

Research Article

Three dimensional structure prediction of panomycocin, a novel Exo- β -1,3-glucanase isolated from *Wickerhamomyces anomalus* NCYC 434 and the computational site-directed mutagenesis studies to enhance its thermal stability for therapeutic applications

Muhammed Tilahun Muhammed^a, Çağdaş Devrim Son^b, Fatih İzgü^{b,*}

^a Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Suleyman Demirel University, Isparta, Turkey

^b Department of Molecular Biology and Genetics, Middle East Technical University, Ankara, Turkey

ARTICLE INFO

Keywords:

Antifungal protein
Panomycocin
Exo- β -1, 3-glucanase
Homology modeling
Computational site-directed mutagenesis
Thermal stability

ABSTRACT

Panomycocin is a naturally produced potent antimycotic/antifungal protein secreted by the yeast *Wickerhamomyces anomalus* NCYC 434 with an exo- β -1,3-glucanase activity. In this study the three dimensional structure of panomycocin was predicted and the computational site-directed mutagenesis was performed to enhance its thermal stability in liquid formulations over the body temperature for topical therapeutic applications. Homology modeling was performed with MODELLER and I-TASSER. Among the generated models, the model with the lowest energy and DOPE score was selected for further loop modeling. The loop model was optimized and the reliability of the model was confirmed with ERRAT, Verify 3D and Ramachandran plot values. Enhancement of the thermal stability of the model was done using contemporary servers and programs such as SPDBViewer, CNA, I-Mutant2.0, Eris, AUTO-MUTE and MUpro. In the region outside the binding site of the model Leu52 Arg, Phe223Arg and Gly254Arg were found to be the best thermostabilizing mutations with 6.26 K, 6.26 K and 8.27 K increases, respectively. In the binding site Glu186Arg was found to be the best thermostabilizer mutation with a 9.58 K temperature increase. The results obtained in this study led us to design a mutant panomycocin that can be used as a novel antimycotic/antifungal drug in a liquid formulation for topical applications over the normal body temperature.

1. Introduction

Panomycocin is a naturally produced antifungal/antimycotic monomeric glycoprotein of 49 kDa with an exo- β -1,3-glucanase activity. It is produced and secreted into the environment by the killer (K +) yeast strain *Wickerhamomyces anomalus* NCYC 434 (formerly known as *Pichia anomala*) and has been shown as a promising potential antifungal agent in biomedicine (Izgü et al., 2007, 2005; Izgü and Altınbay, 2004; Walker, 2011). Panomycocin hydrolyzes O-glycosidic linkage of the β -1,3-linked glucan residues, which are the vital polymers for the integrity of the fungal cell wall, from their non reducing terminus yielding α -glucose. This disrupts the cell wall and leads to the death of the target cells (Izgü et al., 2006). In several studies, the potent *in vitro* antifungal activity of panomycocin against *Candida* spp. and dermatophytes was shown (Izgü et al., 2007). The mammalian cells lack the β -1,3-glucans in their structure, and this highlights the use of panomycocin as a selective antifungal/antimycotic agent in therapy with

improved safety. In solution, panomycocin is stable and active at the pH range between 3.0 and 5.5 up to 37 °C. In lyophilized form, it retains its stability and activity up to 38.5 °C (Izgü and Altınbay, 2004). Recently a liposomal lyophilized powder formulation of panomycocin was developed for topical therapeutic purposes against vulvovaginal candidiasis which affects 80% of women worldwide. For the stability of panomycocin at higher body temperatures in liquid formulations, its thermal stability needs to be enhanced (Izgü et al., 2017).

Thermal stability enhancement can be achieved by several approaches including use of excipients, immobilization, chemical modification and engineering of proteins. The most classical method in protein engineering is the so-called 'rational design' approach which involves 'site-directed mutagenesis' of proteins. With this technique it is possible to change a single amino acid in the sequence of a protein for another amino acid with different physico-chemical properties (Bernal et al., 2018). Rational design is an effective approach when the three dimensional (3D) structure and mechanism of a protein of interest is

* Corresponding author at: Department of Molecular Biology and Genetics, Middle East Technical University, Dumlupınar Bulvarı, No. 1, Ankara, 06800, Turkey.
E-mail address: izgu@metu.edu.tr (F. İzgü).

known. Numerous computational tools such as homology modeling, *ab initio* and threading methods have been developed to predict the 3D structures. There are also computational tools that are used in the investigation of the effects of mutations on the thermal stability of proteins. These in turn allow technically easy, time saving and cost effective protein engineering (Talluri, 2011).

Among the computational 3D prediction methods, the method of homology modeling usually provides the most reliable result as this technique predicts the structure of a protein from its amino acid sequence with high accuracy which is comparable to the results obtained experimentally (Cavasotto and Phatak, 2009). Homologous proteins with related amino acid sequences have similar 3D structure, which is more conserved and changes much slower than the related sequence during evolution (Krieger et al., 2012). Homology modeling, also called comparative or sometimes template based modeling, uses information from one or more related proteins with known 3D structure (template) to generate models for the target protein. The importance of homology modeling is increasing as the number of protein 3D structures determined increases. This makes the drug discovery process faster, easier, cheaper and more practical (Muhammed and Aki-Yalcin, 2018).

The amino acid sequence of panomycocin is similar to the amino acid sequence of the α -1,3-glucanase of *W. anomalus* strain K, which was previously determined and deposited with UniProt databank (accession number AJ222862) (Izgi et al., 2006).

In this study the 3D structure of panomycocin was predicted by homology modeling using MODELLER program (Kuntal et al., 2010a) and I-TASSER (Iterative Threading ASSEMBLY Refinement) server (Yang et al., 2014). The structure built was reliable as ERRAT, Verify 3D and Ramachandran plot gave high scores. The thermostabilizing effect of the computational site-directed mutagenesis was assessed by using SPDBViewer (Guex and Peitsch, 1997), CNA (Constraint Network Analysis) (Krüger et al., 2013), I-Mutant2.0 (Capriotti et al., 2005), AUTO-MUTE (Masso and Vaisman, 2010), Eris (Yin et al., 2007) and MUpPro (Cheng et al., 2005) servers. The substitution of relatively thermostable amino acids at the determined positions in the structure of panomycocin will highly enhance its activity and stability at temperatures higher than the body temperature in liquid formulations for topical applications.

2. Materials and methods

The amino acid sequence in FASTA format, retrieved from UniProt with the accession number AJ222862, was used for the 3D structure prediction of panomycocin (<http://www.uniprot.org/>) (Apweiler et al., 2011).

2.1. Determination of the signal peptide and the KEX2 cleavage site

The signal peptide of the protein was determined using SignalP 4.1 (Petersen et al., 2011), Signal-CF (Chou and Shen, 2007), PrediSi (Hiller et al., 2004) and Signal-3 L (Shen and Chou, 2007) servers. The KEX2 cleavage site was predicted with PROSPER (PROtease Specificity Prediction server) (Song et al., 2012) and SMART (Simple Modular Architecture Research Tool) (Letunic et al., 2015) servers.

