Journal of Engineering Science and Technology
Vol. 15, No. 6 (2020) 4259 - 4272

SEQUENCE ANALYSIS, MULTIPLE TEMPLATES STRUCTURE PREDICTION AND BINDING SITE IDENTIFICATION OF BROMELAIN FROM ANANAS COMOSUS

FATAHIYA M. TAP[1,*], NURUL B. A. KHAIRUDIN[2], FADZILAH A. A. MAJID[3], NUR S. A. RAHMAN[1], ZARINA OMAR[1]

[1]Faculty of Chemical Engineering, Universiti Teknologi MARA Terengganu Branch, Bukit Besi Campus, 23200 Dungun, Malaysia
[2]Malaysia Japan International Institute Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia
[3]Institute of Marine Biotechnology, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia
*Corresponding Author: fatahiya@uitm.edu.my

Abstract

Bromelain is a proteolytic enzyme derived from pineapples and categorized as a cysteine protease. This enzyme is widely used as an anti-inflammatory agent in therapeutic applications. In this study, the molecular interactions of bromelain against phospholipase A2 (Pla2), a target protein in inflammatory diseases, were investigated. The knowledge of the structural properties and mechanism of bromelain is limited because of the unavailability of its structural information. Therefore, a comparative modelling study was conducted using MODELLER 9v14 to predict the three-dimensional (3D) model of stem bromelain and to investigate its structural properties. The 3D model of bromelain was successfully predicted through Homology modelling approach. Pocket detection was conducted, functionally and structurally important residues of the bromelain model were identified, and its cleft and ligand binding site were determined. Protein-protein docking was performed to predict all possible binding modes between bromelain and Pla2. Amino acids involved in the interactions between bromelain and Pla2 were identified. A few functional conserved residues located in the largest cleft of the model were involved in bromelain-Pla2 binding interactions

Keywords: Binding site prediction, Bromelain, Homology modelling, Sequence analysis.

1. Introduction

Bromelain is an enzyme found in pineapples (Ananas comosus) belonging to Bromeliaceace. This enzyme can be extracted from various parts of pineapples including leaves, crown, bark, and stem [1], and can be categorized as a protease because it refers to a group of enzyme that hydrolyses the peptide bonds of proteins. In the presence of oxidizing agents, the sulfhydryl group of cysteine becomes oxidized, and a disulfide bond forms. In addition, bromelain possesses a catalytic-diad that consists of nucleophiles, such as cysteine and histidine that are evolutionarily conserved in all cysteine protease groups.

Cysteine protease are present in all living organisms. This enzyme can be found in viruses, bacteria, fungi and plants. Papain is one of the family of cysteine protease from plants that have been characterized and have known crystal structure. The papain family contains peptidases which are structurally related to the papain. Papain is characterized by a two-domain structure and the active site (Cys25 and His159) is located between the domain [2]. The enzymatic activity of cysteine protease is related to the presence of the catalytic diad formed by the cysteine and histidine. The main activity of this enzyme is metabolic degradation of peptides and proteins. Bromelain is categorized as cysteine protease has been widely investigated because of its exploitable properties and extensive applications in therapeutics and food industries [3-5].

In therapeutic applications, bromelain is used as an anti-inflammatory agent because of its tendency to inhibit inflammatory activities from a selected target protein [6]. In food industries, this enzyme is used as meat tenderizer, anti-browning agent, and food supplement [7-9]. Although bromelain has many uses, its functions and chemical mechanism have yet to be fully understood, especially at molecular levels. To date, a crystal structure of bromelain is yet to be available for researchers. Therefore, this study aimed to conduct sequence analysis, predict the tertiary structure of bromelain, and identify important functional residues in ligand binding sites of bromelain using in silico approaches. Since bromelain shows a great potential as an anti-inflammatory agent, thus the finding from this study could be used for further investigation on the structural and functional properties of this enzymes using experimental and molecular modelling approaches.

2. Materials and Methods

2.1. Sequence analysis and template search

The bromelain sequence with an NCBI accession number of ADY68475.1 (residues 1 to 291) was used as input data for PSI-BLAST against protein data bank (PDB) to identify its homologous structures [10] . In this study, sequence analysis was conducted for two groups by using CLUSTAL OMEGA. The first group involved random bromelain sequences, and the second group covered cysteine proteases derived from different families, namely, human procathepsin K (PDB ID: 1BY8), recombinant barley cysteine endoprotease (PDB ID: 2FO5), Evartamin C (PDB ID: 1O0E), cruzain (PBD ID: 2OZ2), and actinidin-E-64 (PDB ID: 1AEC).

Protparam was used to estimate and determine various properties, such as molecular weight, theoretical isoelectric point (pI), amino acid composition, total number of negatively band positively charged residues, instability index, and aliphatic index. The subcellular localization of protein was predicted by CELLO [11, 12].

2.2. Bromelain structure prediction and model evaluations

The secondary structure of bromelain and the tertiary structure of bromelain were predicted by PSIPRED and MODELLER 9.18, respectively [13, 14]. In our study, the models were generated through multiple template modelling. The models were evaluated based on the lowest energy produced from the modelling. These models were then refined and evaluated using ProSA and ModRefiner [15, 16]. Ramachandran plots were also drawn for each model [17].

