OnLine Journal of Biological Sciences

Original Research Paper

Structural and Substrate Interaction Properties of Alkaline Phosphatase from *Paenibacillus sp*. PSB04: *In-Silico* Analysis

[1]Herman Umbau Lindang and [1,2]Cahyo Budiman

[1]Biotechnology Research Institute, Universiti Malaysia Sabah, Malaysia
[2]Department of Animal Production and Technology, Faculty of Animal Science, IPB University, Indonesia

Article history
Received: 06-03-2022
Revised: 10-06-2022
Accepted: 14-06-2022

Corresponding Author:
Cahyo Budiman
Biotechnology Research Institute, Universiti Malaysia Sabah, Malaysia; Department of Animal Production and Technology, Faculty of Animal Science, IPB University, Indonesia
Email: cahyo82@gmail.com

Abstract: A Phosphate-Solubilising Bacterium (PSB) of *Paenibacillus sp*. PSB04 was previously isolated from the Sabah tropical rainforest in Malaysia. Interestingly, the genome sequence of the PSB04 strain harbored an Alkaline Phosphatase (AP) (EC 3.1.3.1) gene and was hypothesized to have unique structural characteristics. Therefore, this study aims to determine the AP three-Dimensional (3D) model and catalytic mechanism from *Paenibacillus sp*. PSB04 (PAP). To address this, the 3D model of this protein was built and docked into a model substrate of *p*-nitrophenyl phosphate. As a result, the best complex was shown to have the lowest binding energy of -5.9 kcal/moL. Furthermore, the complex showed the atomic coordination of catalytic residues of PAP and the substrate was similar to that of AP from *Escherichia coli* (ECAP), which implies that both APs shared a similar catalytic mechanism. In this mechanism, Ser 94 of PAP acted as nucleophilic residues, which were activated by the Zn ion. Arg 145 is predicted to be mobile due to its location in the loop region, which allows this residue to stabilize the substrate through direction or water-mediated secondary interaction. Docking simulation of pNPP indicated that the putative residues involved in the catalysis mainly are Ser 94, Ser 141, Ala 146, Thr 147, Pro 148, Asp 293, and Glu 294. Glu 294 is considered a unique residue corresponding to Lys 328 ECAP, allowing the PAP to have a better affinity to stabilize the substrate in the binding cavity. Accordingly, a unique catalytic mechanism of PAP was proposed.

Keywords: Bacteria Alkaline Phosphatase, *in Silico*, Molecular Docking, Catalytic Mechanism

Introduction

Alkaline Phosphatase (AP) (EC 3.1.3.1) is a hydrolytic metalloenzyme found abundantly across microorganisms. Studies on AP were mostly related to their important roles in soil Phosphorus (P) availability (Lindang *et al.*, 2021; Millan, 2006). In addition, alkaline AP was also reported to be produced by some Lactic Acid Bacteria (LAB)-based probiotics which lead to the possibility of involvement in food fermentation for health benefits. Afiyah *et al.* (2015) indicated that LAB-based probiotics benefit humans by producing various functional enzymes during fermentation. According to AP from *Escherichia coli* (ECAP), which is the most studied AP, the primary differences between AP, phytase, and acid phosphatase lie in the catalytic mechanism and pH requirement for the catalytic activity to take place (Dong *et al.*, 2021, Liu *et al.*, 2021; Millán, 2006). The phytate enzyme hydrolyses only the ester bond of phytate and releases less phosphorylated myo-inositol and inorganic phosphorus as the by-product (Chen *et al.*, 2015). Meanwhile, Acid Phosphatase (AcP) and AP also attack the ester bond of other P compounds at different pH environments and release inorganic P as the by-product. AcP and AP catalyze optimally in acidic and alkaline conditions, respectively (Coker *et al.*, 2013). Unlike AP, most AcP does not utilize metal ions during catalysis except for tartrate group AcP.

The catalytic residues of AP are highly conserved, consisting of Ser and Arg with the involvement of divalent metal ions of Zn, Mg, Ca, and Cu (Wang and Griffiths, 2009). The structural arrangement of bacteria AP typically consisted of N-Terminal signal peptide, followed by the mature domain and sometimes crown domain. Interestingly, the crown domain was not conserved across the bacterial APs. The most distinctive crown domain structure across reported APs was Vibrio-AP (PDB ID: 3E2D), essential for

the protein's stability and function (Helland *et al.*, 2009; Hjörleifsson *et al.*, 2021; Ásgeirsson *et al.*, 2020).

A mesophilic bacterium, *Paenibacillus sp*. PSB04 was previously isolated from the nutrient-limiting soil rainforest soil of Danum Valley (DV), Sabah, in Malaysia (Lindang *et al.*, 2021). The complete genome of strain PSB04 reveals an alkaline phosphatase encoded in the sequence (Gene Bank ID: WP014096696.1). The specific activity of *Paenibacillus sp*. PSB04 was considerably high and it was hypothesized that the high specific activity of AP of *Paenibacillus sp*. PSB04 (PAP) is due to its unique structural features.

The recent study described the three-dimensional model of PAP, constructed through structural homology modeling and binding properties toward the ligand through docking simulation. Accordingly, the unique structural properties and catalytic mechanism of this protein were then proposed.

Materials and Methods

Structural Homology Modeling of PAP

The amino acid sequence of PAP was retrieved from the complete whole-genome sequence of *Paenibacillus sp*. PSB04, which originated from Danum Valley, Sabah, Malaysia. The gene sequence of PAP is available in NCBI with the accession number WP 014096696.1. The nucleotide sequence was then used for constructing the homology model of PAP using the ab initio protein structure prediction server of Robetta (Waterhouse *et al*., 2018; Kelley *et al*., 2015; Yang *et al*., 2015; Kim *et al*., 2004). The model was then evaluated and validated according to Global Model Quality Estimation (GMQE), QMEAN statistical parameters, and G factor Procheck (Benkert *et al*., 2011, Laskowski *et al*., 1993; Lüthy *et al*., 1992).