2.2. Homology modeling

The sequence just downstream of the KEX2 cleavage site was used for the modeling. Templates to the target sequence were obtained by using BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1997) in NCBI (National Center for Biotechnology Information) (Agarwala et al., 2016). The 3D structures of the templates were downloaded from PDB (Protein Data Bank) (<https://www.rcsb.org/>) (Rose et al., 2017). Then, the target sequence and the templates were loaded to Easy-Modeller 4.0 (Kuntal et al., 2010b). In this program the templates were aligned with each other and the target sequence was aligned with the

templates. Then with MODELLER 9.18 nin. models were generated. Another model was also generated with I-TASSER server for further comparison with MODELLER results.

2.2.1. Comparison of the generated models

DOPE (Discrete Optimized Protein Energy) score, energy, TM (Template Modeling) score and RMSD (Root Mean Square Deviation) values were used to compare the generated models. The DOPE score was obtained from MODELLER, the energy was calculated with GROMOS96 (GRONingen MOlecular Simulation96) (Van Gunsteren W.F., 2017) in the SPDBViewer (Swiss PDB Viewer), the TM score was measured with TM score calculator (Xu and Zhang, 2010) and the RMSD value was calculated with VMD (Visual Molecular Dynamics) (Humphrey et al., 1996) for each model.

2.2.2. Loop modeling

The loop modeling was performed on the best model determined. The DOPE profile of the best model, the structural and alignment analysis between the target and the templates were used to determine the regions that loop modeling might improve the model quality.

Loop modeling was performed separately at five positions in the amino acid sequence (1–5, 233–240, 313–320, 330–337, 368–374). Loop modeling at each position was analysed by the DOPE score, energy, TM score and RMSD value of the resulting five models. The results obtained from each model were compared with the pre-determined best model. The thermal stability of the best model and the best loop model were also determined and compared.

2.2.3. Optimization, verification and validation

Optimization of the best model and its loop model was performed by MODELLER. The optimized resulting models were evaluated by SAVES (Structure Analysis and Verification Server) server (Colovos and Yeates, 1993; Laskowski et al., 1993).

2.3. Determination of binding site

COACH (Yang et al., 2013), COFACTOR (Roy et al., 2012), Meta-Pocket (Binding (2009)), CASTp (Computed Atlas of Surface Topography of Proteins) (Dundas et al., 2006) and DoGSiteScorer (Volkamer et al., 2012) servers were used to predict the binding (active) site of the best model and its loop model.

2.4. Determination of the appropriate sites on the models that would improve the thermal stability

Determination of the sites that would enhance the thermal stability of the models was done with GROMOS96 in the SPDBViewer. The free energy (in KJ/mol) of each position in the models and the energy profiles of the alignment were predicted and investigated. All the amino acids in the unstable regions were substituted with the rest nineteen amino acids. SPDBViewer was used to alter the amino acid residues. The substitutions which gave a lower energy both outside and inside of the binding regions were recorded. The best ten and three amino acid substitutions, outside and inside binding regions respectively, were selected for further analysis. The combinations among the substitutions were also tested and those that gave higher energy were eliminated (five outside and two inside binding region substitutions) from further analysis. The temperature changes of the selected substitutions were then measured (in Kelvin) by CNA server. The accuracy of the predicted sites for the substitutions were further tested with I-Mutant2.0, AUTO-MUTE, Eris and MUpPro servers.

A)

Elem Name	Instances (Matched Sequence)	Positions	View in Jmol	Elem Description	Cell Compartment	Pattern	PHI-Blast Instance Mapping	Structural Filter Info	Probability
CLV_PCSK_KEX2_1	KRG	30-32[A]	-	Yeast kexin2 cleavage site (K-R-l-X or R-R-l-X)	Extracellular, Golgi apparatus	[KR]R	-	-	7.973e-03
CLV_PCSK_PC1ET_1	KRG	30-32[A]	-	NEC1/NEC2 cleavage site (K-R-l-X)	Extracellular, Golgi apparatus, Golgi membrane	KR	-	-	3.903e-03

B)

```

..... M L I S T F I I S S L I I
34 ctctctatagcactagcaaacctatccctcaagaggcggaacacaattttataaacgtggtgattatgggattaccaaagat 120
   L S I A L A N P I P S R G G T Q F Y (K R) G D Y W D Y Q N D   40
121 aaaacccgtggtgtaatttagtggtggtgttttagaaccttcatcactcctcttattgaagctttgaaaatcaagcc 207
   K I R G V N L G G W F V L E P F I T P S L F E A F E N Q G   69
208 caagatgttctctgtgatgaataccattatacaaaccccttgtaaggatttagctaaagagagattggatcaacattggagttca 294

```

Fig. 1. A) KEX2 cleavage site estimation of SMART server. B) KEX2 cleavage site suggested in the previous study (indicated by a question mark).

3. Results

3.1. Determination of the signal peptide and the KEX2 cleavage site

SignalP 4.1, Signal-CF, PrediSi and Signal-3 L server results showed that the first seventeen amino acids constitute the signal peptide.

SMART server estimated the KEX2 cleavage site at positions 30–32 (Fig. 1A). In a previous study on the primary translation product, it was suggested that the KEX2 cleavage site might be at the dibasic Lys-Arg motif at positions 30–31 (Fig. 1B) (Grevesse et al., 2003). The N-terminal amino acid sequencing of panomycocin showed that the secreted mature protein starts with glycine at the position 32 (Izgu et al., 2006). Thus, these findings showed that the cleavage takes place between the arginine and glycine residues in the Lys-Arg-Gly sequence at positions 30–31–32.

3.2. Homology modeling

Although the BLAST results showed that there were proteins with 100% identity and coverage with panomycocin, their 3D structures were unknown. The proteins, 1EQP, 2PB1 and 1CZ1, with known 3D structures were selected as templates. All the templates that were used in this study showed 98% coverage and 66% identity with panomycocin.

3.2.1. Determination of the best model

Nine different models were generated with MODELLER by using the above mentioned templates and then were compared according to their DOPE score, energy, TM score and RMSD values. The TM score and RMSD values of each model were close to each other. Although model 1 had the second lowest DOPE score, it had the lowest energy among others, thus, it was chosen as the best model (Fig. 2A and B).

3.2.2. Loop modeling

Loop modeling was performed at the five different positions in the amino acid sequence as indicated in the methods section. DOPE profile of the best model (model 1) with respect to the templates, the structural and alignment analysis between the target and the templates were utilized for the determination of the regions that would improve the model.

