

Article

Temperature-Dependent Structure–Function Properties of Bacterial Xylose Isomerase Enzyme for Food Applications: An In Silico Study

Maurya Sharma ^{1,2,†}, Naayaa Mehta ^{2,3,†}, Renuka Suravajhala ⁴ , Cynthia Meza ⁵, Shrabana Sarkar ⁶  and Aparna Banerjee ^{6,*} 

¹ Interfaculty Institute of Biochemistry (IFIB), Eberhard Karls University of Tübingen, 72074 Tübingen, Germany

² Bioclues, Hyderabad 500072, India

³ The Young Entrepreneurs Academy, The India Shri School, Mousari, Gurgaon 122002, India

⁴ School of Biotechnology, Amrita Vishwa Vidyapeetam, Amritapuri, Kollam 690525, India

⁵ Doctorado en Biotecnología Traslacional (DBT), Facultad de Ciencias Agrarias y Forestales, Universidad Católica del Maule, Talca 3466706, Chile

⁶ Centro de Investigación de Estudios Avanzados del Maule, Vicerrectoría de Investigación y Posgrado, Universidad Católica del Maule, Talca 3466706, Chile

* Correspondence: abanerjee@ucm.cl

† These authors contributed equally to this work.



Citation: Sharma, M.; Mehta, N.; Suravajhala, R.; Meza, C.; Sarkar, S.; Banerjee, A. Temperature-Dependent Structure–Function Properties of Bacterial Xylose Isomerase Enzyme for Food Applications: An In Silico Study. *Clean Technol.* **2022**, *4*, 1317–1329. <https://doi.org/10.3390/cleantechnol4040081>

Academic Editor: Patricia Luis

Received: 11 November 2022

Accepted: 9 December 2022

Published: 14 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Xylose Isomerase (XI) is an intramolecular oxidoreductase enzyme and catalyzes the reversible conversion of ketoses and aldoses in addition to the bioconversion of ethanol from xylose in the production of bioethanol from hemicellulose. It has a broad range of industrial applications in the food and pharmaceutical sectors, particularly in the production of the sweetener high fructose corn syrup (HFCS). It is one of the most widely used industrial enzymes after protease. Taking this into consideration, four bacterial XI sources were selected based on growth temperature, i.e., psychrophile, mesophile, thermophile, and hyperthermophile, for analyzing Xylose Isomerase's structure-function characteristics. It was found that thermophilic XI was structurally less stable than mesophilic and hyperthermophilic XI, whereas structural plasticity ran opposite towards mesophiles. The interaction of xylose isomerase (XI) with two ligands, namely Amino-2-Hydroxymethyl-Propane-1,3-Diol and (4R)-2-Methylpentane-2,4-Diol, was also studied. Mesophilic XI demonstrated better binding affinity with structurally stabilizing amino acids (Ala, Asp, Gly, Leu, and Arg). In comparison, Thermophilic XI showed nearly similar binding affinity with both Amino-2-Hydroxymethyl-Propane-1,3-Diol and (4R)-2-Methylpentane-2,4-Diol. The results of this investigation suggest that thermophilic XI, followed by mesophilic XI, would be the most appropriate for establishing process stability and sustainability in the food industry.

Keywords: xylose isomerase; temperature dependence; structure–function analyses; food applications

1. Introduction

Though their biochemical basis is not clearly understood, microbial enzymes and their biocatalytic potential have gained considerable interest in the pharmaceutical, food and other biotechnological industries due to their applicability in the production of vinegar, wine, bread, etc. [1]. Their stability and catalytic activity have caused extensive investigation on their isolation, purification, characterization, and applications.

During the last decades, microbial-enzyme-dependent bioprocesses have increased exponentially and while many recombinant enzymes from bacteria and fungi are used commercially, xylose isomerase is one of the most important and is as widely used an industrial enzyme as amylase and protease [2]. The addition of inducer (xylose) and Co^{2+} ions to the fermentation medium helps to increase the thermostability of the engaged

enzyme to achieve optimum isomerization of glucose to fructose and increase the process's commercial feasibility [3]. Consequently, recent analysis has reported that XI has a marketplace value of approximately one billion US dollars [4].

Xylose isomerase (EC no. 5.3.1.5) (glucose isomerase, D-xylose ketol isomerase, xylose isomerase or XI, xylose ketol-isomerase, and xylose ketol-isomerase) is widely dispersed in fungi, bacteria, actinomycetes, and even plants [5]. It is an intramolecular oxidoreductase enzyme that plays a key role in the interconversion of aldoses and ketoses [5]. It is widely used in food industries largely for the industrial production of millions of tons of high fructose corn syrup (HFCS) [1,6] and is thus one of the most studied enzymes recently. HFCS has a broad range of uses as it does not cause crystallization in sucrose, and is ~1.3 times sweeter and ~20% cheaper in comparison to sucrose [7]. As a result, it is used for diabetic patients [3]. Furthermore, XI helps to produce bioethanol from hemicellulose by converting xylose to ethanol [8]. Along with natural substrates, XI can isomerize various sugars (pentoses and hexoses), alcohols, and sugar phosphates [3]. Although the substrate specificity can vary for XIs, it shows broad-spectrum activity against different kinds of sugar substrates, such as L-arabinose, D-ribose, L-rhamnulose, and D-allose [5].

However, recent findings have detected some technical difficulties in the isomerization process. For example, the induction of the *Escherichia coli* K12 XI gene in a recombinant system (pRAC expression vector cloned in *E. coli* BL21 cells) exhibited XI production of 40% of the total protein content, but with low thermostability [9]. Until now, the introduction and expression of the bacterial XI gene (*xylA*) in plant systems has only resulted in a >25 fold increase in enzyme production [10].

The aim of this study is to elucidate the structure–function relationship of bacterial XI enzymes so as to combat the technical difficulties at the time of industrialization, using computational analysis to ascertain feasibility and thus commercial usage. For this, XI was selected from four different bacterial sources (hyperthermophile, thermophile, mesophile, and psychrophile) based on their growth temperature. Physicochemical parameters and primary, secondary, and tertiary structures were detected in order to understand the structural stability and validity of the protein models. The functional analysis and molecular docking of the enzymes with small molecules were performed in order to compare structure–function relationships, which may aid in the successful integration of XI in large-scale food industrial processes.

