Journal of Academia UiTM Negeri Sembilan 5 (2017) 36-43

SUBSTRATE BINDING SITE OF PROTEASES FROM *Bacillus lehensis* G1 BY MOLECULAR DOCKING

N.A. Sulaiman[1,2], N.M. Mahadi[2], and N. Z. Ramly[1]

[1]*Food Biotechnology, Faculty of Science and Technology,*
Islamic Science University of Malaysia (USIM),
Bandar Baru Nilai 71800, Nilai,
Negeri Sembilan, Malaysia

[2]*Malaysia Genome Institute,*
Jalan Bangi, 43000 Kajang. Selangor, Malaysia

Email noorulaini@gmail.com

Abstract

Proteases are abundantly found in all living organisms that are essential to life. Microbial proteases possess a commercial value with various applications in industries. Some of the biotechnological applications are involved detergents, leather, food, pharmaceutical and bioremediation processes. The aim of the study is to analyses the binding cavities in protease, BleG1_1979, from Bacillus lehensis G1, in order to examine the interaction site of protease with other molecules. Protein encoding for BleG1_1979 is 19.2 kDa in size, with a predicted pI value of 5.35 and 42% identity with intracellular protease from Thermococcus onnurineus NA1 from PDB database. Molecular docking was run using Hex v8.0.0. Four substrates were docked to BleG1_1979 pocket cavities, and casein exhibited highest binding affinity as compared to collagen, gelatin and keratin. This has suggested that BleG1_1979 could rapidly hydrolyzed casein and B. lehensis G1 did not contain keratolytic activity. Docking results have provided a basis for further understanding the molecular mechanism of substrates preference for B. lehensis G1 protease. Such knowledge could be explored to expand the usefulness of microbial protease for industrial application.

Keywords: *Bacillus lehensis, protease, homology modelling, molecular docking*

Article history:- Received: xxxxx; Accepted: xxxxx; Published: 15 December 2017

1. Introduction

Enzymes are proteins that catalysed many reactions in all living organisms. Enzymes play important role in the transformation of macromolecules to energy and new materials, for growth, repair and maintenance of the cells (Ibrahim, 2008; Ray, 2012). There are six families in enzyme nomenclature which are oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases. However, only hydrolases enzyme is important in industrial application. Hydrolases is an enzyme that catalyse hydrolysis of chemical bond by the action of water, such as proteases, esterases and glycosidases (Illanes, 2008). By understanding the mechanism of the enzyme such as proteases, this could then be exploited to be used in industry, therapeutics and academia (Li *et al*., 2013).

Microorganism are known to produce proteases and have been dominated the enzyme market worldwide (Gupta *et al*., 2002). There is many studies on the application of microbial protease

in various areas, such as detergent, feather decompose, leather dehairing, silk degumming, and silver recovery (Haddar *et al*., 2010; Sundararajan *et al*., 2010; Sinha and Khare, 2013; Joshi and Satyanarayana, 2013, Rajashekar *et al*., 2011). Due to its wide application, finding new variant of proteases from microbial sources are essential in industrial communities. Among the microbes, *Bacillus* has been known as the major contributor of proteases in industry (Lagzian and Asoodeh, 2012). Species that produces proteases are *B. licheniformis*, *B. subtilis* and *B. pumilus* (Anbu, 2013; Benkiar et al., 2013).

Although there are several *Bacillus*'s proteases that are commercially available worldwide, currently there are no study on Malaysia's *Bacillus* as commercial producers of protease has been reported. Thus, in this paper we reported the study on protease from *B. lehensis* G1 that has been isolated from rubber plantation in Johor, Malaysia (Illias *et al*., 2002). Further in the paper, we will describe the molecular docking of protease BleG1_1979 from *B. lehensis* G1 using four substrates that has never been reported elsewhere. This could then provide new information and opportunities for the discovery, production and modification of novel protease to be used in alkaline conditions.