Amino acid positions 1–5 and 368–374 indicated an inconsistent change in DOPE score and energy as there was no correlation in these values. Positions 233–240 and 330–337 gave higher values than the

best model and the position 313–320 gave lower values of DOPE score and energy when compared to the best model. Thus, only the loop modeling at amino acid positions 313–320 has been selected for further analysis (Fig. 3). The TM score and RMSD value of the selected loop model were also measured. The TM score (0.4083) was lower than that of the best model (0.4099) and the RMSD value (5.5950 Å) was higher than that of the best model (5.5540 Å). As TM score was not higher and the RMSD value was not lower than that of the best model these values were not considered as satisfactory in topology and position measurements. Thus, the computational thermal stability study was performed not only in the loop model but also in the best model.

3.2.3. Optimization, verification and validation of the models

Optimization of the loop model and the best model with MODELLER showed lower DOPE score and energy, higher TM score and lower RMSD value for the models. DOPE score and energy for the loop model was -50808.406250 and -19918.740 KJ/mol respectively and TM score and RMSD value was 0.4083 and 5.5950 Å respectively. DOPE score and energy for the best model was -50360.93359 and -19740.609 respectively and TM score and RMSD value was 0.4099 and 5.5540 Å respectively. The optimized models were verified and validated with SAVES server. The overall quality factor estimated by ERRAT was 89.175. Results obtained from Verify 3D showed that 96.97% of the residues had an average 3D to 1D score greater than or equal to 0.2 and the results of Ramachandran plot showed that 99.70% of the residues were inside the allowed region of the plot (Fig. 4 A, B, C). All the results obtained showed that the loop model which was generated is reliable (Fig. 5). TM score, energy and RMSD values of the I-TASSER model were also calculated and compared with the MODELLER results. The RMSD value obtained (78.0966 Å) for the I-TASSER model was much higher than that of the MODELLER and thus the MODELLER results were preferred and used in this study.

3.3. Binding site determination

COACH, COFACTOR and MetaPocket servers indicated that the binding site pocket included the amino acids Glu23, Phe25, His129, Asn140, Asn185, Glu186, Tyr248, Phe251, Glu285, Trp361 and Trp371. Further binding site predictions by CASTp and DoGSiteScorer servers also gave a similar result but with more amino acid coverage as indicated in Fig. 6.

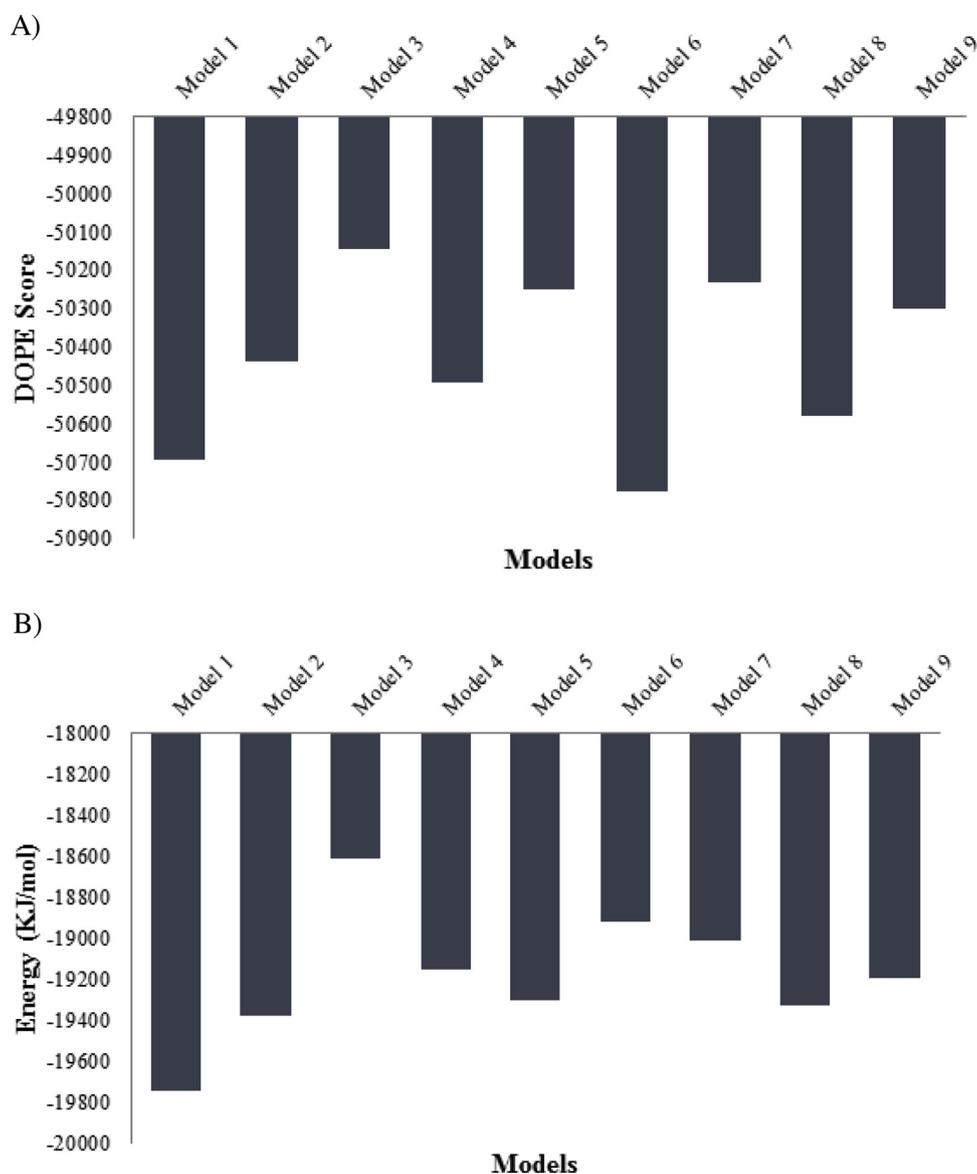


Fig. 2. A) DOPE score of the models (model 1 gave the second lowest score). B) Energy of the models (model 1 gave the lowest value).

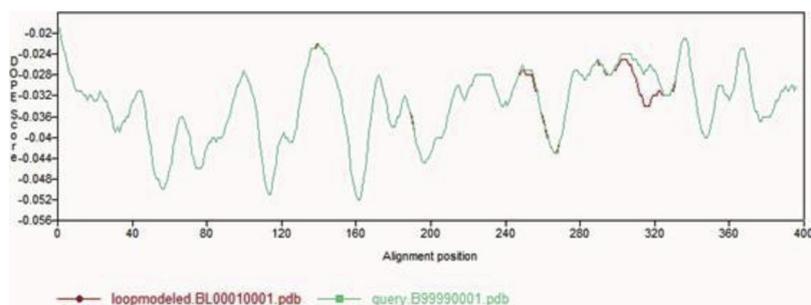


Fig. 3. DOPE profile of the best model (turquoise) and the loop model (red). The loop model showed a lower DOPE score at amino acid positions 313–320.