Molecular Docking

The structural model of PAP obtained above was used for molecular docking study as the receptor. Meanwhile, the known substrate for alkaline phosphatases, *para* Nitrophenyl Phosphate (*p*NPP), was selected as the ligand. Both receptors and ligands were prepared by minimizing the energy and then 3D protonating in MOE (Scholz *et al*., 2015). Further, the docking of ligand onto PAP was performed using yasara Structure software (Yet Another Scientific Artificial Reality Application) (Krieger and Vriend, 2015). Briefly, the receptor and ligand files in .pdb format were used to set target and play macro in the yasara structure software. Then, the docking analyses were performed using local and global docking default macro file dock_runlocal.mcr and dock_run.mcr, respectively. The macro files were used to calculate the interaction energy (kcal/mol) and dissociation constant, K_D (pm), of the PAP protein complexes. Meanwhile, local docking was executed by a pre-define square grid with the size of 10Å × 10Å × 10Å around the catalytic site of the receptor. Further, the receptor file was saved in. yob format as required for the local docking study. Both local and global docking study undergoes 25 VINA docking runs of the ligand to the receptor. Afterward, the PAP complexes were converted into PDB files by YASARA software and visualized for its 2D-3D interactive study with the help of Discovery Studio Visualizer version v19.1.0.18287 (BIOVIA, San Diego, CA, USA) and PyMol Software. To note, the selection of the area around the catalytic site for the binding site during docking was based on the finding of other APs which showed that the substrate docked into this area. The catalytic sites of PAP were identified using sequence alignment as detailed below.

Identification of the Essential Residues for the Catalysis

The identification of active site residues reveals the possible catalytic mechanism of the PAP proteins and the active site residues of alkaline phosphatase are known to be conserved. In this study, the identification of the PAP protein was carried out by multiple sequence alignment and validated by determining its configuration based on the position and distance of the active site.

Multiple sequence alignment of the amino acid sequences was done using T-coffee (Di Tommaso *et al*., 2011). Before amino acid Multiple Sequence Alignment (MSA), the sequences were extracted from the Protein Data Bank (PDB) in FASTA format. The MSA was adapted from the constructed multiple sequence alignment by Wojciechowski *et al*. (2002) and Wende *et al*. (2010). The amino acid sequence of Alkaline phosphatase from *Escherichia coli* (PDB: 1ED8), Human placental (PDB: 1EW2), *Halobacterium salinarum* (PDB: 2), Tab5 AP (PDB: 2IUC), Vibrio (PDB: 3E2D), Shrimp AP (PDB: 1SHN), Parageobacillus (PDB: C1K6P2), *Geobacillus* (PDB: A8WEG), *Bacillus subtillis* (PDB: P1940), Pyro coccus (PDB: Q9UZV2), Thermus (PDB: Q153J0) and *Halomonas* (PDB: 3WBH) were retrieved from PDB database. The generated file from the T-coffee server was then displayed in a shaded box display using the Box-shade tools in the Expasy server to visualize better the aligned residues (Duvaud *et al*., 2021). The well-studied *Escherichia coli* AP (ECAP) was used as the primary reference for studying its possible catalytic mechanism (Holtz and Kantrowitz, 1999).

Molecular Dynamic (MD) Simulation

MD simulation was performed to determine the structural stability of PAP in apo and holo forms. The simulation was done using the YASARA structure package (Krieger *et al.*, 2002). The simulation was carried out in an explicit water environment, at constant pressure, using an AMBER14 force field, in a periodic cell

boundary condition and the model was simulated at 298 K (25°C), pH 7.4. First, the solvated structure was minimized by the steepest descent method for 15,000 steps at a temperature of 298 K and a constant pressure. Then the complex was equilibrated for a 2 ns period. After equilibration, a production MD was run for 100 ns at a constant temperature and pressure. The Root Mean Square Fluctuation (RMSF) values were analyzed for each simulation. The RMSF for the average equilibrium conformation is often used as an indicator of protein rigidity, which is considered a feature related to the resistance to unfolding (Kuzmanic and Zagrovic, 2010). All the prediction and simulation studies were done on a Dell XPS8940 with the processor of Intel®Core™ i7-10700 CPU@2.90GHz.

Results and Discussion

Structural Homology Modeling and Validation of PAP

Robetta modeling structure produced five structural models of PAP (model 01-05). To decide on the most reliable PAP model, further structural evaluation and validation were performed on the five models, as shown in Table 1. All models exhibited more than 30% sequence identity against their template (Table 1). Xiang (2006) reported that a protein sequence with over 30% identity to a known structure could often be predicted with an accuracy equivalent to a low-resolution X-ray structure. This indicated that all models are reliable in terms of their sequence identity. Nevertheless, as shown in Table 1, the QMEAN value of model 02 was 0.2, which is the closest to 0. The score reflected the errors of the model structure based on its geometrical aspects globally (entire structure) and locally (residue level). QMEAN value around zero indicates a good agreement between the model structure and experimental structures of similar size. Scores of -4.0 or below indicate low-quality models (Benkert *et al*., 2011). Further, the Ramachandran plot (Table 1) also revealed that while model 02 has only 94% residues in the favored region, this model has the lowest residues located in the lower region. In addition, the score of Verify 3D was also found to be higher than the requirement (>80%), which indicated the correctness of the model. This score evaluates segments of the model based on how well the environments of the residues in that segment (e.g., burial, secondary structure) correlate with their observed propensities for being in those environments (Eisenberg *et al*., 1997). In addition, the overall average G-factor of dihedral angles and main-chain covalent forces of model 02 was 0.18, which was greater than the acceptable cut-off of -0.5. The G-factor provides a measure of the plausibility of a stereochemical property and a high G-factor means the property corresponds to a high probability of conformation (Aslanzadeh and Ghaderian, 2012). Altogether, model 02 was considered the best model of PAP.

As shown in Fig. 1, the best structural model of PAP forms a dimeric structure, in which each monomer consists of 61 (14.6%) α-helix and 124 (29.6%) β-sheets (Fig. 2). There were fourteen predicted α-helical structures; α1 span from Lys 55 to Va l65, followed by α2 from Gln 73 to Asp 76, α 3 span from Ser 94-Ser 103, α4 span from Leu 124-Glu 130, α5 span from Ala 146-Ala 152, α6 span from Gln 161-Glu 170, α 7 span from Glu 181 E-Trp 184, α 8 span from Leu 212-Gln 218, α9 span from Lys 227-Ser 231, α 10 spans from Leu 265-Leu 276, α11 spans from Gly 291-His 297, α 12 spans from Ala 301-Lys 323, α 13 spans from Ala 395-he 398 and lastly, α 14 spans from Tyr 406-Trp 415 and ten β-sheaths; β1which span from Gln 44-Gly 50, β2 was from Tyr 79-Arg 84, β3 at Ser 134-Asn 142, β4 at Val 175-Gly 179, β5 at Asn 222-Val 224, β6 at Ile 237-Leu 241, β7 at Phe 283-Glu 289, β8 at Thr 327-Leu 333, β9 at Val 386-Thr 391 and lastly β 10 at Gly 400-Glu 403, that build up the structural component of PAP protein.