2. Materials and Methods

2.1 Sequence Analysis and Phylogenetic tree

Amino acid sequence of protease BleG1_1979 (previously known as BleG1_1940) from *B. lehensis* G1 was retrieved from Malaysia Genome Institute (http://www.mgi-nibm.my/bacillus_lehensis_g1/). BleG1_1979 physiochemical properties which includes amino acid composition, isoelectric point (pI), molecular weight (kDa), number of charges residues, instability and aliphatic index, and Grand Average of Hydropathy (GRAVY) was analysed using Protparam server (Gasteiger *et al*., 2005). SCFBio (http://www.scfbio-iitd.res.in/)(Singh et al., 2011). COACH (http://zhanglab.ccmb.med.umich.edu/COACH/) servers was used to predict the binding site for BleG1_1979. Phylogenetic tree was generated using cluster omega (http://www.ebi.ac.uk/Tools/msa/clustalo/)(Sievers *et al*., 2011).

2.2 Molecular Docking

The Docking of the BleG1_1979 with different substrates was carried out using Hex v8.0.0 (Macindoe *et al*., 2010) with the purpose to analyze the substrate specificity for BleG1_1979. Substrates structure was retrieved from PDB and PubChem database and all images was visualized using Chimera v1.11.2 (Pettersen *et al*., 2004).

3. Result and Discussion

Physical properties play a vital role in activity of an enzyme. Analysis of physicochemical properties using ProtParam revealed that BleG1_1979 has 174 amino acids, 19196.3 molecular weight, 5.35 pI, 41.88 instability index, 94.25 aliphatic index, and -0.257 grand average of hydrophobicity (GRAVY) with more negatively charged residues than positively charged amino acids. According to the Protparam results, BleG1_1979 is unstable protein with value more than 40 in the instability index. The negative GRAVY indicates the sum of hydrophilicity and hydrophobicity of all amino acids in protein. Analysis of GRAVY exhibited high aliphatic index in BleG1_1979 which point out as a positive factor for increasing thermos-stability in protein (Gasteiger *et al*., 2005).

BleG1_1979 amino acid sequences used in the phylogenetic tree were selected based on the sequences deposited in NCBI database. The phylogenetic tree of BleG1_1979 was rooted to *Domibacillus* as an outgroup depicted in Figure 1. The cladogram was divided into three descendent taxa and BleG1_1979 was clustered with *B. wakoensis* and *B. plakortidis*. BleG1_1979 was very clearly integrated with *B. plakortidis*. Since *B. lehensis* was not clustered together with *B. gottheilii*, this indicates diversity and genetic variability among the different clusters. Comparison of the BleG1_1979 with data in the NCBI database has indicated that all of the identified isolates belonged to the genera *Bacillus*, *Domibacillus, Jeotgalibacillus, Oceanobacillus* and *Paraliobacillus.*

Figure 1. Unrooted neighbor-joining tree showing the relationship between BleG1_1979 and other several closely related glutamine amidotransferase. The tree contains 11 different bacterial species from NCBI reference sequence are indicated with strains.

Docking is a technique that is generally used for drug screening and design. However, there are many parameter need to be considered while undertaking a docking study, such as the incorrect binding site, unsuitable small-molecule database, choice of docking pose, and lack of clarity over whether the compound is an inhibitor or agonist (Chen, 2015). In this study, the 3D model of BleG1_1979 was previously predicted in Sulaiman *et al*., (2017) was used to assess the binding ability for substrates. The active site prediction for BleG1_1979 was carried out to evaluate the precise binding site. Two Web servers were used to produce the most reliable results. Among the freely accessible computational tools for active site prediction in this study are SCFBio and COACH. SCFBio results showed 22 possibilities for binding cavities in BleG1_1979 with the highest cavity volume was 920 and the smallest cavity was 146. The residues involve are Glu12, Lys46, Trp74, Asp77, Lys78, Arg80, Arg81, His104, Ile110, Gly125, Asp128, Asp129, Asn132 and Ala133. Whilst, from COACH web server, there are two sets combination of ligand binding site prediction which are TM-SITE and S-SITE. TM-SITE is the structure-based algorithm whilst S-SITE is based on the binding-specific sequence alignment. The highest C-score from COACH analysis was belong to three residues Glu15, Cys106, and Pro155. The C-score showed the confidence score of the prediction as the high score specifies a more reliable prediction (Yang *et al*., 2013).

Hex v8.0.0 was used for calculating and displaying docking modes between BleG1_1979 and substrate molecules. Hex's algorithm uses spherical polar Fourier (SPF) that could address the issues with the Cartesian-grid Fast Fourier Transform (FFT) approaches. The advantage of Hex's SPF is the rotational correlations and short execution times. Hex's SPF algorithm has been validated in the CAPRI (Critical Assessment of PRedicted Interactions) in blind docking

experiment, and an acceptable Hex prediction has frequently been found in current CAPRI scoring (Ritchie, 2003). Thus, Hex software offers fast and convenient way to produce high quality docking predictions (Macindoe *et al*., 2010).

