longest LA loop of our protease structure, plays an important role in regulating the catalytic activity and oligomerization state of HtrA like proteases [57]. The minimum functional unit which can carry out proteolytic activity in cases of HtrA-like proteases is a trimer [58] and the interaction of the protease domain with neighbouring monomers are responsible for maintaining the trimer state. The LA loop thus contributes to maintaining thermal stability and the oligomeric state of the protease, as it is believed to play a major role in these interactions [59,60]. Ala29 in the LA loop is the least conserved at this position across all homologous sequences and FireProt predicted a high stabilizing impact of glycine at this position, based on the FoldX and Rosetta scores [61,62]. A closer look at the protein structure reveals the formation of one additional hydrogen bond (Fig. S6A) by the substituted glycine and a change of 2 kcal/mol could be a potential reason for its enhanced thermostability [39,63,64]. On the basis of these observations, we propose that A29G substitution contributed to the stabilization of the oligomeric state of the protease that is required for the functional stability of the enzyme.

FireProt predicted the V336I substitution by the evolution-based approach which was also coupled with FoldX energy values to rule out possible destabilizing mutations [28]. This substitution resulted in an increase of 5 °C for its T_m and improved catalytic activity when compared to the WT and A29G. Val336 has located in the loop of the PDZ1 domain and its substitution with Ile resulted in the formation of five new hydrophobic interactions and two hydrogen bonds without upsetting important stabilizing features such as the ionic interaction between Lys314 and Asp317 (Fig. S6B). Additionally, hydrophobic interactions play important roles in maintaining the globular conformation of proteins and would be expected to contribute to the overall stability [65].

It is known that the PDZ1 domain is required for substrate recognition and therefore essential for catalytic activity. Hence, any beneficial mutations in this region would have the greatest impact on catalytic activity [66]. According to the “hold and bite mechanism” of these type of serine proteases, PDZ1 domains tether the substrate and presents it to catalytic domain for proteolytic activity [67–69] which infers the

possibility that stable interactions of the protease domain with PDZ1 is important to maintain the catalytic activity and destabilization of its interdomain interface can have a negative impact on its structure and activity. Previous studies reported mutants constructed in a way that impaired hydrophobic interactions between the protease and PDZ domains, resulting in the total loss of catalytic activity [45,70]. In our case, V336, seen at the interface of the domains, when substituted with isoleucine stabilized the interface by making new hydrophobic interactions within the PDZ domain and as well as with the Leu247 of the protease domain. This observation in view of previously reported studies helps us to infer that this might be the possible reason for high catalytic activity of the mutant as compared to the wild-type [45]. The mutant experimental kinetic data and molecular docking simulation also supports this hypothesis. The double mutant A29G/V336I was constructed to assess a potential synergetic effect. Surprisingly, the thermostability of the double mutant was far less than its single mutant counterparts as well as the WT (Supplementary file). The result, therefore, demonstrates that beneficial mutations do not always produce synergistic effects [10,71–73].

Our kinetic studies show lower K_m values for both the mutants, indicating improved substrate affinity [74]. The molecular docking simulation of the wild-type and the mutants also predicted better affinities of the substrate with both mutants. As indicated in Table-2, there is not any significant difference among Kcat values of the wild-type and the mutants but there is a significant increase in the K_{cat}/K_m ratio for the mutants, showing increased catalytic efficiency [75]. This data suggests that a better affinity of the substrate with the mutants leads to better catalytic activity. From an industrial perspective, high thermostability and catalytic efficiency are desirable features [7] that support improved enzyme cost and reaction time. Thus, the improved thermostability and activity of the A29G and V336I mutants demonstrate their potential as future candidates for further development for industrial applications. In this study, energy-based and evolution-based rational design was used to increase thermostability and catalytic efficiency of SP. Although the rational design has many advantages over directed design approach, accurate prediction of stable mutants using various computational algorithm still needs a lot of improvement as in our case, 75% of the predicted mutants proved to be unstable. The coupling of prediction methodologies with MD simulations or Constrained Network Analysis (CNA) can reduce the false positive predictions. Moreover, there is also a need to develop strategies to predict the synergetic impact of stable mutants.

Acknowledgment

This project is supported by a Higher Education Commission of Pakistan grant (NRPU-4571). We are also thankful to the International Research Support Initiative Program (IRSIP), HEC for providing a travel grant to NMA to work in the laboratory of KS.