The dimeric structure of PAP follows the typical quaternary structure of other bacterial APs, which were widely reported to be homo-dimeric structures in crystal structure and solution (Aiba *et al*., 2017; Sharma *et al*., 2014). The active site of PAP (Ser 94 and Arg 158) is located between α 5 and β 7 (Fig. 2). These residues are well conserved to the other well-reported bacterial APs. Structural comparison with other APs indicated that PAP has the closest similarity to human placenta AP (PDB ID: 1EW2), with an RMSD of 1.04 Å, as shown in Fig. 1C. Meanwhile, the structural comparison with *E. coli* AP (ECAP), the most extensively studied AP, revealed a moderate structural similarity between both structures, with an RMSD of 2.12 Å (Fig. 1D). The RMSD reflected the similarity between two proteins, which was computed between aligned pairs of their backbone Cα atoms. The lower RMSD indicated a small deviation between the Cα; hence both structures are considered similar (Reva *et al*., 1998). While there is no definite RMSD threshold, an RMSD of less than two is often used to consider the high similarity of the aligned structures. Further detailed analysis of the dimeric interface revealed some hydrophobic interactions that are essential for dimerization stability. The critical involvement of hydrophobic interactions in the dimeric interface was found to be common in some proteins (Budiman *et al*., 2011).

Identification of the Catalytic Active Sites and Metal-binding Sites of PAP

A multiple amino acid sequence alignment of PAP with other selected APs is shown in Fig. 3. The alignment revealed that Ser 94 and Arg 158 of PAP were well aligned to the active sites of ECAP (Ser 102 and Arg 166), which implied that these two residues are likely to be the catalytic sites of PAP. In addition, almost all the residues with essential roles in the metal-binding and active sites

were well conserved. This indicated that PAP might utilize the canonical catalytic mechanism of APs with the involvement of similar metal ions. This, nevertheless, remain to be experimentally confirmed.

Molecular Docking of PAP

Docking simulation was performed between PAP and *p*NPP as the ligand, using global and local docking approaches. In total, seven complex models were obtained, where the best model was depicted based on the consideration of its binding energy, dissociation energy, and orientation towards the ligand. Figure 4(A) showed the best complex of PAP with *p*NPP as a ligand. The complex has a binding energy of -5.9 kcal/moL with dissociation energy of 48.719 um.

Further, MD simulation revealed that the average RMSF values of the apo- and holo forms of PAP were 2.27 and 2.21 Å, respectively (Fig. 5). This is in the range of acceptable RMSF value (1-3 Å) for protein to be considered structurally stable (Parikesit and Nurdiansyah, 2021). To note, some segment residues showed fluctuation higher than 3. The segments include 65-70, 83-92, 105-120, 150-165, 182-210, 250-260, and 340-375 (Fig. 2). These segments are located in the loop region, therefore acceptable for high flexibility during the simulation. Cheng and Ivanov (2012) indicated that RMSF reflects the flexibility of different regions of a protein, which can be related to crystallographic B factors. Accordingly, it is acceptable for the loop region to have high RMSF values during the simulation due to high B factors. Altogether, the MD simulation convincingly indicated that the model of PAP, in the absence or the presence of substrate (pNPP), is considered stable and reliable for the analysis.

The complex showed that the Ser 94 and Arg 158 catalytic sites were in close contact with the ligand and the distance between Ser 94 and the phosphoryl group of pNPP was 3.4 Å. The side chain of Ser 94 facing toward the phosphate group of the ligand indicates the possibility for the first catalysis reaction to occur. In the case of ECAP, the nucleophilicity of Ser catalytic residue is activated by the Zn^{2+} ion. While our model was unable to depict the presence of any metal ion, nevertheless, the sequence alignment (Fig. 3) indicated the possibility of a Zn binding site along with the sequences. To note, the possibility of the presence of a Zn binding site in PAP is solely based on the conservation residues with the known Zn binding site of ECAP (Fig. 3), which indeed requires experimental confirmation. Prediction of the zinc-binding site through sequence alignment with known Zn binding proteins is considered a useful tool for discovering new zinc-binding sites (Furukawa *et al*., 2018; Hubbard *et al*., 1991). Figure 4(B) showed one of the coordination of metal-binding sites around Ser 94 and Arg 158 catalytic sites, which indicated that the site formed by Asp 335 or His 336 or Thr 147 or Asp 293 is in proximity to Ser 94. This implied that the activation of Ser 94 nucleophilicity by Zn^{2+}, coordinated by any of those residues, is an acceptable route.

On the other hand, Fig. 4(A) indicates the distance between Arg 158 of PAP and PO_4 moiety of the ligand is in the range of 5.9-7.0 Å. Holtz and Kantrowitz (1999) reported that the catalytic site of Arg 166 of ECAP directly interacts with the PO_4 substrate through its side chain. During the catalysis, its guanidinium group is vital to stabilize the transition state of the reaction. Notably, the distance between Arg 158 of PAP and PO_4 is unfavorable for the direct interaction (5.9-7.0 Å). Nevertheless, the position of Arg 158 is located at the loop region (between Ser 153 and Lys 160), which indicates that the 5.9-7.0 Å-distance is quite dynamic. The Arg158 can be at a favorable distance to the PO_4 for catalysis due to the mobility of the loop. Binding to the ligand may also stabilize the loop to be an inappropriate distance for the catalysis. A similar phenomenon was also reported for a small cysteine protease of pineapple MD2, harboring mobile catalytic site Cys at the loop region. Structural modeling of the Cys is quite distant from the ligand, yet it exhibits remarkable activity (Razali *et al*., 2020). Further, it was also known that the loop structure is highly flexible and mobile and the dynamic loop structure facilitates the catalytic activity of PAP (Hjörleifsson *et al*., 2021; Razali *et al*., 2020). Alternatively, a water molecule might mediate the interaction between the guanidinium group of Arg158 to PO_4. To note, the 5.9-7.0 Å-distance is sufficient to accommodate a water molecule. Interestingly, 4B also showed some metal-binding sites close to Arg 158, mainly Thr 147 and Asp 293. This leads to a possibility of metal ion-mediate interaction of Arg 158 to the substrate.