3.4. Determination of the appropriate sites that would improve the thermal stability

The GROMOS96 in SPDBViewer showed that amino acid substitutions at positions 52 Leu→Arg, 120 Tyr→Arg, 130 Gly→Ala, 195 Gly→Pro, 204 Gly→Arg, 223 Phe→Arg, 243 Ile→Arg, 254 Gly→Arg, 293 Cys→Arg and 299 Gly→Arg were the best ten positions that might increase the thermal stability of panomycocin outside of the binding

region.

The combinations of the substitutions between the amino acids, which were close to each other in the model, were also tested for their contribution to the thermal stability of the protein. Among the combinations that were performed, the amino acid positions 52, 120, 204, 223 and 254 gave a lower energy. When all the substitutions at those five positions were combined, the total energy was -21047.369 KJ/mol which was lower than that of the uncombined substitution. This showed

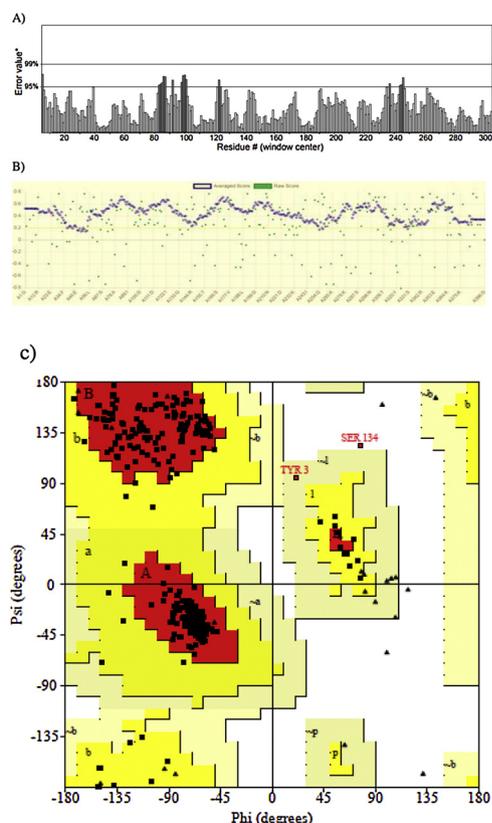


Fig. 4. Verification and validation of the loop model. A) ERRAT results (95% indicates rejection limit). B) Verify 3D (96.97% of the residues has 3D to 1D \geq 2). C) Ramachandran plot. The red region in the plot indicates energetically the most favored region; the yellow region represents the allowed region, and the white field represents the disallowed region.

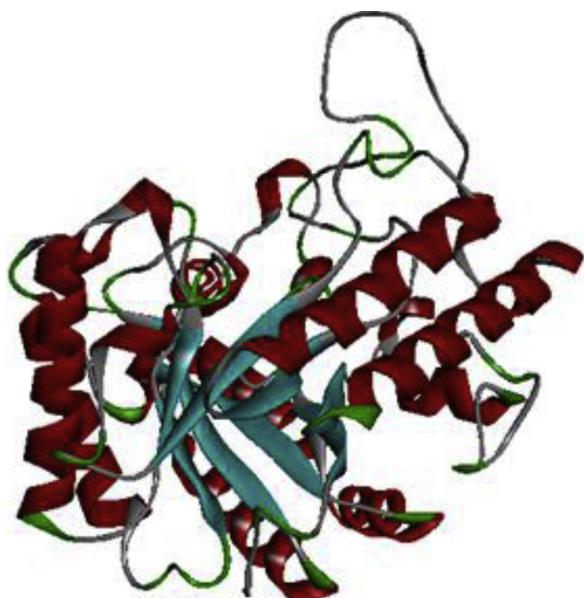


Fig. 5. 3D structure of the loop model of panomycocin generated with MODELLER.

an energy minimization of -1182.670 KJ/mol in the entire model. When all the substitutions at the above mentioned ten best thermostabilizing positions were combined, the energy of the entire model was higher (-20993.869 KJ/mol). Thus, the amino acid positions at 52, 120, 204, 223 and 254 were determined as the best substitution positions outside

the binding region for the thermostabilization of the protein. GROMOS96 in SPDViewer also showed the best three amino acid substitution positions inside the binding region as 25 Phe→Arg, 186 Glu→Arg and 371 Trp→Arg.

Substitution mutations in both outside and inside the binding region were estimated in terms of temperature (Kelvin) by the CNA server. Before any substitution mutation, the cluster configuration entropy type 2 (H type 2), which is related with the melting point of the protein, was -4.84 kcal/mol that is equivalent to 396.74 K. This value was compared to the values obtained after substitution mutations. Mutations at positions 52, 223 and 254 outside the binding region and mutation 186 inside the binding region increased the thermal stability of the protein (Table 1, Fig. 7). The accuracy of the results obtained with GROMOS96 and CNA server were also confirmed with I-Mutant2.0, Eris, AUTOMUTE and MUpro servers.

4. Discussion

In this study we have determined the 3D structure of panomycocin which is a naturally produced potential antifungal protein with an exo- β -1,3-glucanase activity. We have also determined the amino acid positions for substitutions that would enhance its thermal stability over the normal body temperatures in liquid formulations for topical therapeutic applications. Panomycocin which is produced and secreted by *W. anomalous* NCYC 434 has exactly the same amino acid sequence as the amino acid sequence of the exo- β -1,3-glucanase produced by *W. anomalous* strain K, which has been previously deposited with UniProt databank with accession number AJ222862. This amino acid sequence contains the signal peptide and the site for cleavage. In the literature the dibasic amino acids Lys-Arg at positions 30–31 was suggested as the probable KEX2 cleavage sites (Grevesse et al., 2003). The SMART server indicated the KEX2 cleavage sites as the amino acids Lys-Arg-Gly (30–32) with Lys-Arg cleavage pattern. The N-terminal amino acid of the secreted panomycocin is glycine and this proved that the KEX2 cleavage site is Lys-Arg-Gly-(-31-|-32-). Homology modeling was performed with the secreted peptide sequence which is just downstream of the KEX2 cleavage site using MODELLER 9.18. Nine models were generated and compared with respect to their DOPE score, energy, TM score and RMSD value. All the models showed similar TM score and RMSD value. Among them the model with the lowest energy and relatively low in DOPE score was selected for further computational thermal stability studies as low energy and low DOPE score values indicate a more stable 3D structure (Pucci et al., 2016).