2.3. Bromelain binding site predictions and docking simulation

Relevant amino acids involved in ligand binding were predicted by using the program Cofactor [18]. Pockets were predicted and the druggability of bromelain was assessed in DogSiteScorer [19]. The cleft for the detection of local binding site similarities was analysed by IsoCleftFinder [20]. The sequence of bromelain functioned as an input file in ConSurf to determine conserved functional and structural amino acids [21]. The functional site of bromelain was predicted by utilizing POOL [22]. Protein-protein docking between bromelain and phospholipase A2 (PLA2) was performed using PIPER [23] from Schrodinger.

3. Results and Discussion

Previously, the prediction of bromelain model was conducted based on a single template modeling [24]. However, the model that was predicted using single template needed further refinement and quality assessment in order to obtain the most accurate model. A consensus has been developed that the protein structure prediction via multiple templates is more accurate than the single template because it will cover most of the target sequence. Thus, in this study, the secondary and tertiary structures of bromelain were predicted using multiple templates, and the accuracy of prediction depended on the degree of sequence similarity.

3.1. Sequence analysis and template search

The amino acid sequence of bromelain that contained 291 residues was saved in FASTA format. PSI-BLAST search revealed that all the potential homologous structures consisted of protease-conserved domains belonging to a putative flagellar system-associated. From the result of PSI-BLAST, the first hit with the highest score was selected as a template for homology modelling. The top four hits (Table 1) were protein with PDB code, 3TNX (42% identity), 1S4V (53% identity), 2FO5 (50% identity), and 1IWD (49% identity).

Twelve bromelain sequences from different species with amino acid lengths of 250-350 were subjected to multiple sequence alignment to identify the conserved residues. Thirty conserved residues were obtained (Fig. 1). All of the conserved residues consisted of polar, non-polar, acidic, and basic groups of the protein. Most of the conserved residues, such as Tyr, Asn, Gly, Thr, Gln, and Cys, belong to the polar group, followed by the non-polar (Phe, Trp, Ala, Val, and Pro), acidic (Glu and Asp), and basic (Arg and His) groups of the protein.

Table 1. Top four hits for bromelain templates from PSI-BLAST results.

Template	Sequence identity %	Prosa z-score
3TNX	42	-8.83
1S4V	53	-7.98
2FO5	50	-7.07
1IWD	49	-6.75

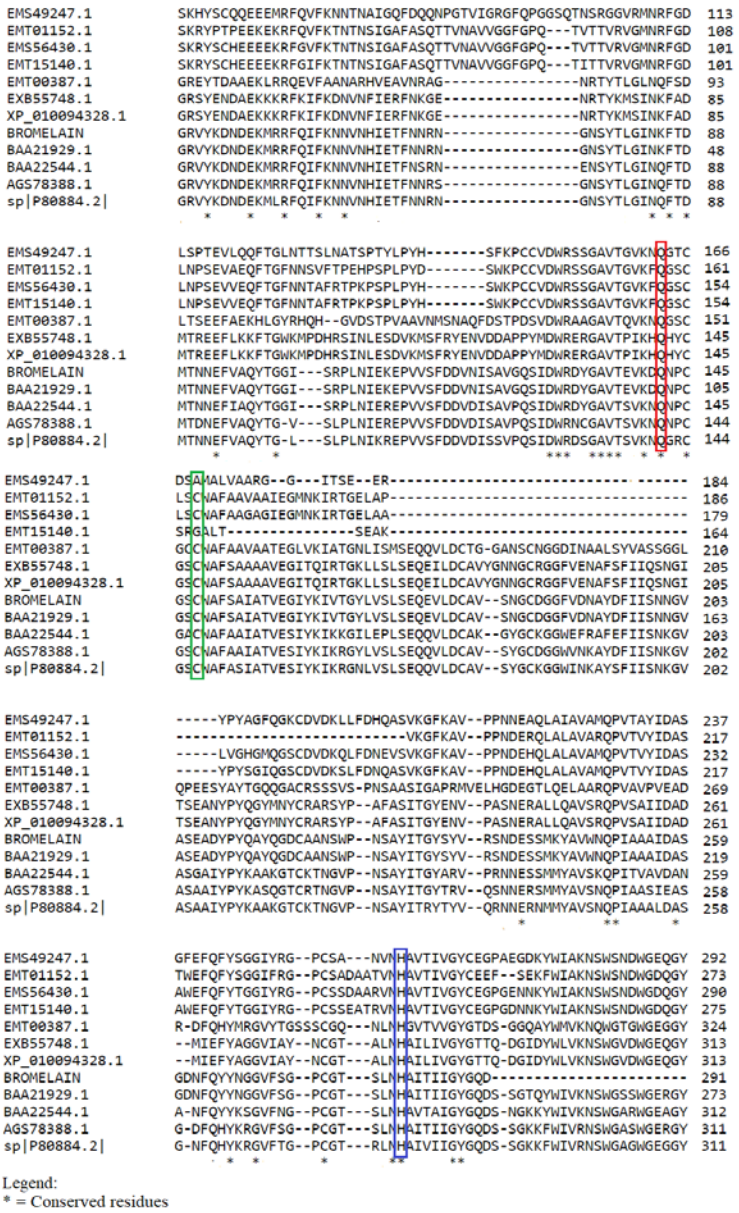


Fig. 1. Multiple sequence alignment of the 12 bromelain enzymes. The highlighted blocks showed the conserved catalytic residues, such as Gln142 (red block), Cys148 (green block), and His281 (blue block).