Further, a 2D interaction map between PAP and pNPP (Fig. 6) showed that the interaction between PAP and *p*NPP is dominated by non-polar interaction, facilitated by Ser 94, Ser 141, Ala 146, Thr 147, Pro 148, Asp 293, and Glu 294. Interestingly, Arg158 was not depicted in the 2D map, which indicated that the interaction of this residue to the ligand might not be in a direct way. To note, the docking simulation in this study could only depict the direct interaction, with no water molecule involved. Further, the role of Arg 158 in catalysis is an exciting avenue to be experimentally investigated.

Further, when the essential residues for catalysis of ECAP were aligned to the residues of PAP, most of the residues were highly conserved, except for Glu 294 of PAP (corresponds to Lys 328 ECAP) (Fig. 7). To note, Asp 101, Asp 153 and Lys 328 of ECAP were reported to form secondary interactions with phosphate through either a water molecule or Arg-166. Regarding Lys 328 of ECAP, a salt link between Lys 328 and Asp 153 is formed, which further orients Asp 153 to interact with Mg^{2+} for stabilizing the PO_4 moiety in the substrate cavity. Lys 328 could also interact with water molecules which further protonates the oxygen atom of PO_4 moiety for catalysis.

In PAP, Asp 153 of ECAP corresponds to Asp145, located close to Glu 294. Nevertheless, Glu 294 and Asp 145 would not form a salt bridge due to the repulsion between these residues. The absence of the salt link might lead to a "swing" position of Glu 294, which allows this residue to interact with the phosphate substrate. This assumption is supported by the 2D map (Fig. 6), which showed direct interaction between Glu 294 to the ligand. This direct interaction differs from Lys 328 of ECAP, which interacts with the ligand through a water molecule. In addition, a "swing" orientation was also observed when Lys 328 was mutated to His in ECAP, which further increased the ligand-binding stability (Holtz and Kantrowitz, 1999). This might be the reason for the high catalytic activity of PAP, as reported earlier (Lindang *et al*., 2021). In addition, the replacement of Lys 328 with Glu 294 might also change the charge of the substrate-binding cavity of PAP. Koutsioulis *et al.* (2010) and O'Brien *et al.* (2008) reported that the charge of the residues within the catalytic binding pocket was found to have a serious effect on the catalytic activity of AP. The ZnI metal biding residue is located next to Glu 294 and it is postulated that Asp 293 coordinates the Zn^{2+} and facilitates the interaction of these two residues. Nevertheless, the proposed indirect interaction can be experimentally confirmed by an x-ray crystallography study.

To note, it is not only PAP that has different residues at the Lys 328 position. The AP of *Geobacillus thermodenitrificus* strain T2, GtAP (A8WEG4), was also reported to have Glu at the position corresponding to Lys 328 of ECAP. In addition, the replacement of Lys 328 of ECAP to His can be seen in the human placenta (1EW2), halobacterium (2 X 98), ShrimpAP (1SHN), and Pyrococcus (Q9UZV2). Meanwhile, replacement of the Lys 328 to Trp was seen in *Antarctic bacteria* Tab5 (2IUC), *Parageobacillus* (C1K6P2), *Bacillus subtillis* (P1940), Vibrio (3E2D), and *Halomonas* (3WBH). This indicated that another residue could take over the role of Lys 328 in the catalysis, which further leads to a unique catalytic mechanism due to the specific part of the non-Lys residue at this position.

Table 1: Evaluation and validation of the generated structural quality of PAP

Model	01	02	03	04	05
QMEAN	-0.50	-0.20	-1.30	-0.90	-0.50
Ramachandran residues in favoured region (%)	96.76	94.80	97.10	95.92	97.10
Ramachandran residues in outlier region (%)	0.20	0.10	0.72	0.40	0.20
Ramachandran residues in allowed region (%)	3.00	5.00	2.20	3.80	2.60
Sequence identity	30.29	30.29	30.29	30.29	30.29
Verify3D (%)	89.74	90.69	86.87	93.56	89.02
G factor	0.19	0.18	0.16	0.16	0.53

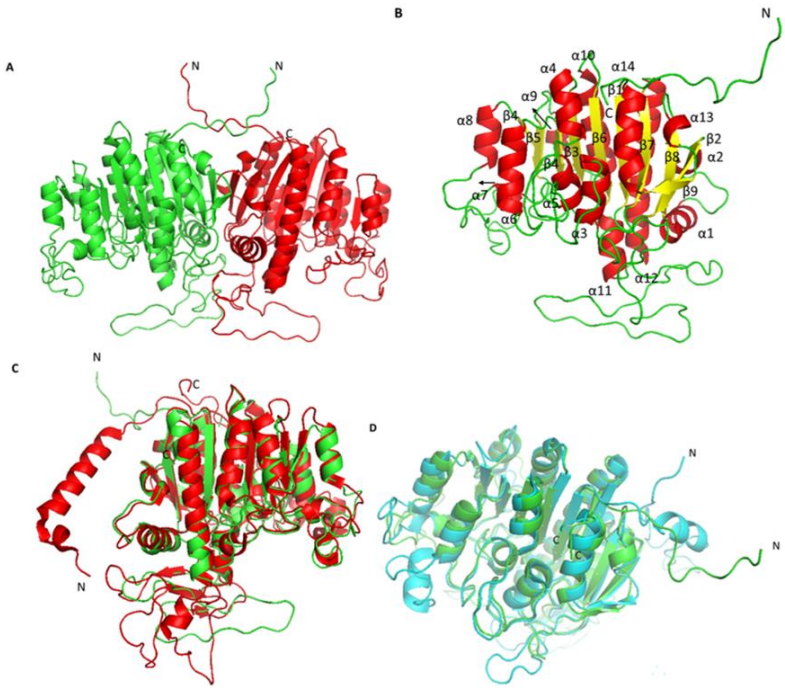


Fig. 1: (A) Three-dimensional model of PAP in its dimeric form. (B). The monomeric form of PAP. The α-helix and β-sheets secondary structures are indicated by red and yellow, respectively, and numbered sequentially. (C) Structural alignment of PAP and human placental AP (PDB ID: 1EW2), colored in green and red, respectively. (D). Structural alignment of PAP and *E. coli* AP (PDB ID: 1ALK), colored in green and cyan, respectively. The N-and C-terminals are also indicated for clarity