The specificity of the function of a protein structure is often determined by its loops. Accuracy of loop modeling is an important factor which determines the value of the generated models for further applications (Kmieciak et al., 2016). We have performed loop modeling at five different positions in the amino acid sequence by using DOPE profile of the best model with respect to the templates, the structural and alignment analysis between the query and the templates. Only the loop modeling at position 313–320 gave lower values of DOPE score and energy when compared to the best model. This showed that the loop modeling at this position is stable and thus it was selected for the further computational thermal stability study. Optimization of the loop model with MODELLER showed lower DOPE score, lower energy values, higher TM score and lower RMSD value than that of the model before optimization. Thus, by optimization, the quality of the loop model has been improved. The optimized model was then verified and validated with SAVES. We have obtained values of 89.175, 96.97% and 99.70% for ERRAT, Verify 3D and Ramachandran plot respectively. These high values proved that the generated model was reliable (Khor et al., 2014). We have also compared the quality of the model generated with MODELLER using I-TASSER server which is among the best performing servers in the CASP (Critical Assessment of protein Structure Prediction) experiments. Although the energy and TM score of the models built with both MODELLER and I-TASSER were similar, the

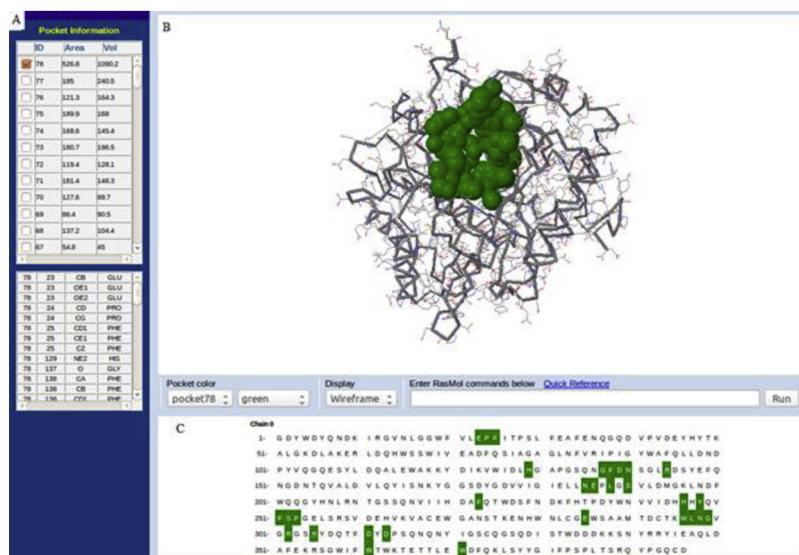


Fig. 6. A) The table of the area and the volume for different binding sites of the model. B) The three dimensional structure of the best binding site. C) Binding site analysis by CASTp server. Green color illustrates the binding site position (Dundas et al., 2006).

Table 1

CNA results of the best thermostabilizing amino acid substitutions inside and outside of the binding region of panomycocin.

Substitutions	Non mutated	Outside the binding region						Inside the binding region	
		Leu52Arg	Tyr120Arg	Gly204Arg	Phe223Arg	Gly254Arg	Phe25Arg	Glu186Arg	Trp371Arg
H type 2 (kcal/mol)	-4.84	-5.15	-4.84	-4.83	-5.15	-5.15	-4.89	-5.32	-4.78
Melting point (K)	396.74	403.00	396.74	396.68	403.00	405.01	397.72	406.32	395.66

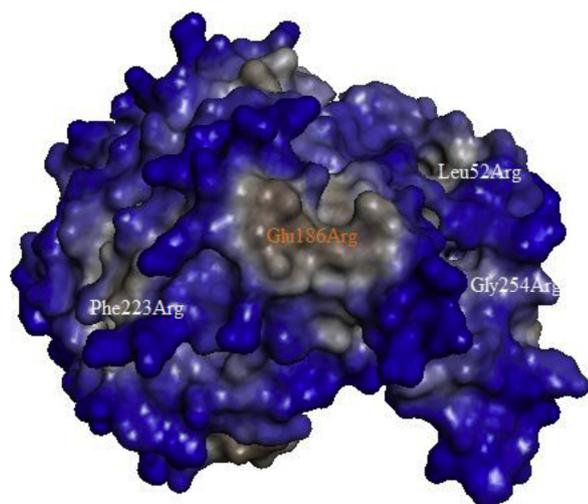


Fig. 7. Surface of the model that was generated for panomycocin and the exact places of the best amino acid positions that enhanced its thermal stability (letters in white color indicates the amino acids outside and in orange color indicates the amino acids inside the binding site). As the color changes from grey to blue, the hydrophilicity increases.

RMSD value of the I-TASSER model was higher. This indicates that the model generated by MODELLER is more similar to the templates than the I-TASSER (Kufareva and Abagyan, 2012).

The amino acids inside and outside of the binding region of a protein have different roles and the thermal instability may result from the deformation of the 3D structure resulting in deactivation of the binding site (Lee et al., 2017). Therefore, the computational thermal stability study was performed separately both inside and outside of the binding region. COACH, COFACTOR, MetaPocket, CASTp and DoGSiteScorer

servers indicated the amino acids inside the binding region as Glu23, Phe25, His129, Asn140, Asn185, Glu186, Tyr248, Phe251, Glu285, Trp361 and Trp371. UniProt data bank showed that α -D-glucanases have Glu at positions 217 and 316 in their active sites. In our model this corresponds to the positions 186 and 285 in the binding region as panomycocin was modeled after the amino acid sequence (31 aa) just upstream to the KEX2 cleavage site was removed.

Determination of the amino acid substitution positions on the 3D structure of panomycocin models that would enhance the thermal stability was performed with GROMOS96 in the SPDBViewer. In addition to the free energy calculation of the proteins, GROMOS has the advantage of providing energy minimization to refine the 3D structure of the proteins. In our case we have substituted an amino acid position with another amino acid so that refinement of the new structure was needed to avoid the possible unstable conditions arising from these mutations (Schmid et al., 2011). Furthermore, as mutations were performed with SPDBViewer, using GROMOS96 inside this viewer made the study more practical. Ten positions for amino acid substitutions outside and three positions for amino acid substitutions inside the binding region, which gave lower energy, were detected. The combinations among these thermostabilizing amino acid substitutions were also tested as combinations have been reported to have a non-additive effect on the thermal stability of the protein (Strub et al., 2004). Modern force fields like GROMOS that are currently used in energy calculations has some limitations. For example, they rely on a fixed charge model and atom types (Cournia et al., 2017). As a result, CNA server, which is rigidity theory-based thermal unfolding simulations of proteins for linking structure and thermostability, was used to increase the reliability of the study. In relative to other web servers that allow performing and analyzing thermal unfolding simulations of proteins, CNA server provides the most detailed information as it performs rigid cluster decompositions, simulates thermal unfolding and computes global and local flexibility indices (Mortazavi and Hosseinkhani, 2011).