The multiple sequence alignment between bromelain and cysteine proteases from different families was also analysed to investigate the conserved catalytic residues (Fig. 2). The sequence alignments shown in Figs. 1 and 2 indicated that the catalytic residues of stem bromelain consisted of His281, Gln142, and Cys148 that acted as nucleophiles. These residues were highly conserved (Figs. 1 and 2).

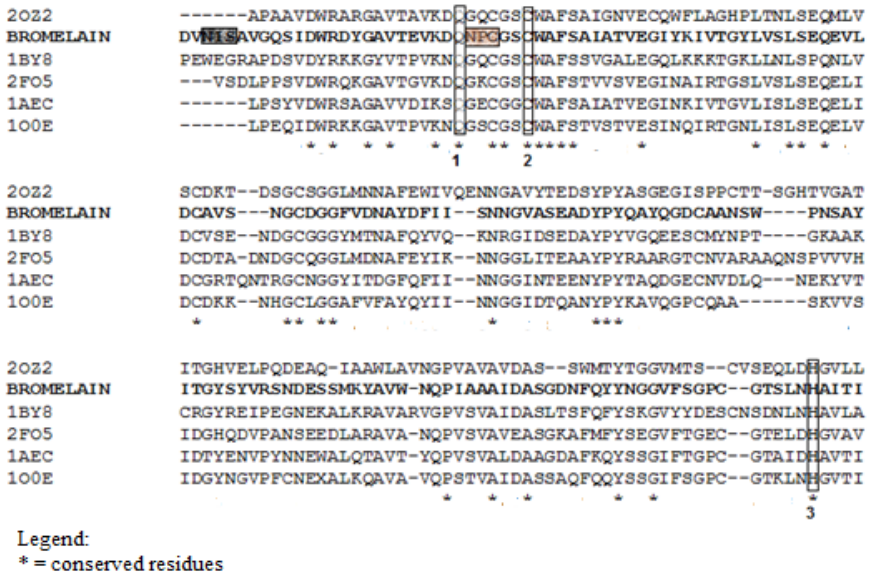


Fig. 2. Multiple sequence alignment of the 12 bromelain enzymes. The highlighted blocks showed the conserved catalytic residues, such as Gln142 (block 1), Cys148 (block 2), and His281 (block 3).

Table 2 showed the predicted chemical and physical properties of bromelain. These predicted parameters were important for development of experimental study. The aliphatic index of the protein is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine). Instability index provides an estimate of the stability of proteins in the application of using this protein in laboratory work. The negative GRAVY value shows that bromelain had a hydrophilic property.

Table 2. Physical and chemical properties of bromelain

	Residue	Molecular weight	Instability Index	Aliphatic index	GRAVY
Bromelain	291	32299.77	33.47 (stable)	68.04	-0.295

3.2. Structural prediction and model evaluation

The secondary structure of the bromelain model was predicted using PSIPRED (Fig. 3). This bromelain model consisted of 36% of helix, 6% of strand, and 60% of coil.

The top four protein structures with the highest scores were selected as templates for homology modelling. These protein structures were 3TNX (42%), 1S4V (53%), 2FO5 (50%), and 1IWD (49%), and 200 models were developed using MODELLER. The best model was selected based on the lowest energies of DOPE and Molpdf. Model101 with DOPE (−3009.63) and Molpdf (9190.44) was

selected as the best model (Fig. 4). Loop modelling was performed on residues between Met1 and Pro29 because this region was modelled without any templates. Our results revealed that 100 conformations were produced, and ten models were chosen for further evaluation based on the lowest DOPE energy [25].