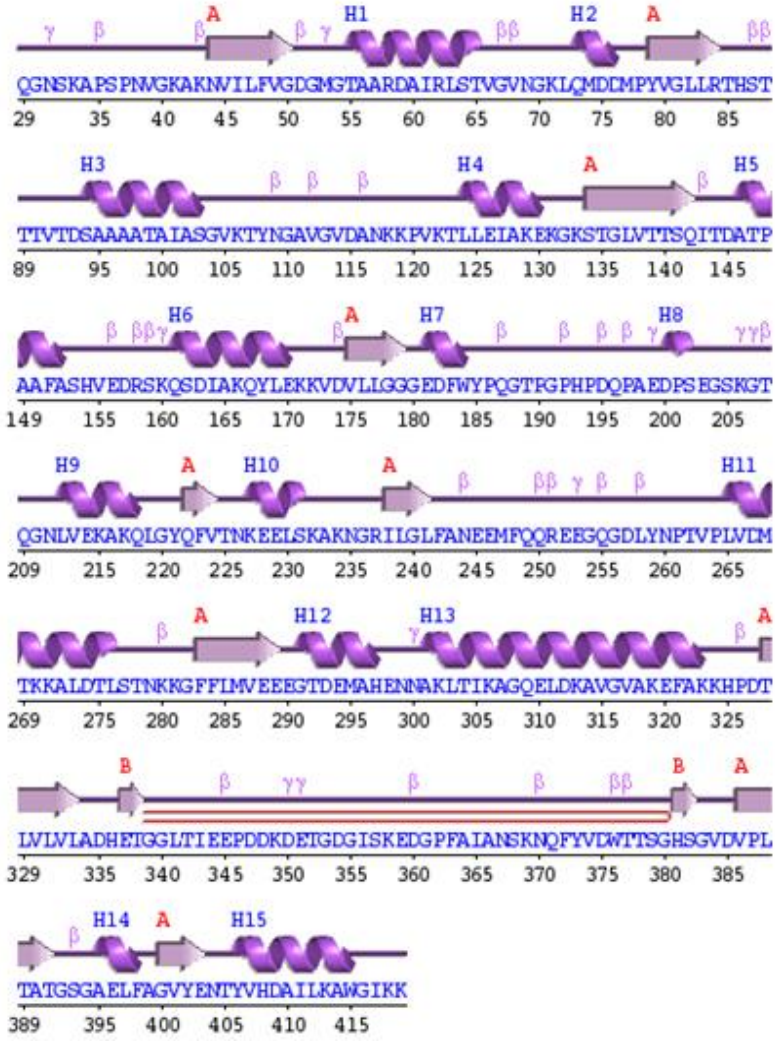
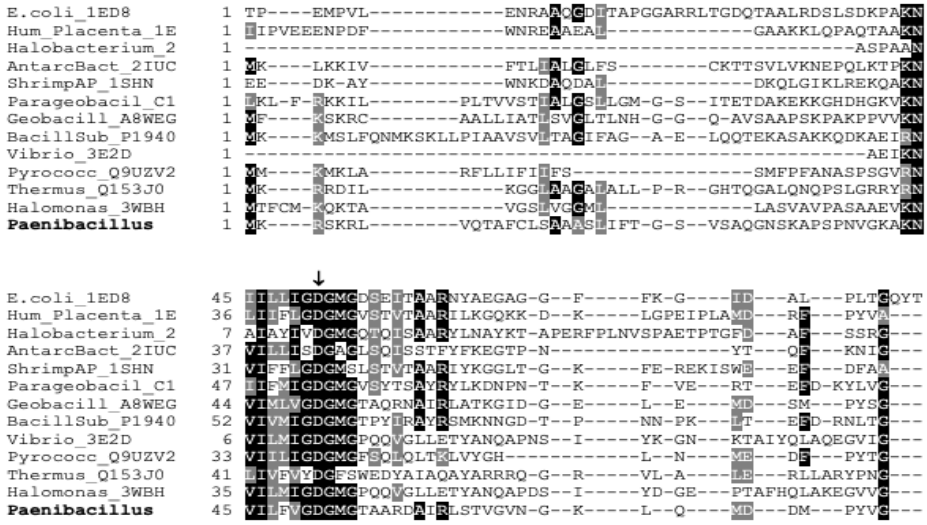
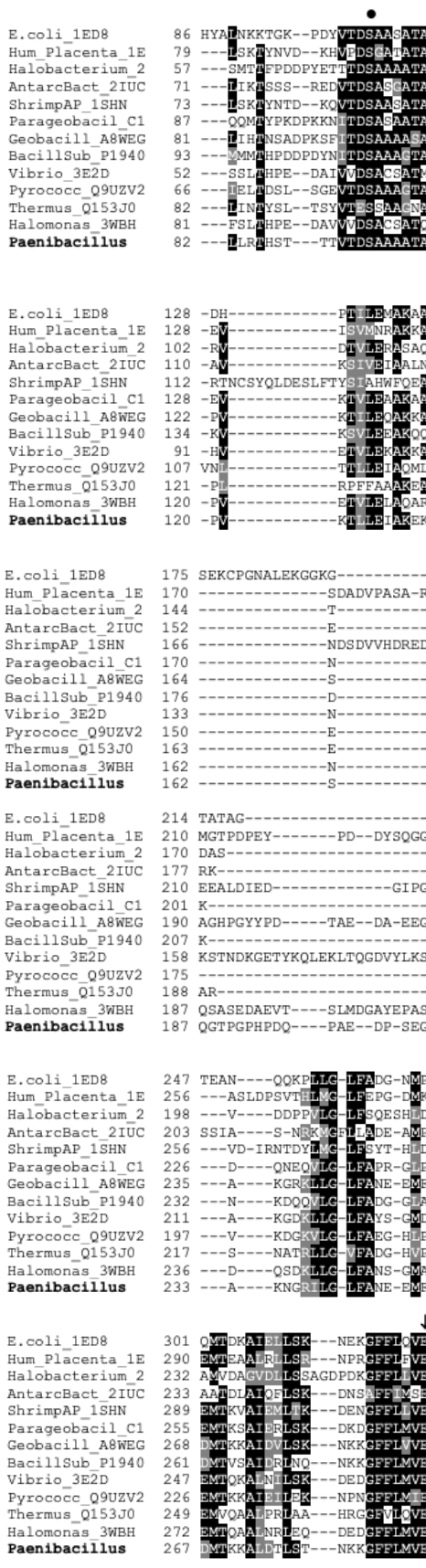


Fig. 2: Secondary structure prediction of PAP generated by PDBSum server




E.coli_1ED8
Hum_Placenta_1E
Halobacterium_2
AntarcBact_2IUC
ShrimpAP_1SHN
Parageobacil_C1
Geobacill_A8WEG
BacillSub_P1940
Vibrio_3E2D
Pyrococc_Q9UZV2
Thermus_Q153J0
Halomonas_3WBH
Paenibacillus