Hex assumes rigid BleG1_1979 and binds substrate/ligand into the active sites. Four substrates were used in this study to test the binding mode for substrate preference. The substrates were keratin, casein, gelatin and collagen;. In molecular docking, the scoring function is the key component of a protein-ligand docking algorithm, since it directly determined the reliability of the algorithm (Huang and Zou, 2010). Docking software usually calculate the free energy of binding (ΔG) between a ligand and a receptor, that is based on evaluations of the total energy of intermolecular forces, which includes Van der Waals, hydrogen, electrostatic, and hydrophobic bond. From the calculation of ΔG value, the lower the ΔG values, the more favourable ligand binding, and vice versa (Jacob *et al*., 2012).

In Hex docking score, casein showed lowest energy of -542.44 followed by collagen -450.76. The high-energy score is recorded by keratin with -32.95 of E-score (Table 1). These results suggested that casein is an ideal substrate with high binding mode for protease BleG1_1979 as compared to keratin, gelatin and collagen. This result is different from serine protease from *B. circulans* strain DZ100 which showed similar broad substrates specificity but with high keratinolytic activity, which made it possible to apply in feather-biodegradation process (Benkiar *et al*., 2013). Hex E-score from *B. pumilus* strain MP 27 serine protease revealed different arrangement for binding with highest score for gelatine (-609.97), followed by collagen (-602.69), casein (-427.67) and keratin (-341.53) (Baweja et al., 2016). In this study, the low E-score docking for keratin signify that BleG1_1979 protease probably did not contain keratinolytic activity. The Hex docking schematic was showed in Figure 2. BleG1_1979 docked has demonstrated that the substrates did not directly bind to the conserved catalytic residue, Cys106. This finding is different with the result showed by cysteine proteases from *Xanthomonas campestris*. In molecular docking of *X. campestris*, the cysteine protease degrades the patch of Aβ peptide with direct interaction between hydrogen atom of sulfhydryl group in the catalytic site, Cys17, with carboxylic site of Lys16. However, in *B. lehensis* G1, the obtain results suggested that the large size of the substrate could possibly require for a proper folding in the binding groove in order to access into the hydrophobic region inside BleG1_1979.

Table 1. Docking scores of BleG1_1979 structure using different substrate.

Ligand docking with BleG1_1979	Hex score (E-total)
Keratin	-32.95
Casein	-542.44
Gelatin	-434.44
Collagen	-450.76