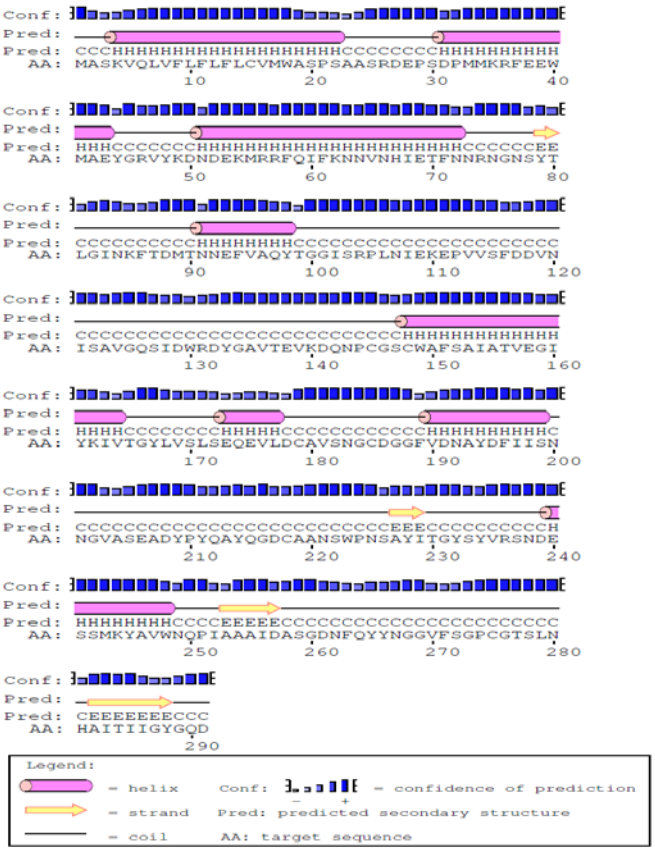


Fig. 3. Components of secondary structure prediction in the bromelain model.

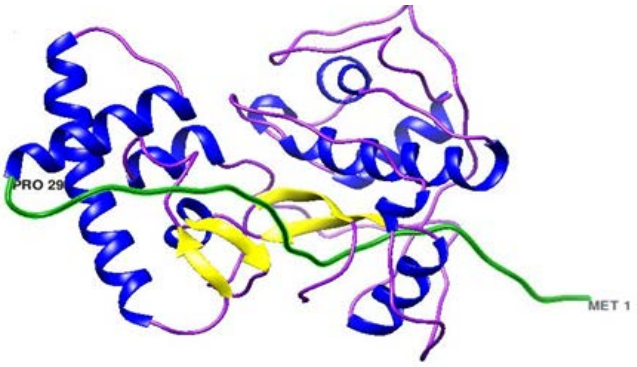


Fig. 4. Model from multiple templates modelling highlighting residues from Met1 to Pro29 (green coil) that should be further refined. The helix, strand, and coil represent in blue, yellow, and purple, respectively.

3.3. Model evaluation and refinement

Ten selected models were further evaluated using ProSa-web and Ramachandran plot. The z-scores of the models were found to be within the range of scores typically observed in native proteins with a similar size [26]. The results of ProSa z-score and Ramachandran plot analyses are presented in Table 3. $Model_{58}$ yielded the best scores for analyses.

Table 3: Top 10 predicted models after loop modelling based on the lowest DOPE value.

Model	Favourable%	Allowed %	Disallowed %	ProSa z-score
$Model_{9}$	91.7	6.2	2.1	-5.79
$Model_{14}$	92.7	5.2	2.1	-5.75
$Model_{17}$	92.0	5.9	2.1	5.71
$Model_{38}$	93.1	5.2	1.7	-5.9
$Model_{48}$	93.1	5.2	1.7	-5.95
$Model_{58}$	93.8	4.5	1.7	-6.12
$Model_{60}$	93.8	4.5	1.7	-6.08
$Model_{89}$	93.1	5.2	1.7	-5.95
$Model_{91}$	92.7	5.5	1.7	-6.06
$Model_{94}$	92.7	5.5	1.7	-6.09

The selected model (model58) was further refined using ModRefiner because it has the lowest z-score and highest percentage of favourable region which is 93.8. This means all residues for this selected model located in the most stable region. The Ramachandran plots of the initial and final models after refinement were compared, and 93.8% and 96.2% of the residues were located in the favourable region for the raw model and the refined model, respectively. The percentages of the residues located in the disallowed region were 1.7% and 1.0% for the initial and final models, respectively.

3.4. Bromelain binding site prediction

Our binding site analysis in Cofactor predicted that residues Gln142-Trp149, Asp187-Gly189, Ala255, and Thr279-His281 were found to be involved in the binding site of bromelain (Fig. 5). The three active residues, namely, Gln142, Cys148, and His281, were in the predicted binding site (Fig. 5). The calculated binding site score of the predicted binding site was 1.32. Binding site score is a measure of local similarity (sequence and structure) between a template binding site and a predicted binding site in the query structure. Based on large-scale benchmarking analysis, a binding site score > 1 reflects a significant local match between the predicted and template binding site [26]. Binding pockets and descriptors were calculated using DogSiteScorer and the results are shown in Fig. 6 and Table 4, respectively. Figure 6, there were two potential binding pockets were

predicted which were pockets one and two are shown in yellow and purple, respectively. Pocket one (yellow) was selected as the optimum pocket because of its higher volume and larger surface area than those of Pocket two.

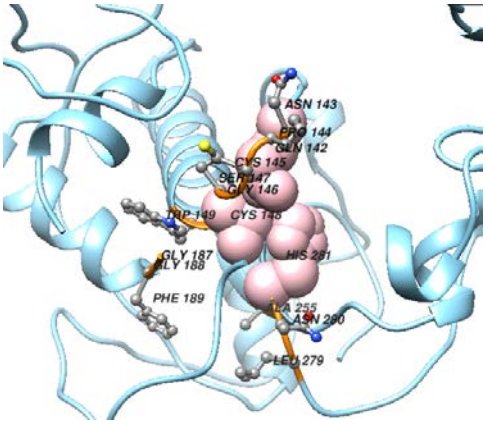


Fig. 5. Bromelain binding site prediction using Cofactor software. The query and catalytic residues involved in the binding interactions are shown as stick and sphere representation, respectively.