2. Materials and Methods

2.1. Retrieval of the Experimental Sequences

The RCSB PDB Database has protein sequences with crystallographic structure and other information. In this database, ~500 XI sequences of bacterial origin are available based on refinement resolution (Å), classification, and symmetry type. The present study focuses on the structure–function comparison between XIs from different bacteria grown at different temperatures. Therefore, as representatives of each growth temperature, XI enzyme sequences were retrieved from four different bacterial sources: psychrophilic *Paenibacillus* sp (6INT), mesophilic *Streptomyces olivochromogenes* (1XYC), thermophilic *Thermus thermophilus* (1BXB), and hyperthermophilic *Thermotoga neapolitana* (1A0E) from the RCSB PDB open-source database in both FASTA and PDB format.

2.2. Physicochemical Characterization

The online ExPASy-ProtParam server [10] was used for the determination of the physicochemical parameters of the protein, to understand the nature of the protein. In this case, analysis of the physicochemical attributes of XIs was performed by studying the different characters, viz., amino acid (AA) constitution, molecular weight (MW), theoretical isoelectric point (pI), extinction coefficient, instability index (II), aliphatic index (AI), and grand average hydropathy (GRAVY), using ExPASy-ProtParam [10].

2.3. Structural Analysis

2.3.1. Primary Structure Prediction

Primary structure analysis gives an idea of the amino acid composition where AAs are joined by peptide bonds to form a peptide chain [11,12]. To understand this primary protein structure in the laboratory, different qualitative and quantitative methods are used, such as ion exchange chromatography, biuret test, etc. [13]. The computational platform ExPASy-ProtParam [11] was used to determine the amino acid compositions and to calculate the total number of negatively charged and positively charged amino acid residues.

2.3.2. Secondary Structure Analysis

The secondary structure of a protein is determined by the number of hydrogen bonds in the peptide chain forming different secondary structures. Usually, in the laboratory, the secondary structure of the protein is predicted by NMR (Nuclear Magnetic Resonance) spectroscopic study [14]. Here in the computational analysis, different secondary structural forms such as α -helices, β -sheets, β -turns, and random coils were predicted using the online SOPMA server [15] and the FASTA format of the protein sequences.

2.3.3. Tertiary Structure Analysis

Salt bridges within the tertiary structure play an important role in the structural folding and functionality of the protein. In the laboratory, tertiary structure is normally measured by X-ray crystallography. However, computational analysis can be achieved using in silico tools. ESBRI, a web server [16] was used to assess the combination of salt bridges along with the mean distances of the amino acid pairs present in the salt bridge of the tertiary structure of the retrieved protein sequences. In addition, one online tool, TMHMM [17], was used to predict the presence of transmembrane helices in the tertiary structure of the proteins.

2.4. Homology Modeling and Structural Validation

The tertiary structure provides valuable insight into the functionality of the protein on the molecular level. After tertiary structure evaluation and before docking analyses, homology modeling and validation were performed using various online platforms. The SWISS-MODEL workspace [18] was used to perform 3D homology modeling of the recovered proteins. To find the most suitable and best-fitting protein structures based on model quality, three different scoring approaches were used: QMEAN [19], QMEANDisCo [20], and QMEANBrane [21]. The computational server ERRAT [22] was used to confirm the retrieved crystallographic protein structures by statistically analyzing the non-bonded interaction among different atoms present in the protein sequence [23]. Three-dimensional model structures of the studied proteins were also validated by generating Ramachandran plots using a web-based server, PROCHECK, which helps to find energetically allowed regions for backbone dihedral angles ψ against ϕ of the amino acids in the peptide chain [24]. RAMPAGE was used to calculate the percentage of amino acids present in the favorable region, allowed region, and outlier region of the proteins [25].

2.5. Functional Analysis

Motifs of the proteins are important to predict their functionality. Motifs present in the presently studied proteins were identified with the help of the online web-based tool MEME suite (Multiple Extraction-Maximization for Motif Elicitation) [26]. All the selected sequences of xylose isomerase were analyzed using the online tool InterProScan to classify them according to the relationship between protein family and important sites of present functional domains [27].

2.6. Molecular Docking Analysis

2.6.1. Preparation of Amino-2-Hydroxymethyl-Propane-1,3-Diol and (4R)-2-Methylpentane-2,4-Diol Ligands

Protein Data Bank in Europe (PDBe), a global repository of macromolecular structural models (<https://www.ebi.ac.uk/pdbe/>; accessed on 16 July 2022), was used to find enzyme-specific ligands for molecular docking analysis. Three-dimensional structures of the two selected ligands—Amino-2-Hydroxymethyl-Propane-1,3-Diol ($C_4H_{12}NO_3$) with a molecular weight of 122.143 Da, and (4R)-2-Methylpentane-2,4-Diol ($C_6H_{14}O_2$) with a molecular weight of 118.174 Da—were retrieved in PDB format. These two ligands were chosen for docking analysis based on their binding affinity with proteins as large molecules.

2.6.2. Molecular Docking to Investigate Protein–Ligand Interaction

For cavity-detection guided blind docking, the CB-Dock2 server [28,29] was used. It helped to predict the binding sites of selected proteins by calculating the sizes of the center with a novel curvature-based cavity detection technique [30]. To perform the analyses, both the protein and ligand files were used in PDB format and five possible coupling cavities were identified. The cavity with the lowest binding energy was selected for each protein for the docking analysis. The ligands and proteins were then visualized with the Ball, Stick, and Surface options, respectively.

3. Results and Discussions

3.1. Retrieval of the Experimental Sequences

Sequences of XI enzymes isolated from psychrophilic, mesophilic, thermophilic, and hyperthermophilic bacteria were retrieved from the RCSB-PDB open-source data bank in FASTA and PDB format. These enzymes are from psychrophile *Paenibacillus* sp. R4 (6INT), mesophile *Streptomyces olivochromogenes* (1XYC), thermophile *Thermus thermophilus* HB8 (1BXB), and hyperthermophile *Thermotoga neapolitana* (1AOE).