By using CNA server, we have determined three (52, 223, 254) outside and one (186) inside the binding region amino acid substitution positions among the others that would enhance the thermal stability of the protein the CNA server gave a consistent result with the energy values calculated by GROMOS96 in SPDBViewer for all these positions. This result was also confirmed by using I-Mutant2.0, Eris, AUTO-MUTE and MUp servers.

The computational site-directed mutagenesis study showed that the mutations that highly enhanced the thermal stability of panomycocin were the substitutions of a residue by arginine (Leu52 Arg, Glu186Arg, Phe223Arg and Gly254Arg). In several studies it was reported that the change of hydrophobicity to hydrophilicity of the amino acid residues on the solvent exposed surface is a good strategy in the thermostabilization of a protein. Thus, substituting a hydrophobic residue by arginine, which is a positively charged hydrophilic amino acid, is expected to increase the thermostability of proteins. Strub et al. increased the stability of acetylcholinesterase by substituting solvent exposed hydrophobic residues by arginine (Sokalingam et al., 2012). Mortazavi et al. enhanced the thermostability of firefly luciferases by substituting solvent exposed hydrophobic residues by arginine (Zhou et al., 2013). In our study we have substituted the hydrophobic amino acids leucine, phenylalanine and glycine with arginine and the thermal stability has been increased by 6.26 K, 6.26 K and 8.27 K respectively. Even substitutions of hydrophilic residues on the surface by arginine may increase the thermostability of proteins. Sokalingam et al. increased the stability of green fluorescent protein (GFP) by substituting solvent exposed surface lysines by arginines (Zhou et al., 2008). Although lysine and arginine are both positively charged basic amino acids, the guanidinium group of arginine permits interactions in three directions that enables it to form a higher number electrostatic interactions and its basic residue has higher pKa that may generate more stable ionic interactions. Zhou et al. increased the thermostability of xylanase II from *Aspergillus usami* E001 by replacing serines and threonines on the solvent exposed surface of the enzyme with arginines (Kumwenda et al., 2013). When glutamic acid, which has a charged side chain, has been substituted by arginine, the thermal stability of panomycocin increased by 9.58 K. Moreover, bioinformatics analysis showed that one of the most striking features of thermostable proteins is the higher proportion of arginine in the exposed surfaces (Mclachlan et al., 2008). For example, Kumwenda et al. observed high frequency of arginine (on the surfaces) and alanine (in well buried areas) in thermostable proteins of *Thermus thermophilus* HB27 (Cournia et al., 2017). Similarly all the best thermostabilizing positions are at water exposed surface of panomycocin (Fig. 7). Thus, the results that we have obtained in this study are in accordance with the computational and experimental studies mentioned above that was conducted previously.

The aim of the computational thermal stability work is to design a thermostable protein without changing its activity. In this study we have found that the 3D structure and the binding site of panomycocin has not changed after the substitutions that were performed. Thus, the mutant protein is expected to bind to the same substrate and show the same activity with the wild type panomycocin.

5. Conclusion

The 3D structure of panomycocin was predicted with MODELLER and the best model was optimized and loop modeling was performed. Verification and validation showed the reliability of the model that was generated.

Enhancement of the thermal stability of the model was done using contemporary servers and programs. In the region outside the binding site Leu52 Arg, Phe223Arg and Gly254Arg were found to be the best thermostabilizing mutations with 6.26 K, 6.26 K and 8.27 K increases respectively. In the binding site Glu186Arg was found to be the best thermostabilizer mutation with a 9.58 K temperature increase.

This mutant exo- β -1,3-glucanase can then be used as a novel

antimycotic/antifungal drug in a liquid formulation for topical applications over normal body temperatures. The above mentioned methods can be used in future studies for the detection of the amino acid point mutations that could increase the thermal stability of similar proteins.

Conflict of interest

The authors declare that there is no conflict of interest in this work.

Acknowledgment

This work was supported by a grant from the Scientific and Technological Research Council of Turkey (TUBITAK) (Project no: 115Z376).

References

- Agarwala, R., Barrett, T., Beck, J., Benson, D.A., Bollin, C., Bolton, E., Bourexis, D., Brister, J.R., Bryant, S.H., Canese, K., Charowhas, C., Clark, K., Dicuccio, M., Dondoshansky, I., Federhen, S., Feolo, M., Funk, K., Geer, L.Y., Gorenkov, V., Hoepfner, M., Holmes, B., Johnson, M., Khotomlianski, V., Kimchi, A., Kimelman, M., Kitis, P., Klimke, W., Krasnov, S., Kuznetsov, A., Landrum, M.J., Landsman, D., Lee, J.M., Lipman, D.J., Lu, Z., Madden, T.L., Madej, T., Marchler-Bauer, A., Karsch-Mizrachi, I., Murphy, T., Orris, R., Ostell, J., O'Sullivan, C., Panchenko, A., Phan, L., Preuss, D., Pruitt, K.D., Rodarmar, K., Rubinstein, W., Sayers, E., Schneider, V., Schuler, G.D., Sherry, S.T., Sirotkin, K., Siyan, K., Slotta, D., Soboleva, A., Soussov, V., Starchenko, G., Tatusova, T.A., Todorov, K., Trawick, B.W., Vakarov, D., Wang, Y., Ward, M., Wilbur, W.J., Yaschenko, E., Zbicz, K., 2016. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 44, D7–D19. <https://doi.org/10.1093/nar/gkv1290>.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
- Apweiler, R., Martin, M.J., O'Donovan, C., Magrane, M., Alam-Faruque, Y., Antunes, R., Barrell, D., Bely, B., Bingley, M., Binns, D., Bower, L., Browne, P., Chan, W.M., Dimmer, E., Eberhardt, R., Fazzini, F., Fedotov, A., Foulger, R., Garavelli, J., Castro, L.G., Huntley, R., Jacobsen, J., Kleen, M., Laiho, K., Legge, D., Lin, Q., Liu, W., Luo, J., Orchard, S., Patient, S., Pichler, K., Poggioli, D., Pontikos, N., Pruess, M., Rosanoff, S., Sawford, T., Sehra, H., Turner, E., Corbett, M., Donnelly, M., Van Rensburg, P., Xenarios, I., Bougueleret, L., Auchincloss, A., Argoud-Puy, G., Axelsen, K., Bairoch, A., Baratin, D., Blatter, M.C., Boeckmann, B., Bolleman, J., Bollondi, L., Boutet, E., Quintaje, S.B., Breuza, L., Bridge, A., De Castro, E., Coudert, E., Cusin, I., Doche, M., Dornevil, D., Duvaud, S., Estreicher, A., Famiglietti, L., Feuermann, M., Gehant, S., Ferro, S., Gasteiger, E., Gateau, A., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., Hulo, N., James, J., Jimenez, S., Jungo, F., Kappler, T., Keller, G., Lara, V., Lemercier, P., Lieberherr, D., Martin, X., Masson, P., Moinat, M., Morgat, A., Paesano, S., Pedruzzi, I., Pilboud, S., Poux, S., Pozzato, M., Redaschi, N., Rivoire, C., Roehert, B., Schneider, M., Sigrist, C., Sonesson, K., Staehli, S., Stanley, E., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Veuthey, A.L., Wu, C.H., Arighi, C.N., Arminski, L., Barker, W.C., Chen, C., Chen, Y., Dube, P., Huang, H., Mazumder, R., McGarvey, P., Natale, D.A., Natarajan, T.G., Nchoutmoube, J., Roberts, N.V., Suzek, B.E., Ugochukwu, U., Vinayaka, C.R., Wang, Q., Wang, Y., Yeh, L.S., Zhang, J., 2011. Ongoing and future developments at the universal protein resource. *Nucleic Acids Res.* 39, 214–219. <https://doi.org/10.1093/nar/gkq1020>.
- Bernal, C., Rodríguez, K., Martínez, R., 2018. Integrating enzyme immobilization and protein engineering: an alternative path for the development of novel and improved industrial biocatalysts. *Biotechnol. Adv.* 36, 1470–1480. <https://doi.org/10.1016/j.biotechadv.2018.06.002>.
- Bingding, H., 2009. MetaPocket: a meta approach to improve protein ligand binding site prediction. *OMICS* 13, 325–330. <https://doi.org/10.1089/omi.2009.0045>.
- Capriotti, E., Fariselli, P., Casadio, R., 2005. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* 33, 306–310. <https://doi.org/10.1093/nar/gki375>.
- Cavasotto, C.N., Phatak, S.S., 2009. Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today* 14, 676–683. <https://doi.org/10.1016/j.drudis.2009.04.006>.
- Cheng, J., Randall, A., Baldi, P., 2005. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins Struct. Funct. Bioinf.* 62, 1125–1132. <https://doi.org/10.1002/prot.20810>.
- Chou, K.C., Shen, H.B., 2007. Signal-CF: a subsite-coupled and window-fusing approach for predicting signal peptides. *Biochem. Biophys. Res. Commun.* 357, 633–640. <https://doi.org/10.1016/j.bbrc.2007.03.162>.
- Colovos, C., Yeates, T., 1993. Verification of Protein Structures: Patterns of Nonbonded Atomic Interactions. pp. 1511–1519.
- Cournia, Z., Allen, B., Sherman, W., 2017. Relative binding free energy calculations in drug discovery: recent advances and practical considerations. *J. Chem. Inf. Model.* 57, 2911–2937. <https://doi.org/10.1021/acs.jcim.7b00564>.
- Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., Liang, J., 2006. CASTp: computed atlas of surface topography of proteins with structural and topographical