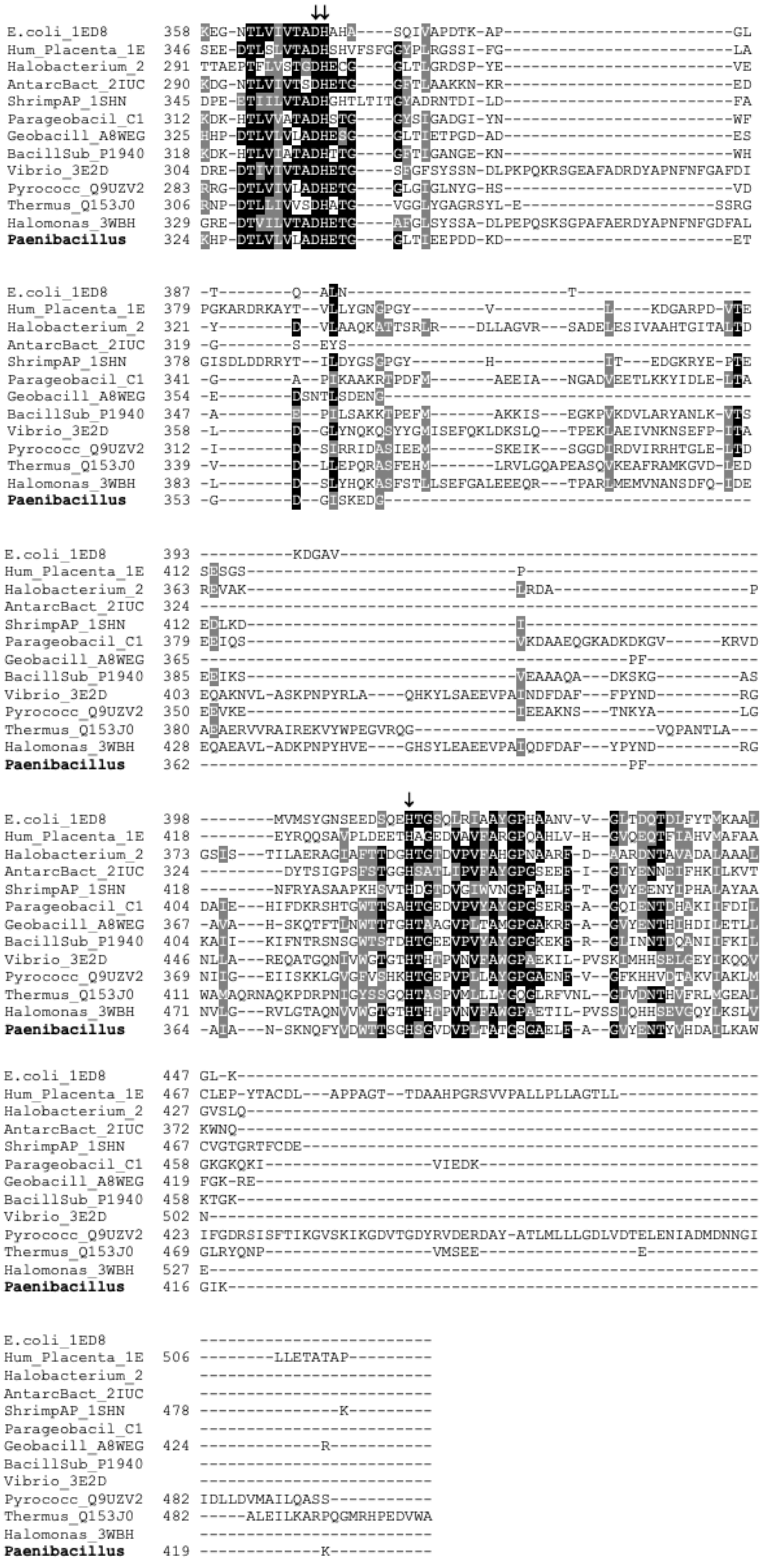


Fig. 3: Multiple Sequence Alignment of PAP amino acid sequence with previously reported APs. Symbols above the sequences indicate the conserved residues which are involved in the active site, according to the following legend: ↓, metal ion coordination (ZnI, ZnII, Mg primary ligand); ⇓, metal ion coordination (Mg secondary ligand; •, phosphoryl intermediate formation; **O**, phosphate coordination. The peptide sequence (Met1-Xx) was omitted in this alignment for clarity

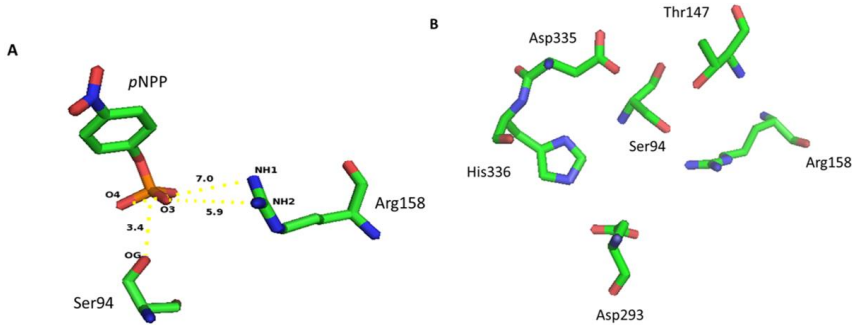


Fig. 4(A): The orientation between catalytic sites (Ser 94 and Arg 158) and pNPP substrate. The yellow dotted line indicated the distance between the atoms in the Angstrom unit (Å). 4(B)**.** Location of the putative metal-binding sites (Thr 147, Asp 335, His 336, and Asp 293) around the catalytic sites (Ser 94 and Arg 158)

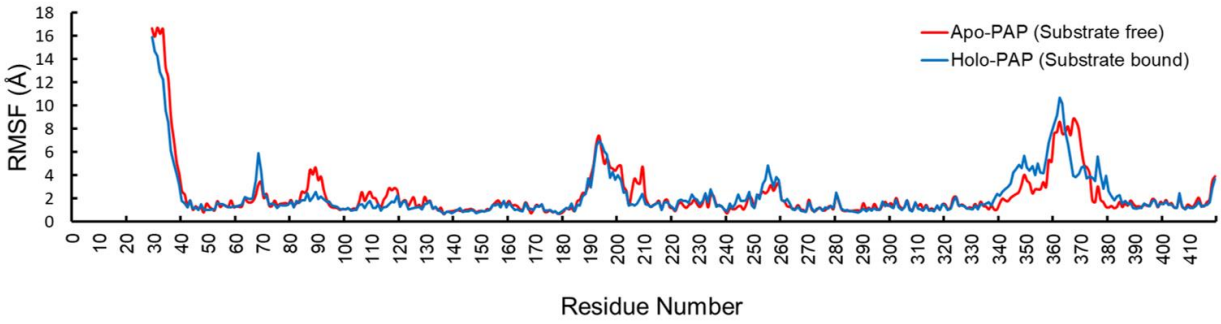


Fig. 5: Fluctuation (RMSF) of holo- and Apo-forms of PAP during 100 ns simulation