3.2. Physicochemical Characterization

Understanding of physicochemical characters along with a comparison between negatively charged amino acids and positively charged amino acids are preliminary steps needed to predict the behavior and nature of the studied protein. In our study, the XI enzyme isolated from four different bacterial sources (psychrophile, mesophile, thermophile, and hyperthermophile) was characterized by distinct physicochemical parameters (Table 1). For proteins alkaline in nature, the theoretical isoelectric point (pI) value will be more than 7; whereas it will be less than 7 for a protein acidic in nature [31]. In this study, all four bacterial isolates have a PI value of less than 7 [Mesophile (4.98) < thermophile (5.33) < psychrophile (5.34) < hyperthermophile (5.47)] indicating that the protein contains more acidic amino acids (aspartic acids or glutamic acids). This result highlights that the mesophile is the most acidic in nature among all studied XI. The stable nature of the studied protein can be predicted from the Instability index value below 40 [32]. XI isolated from the hyperthermophile (29.61) was found to be the most stable with the lowest instability index followed by the thermophile (30.01), psychrophile (30.77), and mesophile (32.95). A higher aliphatic index indicates the thermostable nature of globular proteins based on the amino acid composition of the alanine, valine, and leucine amino acids. XI isolated from the psychrophile is the most thermostable, because of the presence of more aliphatic amino acids. A negative GRAVY value (−0.272, −0.381, −0.411, and −0.377 for the psychrophile, mesophile, thermophile, and hyperthermophile, respectively) indicates the hydrophilic nature of proteins and confirms a better interaction between protein and water [33]. Both the number of positively charged amino acids as well as negatively charged amino acids were found to be highest in hyperthermophile. In addition, the total number of positively charged amino acids was much less than the total number of negatively charged amino acids in all four proteins. Thus, considering the physicochemical parameters overall, the hyperthermophilic XI showed the best structural stability.

Table 1. Comparative physicochemical characteristics of four chosen bacterial xylose isomerase (XI) enzymes from a psychrophile, mesophile, thermophile, and hyperthermophile.

Serial No.	Physicochemical Characters							
	PDB ID	Bacterial Isolates	Number of AA	Theoretical PI	MW (Da)	II	AI	GRAVY
1.	Psychrophile (6INT)	<i>Paenibacillus</i> sp. R4	438	5.34	48,880.99	30.77	84.95	−0.272
2.	Mesophile (1XYC)	<i>Streptomyces olivochromogenes</i>	386	4.98	42,791.95	32.95	78.24	−0.381
3.	Thermophile (1BXB)	<i>Thermus thermophilus</i> HB8	387	5.33	43,906.75	30.01	80.26	−0.411
4.	Hyperthermophile (1A0E)	<i>Thermotoga neapolitana</i>	443	5.47	50,761.77	29.61	79.53	−0.377

AA—Amino Acids; MW—Molecular Weight; II—Instability Index; AI—Aliphatic Index; GRAVY—Grand Average Hydropathy.

3.3. Structural Analysis

3.3.1. Primary Structure Prediction

From the primary structural analysis, it was found that the aliphatic amino acids leucine (11.1% in thermophile, 10.5% in psychrophile, 3.8% in mesophile, 3% in hyperthermophile) and alanine (11.3% in thermophile, 10.5% in psychrophile, 12.7% in mesophile, 7.3% in hyperthermophile) were present in all four XI sequences. The difference in percentages of amino acids contributing to protein structure is represented in Figure 1A. Both positively and negatively charged amino acids were the highest in the hyperthermophile. In addition, the total number of negatively charged amino acids (Asp + Glu) was much higher than the total number of positively charged amino acids (Arg + Lys) in all four proteins (Figure 1B). Ala, Asp, Gly, Leu, and Arg amino acids were present in all protein structures, hence stabilizing the protein structures. Cysteine is hydrophobic and polar in nature [13]. The results confirmed that the studied proteins contain high amounts of aliphatic amino acids with net negative charge and contain no cysteine residues at all. This finding indicates that all four studied proteins are intracellular, nonpolar, and hydrophilic in nature.

3.3.2. Secondary Structure Prediction

In the computational study, the secondary structures of the four bacterial XIs were predicted using the web-based server SOPMA [16]. All the proteins were rich in α -helices (50.68%, 46.89%, 47.03%, and 50.34%, respectively, for the psychrophile, mesophile, thermophile, and hyperthermophile) followed by random coils (32.88%, 35.23%, 36.95%, and 32.05% for the psychrophile, mesophile, thermophile, and hyperthermophile, respectively). The percentage of the extended strand was 10.50%, 12.95%, 11.37%, and 10.84% for the psychrophile, mesophile, thermophile, and hyperthermophile, respectively (Figure 1C). This result predicts that the selected proteins are mainly rich in α -helices and random coils with beta strands having the least presence. A higher percentage of amino acids forming α -helices indicates the temperature stability of any enzyme [34]. On the other hand, a higher percentage of random coils denotes that the enzyme is functionally active [13]. According to this study, thermophilic-origin XI was found to be most functionally active, whereas psychrophilic-origin XI had the most α -helices, i.e., the best temperature stability [34–36].