Contribution

NMA and AK carried out the experimental analysis and wrote the manuscript. AH isolated and identified the bacterial strain. DK carried out MD analysis. AMMS and HYM helped in CD spectroscopy and data analysis. NZ and KS supervised the study and finalized the manuscript. All the authors have read and revised the manuscript.

Competing interests

Authors do not have any competing interests.

Availability of material and data

Recombinant strain can be provided upon request to the corresponding author for reproducibility. No data set is generated in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2018.12.218.

References

[1] B. Turk, D. Turk, V. Turk, Protease signalling: the cutting edge, EMBO J. 31 (7) (2012) 1630–1643.
[2] D. Agrawal, et al., Production of alkaline protease by *Penicillium* sp. under SSF conditions and its application to soy protein hydrolysis, Process Biochem. 39 (8) (2004) 977–981.
[3] R. Gupta, et al., An overview on fermentation, downstream processing and properties of microbial alkaline proteases, Appl. Microbiol. Biotechnol. 60 (4) (2002) 381–395.
[4] C.S. Craik, M.J. Page, E.L. Madison, Proteases as therapeutics, Biochem. J. 435 (1) (2011) 1–16.
[5] C. Silva, et al., Practical insights on enzyme stabilization, Crit. Rev. Biotechnol. 38 (3) (2018) 335–350.
[6] T. Lonhienne, C. Gerday, G. Feller, Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility, Biochim. Biophys. Acta 1543 (1) (2000) 1–10.
[7] W. Xia, et al., Engineering a highly active thermophilic β-glucosidase to enhance its pH stability and saccharification performance, Biotechnol. Biofuel. 9 (1) (2016) 147.
[8] T. Hasunuma, et al., A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology, Bioresour. Technol. 135 (2013) 513–522.
[9] D. Klein-Marcuschamer, et al., The challenge of enzyme cost in the production of lignocellulosic biofuels, Biotechnol. Bioeng. 109 (4) (2012) 1083–1087.
[10] X. Chen, et al., Engineering the conserved and noncatalytic residues of a thermostable β-1, 4-endoglucanase to improve specific activity and thermostability, Sci. Rep. 8 (1) (2018) 2954.
[11] I. Wu, F.H. Arnold, Engineered thermostable fungal Cel6A and Cel7A cellobiohydrolases hydrolyze cellulose efficiently at elevated temperatures, Biotechnol. Bioeng. 110 (7) (2013) 1874–1883.
[12] J.M. Woodley, Integrating protein engineering with process design for biocatalysis, Phil. Trans. R. Soc. A 376 (2110) (2018) 20170062.
[13] C. Wang, et al., Improving the thermostability of alpha-amylase by combinatorial coevolving-site saturation mutagenesis, BMC Bioinf. 13 (1) (2012) 263.
[14] M. Musil, et al., FireProt: web server for automated design of thermostable proteins, Nucleic Acids Res. (2017) gkx285.
[15] J.K. Blum, M.D. Ricketts, A.S. Bommarius, Improved thermostability of AEH by combining B-FIT analysis and structure-guided consensus method, J. Biotechnol. 160 (3–4) (2012) 214–221.
[16] B. Jaouadi, et al., Enhancement of the thermostability and the catalytic efficiency of *Bacillus pumilus* CBS protease by site-directed mutagenesis, Biochimie 92 (4) (2010) 360–369.
[17] N.Z. Jaouadi, et al., Probing the crucial role of Leu31 and Thr33 of the *Bacillus pumilus* CBS alkaline protease in substrate recognition and enzymatic depilation of animal hide, PLoS One 9 (9) (2014), e108367.
[18] H. Takagi, et al., Mutant subtilisin E with enhanced protease activity obtained by site-directed mutagenesis, J. Biol. Chem. 263 (36) (1988) 19592–19596.
[19] H. Wang, et al., Screening and mutagenesis of a novel *Bacillus pumilus* strain producing alkaline protease for dehairing, Lett. Appl. Microbiol. 44 (1) (2007) 1–6.
[20] A. Dereeper, et al., Phylogeny. fr: robust phylogenetic analysis for the non-specialist, Nucleic Acids Res. 36 (suppl_2) (2008) W465–W469.
[21] C.E. Lane, et al., CEMAsuite: open source degenerate PCR primer design, Bioinformatics 31 (22) (2015) 3688–3690.
[22] E. Gasteiger, et al., Protein identification and analysis tools on the ExPASy server, The Proteomics Protocols Handbook, Springer 2005, pp. 571–607.
[23] T.N. Petersen, et al., SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods 8 (10) (2011) 785.
[24] R.D. Finn, et al., InterPro in 2017—beyond protein family and domain annotations, Nucleic Acids Res. 45 (D1) (2016) D190–D199.
[25] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinf. 9 (1) (2008) 40.
[26] D. Xu, Y. Zhang, Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization, Biophys. J. 101 (10) (2011) 2525–2534.
[27] C. Wiedemann, P. Bellstedt, M. Görlach, CAPITO—a web server-based analysis and plotting tool for circular dichroism data, Bioinformatics 29 (14) (2013) 1750–1757.
[28] D. Bednar, et al., FireProt: energy-and evolution-based computational design of thermostable multiple-point mutants, PLoS Comput. Biol. 11 (11) (2015), e1004556.
[29] M. Niklasson, et al., Robust and convenient analysis of protein thermal and chemical stability, Protein Sci. 24 (12) (2015) 2055–2062.
[30] C. Cupp-Enyard, Sigma's non-specific protease activity assay-casein as a substrate, J. Vis. Exp. 19 (2008).
[31] J.C. Phillips, et al., Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (16) (2005) 1781–1802.
[32] J.V. Ribeiro, et al., QwikMD—integrative molecular dynamics toolkit for novices and experts, Sci. Rep. 6 (2016) 26536.
[33] A.D. MacKerell Jr., et al., All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (18) (1998) 3586–3616.