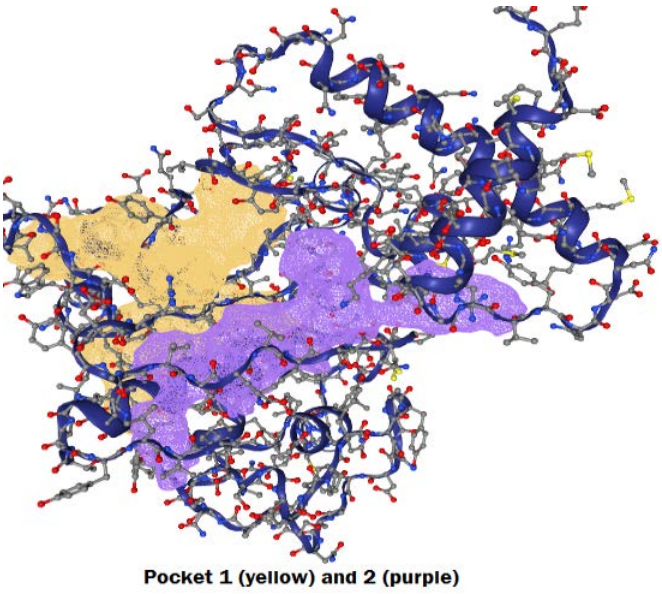


Fig. 6. Representation of binding pocket prediction using DogSiteScorer at pocket one (yellow) and two (purple).

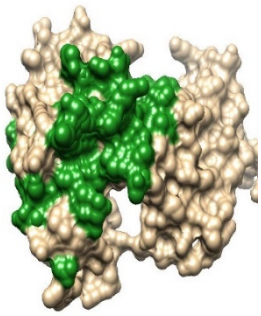
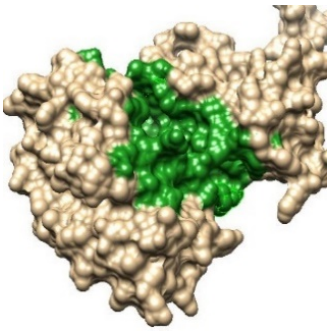
Table 4. Description of bromelain pocket prediction

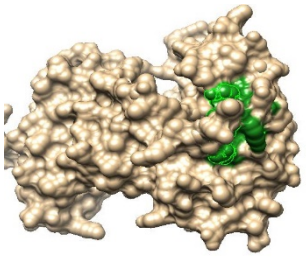
Pocket	Volume Å^3	Surface Å^2	Depth Å	Enclosure	Drug score
1	2004.21	2358.61	23.27	0.16	0.81
2	1144.97	1451.27	36.33	0.16	0.81

3.5. Bromelain cleft analysis

Three important clefts were identified on the protein surface. Identifying the cleft is crucial because the interactions of protein with other molecules lead to the protein's function and the key factor of interactions lies on the shape and chemical properties of the protein's surface. On the protein's surface containing many clefts and these clefts are relevance to the binding site of the protein [27]. The detected clefts and the corresponding residues are presented in Table 5. All residues that contribute at least one atom to the cleft are shown in green. The active site of the protein is usually located in the largest clefts or cavities, and its ligand is bound to the largest cleft in more than 83% of single-chain proteins [24]. Val95, Ala 96, Gln97, Thr99, Gly100, Gly101, Ile102, Ser103, Arg104, Pro105, Leu106, Glu109, Val114, Phe116, Asp117, Val119, Pro144, Cys145, Gly146, Thr156, Ile160, Tyr161, Val164, Thr165, Tyr167, Val169, Val181, Ser182, Asn183, Cys185, Asp186, Gly187, Phe189, Val190, Asp191, Asn192, Tyr194, Ser226, Ala227, Tyr228, Ile229, Thr230, Gly231, Tyr232, Ser233, Tyr234, Val235, Ser237, Asn238, Met243, Tyr245, Gln250, Ile252, Ala252, Ala254, Ala255, Thr277, Ser278, Leu279, and Asn280 were the residues forming the largest cleft on the surface of bromelain.

Table 5. Three important clefts (green surface) identified on the bromelain surface by using IsoCleft Finder.

Cleft model	Residues
	Val95, Ala96, Gln97, Thr99, Gly100, Gly101, Ile102, Ser103, Arg104, Pro105, Leu106, Glu109, Val114, Phe116, Asp117, Val119, Pro144, Cys145, Gly146, Thr156, Ile160, Tyr161, Val164, Thr165, Tyr167, Val169, Val181, Ser182, Asn183, Cys185, Asp186, Gly187, Phe189, Val190, Asp191, Asn192, Tyr194, Ser226, Ala227, Tyr228, Ile229, Thr230, Gly231, Tyr232, Ser233, Tyr234, Val235, Ser237, Asn238, Met243, Tyr245, Gln250, Ile252, Ala252, Ala254, Ala255, Thr277, Ser278, Leu279, Asn280
	Tyr44, Arg46 Asn84 Lys85, Phe86, Met89, Glu93, Phe94, Gln97, Tyr98, Arg131, Ala135, Val136, Thr137, Glu138, Val139, Lys140, Asp141, Gln142, Asn143, Pro144, Phe151, Ser152, Ala155, Glu158, Leu168, Ser170, Glu173, Glu206, Glu240, Met243, Lys244, Val247, Pro251, Ile256, Asp257, Phe263, Gln264, Tyr266, Gly269, Val270, Phe271, Ser272, Cys275, His281, Ala282, Ile283, The284, Ile285, Ile286, Gly287, Tyr288, Gly289, Gln290, Asp291