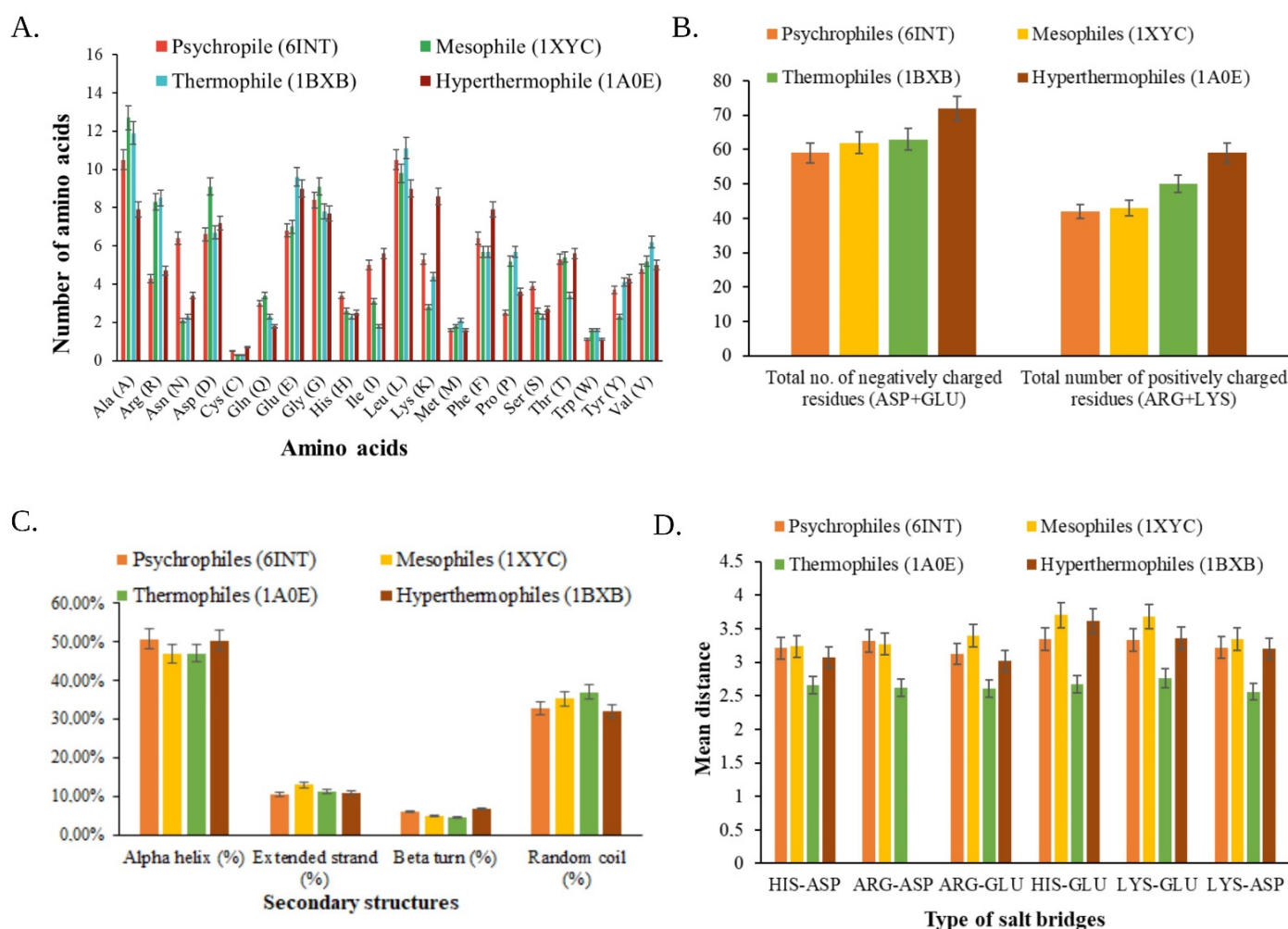


Figure 1. Diagrammatic representation of comparison of the primary, secondary, and tertiary structure between xylose isomerase from four different classes (psychrophile, mesophile, thermophile, and hyperthermophile). (A) percentages of amino acids contributing to structural configuration, (B) number of positively charged (ARG + LYS) and negatively charged (ASP + GLU) amino acid residues, (C) different kinds of secondary structures, (D) salt bridge composition along with mean distances.

3.3.3. Analysis of Tertiary Structure

The web-based ESBRI platform was used to compute the number of salt bridges as well as to calculate the mean distance between the salt-bridge-forming amino acid residues (Figure 1D). Salt bridges are the hydrogen bonding between positively charged (Arg, Lys, His) and negatively charged amino acids (Glu, Asp) with a maximum interatomic distance of 7.0 Angstrom (Å) [37]. As salt bridges have an important role in structural stability, the presence of a sufficient number of different salt bridges in all the chosen XIs proves that they are structurally stable [13]. The most common salt bridges for the mesophile were HIS-GLU (3.7 Å), LYS-GLU (3.68 Å), and ARG-GLU (3.4 Å); for the hyperthermophile HIS-GLU (3.62 Å); and in the case of the thermophile and psychrophile, all possible salt bridges (HIS-ASP, ARG-ASP, ARG-GLU, HIS-GLU, LYS-GLU, and LYS-ASP) were present with a mean distance between 2.6 and 3.7 Å. From the results, it can be inferred that the mesophile is the most stable followed by the hyperthermophile, psychrophile, and thermophile. Transmembrane helix determination helps to understand the location of the protein, i.e., the presence of the transmembrane helix confirms the transmembrane nature whereas the absence of it determines the cytoplasmic nature of the selected protein [13,17]. From the results, it was found that all the selected XI have no transmembrane helices

present in their tertiary structure, thus confirming the cytoplasmic location of the selected XI enzymes.

3.4. Homology Modeling and Structural Validation

The model quality score was analyzed using the ERRAT server for evaluating the error region in crystallographic structure through statistically investigating non-bonded atomic interactions. An error value >95 denotes a crystallographic, structurally stable protein model [22]. Interestingly, overall quality was found to be better for the mesophilic XI (98.95) followed by the thermophile (97.62), hyperthermophile (96.49), and psychrophile (96.15) (Table 2). The stereochemical validation of the selected protein structures was performed by generating Ramachandran plots through the PROCHECK platform [24]. ϕ - ψ torsion angles were plotted against each other in the Ramachandran plot where, among four quadrants, quadrant 1 contains conformations in the allowed region with rare left-handed α -helices, quadrant 2 has the most favorable sterically allowed regions with the presence of β -strands, quadrant 3 of the plot contains right handed α -helices, and the disfavored region is located in quadrant 4 of the plot.