[34] B.R. Brooks, et al., CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. 4 (2) (1983) 187–217.
[35] R.B. Best, et al., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ1 and χ2 dihedral angles, J. Chem. Theory Comput. 8 (9) (2012) 3257–3273.
[36] W.L. Jorgensen, et al., Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (2) (1983) 926–935.
[37] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N· log (N) method for Ewald sums in large systems, J. Chem. Phys. 98 (12) (1993) 10089–10092.
[38] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1) (1996) 33–38.
[39] S. Albeck, R. Unger, G. Schreiber, Evaluation of direct and cooperative contributions towards the strength of buried hydrogen bonds and salt bridges1, J. Mol. Biol. 298 (3) (2000) 503–520.
[40] D. Schneidman-Duhovny, et al., PatchDock and SymmDock: servers for rigid and symmetric docking, Nucleic Acids Res. 33 (suppl_2) (2005) W363–W367.
[41] J. Gasteiger, C. Rudolph, J. Sadowski, Automatic generation of 3D-atomic coordinates for organic molecules, Tetrahedron Comput. Methodol. 3 (6) (1990) 537–547.
[42] C. Zhang, et al., Determination of atomic desolvation energies from the structures of crystallized proteins1, J. Mol. Biol. 267 (3) (1997) 707–726.
[43] D. Duhovny, R. Nussinov, H.J. Wolfson, Efficient unbound docking of rigid molecules, International Workshop on Algorithms in Bioinformatics, Springer, 2002.
[44] N.D. Rawlings, et al., The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database, Nucleic Acids Res. 46 (D1) (2017) D624–D632.
[45] D. Zurawa-Janicka, et al., Temperature-induced changes of HtrA2 (Omi) protease activity and structure, Cell Stress Chaperones 18 (1) (2013) 35–51.
[46] J. Skorko-Glonek, et al., Comparison of the structure of wild-type HtrA heat shock protease and mutant HtrA proteins. A Fourier transform infrared spectroscopic study [published erratum appears in J Biol Chem 1995 Dec 29; 270 (52): 31413], J. Biol. Chem. 270 (19) (1995) 11140–11146.
[47] N. Singh, R.R. Kuppili, K. Bose, The structural basis of mode of activation and functional diversity: a case study with HtrA family of serine proteases, Arch. Biochem. Biophys. 516 (2) (2011) 85–96.
[48] J. Jiang, et al., Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins, Proc. Natl. Acad. Sci. 105 (33) (2008) 11939–11944.
[49] Q. Li, et al., Commercial proteases: present and future, FEBS Lett. 587 (8) (2013) 1155–1163.
[50] M. Dodia, et al., Purification and stability characteristics of an alkaline serine protease from a newly isolated Haloalkaliphilic bacterium sp. AH-6, J. Ind. Microbiol. Biotechnol. 35 (2) (2008) 121–131.
[51] R.R. Rao, et al., Alkaline protease production from *Brevibacterium luteolum* (MTCC 5982) under solid-state fermentation and its application for sulfide-free unhairing of cowhides, Appl. Biochem. Biotechnol. 182 (2) (2017) 511–528.
[52] S. Singh, P. Gupta, B.K. Bajaj, Characterization of a robust serine protease from *Bacillus subtilis* K-1, J. Basic Microbiol. 58 (1) (2018) 88–98.
[53] H. Takagi, et al., Enhancement of the thermostability of subtilisin E by introduction of a disulfide bond engineered on the basis of structural comparison with a thermophilic serine protease, J. Biol. Chem. 265 (12) (1990) 6874–6878.
[54] W.S. Mak, J.B. Siegel, Computational enzyme design: transitioning from catalytic proteins to enzymes, Curr. Opin. Struct. Biol. 27 (2014) 87–94.