	Glu109, Lys110, Pro112, Asp191, Asn192, Tyr232, Tyr234

3.6. Identification of functionally and structurally important residues in bromelain

ConSurf results indicated that the majority of bromelain residues were predicted to be highly conserved and exposed. The important residues were located in the middle of the sequence and were exposed on the surface. This result revealed that 54 and 20 residues were predicted to be functionally and structurally important, respectively (Fig. 7).

The predicted residues that functionally and structurally important are all the residues label by f (red) and s (blue), respectively. POOL program predicted ten residues as the most functional at the protein structure. These residues were Tyr44, Arg46, Asn84, Tyr98, Lys140, Gln142, Gln173, His281, Ala282, and Ile283. ConSurf suggested that these residues (Arg46, Asn84, Lys140, Gln142, Gln173, His281, Ala282, and Ile283) were described as functional, highly conserved, and exposed.

Conservation analysis is an effective indicator to identify the functional relevance of proteins, and this technique has been used in conjunction with structural information, such as prediction of functionally important residue, detection of residues involved in ligand binding and protein-protein interaction, and determination of functionally specific proteins [19].

The catalytic and functional relevance of residues in proteins can be often mutated to yield a stable protein. This result can be achieved through site-directed mutagenesis to provide the potential production of desirable protein functionalities.

3.7. Protein-protein docking

Binding site analysis, cleft analyses and functionally important residue identification denote that some functionally conserved residues, especially Gln142, Cys145, Gly146, Gly187, Asn280, and His281, are located in the first and second largest clefts that participate in the ligand binding site, and these amino acids may be involved in bromelain-PLA2 interactions (Fig. 5, Table 5, and Fig. 7).

PIPER from Schrodinger was used to predict the possible binding modes in the translational and rotational space between the selected region of bromelain (ligand) and PLA2 (receptor). Each pose of the complex was evaluated using an energy-based scoring function, which is depending on the binding energy calculated using the force field of the software. The model with the lowest energy, the top five predictions are shown in Fig. 8. All the residues involved in the binding site are similar to those predicted using Cofactor and ConSurf server.

The structure of bromelain and its binding site properties should be characterized because interactions between bromelain and PLA2 occur at atomic levels, and characterization results will help enhance our understanding of active mechanisms relevant to designing new anti-inflammatory agents with improved properties. Computational models of a bromelain-PLA2 complex may be practical to introduce a new anti-inflammatory agent.

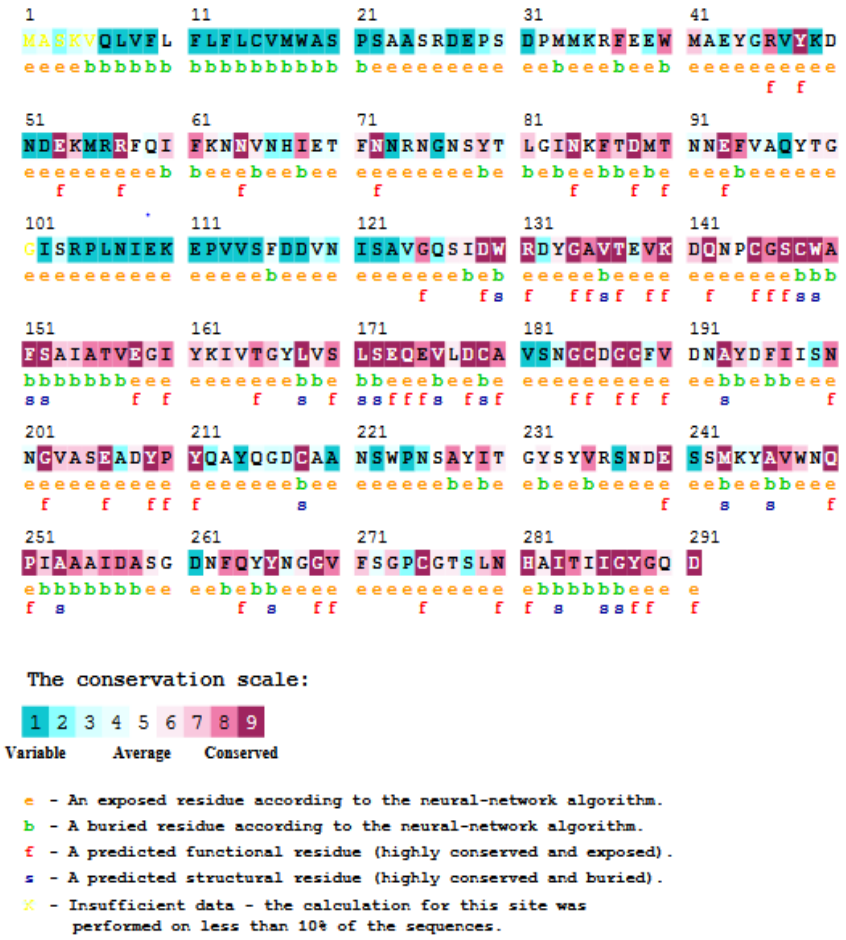


Fig. 7. Coloured illustration of the conserved functional and structural residues predicted by ConSurf server.