Table 2. Comparative assessment of structural quality of four chosen bacterial xylose isomerase (XI) enzymes from a psychrophile, mesophile, thermophile, and hyperthermophile.

Serial No.	Quality Assessment Scores					
	PDB ID	Bacterial Isolates	3D-1D Score (%)	ERRAT Quality Factor	QMEAN Z-Score	AA in FR of Ramachandran Plot (%)
1.	Psychrophile (6INT)	<i>Paenibacillus</i> sp. R4	88.70	96.15	−0.30	91.4
2.	Mesophile (1XYC)	<i>Streptomyces olivochromogenes</i>	91.06	98.95	1.25	92.6
3.	Thermophile (1BXB)	<i>Thermus thermophilus</i> HB8	87.47	97.62	0.43	89.9
4.	Hyperthermophile (1A0E)	<i>Thermotoga neapolitana</i>	88.71	96.49	−0.28	92.4

FR—Favorable region.

By studying the generated Ramachandran plots it was found that >95% of residues were present in the favored region of all the selected XI structures, i.e., 92.4% for the hyperthermophilic XI, 89.9% for the thermophilic XI, 91.4% for the psychrophilic XI, and 92.6% for the mesophilic XI (Table 2). This result supports the satisfactory quality of the selected protein models. These selected four models are reliable, as they had no or very few amino acid residues in the disallowed region of the Ramachandran plot. These stereochemical properties were confirmed in RAMPAGE via generating the plot in PstScript format [38]. Examining the plot, it was found that the hyperthermophilic XI had 96.82%, the thermophilic XI had 96.56%, the psychrophilic XI had 96.31%, and the mesophilic XI had 95.96% residues in the highly preferred region of the protein (Figure 2). In the case of all four studied models, >90% of residues lie in favorable regions of the Ramachandran plot, confirming the stable nature of the enzymes [38].

The SWISS-MODEL QMEAN tool was used to estimate the global model quality of the proteins. The target protein structures were compared with non-redundantly available PDB structures to predict the overall quality (Figure 2). The QMEAN Z-score value was −0.30 for the psychrophile, 1.25 for the mesophile, 0.43 for the thermophile, and −0.28 for the hyperthermophile. The expected QMEAN score and Z-average value of a protein should lie within a range of 0–1 [39] and <1 [40], respectively. The QMEAN result accurately showed that all the target proteins lie in the expected region except for the mesophilic XI. QMEAN DisCo scores normally predict the structural model's consistency based on statistical calculation; a score within the range of 0–1 denotes better model quality [20]. The QMEAN DisCo scores for the present study (0.86 ± 0.05 for the psychrophile, 0.88 ± 0.05 for the mesophile, 0.92 ± 0.05 for the thermophile, and 0.86 ± 0.05 for the hyperthermophile) showed good local model quality and reliability with the residues along with XI sequences.

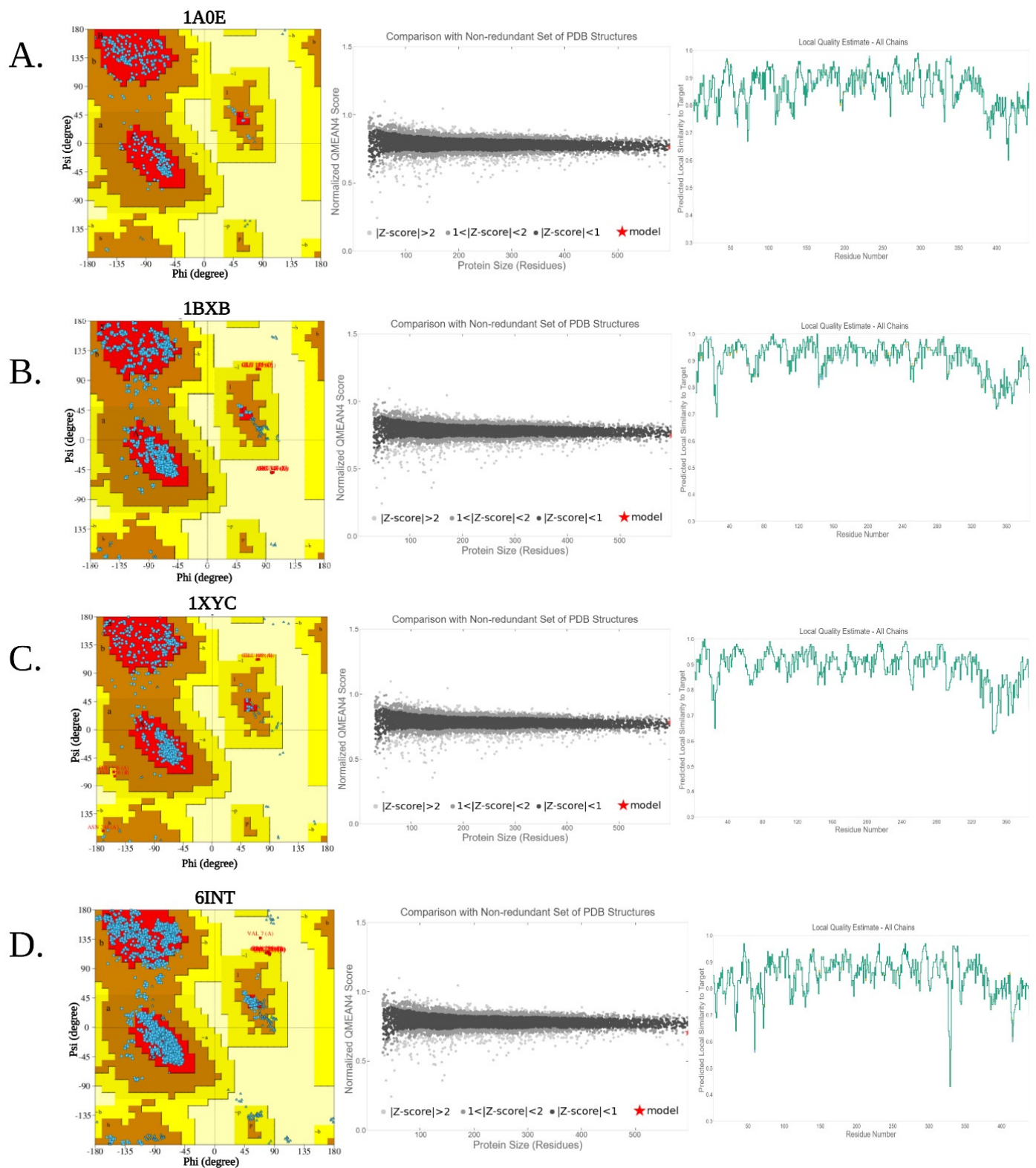


Figure 2. Quality assessment from Ramachandran plot, evaluation of Z-score, and local model quality estimation of (A) hyperthermophile, (B) thermophile, (C) mesophile, and (D) psychrophile.