- mapping of functionally annotated residues. *Nucleic Acids Res.* 34, 116–118. <https://doi.org/10.1093/nar/gkl282>.
- Grevesse, C., Lepoivre, P., Jijakli, M.H., 2003. Characterization of the exoglucanase-encoding gene PaEXG2 and study of its role in the biocontrol activity of *Pichia anomala* strain K. *Phytopathology* 93, 1145–1152. <https://doi.org/10.1094/PHYTO.2003.93.9.1145>.
- Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18, 2714–2723. <https://doi.org/10.1002/elps.1150181505>.
- Hiller, K., Grote, A., Scheer, M., Münch, R., Jahn, D., 2004. PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* 32, 375–379. <https://doi.org/10.1093/nar/gkh378>.
- Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
- Izgu, F., Altunbay, D., 2004. Isolation and characterization of the K5-Type yeast killer protein and its homology with an Exo- β -1,3-glucanase. *Biosci. Biotechnol. Biochem.* 68, 685–693. <https://doi.org/10.1271/bbb.68.685>.
- Izgu, F., Altinbay, D., Sertkaya, A., 2005. Enzymic activity of the K5-type yeast killer toxin and its characterization. *Biosci. Biotechnol. Biochem.* 69, 2200–2206. <https://doi.org/10.1271/bbb.69.2200>.
- Izgu, F., Altinbay, D., Acun, T., 2006. Killer toxin of *Pichia anomala* NCYC 432; purification, characterization and its exo- β -1,3-glucanase activity. *Enzyme Microb. Technol.* 39, 669–676. <https://doi.org/10.1016/j.enzmictec.2005.11.024>.
- Izgu, F., Altinbay, D., Türeli, A.E., 2007. In vitro susceptibilities of *Candida* spp. To Panomycocin, a novel exo- β -1,3-glucanase isolated from *Pichia anomala* NCYC 434. *Microbiol. Immunol.* 51, 797–803. <https://doi.org/10.1111/j.1348-0421.2007.tb03975.x>.
- Izgu, F., Bayram, G., Tosun, K., Izgu, D., 2017. Stratum corneum lipid liposome-encapsulated panomycocin: preparation, characterization, and the determination of antimycotic efficacy against candida spp. Isolated from patients with vulvovaginitis in an in vitro human vaginal epithelium tissue model. *Int. J. Nanomed.* 12, 5601–5611. <https://doi.org/10.2147/IJN.S141949>.
- Khor, B.Y., Tye, G.J., Lim, T.S., Noordin, R., Choong, Y.S., 2014. The structure and dynamics of BmR1 protein from *Brugia malayi*: in silico approaches. *Int. J. Mol. Sci.* 15, 11082–11099. <https://doi.org/10.3390/ijms150611082>.
- Kmiecik, S., Gront, D., Kolinski, M., Wieteska, L., Dawid, A.E., Kolinski, A., 2016. Coarse-grained protein models and their applications. *Chem. Rev.* 116, 7898–7936. <https://doi.org/10.1021/acs.chemrev.6b00163>.
- Krieger, Elmar, Nabuurs, Sander B., Vriend, G., 2012. Homology modeling. In: Bourne, E., Philip, Weissig, H. (Eds.), *Structural Bioinformatics*. Wiley-Liss, pp. 507–520. <https://doi.org/10.1007/978-1-61779-588-6>.
- Krüger, D.M., Rathi, P.C., Pfeleger, C., Gohlke, H., 2013. CNA web server: rigidity theory-based thermal unfolding simulations of proteins for linking structure, (thermo)-stability, and function. *Nucleic Acids Res.* 41, 340–348. <https://doi.org/10.1093/nar/gkt292>.
- Kufareva, I., Abagyan, R., 2012. Methods of protein structure comparison. *Methods Mol. Biol.* 857, 231–257. <https://doi.org/10.1007/978-1-61779-588-6>.
- Kumwenda, B., Lithauer, D., Bishop, Ö.T., Reva, O., 2013. Analysis of protein thermostability enhancing factors in industrially important *Thermus* bacteria species submit a paper. *Evol. Bioinf.* 2013, 327–342. <https://doi.org/10.4137/EBO.S12539>.
- Kuntal, B.K., Aparoy, P., Reddanna, P., 2010a. EasyModeller: a graphical interface to modeller. *BMC Res. Notes* 3, 226. <https://doi.org/10.1186/1756-0500-3-226>.
- Kuntal, B.K., Aparoy, P., Reddanna, P., 2010b. EasyModeller: a graphical interface to EasyModeller: a graphical interface to MODELLER. *BMC Res. Notes* 3, 1–5. <https://doi.org/10.1186/1756-0500-3-226>.
- Laskowski, R.A., Macarthur, M.W., Moss, D.S., T.J., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283–291. <https://doi.org/10.1107/S0021889892009944>.
- Lee, J., Konc, J., Janežič, D., Brooks, B.R., 2017. Global organization of a binding site network gives insight into evolution and structure-function relationships of proteins. *Sci. Rep.* 7, 1–11. <https://doi.org/10.1038/s41598-017-10412-z>.
- Letunic, I., Doerks, T., Bork, P., 2015. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res.* 43, D257–D260. <https://doi.org/10.1093/nar/gku949>.
- Masso, M., Vaisman, I.I., 2010. AUTO-MUTE: web-based tools for predicting stability changes in proteins due to single amino acid replacements. *Protein Eng. Des. Sel.* 23, 683–687. <https://doi.org/10.1093/protein/gzq042>.