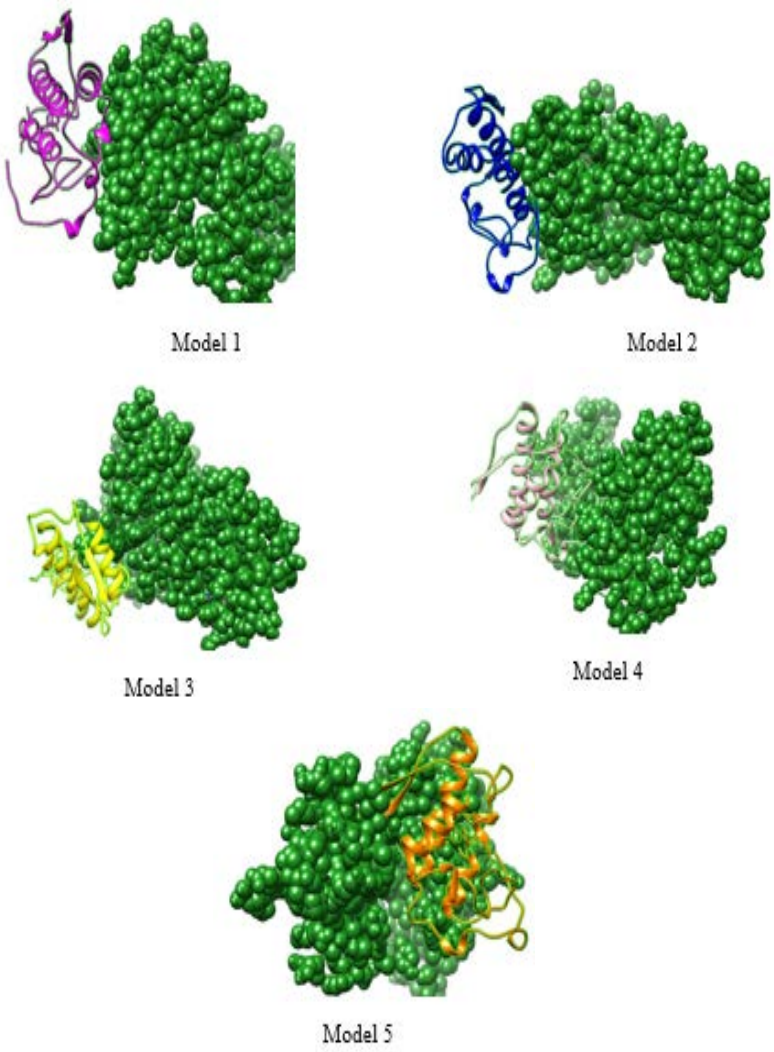


Fig. 8. PLA2 binds to a particular region of the bromelain model. Bromelain and PLA2 were presented as surface and cartoon representation, respectively

4. Conclusions

In conclusion, the sequence analysis between stem bromelain and other cysteine proteases was performed using multiple sequence alignment method. This analysis highlights that catalytic residues Gln143, Cys148, and His 281 are conserved with other cysteine proteases. The binding pockets were predicted using comparative method and all catalytic residues were found to be located in the binding pocket. Results of the tertiary structure model and ligand binding site analysis, cleft analyses and functionally important residue identification denote that some functionally conserved residues, especially Gln142, Cys145, Gly146, Gly187, Asn280, and His281, are located in the first and second largest clefts that participate in the ligand binding site. These residues are also involved in the protein-protein

interactions which are between bromelain and PLA2. This study proposed a few functionally conserved residues, especially Gln 142, Cys 145, Gly 146 Gly 187, Asn 280, and His 281, which are located in the largest cleft that may participate in ligand binding activities. The identification of the predicted binding residues is important for further used in molecular docking study.