3.5. Functional Analysis

In silico functional analysis was performed by finding the functional motifs present in the XI sequences. The MEME suite helps find the functional motifs required for the biological activity of any enzyme [13]. In order to identify any protein, motifs are used

as signature sequences. The e-value of the predicted motif helps understand the level of functional accuracy. If the e-value is low, then the predicted motifs are more accurate and vice versa [41]. MEME additionally helps predict the consensus sequence, occurrences of sites, and the level of conservation at each position in the pattern [26]. In our present study, XI of all four different bacterial isolates had motifs with functionally accurate e-values i.e., 1.3×10^1 for the thermophile, 1.4×10^3 for the hyperthermophile, 1.5×10^2 for the mesophile, and 1.7×10^2 for the psychrophile.

Interproscan helps to classify protein families based on functional genomes [42]. In the case of XI, it was observed that all four chosen proteins have similar sequences. For example, the signature sequence for the bacterial XI family, which confirms the classification of the protein family, is the XI-like TIM barrel domain. This confirms the functionality of the enzymes in catalyzing the isomerization reaction between ketoses and aldoses, genes for XI-like homologous superfamily, and few other unintegrated genomes. The hyperthermophilic XI played a role in the carbohydrate (D-xylose) metabolic process to form ethanol and was present in the bacterial XI family. Being present within the cytoplasm, it promoted intramolecular oxidoreductase activity followed by isomerase activity and catalytic activity, leading to its molecular function. Magnesium ions bind with the enzyme to help in catabolic reactions. XI of thermophilic bacteria was found to actively participate in the D-xylose metabolic process and carbohydrate metabolic process, with molecular xylose isomerase activity similar to that of the previous one. XI of mesophilic bacteria also had a similar kind of biological process and molecular function in the cell cytoplasm to that of thermophilic bacteria. The psychrophilic XI had similar functional properties to those of hyperthermophilic xylose isomerase.

3.6. Molecular Docking to Investigate Protein–Ligand Interaction

Molecular docking helps to effectively predict the relationship between protein structure and ligand molecules [43]. The biological activity of the small molecules with proteins can be determined with docking analysis [44]. The results of CB-DOCK2 are presented in Figure 3. Before docking, the small molecules/ligands [Amino-2-Hydroxymethyl-Propane-1,3-Diol and (4R)-2-Methylpentane-2,4-Diol] were prepared. Substrate binding affinity is an important characteristic for determining binding capacity with the specific enzymes. Four parameters in the field of chemoinformatics, namely, absorption, distribution, metabolism, and excretion (ADME), are required for targeted binding assessment with ligands [45,46]. Enzymes play important roles in cellular biological processes with high industrial value, such as XI in bioethanol production. Hence, mechanistically appropriate binding affinity of the substrate with the enzyme need to be aptly predicted for its future industrial usage [47].

Five possible binding sites of the proteins to interact with the small molecules (ligands) were predicted by calculating the centers and sizes with a novel curvature-based cavity-detection approach using Autodock Vina [30]. The curvature-dependent surface-area model [48] of large molecules (XIs) with small molecules [(4R)-2-Methylpentane-2,4-Diol and Amino-2-Hydroxymethyl-Propane-1,3-Diol] was made using predicted binding affinity interactions between proteins and ligands. Among all predicted binding scores (Vina scores), the interactions with the lowest binding energy were selected (Figure 3). The binding affinity of all the interactions was found to be good. The hyperthermophile had a binding energy of -4.5 KJ/mol for (4R)-2-Methylpentane-2,4-Diol and -5.3 KJ/mol for Amino-2-Hydroxymethyl-Propane-1,3-Diol. The thermophile had a binding energy of -5.1 KJ/mol for both (4R)-2-Methylpentane-2,4-Diol and Amino-2-Hydroxymethyl-Propane-1,3-Diol. The mesophile had a binding energy of -5.4 KJ/mol for (4R)-2-Methylpentane-2,4-Diol and -4.4 KJ/mol for Amino-2-Hydroxymethyl-Propane-1,3-Diol. Finally, the psychrophile had a binding energy of -4.5 KJ/mol for (4R)-2-Methylpentane-2,4-Diol and -5.1 KJ/mol for Amino-2-Hydroxymethyl-Propane-1,3-Diol.

(Figure 3). The ligand was colored by element (blue: N, red: O, grey: C, and white: H) attached to O or N atoms (Figure 3). The ligand formed a closed cavity-like structure with the help of hydrogen bonds near the end of the coil and atoms of the aromatic ring, which disappear at the time of protein–ligand docking [12].