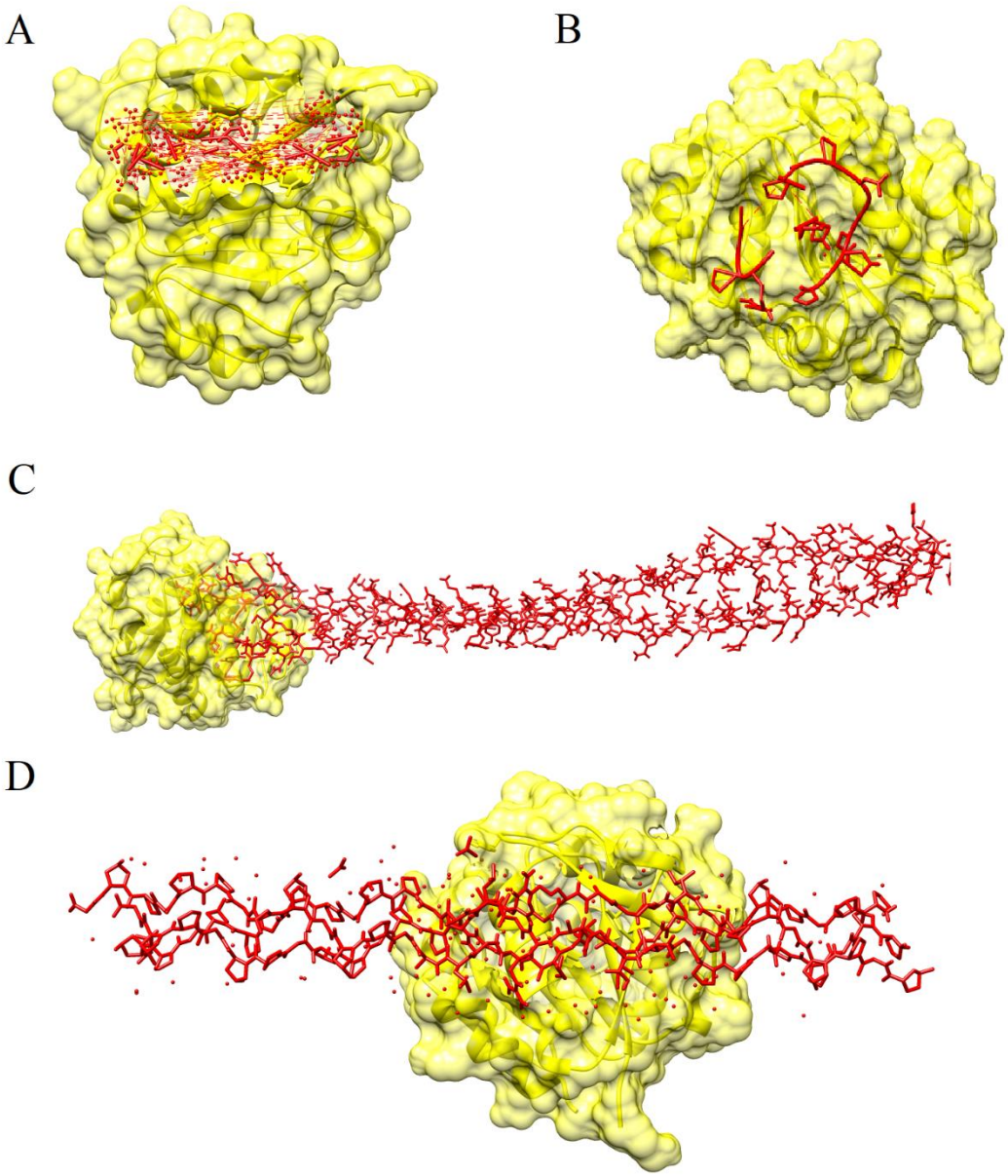

Figure 2. Model structure of BleG1_1979 docking with different substrates. (A) Casein, (B) Gelatine, (C) Keratin and (D) Collagen. Protease BleG1_1979 were showed as 60% surface transparency in yellow whilst substrates as lines.

4. Conclusion

In this paper, the modeled structure of BleG1_1979 docking using four substrates were presented and this study has allowed us to recognize the binding cavities in BleG1_1979. Casein showed the highest binding affinity with BleG1_1979 with keratin was the lowest. According to SCFBio and COACH, there are seventeen residue that were predicted to be involved in the BleG1_1979 binding pocket, however, only three hydrogen bond were visibly form between

carboxylate ion of acidic residue (Glu) with an ammonium ion of basic residue (Lys). Whilst, the rest of the predicted amino acid could be formed as close contact interaction with casein. Currently, the stability of BleG1_1979 docked with casein at different temperature using Molecular Dynamic simulation are in pursued. These results could improve our understanding in the enzyme-substrate interaction and thus, facilitate the application of proteases in industry.

Acknowledgements

We would like to thank Science Fund from the Malaysia Ministry of Science, Technology and Inovation (MOSTI) (Grant Number: USIM/SF-UKM/FST/30/40214) for financial support.

References

Anbu, P. (2013). "Characterization of solvent stable extracellular protease from *Bacillus koreensis* BK-P21A," *International Journal of Biological Macromolecules*, **56**, 162-168.

Baweja, M., Tiwari, R., Singh, P. K., Nain, L. and Shukla, P. (2016). "An alkaline protease from *Bacillus pumilus* MP 27: functional analysis of its binding model toward its applications as detergent additive," *Frontiers in Microbiology*, **7**, 1-14.

Benkiar, A., Jaouadi, Z., Badis, A., Rebzani, F., Touioui, B., Rekik, H., Naili, B., Zohra, F. Bejar, S. and Jaouadi, B. (2013). "International biodeterioration biodegradation biochemical and molecular characterization of a thermo- and detergent-stable alkaline serine keratinolytic protease from *Bacillus circulans* strain DZ100 for detergent formulations and feather-biodegradation," *International Biodeterioration Biodegradation.* **83**, 129-138.

Chen, Y.C. (2015). "Beware of docking!," *Trends Pharmacol Sci*,. **36**,78-95.