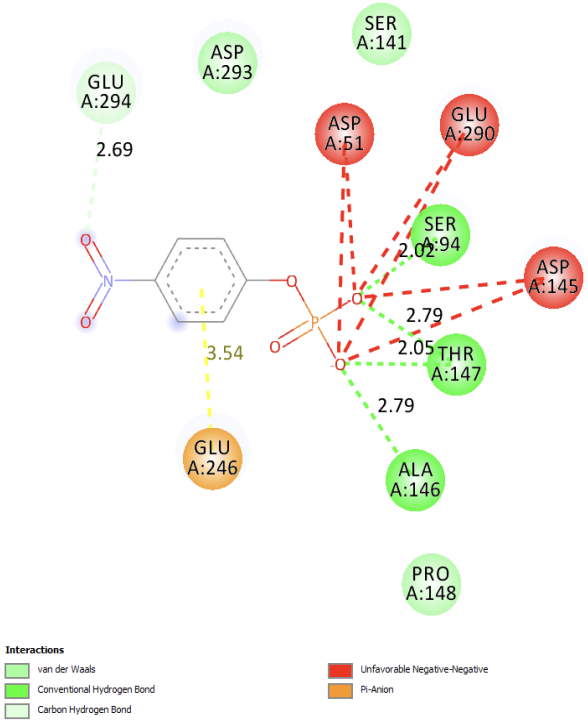


Fig. 6: 2D visualization of residues within the catalytic binding pocket of PAP with pNPP as the ligand

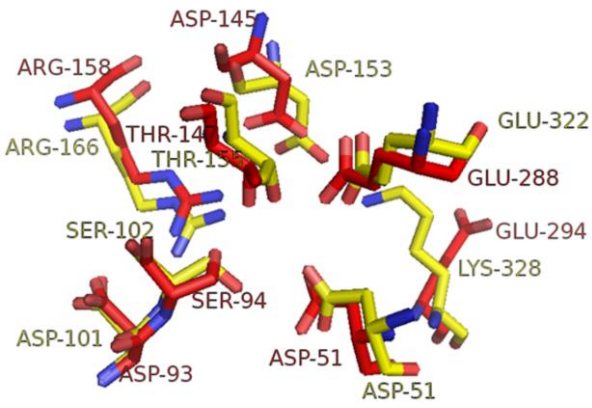


Fig. 7: Structural alignment of the important residues in PAP (red color) and ECAP (yellow color)

Conclusion

Structural homology modeling showed that PAP forms a dimeric structure with canonical catalytic sites of Ser 94 and Arg 158. The sequence alignment of PAP revealed the binding sites for Zn^{2+} and Mg^{2+} ions, which are suspected of playing an essential role in catalysis. The docking simulation of PAP to the *p*NPP substrate and comparison with ECAP revealed that the uniqueness of Glu 294, which correspond to Lys 328 of

ECAP, might account for the unique catalysis mechanism employed by PAP. This uniqueness may also contribute to the high activity of PAP and is promising to be further harnessed for further applications.

Acknowledgment

H.U.L is a recipient of a Ph.D. scholarship under the Malaysia Public Service Department (PPC 2018).

Funding Information

This research is under UMSGreat Research Grant of GUG0277-2/2018.

Author's Contributions

Herman Umbau Lindang: Conceptualization, methodology and funding acquisition, Data collection, writing-original draft preparation.

Cahyo Budiman: Conceptualisation, methodology and funding acquisition, Data analysis and validation, writing-review and finalization of the manuscript.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all other authors have read and approved the manuscript and no ethical issues are involved.

References

Afiyah, D. N., Arief, I. I., & Budiman, C. (2015). Proteolytic characterization of trimmed beef fermented sausages inoculated by Indonesian probiotics: Lactobacillus plantarum IIA-2C12 and Lactobacillus acidophilus IIA-2B4. Advance Journal of Food Science and Technology. Advance Journal of Food Science and Technology, 8(1), 27-35. doi.org/10.19026/ajfst.8.1459

Aiba, H., Nishiya, Y., Ojima, Y., & Azuma, M. (2017). Over-expression, characterization and modification of highly active alkaline phosphatase from a Shewanella genus bacterium. Bioscience, Biotechnology and Biochemistry, 81(10), 1994-2001. doi.org/10.1080/09168451.2017.1356217

Ásgeirsson, B., Markússon, S., Hlynsdóttir, S. S., Helland, R., & Hjörleifsson, J. G. (2020). X-ray crystal structure of Vibrio alkaline phosphatase with the non-competitive inhibitor cyclohexylamine. Biochemistry and Biophysics Reports, 24, 100830. doi.org/10.1016/j.bbrep.2020.100830

Aslanzadeh, V., & Ghaderian, M. (2012). Homology modeling and functional characterization of PR-1a protein of Hordeum vulgare subsp. Vulgare. Bioinformation, 8(17), 807. doi.org/10.6026/97320630008807

Benkert, P., Biasini, M., & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics, 27(3), 343-350. doi.org/10.1093/bioinformatics/btq662

Budiman, C., Angkawidjaja, C., Motoike, H., Koga, Y., Takano, K., & Kanaya, S. (2011). Crystal structure of N-domain of FKBP22 from *Shewanella sp*. SIB1: Dimer dissociation by disruption of Val-Leu knot. Protein Science, 20(10), 1755-1764. doi.org/10.1002/pro.714

Chen, C. C., Cheng, K. J., Ko, T. P., & Guo, R. T. (2015). Current progress in phytase research: Three-dimensional structure and protein engineering. Chem Bio Eng Reviews, 2(2), 76-86. doi.org/10.1002/cben.201400026

Cheng, X., & Ivanov, I. (2012). Molecular Dynamics. In A. N. Reisfeld Brad and Mayeno (Ed.), Computational Toxicology. Humana Press. pp, 243-285. doi.org/10.1007/978-1-62703-050-2_11