4. Conclusions

The present study suggests that the chosen four model proteins of XI from different bacterial origins based on growth temperatures are of good quality with stable crystallographic structure. Structural analysis confirms the strict cytoplasmic nature of the bacterial XI enzymes. In addition, they are also thermostable and hydrophilic. From the overall comparison, it was found that the psychrophile is more much structurally stable than the other three bacterial XI. All four enzymes have more than 95% amino acid residues in the energetically favorable region of the Ramachandran plot. From the present in silico study, it was found that the psychrophilic and hyperthermophilic XI are more thermally stable than the mesophile, whereas, because of the presence of random coils, the thermophile and mesophile are more functionally active compared to the other two XIs. From structural validation, it was found that all the XIs are structurally stable with very little variation in their values. Mesophilic XI showed minimal binding affinity with the small molecules [only (4R)-2-Methylpentane-2,4-Diol] denoting the function of the active site of the protein. In addition, the thermophile also has minimum binding affinity for both the ligands [(4R)-2-Methylpentane-2,4-Diol and Amino-2-Hydroxymethyl-Propane-1,3-Diol respectively]. Hence, from the functional point of view, the mesophilic and thermophilic XI can be efficiently used in the food industry in the future. Finally, this comparative in silico structure–function analysis of bacterial XIs will help in designing the necessary laboratory setup to conduct industrial experiments properly in the future by choosing the right candidates.

Author Contributions: Conceptualization, A.B.; methodology, M.S. and N.M.; formal analysis, M.S., N.M. and S.S.; data curation, M.S. and N.M.; writing—original draft preparation, A.B. and S.S.; writing—review and editing, S.S., A.B., C.M. and R.S.; supervision, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: M.S., N.M. and A.B. are thankful to the Bioclues Organization, India for the work carried out as a part of the mentor–mentee program, and to Bioinformatics for School children (BIXS).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Singh, R.; Kumar, M.; Mittal, A.; Mehta, P.K. Microbial enzymes: Industrial progress in 21st century. *3 Biotech* **2016**, *6*, 174. [[CrossRef](#)] [[PubMed](#)]
2. Deshpande, V.; Rao, M. Glucose Isomerase. In *Enzyme Technology*; Springer: New York, NY, USA, 2006; pp. 239–252. [[CrossRef](#)]
3. Bhosale, S.H.; Rao, M.B.; Deshpande, V.V. Molecular and industrial aspects of glucose isomerase. *Microbiol. Rev.* **1996**, *60*, 280–300. [[CrossRef](#)] [[PubMed](#)]
4. Al-Dhabi, N.A.; Esmail, G.A.; Ghilan, A.K.M.; Arasu, M.V. Isolation and screening of *Streptomyces* sp. Al-Dhabi-49 from the environment of Saudi Arabia with concomitant production of lipase and protease in submerged fermentation. *Saudi J. Biol. Sci.* **2020**, *27*, 474–479. [[CrossRef](#)] [[PubMed](#)]
5. Nam, K.H. Glucose isomerase: Functions, structures, and applications. *Appl. Sci.* **2022**, *12*, 428. [[CrossRef](#)]
6. Kilara, A.; Shahani, K.M.; Shukla, T.P. The use of immobilized enzymes in the food industry: A review. *Crit. Rev. Food Sci. Nutr.* **1979**, *12*, 161–198. [[CrossRef](#)] [[PubMed](#)]

7. Staudigl, P.; Haltrich, D.; Peterbauer, C.K. L-Arabinose isomerase and D-xylose isomerase from *Lactobacillus reuteri*: Characterization, coexpression in the food grade host *Lactobacillus plantarum*, and application in the conversion of D-galactose and D-glucose. *J. Agric. Food Chem.* **2014**, *62*, 1617–1624. [\[CrossRef\]](#)
8. Singh, R.S.; Singh, T.; Pandey, A. Microbial Enzymes—An Overview. In *Advances in Enzyme Technology*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 1–40.
9. Rozanov, A.S.; Zagrebelyi, S.N.; Beklemishchev, A.B. Cloning of *Escherichia coli* K12 xylose isomerase (glucose isomerase) gene and studying the enzymatic properties of its expression product. *Prikl. Biokhim. Mikrobiol.* **2009**, *45*, 38–44. [\[CrossRef\]](#)
10. Haldrup, A.; Noerremark, M.; Okkels, F.T. Plant selection principle based on xylose isomerase. *Vitr. Cell. Dev. Biol.—Plant* **2001**, *37*, 114–119. [\[CrossRef\]](#)
11. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; John, M.W., Ed.; Humana Press: Totowa, NJ, USA, 2005; pp. 571–607.
12. Flores-Castañón, N.; Sarkar, S.; Banerjee, A. Structural, functional, and molecular docking analyses of microbial cutinase enzymes against polyurethane monomers. *J. Hazard. Mater. Lett.* **2022**, *3*, 100063. [\[CrossRef\]](#)
13. Sarkar, S.; Banerjee, A.; Chakraborty, N.; Soren, K.; Chakraborty, P.; Bandopadhyay, R. Structural-functional analyses of textile dye degrading azoreductase, laccase and peroxidase: A comparative in silico study. *Electron. J. Biotechnol.* **2020**, *43*, 48–54. [\[CrossRef\]](#)
14. Pramanik, K.; Kundu, S.; Banerjee, S.; Ghosh, P.K.; Maiti, T.K. Computational-based structural, functional and phylogenetic analysis of *Enterobacter phytases*. *3 Biotech* **2018**, *8*, 262. [\[CrossRef\]](#)
15. Geourjon, C.; Deleage, G. SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics* **1995**, *11*, 681–684. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Costantini, S.; Colonna, G.; Facchiano, A.M. ESBRI: A web server for evaluating salt bridges in proteins. *Bioinformation* **2008**, *3*, 137. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Sonnhammer, E.L.; Von Heijne, G.; Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1998**, *6*, 175–182. [\[PubMed\]](#)
18. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Benkert, P.; Biasini, M.; Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **2011**, *27*, 343–350. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Studer, G.; Rempfer, C.; Waterhouse, A.M.; Gumienny, R.; Haas, J.; Schwede, T. QMEANDisCo—Distance constraints applied on model quality estimation. *Bioinformatics* **2020**, *36*, 1765–1771. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Studer, G.; Biasini, M.; Schwede, T. Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics* **2014**, *30*, i505–i511. [\[CrossRef\]](#)
22. Colovos, C.; Yeates, T.O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* **1993**, *2*, 1511–1519. [\[CrossRef\]](#)
23. Caimano, M.J.; Sivasankaran, S.K.; Allard, A.; Hurley, D.; Hokamp, K.; Grassmann, A.A.; Hinton, J.C.D.; Nally, J.E. A model system for studying the transcriptomic and physiological changes associated with mammalian host-adaptation by *Leptospira interrogans* serovar Copenhageni. *PLoS Pathog.* **2014**, *10*, e1004004. [\[CrossRef\]](#)
24. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **1993**, *26*, 283–291. [\[CrossRef\]](#)
25. Lovell, S.C.; Davis, I.W.; Arendall, W.B., III; de Bakker, P.I.W.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by C α geometry: ϕ , ψ and C β deviation. *Proteins* **2003**, *50*, 437–450. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME suite. *Nucleic Acids Res.* **2015**, *43*, W39–W49. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Quevillon, E.; Silventoinen, V.; Pillai, S.; Harte, N.; Mulder, N.; Apweiler, R.; Lopez, R. InterProScan: Protein domains identifier. *Nucleic Acids Res.* **2005**, *33*, W116–W120. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Liu, Y.; Yang, X.; Gan, J.; Chen, S.; Xiao, Z.-X.; Cao, Y. CB-Dock2: Improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res.* **2022**, *50*, W159–W164. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Yang, X.; Liu, Y.; Gan, J.; Xiao, Z.-X.; Cao, Y. FitDock: Protein–ligand docking by template fitting. *Brief. Bioinform.* **2022**, *23*, bbac087. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Liu, Y.; Grimm, M.; Dai, W.-T.; Hou, M.-C.; Xiao, Z.-X.; Cao, Y. CB-Dock: A web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacol. Sin.* **2020**, *41*, 138–144. [\[CrossRef\]](#)
31. Hasan, R.; Rony, M.N.H.; Ahmed, R. In silico characterization and structural modeling of bacterial metalloprotease of family M4. *J. Genet. Eng. Biotechnol.* **2021**, *19*, 25. [\[CrossRef\]](#)
32. Gamage, D.G.; Gunaratne, A.; Periyannan, G.R.; Russell, T.G. Applicability of Instability Index for In vitro Protein Stability Prediction. *Protein Pept. Lett.* **2019**, *26*, 339–347. [\[CrossRef\]](#)
33. Pramanik, K.; Soren, T.; Mitra, S.; Maiti, T.K. In silico structural and functional analysis of Mesorhizobium ACC deaminase. *Comput. Biol. Chem.* **2017**, *68*, 12–21. [\[CrossRef\]](#)
34. Petukhov, M.; Kil, Y.; Kuramitsu, S.; Lanzov, V. Insights into thermal resistance of proteins from the intrinsic stability of their α -helices. *Proteins* **1997**, *29*, 309–320. [\[CrossRef\]](#)