- McLachlan, M.J., Johannes, T.W., Zhao, H., 2008. Further improvement of phosphite dehydrogenase thermostability by saturation mutagenesis. *Biotechnol. Bioeng.* 99, 268–274. <https://doi.org/10.1002/bit>.
- Mortazavi, M., Hosseinkhani, S., 2011. Design of thermostable luciferases through arginine saturation in solvent-exposed loops. *Protein Eng. Des. Sel.* 24, 893–903. <https://doi.org/10.1093/protein/gzr051>.
- Muhammed, M.T., Aki-Yalcin, E., 2018. Homology modeling in drug discovery: overview, current applications and future perspectives. *Chem. Biol. Drug Des.* 1–9. <https://doi.org/10.1111/cbdd.13388>.
- Petersen, T.N., Brunak, S., von Heijne, G., N.H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786.
- Pucci, F., Bourgeois, R., Rooman, M., 2016. Predicting protein thermal stability changes upon point mutations using statistical potentials: introducing HoTMuSIC. *Sci. Rep.* 6, 1–9. <https://doi.org/10.1038/srep23257>.
- Rose, P.W., Prlić, A., Altunkaya, A., Bi, C., Bradley, A.R., Christie, C.H., Di Costanzo, L., Duarte, J.M., Dutta, S., Feng, Z., Green, R.K., Goodsell, D.S., Hudson, B., Kalro, T., Lowe, R., Peisach, E., Randle, C., Rose, A.S., Shao, C., Tao, Y.P., Valasatava, Y., Voigt, M., Westbrook, J.D., Woo, J., Yang, H., Young, J.Y., Zardecki, C., Berman, H.M., Burley, S.K., 2017. The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res.* 45, D271–D281. <https://doi.org/10.1093/nar/gkw1000>.
- Roy, A., Yang, J., Zhang, Y., 2012. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Res.* 40, 471–477. <https://doi.org/10.1093/nar/gks372>.
- Schmid, N., Allison, J.R., Gunsteren, W.F.Van, 2011. Biomolecular structure refinement using the GROMOS simulation software. *J. Biomol. NMR* 51, 265–281. <https://doi.org/10.1007/s10858-011-9534-0>.
- Shen, H.B., Chou, K.C., 2007. Signal-3L: a 3-layer approach for predicting signal peptides. *Biochem. Biophys. Res. Commun.* 363, 297–303. <https://doi.org/10.1016/j.bbrc.2007.08.140>.
- Sokalingam, S., Raghunathan, G., Soundararajan, N., Lee, S.-G., 2012. A study on the effect of surface lysine to arginine mutagenesis on protein stability and structure using green fluorescent protein. *PLoS One* 7, e40410. <https://doi.org/10.1371/journal.pone.0040410>.
- Song, J., Tan, H., Perry, A.J., Akutsu, T., Webb, G.I., Whisstock, J.C., Pike, R.N., 2012. PROSPER: an integrated feature-based tool for predicting protease substrate cleavage sites. *PLoS One* 7. <https://doi.org/10.1371/journal.pone.0050300>.
- Strub, C., Alies, C., Lougarre, A., Ladurantie, C., Czapllicki, J., Fournier, D., 2004. Mutation of exposed hydrophobic amino acids to arginine to increase protein stability. *BMC Biochem.* 5, 1–6. <https://doi.org/10.1186/1471-2091-5-9>.
- Talluri, S., 2011. Advances in engineering of proteins for thermal stability. *Int. J. Adv. Biotechnol. Res.* 2, 190–200.
- Van Gunsteren, W.F., B.H.J.C., 2017. The GROMOS Software for (Bio) Molecular Simulation. The GROMOS Software for (Bio) Molecular Simulation.
- Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F., Rarey, M., 2012. Combining global and local measures for structure-based druggability predictions. *J. Chem. Inf. Model.* 52, 360–372. <https://doi.org/10.1021/ci200454v>.
- Walker, G.R., 2011. *Pichia anomala*: cell physiology and biotechnology relative to other yeasts. *Antonie Van Leeuwenhoek* 99, 25–34. <https://doi.org/10.1007/s10482-010-9491-8>.
- Xu, J., Zhang, Y., 2010. How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics* 26, 889–895. <https://doi.org/10.1093/bioinformatics/btq066>.
- Yang, J., Roy, A., Zhang, Y., 2013. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics* 29, 2588–2595. <https://doi.org/10.1093/bioinformatics/btt447>.
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2014. The I-TASSER suite: protein structure and function prediction. *Nat. Methods* 12, 7–8. <https://doi.org/10.1038/nmeth.3213>.
- Yin, S., Ding, F., Dokholyan, N.V., 2007. Eris: an automated estimator of protein stability. *Nat. Methods* 4, 466–467. <https://doi.org/10.1038/nmeth0607-466>.
- Zhou, X.X., Wang, Y.B., Pan, Y.J., Li, W.F., 2008. Differences in amino acids composition and coupling patterns between mesophilic and thermophilic proteins. *Amino Acids* 34, 25–33. <https://doi.org/10.1007/s00726-007-0589-x>.
- Zhou, C., Zhang, M., Wang, Y., Guo, W., Liu, Z., Wang, Y., Wang, W., 2013. Enhancement of the thermo-alkali-stability of xylanase II from *Aspergillus usami* E001 by site-directed mutagenesis. *Afr. J. Microbiol. Res.* 7, 1535–1542. <https://doi.org/10.5897/AJMR12.1561>.