[55] H.J. Wijma, et al., Computationally designed libraries for rapid enzyme stabilization, Protein Eng. Des. Sel. 27 (2) (2014) 49–58.
[56] R.S. Komor, et al., Highly thermostable fungal cellobiohydrolase I (Cel7A) engineered using predictive methods, Protein Eng. Des. Sel. 25 (12) (2012) 827–833.
[57] D. Figaj, et al., The LA loop as an important regulatory element of the HtrA (DegP) protease from *Escherichia coli*; structural and functional studies, J. Biol. Chem. 289 (2014) 15880–15893.
[58] A. Jomaa, et al., The inner cavity of *Escherichia coli* DegP protein is not essential for molecular chaperone and proteolytic activity, J. Bacteriol. 189 (3) (2007) 706–716.
[59] N. Rai, A. Ramaswamy, Temperature dependent dynamics of DegP-trimer: a molecular dynamics study, Comput. Struct. Biotechnol. J. 13 (2015) 329–338.
[60] T. Koper, et al., Analysis of the link between the redox state and enzymatic activity of the HtrA (DegP) protein from *Escherichia coli*, PLoS One 10 (2) (2015), e0117413.
[61] J. Huang, D.-F. Xie, Y. Feng, Engineering thermostable (R)-selective amine transaminase from *Aspergillus terreus* through in silico design employing B-factor and folding free energy calculations, Biochem. Biophys. Res. Commun. 483 (1) (2017) 397–402.
[62] G. Li, et al., Identification of a hot-spot to enhance *Candida rugosa* lipase thermostability by rational design methods, RSC Adv. 8 (4) (2018) 1948–1957.
[63] C. Nick Pace, J.M. Scholtz, G.R. Grimsley, Forces stabilizing proteins, FEBS Lett. 588 (14) (2014) 2177–2184.
[64] H. Yu, et al., Two strategies to engineer flexible loops for improved enzyme thermostability, Sci. Rep. 7 (2017) 41212.
[65] C.N. Pace, et al., Contribution of hydrophobic interactions to protein stability, J. Mol. Biol. 408 (3) (2011) 514–528.
[66] J. Iwanczyk, et al., Role of the PDZ domains in *Escherichia coli* DegP protein, J. Bacteriol. 189 (8) (2007) 3176–3186.
[67] T. Krojer, et al., Interplay of PDZ and protease domain of DegP ensures efficient elimination of misfolded proteins, Proc. Natl. Acad. Sci. 105 (22) (2008) 7702–7707.
[68] Y. Masato, et al., Binding of proteins to the PDZ domain regulates proteolytic activity of HtrA1 serine protease, Biochem. J. 381 (3) (2004) 895–904.
[69] T. Clausen, C. Southan, M. Ehrmann, The HtrA family of proteases: implications for protein composition and cell fate, Mol. Cell 10 (3) (2002) 443–455.
[70] W. Li, et al., Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi, Nat. Struct. Mol. Biol. 9 (6) (2002) 436.
[71] X. Liu, et al., Directed evolution and secretory expression of a pyrethroid-hydrolyzing esterase with enhanced catalytic activity and thermostability, Microb. Cell Factories 16 (1) (2017) 81.
[72] C. Niu, et al., Engineering the residual side chains of HAP phytases to improve their pepsin resistance and catalytic efficiency, Sci. Rep. 7 (2017) 42133.
[73] W. Yang, et al., Improved thermostability of an acidic xylanase from *Aspergillus sulphureus* by combined disulphide bridge introduction and proline residue substitution, Sci. Rep. 7 (1) (2017) 1587.
[74] J.M. Berg, T.J. Tymoczk, L. Stryer, Biochemistry, 5th ed. W H Freeman, New York, 2002.
[75] A.D. Mesecar, B.L. Stoddard, D.E. Koshland, Orbital steering in the catalytic power of enzymes: small structural changes with large catalytic consequences, Science 277 (5323) (1997) 202–206.