35. Kumar, S.; Tsai, C.-J.; Nussinov, R. Factors enhancing protein thermostability. *Protein Eng. Des. Sel.* **2000**, *13*, 179–191. [[CrossRef](#)] [[PubMed](#)]
36. Escobedo, A.; Topal, B.; Kunze, M.B.A.; Aranda, J.; Chiesa, G.; Mungianu, D.; Bernardo-Seisdedos, G.; Eftekharzadeh, B.; Gairí, M.; Pierattelli, R.; et al. Side chain to main chain hydrogen bonds stabilize apolyglutamine helix in a transcription factor. *Nat. Commun.* **2019**, *10*, 2034. [[CrossRef](#)]
37. Kumar, S.; Nussinov, R. Salt bridge stability in monomeric proteins. *J. Mol. Biol.* **1999**, *293*, 1241–1255. [[CrossRef](#)] [[PubMed](#)]
38. Notredame, C.; Higgins, D.G.; Heringa, J. T-coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **2000**, *302*, 205–217. [[CrossRef](#)] [[PubMed](#)]
39. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)]
40. Benkert, P.; Künzli, M.; Schwede, T. QMEAN server for protein model quality estimation. *Nucleic Acids Res.* **2009**, *37*, W510–W514. [[CrossRef](#)]
41. Behbahani, M.; Rabiei, P.; Mohabatkar, H. A Comparative Analysis of Allergen Proteins between Plants and Animals Using Several Computational Tools and Chou's pseAAC Concept. *Int. Arch. Allergy Immunol.* **2020**, *181*, 813–821. [[CrossRef](#)]
42. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)]
43. Abdelhedi, O.; Nasri, R.; Mora, L.; Jridi, M.; Toldrá, F.; Nasri, M. In silico analysis and molecular docking study of angiotensin I-converting enzyme inhibitory peptides from smooth-hound viscera protein hydrolysates fractionated by ultrafiltration. *Food Chem.* **2018**, *239*, 453–463. [[CrossRef](#)]
44. Gezen, H.; Gürdere, M.B.; Dinçer, A.; Özbek, O.; Koçyiğit, Ü.M.; Taslimi, P.; Tüzün, B.; Budak, Y.; Ceylan, M. Synthesis, molecular docking, and biological activities of new cyanopyridine derivatives containing phenylurea. *Arch. Pharm.* **2021**, *354*, 2000334. [[CrossRef](#)]
45. Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 42717. [[CrossRef](#)] [[PubMed](#)]
46. Bartuzi, D.; Kaczor, A.A.; Targowska-Duda, K.M.; Matosiuk, D. Recent advances and applications of molecular docking to G protein-coupled receptors. *Molecules* **2017**, *22*, 340. [[CrossRef](#)] [[PubMed](#)]
47. Das, S.; Shimshi, M.; Raz, K.; Eliaz, N.N.; Mhashal, A.R.; Ansbacher, T.; Major, D.T. Enzydock: Protein–ligand docking of multiple reactive states along a reaction coordinate in enzymes. *J. Chem. Theory Comput.* **2019**, *15*, 5116–5134. [[CrossRef](#)] [[PubMed](#)]
48. Cao, Y.; Li, L. Improved protein–ligand binding affinity prediction by using a curvature-dependent surface-area model. *Bioinformatics* **2014**, *30*, 1674–1680. [[CrossRef](#)] [[PubMed](#)]