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*" The Proteomics Protocols Handbook*, Humana Press 571-607.

Gupta, R., Beg, Q. K. and Lorenz, P. (2002). "Bacterial alkaline proteases: molecular approaches and industrial applications," *Applied Microbiology and Biotechnolog*., **59**, 15-32.

Haddar, A., Sellami-Kamoun, A., Fakhfakh-Zouari, N., Hmidet, N. and Nasri, M. (2010). "Characterization of detergent stable and feather degrading serine proteases from *Bacillus mojavensis* A21," *Biochemical Engineering Journal*, **51**, 53-63.

Huang, S. Y., and Zou, X. (2010). "Advances and challenges in protein-ligand docking," *International Journal of Molecular Sciences*, **11**, 3016-3034.

Ibrahim, C. O. (2008)."Development of applications of industrial enzymes from Malaysian indigenous microbial sources," *Bioresource Technology*, **99**, 4572–82.

Illanes, A., (2008). *Enzyme Biocatalysis: Principles and Applications*, Springer, Chile.

Illias, R.M., Fen, T.S., Abdulrashid, N.A., Yusoff, W.M.W., Hamid, A.A., Hassan, O. and Kamaruddin, K. (2002). "Cyclodextrin glucanotransferase producing alkalophilic *Bacillus* Sp. G1: its cultural condition and partial characterization of the enzyme," *Pakistan Journal of Biological Sciences*. **5**, 688-692.

Jacob, R. B., Andersen, T. and McDougal, O. M. (2012). "Accessible high-throughput virtual screening molecular docking software for students and educators," *PLoS Computational Biology*, **8**, 1-5.

Joshi, S. and Satyanarayana, T. (2013). "Characteristics and applications of a recombinant alkaline serine protease from a novel bacterium *Bacillus lehensis*," *Bioresource Technology.***131**, 76-85.

Lagzian, M. and Asoodeh, A. (2012). "An extremely thermotolerant, alkaliphilic subtilisin-like protease from hyperthermophilic *Bacillus* sp. MLA64," *International Journal of Biological Macromolecules*. **15**, 960-7.

Li, Q., Yi, L., Marek, P. and Iverson, B. L. (2013). "Commercial Proteases: present and future," *FEBS Letters*, **587**,1155-1163.

Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M. and Ritchie, D. W. (2010). "HexServer : an FFT-based protein docking server powered by graphics processors," *Nucleic Acids Research*, **38**, 445–449.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004). "UCSF Chimera - A visualization system for exploratory research and analysis," *Journal of Computational Chemistry*, **25**, 1605-1612.

Rajasekhar, A., Ravi, V., Reddy, M. N. and Rao, K. R. S. S. (2011). "Thermostable bacterial protease - a new way for quality silk production," *International Journal of Bio-Science and Bio-Technology*, **3**(4), 43-58.

Ray, A. (2003). "Protease Enzyme- Potential Industrial Scope," *Int. J. Tech*., **2**, 1-4.

Ritchie, D. W. (2003)."Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2," *Proteins: Structure, Function and Genetics*," **52**, 98-106.

Sinha, R. and Khare, S. K. (2012). "Characterization of detergent compatible protease of a halophilic *Bacillus* Sp. EMB9: differential role of metal ions in stability and activity," *Bioresource Technology*. **145**, 357-361.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopezs, R.,

Mcwilliam, H., Remmet, M., Söding, J., Thompson, J. D. and Higgins, D. G. (2014). "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega," *Molecular Systems Biology*, **7**, 539–539.

Singh, T., Biswas, D., and Jayaram, B. (2011). "AADS - An automated active site identification, docking, and scoring protocol for protein targets based on physicochemical descriptors," *Journal of Chemical Information and Modeling*, **51**, 2515–2527.

Sundararajan, S., Kannan, C. N. and Chittibabu, S. (2010). "Alkaline protease from *Bacillus cereus* VITSN04: potential application as a dehairing agent," *Journal of Bioscience and Bioengineering*, **111**,128-33.

Sulaiman, N.A. Mahadi, N. M. and Ramly, N. Z. (2017). "Identification of proteolytic genes from *Bacillus lehensis* G1," *Journal of Engineering and Science Research*, **1**.

Yang, J., Roy, A. and Zhang, Y. (2013). "Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment," *Bioinformatics*, **29**, 2588-2595.