References

1. Gautam, S.S.; Mishra, S.K.; Dash, V.; Goyal, A.K.; and Rath, G. (2010). Comparative study of extraction, purification and estimation of bromelain from stem and fruit of pineapple plant. *The Thai Journal of Pharmaceutical Sciences*, 34(2), 67-76.
2. Grzonka, Z.; Jankowska, E.; Kasprzykowski, F.; Kasprzykowska, R.; Lankiewicz, L.; Wiczk, W.; Wieczerzak, E.; Ciarkowski, J.; Drabik, P.; Janowski, R.; Kozak, M.; Jaskólski, M; and Grubb, A. (2001). Structural studies of cysteine proteases and their inhibitors. *Acta Biochim Polonica*, 48(1), 1-20.
3. Arshad, Z.I.M.; Amid, A.; Yusof, F.; Jaswir, I.; Ahmad, K.; and Loke, S.P. (2014). Bromelain: an overview of industrial application and purification strategies. *Applied Microbiology and Biotechnology*, 98, 7283-7297.
4. Agostinis, C.; Zorzet, S.; Leo, R. De; Zauli, G.; Seta, F. De; and Bulla, R. (2015). The Combination of N-Acetyl Cysteine, Alpha-Lipoic Acid, and Bromelain Shows High Anti-Inflammatory Properties in Novel In Vivo and In Vitro Models of Endometriosis. *Mediators of Inflammation*, 2015, 1-9.
5. Pang, W.C.; Ramli, A.N.M.; and Hamid, A.A.A. (2020). Comparative modelling studies of fruit bromelain using molecular dynamics simulation. *Journal of Molecular Modelling*, 26(6), 142.
6. Lehra, K.S.; Kaur, R.; Sharma, S.; Kapoor, A.; and Singh, S. (2014). Anti-inflammatory agents from plants - Part III. *Indian Journal Natural Products and Resources*, 5(2), 121-128.
7. Chaurasiya, R.S.; Sakhare, P.Z.; Bhaskar, N.; and Hebbar, H.U. (2014). Efficacy of reverse micellar extracted fruit bromelain in meat tenderization. *Journal of Food Science and Technology*, 52(6), 3870-3880.
8. Pavan, R.; Jain, S.; and Kumar, A. (2012). Properties and therapeutic application of bromelain: A review. *Biotechnology Research International*, Volume 2012 |Article ID 976203, 1-6.
9. Ramalingam, C.; Srinath, R.; and Islam, N.N. (2012). Isolation and characterization of Bromelain from pineapple (Ananas Comosus) and comparing its anti-browning activity on apple juice with commercial anti-browning agents. *Elixir Food Science*, 45(2012), 7822-7826.
10. Berman, H.M.; Westbrook. J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; and Bourne, P.E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235-242.
11. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. *Proteomics Protocol Handbook, Humana Press*, 571-607.
12. Yu, Chin-Seng.; Chen, Yu-Ching.; Lu, Chih-Chao.; and Hwang, Jenn-Kang. (2006). Prediction of protein subcellular localization. *Proteins*, 64(3), 643-651.

13. Buchan, D.W.A.; Minneci, F.; Nugent, T.C.O.; Bryson, K.; and Jones, D.T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research 2013 July*, 41(W1), 349-357.
14. Webb, B.; and Sali, A. (2017). Protein Structure Modeling with MODELLER. *Functional Genomics*, 1654, 39-54.
15. Wiederstein, M.; and Sippl, M.J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35(2), 407-410.
16. Xu, D.; and Zhang, Y. (2011). Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-Step Atomic-Level Energy Minimization. *Biophyssical Journal*, 101(10), 2525-2534.
17. Lovell, S.C.; Davis, I.W.; Adrendall, W.B.; de Bakker, P.I.W.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; and Richardson, D.C. (2003). Structure validation by C alpha geometry: phi,psi and C beta deviation. *Proteins: Structure, Function and Biometrics*, 50(3), 437-450.
18. Roy, A.; and Zhang, Y. (2012). Recognizing protein-ligand binding sites by global structural alignment and local geometry refinement. *Structure*, 26(6), 987-997.
19. Volkamer, A.; Kuhn, D.; Rippmann, F.; and Rarey, M. (2012). Dogsitescorer: A web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*, 28(15), 2074-2075.
20. Kurbatova, N.; Chartier, M.; Zylber, M.I.; and Najmanovich, R. (2013). IsoCleft Finder - a web-based tool for the detection and analysis of protein binding-site geometric and chemical similarities. *F1000Research*, 2, 1-13.
21. Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; and Ben-Tal, N. (2010). ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research,* 38(2), 529-533.
22. Tong, W.; Wei, Y.; Murga, L.F.; Ondrechen, M.J.; and Williams, R.J. (2009). Partial Order Optimum Likelihood (POOL): Maximum likelihood prediction of protein active site residues using 3D structure and sequence properties. *PLoS Computational Biology*, 5(1).
23. Chuang, G-Y.; Kozakov, D.; Brenke, R.; Comeau, S.R.; and Vajda, S. (2008). DARS (Decoys As the Reference State) Potentials for Protein-Protein Docking. *Biophysical Journal*, 95(9), 4217-4227.
24. Tap, F.M.; Majid, F.A.A.; and Khairudin, N.B.A. (2016). Structure prediction of Stem Bromelain from pineapples (Ananas Comosus) using procaricain enzyme as a modelling template. *Int J Appl Eng Res*, 11(9), 6109-6111.
25. Karami, Y.; Guyon, F.; De Vries, S.; and Tufféry, P. (2018). DaReUS-Loop: accurate loop modeling using fragments from remote or unrelated proteins. *Scientific Reports 8*, 13673.
26. Wiederstein, M.; and Sippl, M.J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35(2), W407-W410.
27. Laskowski, R.A.; Luscombe, N.M.; Swindells, M.B.; and Thornton, J.M. (1996). Protein clefts in molecular recognition and function. *Protein Science, 5*(12), 2438-2452.