Coker, O. O., Warit, S., Rukseree, K., Summpunn, P., Prammananan, T., & Palittapongarnpim, P. (2013). Functional characterization of two members of histidine phosphatase superfamily in Mycobacterium tuberculosis. BMC Microbiology, 13(1), 1-12. doi.org/10.1186/1471-2180-13-292

Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J. M., & Notredame, C. (2011). T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Research, 39 (suppl_2), W13-W17. doi.org/10.1093/nar/gkr245

Dong, Y., Xia, Y., Yin, J., Zhou, D., Sang, Y., Yan, S., & Tu, Q. (2021). Optimization, characteristics and functions of alkaline phosphatase from *Escherichia coli*. Frontiers in Microbiology, 12, 761189-761189. doi.org/10.3389/fmicb.2021.761189

Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V., & Durinx, C. (2021). Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. Nucleic Acids Research, 49(W1), W216-W227. doi.org/10.1093/nar/gkab225

Eisenberg, D., Lüthy, R., & Bowie, J. U. (1997). VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods in Enzymology, 277, 396–404. doi.org/10.1016/s0076-6879(97)77022-8

Furukawa, Y., Lim, C., Tosha, T., Yoshida, K., Hagai, T., Akiyama, S., & Shiro, Y. (2018). Identification of a novel zinc-binding protein, C1orf 123, as an interactor with a heavy metal-associated domain. PloS One, 13(9), e0204355. doi.org/10.1371/journal.pone.0204355

Helland, R., Lie, R., & Ásgeirsson, B. (2009). The 1. 4 Å crystal structure of the large and cold-active Vibrio sp. alkaline phosphatase. BBA-Proteins and Proteomics, 2, 297–308. doi.org/10.1016/j.bbapap.2008.09.020

Hjörleifsson, J. G., Helland, R., Magnúsdóttir, M., & Ásgeirsson, B. (2021). The high catalytic rate of the cold-active Vibrio alkaline phosphatase requires a hydrogen-bonding network involving a large interface loop. FEBS Open Bio, 11(1), 173-184. doi.org/10.1002/2211-5463.13041

Holtz, K. M., & Kantrowitz, E. R. (1999). The mechanism of the alkaline phosphatase reaction: Insights from NMR, crystallography and site-specific mutagenesis. FEBS Letters, 462(1-2), 7-11. doi.org/10.1016/S0014-5793(99)01448-9

Hubbard, S. R., Bishop, W. R., Kirschmeier, P., George, S. J., Cramer, S. P., & Hendrickson, W. A. (1991). Identification and characterization of zinc-binding sites in protein kinase C. Science, 254(5039), 1776-1779. doi.org/10.1126/science.1763327

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols, 10(6), 845-858. doi.org/10.1038/nprot.2015.053

Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. Nucleic Acids Research, 32(suppl_2), W526-W531. doi.org/10.1093/nar/gkh468

Koutsioulis, D., Lyskowski, A., Mäki, S., Guthrie, E., Feller, G., Bouriotis, V., & Heikinheimo, P. (2010). The coordination sphere of the third metal site is essential to the activity and metal selectivity of alkaline phosphatases. Protein Science, 19(1), 75-84. doi.org/10.1002/pro.284

Krieger, E., Koraimann, G., & Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. Proteins, 47(3), 393–402. doi.org/10.1002/prot.10104

Krieger, E., & Vriend, G. (2015). New ways to boost molecular dynamics simulations. Journal of Computational Chemistry, 36(13), 996-1007. doi.org/10.1002/jcc.23899

Kuzmanic, A., & Zagrovic, B. (2010). Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. Biophysical Journal, 98(5), 861–871. doi.org/10.1016/j.bpj.2009.11.011

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). Procheck: A program to check the stereochemical quality of protein structures. Journal of Applied Crystallography, 26(2), 283-291. doi.org/10.1107/S0021889892009944

Lindang, H. U., Subbiah, V. K., Rodrigues, K. F., & Budiman, C. (2021). Isolation, identification and characterization of phosphate solubilizing bacteria, *Paenibacillus sp*., from the soil of danum valley tropical rainforest, Sabah, Malaysia. Biodiversitas, 22(10), 4370–4378. doi.org/10.13057/biodiv/d221030

Liu, C., Yang, C., Yang, Q., Wang, J., & Liu, Y. (2021). Effect of divalent metals on phytase activity and enzymatic kinetics of As-hyperaccumulation ferns. Acta Agriculture Scandinavica, Section B-Soil and Plant Science, 71(2), 112-123. doi.org/10.1080/09064710.2020.1856917

Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. Nature, 356(6364), 83-85. doi.org/10.1016/S0076-6879(97)77022-8

Millán, J. L. (2006). Alkaline phosphatases. Purinergic Signaling, 2(2), 335-341. doi.org/10.1007/s11302-005-5435-6

O'Brien, P. J., Lassila, J. K., Fenn, T. D., Zalatan, J. G., & Herschlag, D. (2008). Arginine coordination in enzymatic phosphoryl transfer: Evaluation of the effect of Arg 166 mutations in *Escherichia coli* alkaline phosphatase. Biochemistry, 47(29), 7663-7672. doi.org/10.1021/bi800545n

Parikesit, A. A., & Nurdiansyah, R. (2021). Natural products repurposing the H5N1-based lead compounds for the most fit inhibitors against 3C-like protease of SARS-COV-2. J Pharm Pharmacogn Res, 9(5), 730-745. doi.org/10.2174/2666796702666210222105547

Razali, R., Kumar, V., & Budiman, C. (2020). Structural insights into the enzymatic activity of protease bromelain of MD2 pineapple. Pak J Biol Sci, 23(6), 829-838. doi.org/10.3923/pjbs.2020.829.838

Reva, B. A., Finkelstein, A. V., & Skolnick, J. (1998). What is the probability of a chance prediction of a protein structure with an rmsd of 6 Å? Folding and Design, 3(2), 141-147. doi.org/10.1016/S1359-0278(98)00019-4

Scholz, C., Knorr, S., Hamacher, K., & Schmidt, B. (2015). Docktite is a highly versatile step-by-step workflow for covalent docking and virtual screening in the molecular operating environment. Journal of Chemical Information and Modeling, 55(2), 398-406. doi.org/10.1021/ci500681r

Sharma, U., Pal, D., & Prasad, R. (2014). Alkaline phosphatase: An overview. Indian Journal of Clinical Biochemistry, 29(3), 269-278. doi.org/10.1007/s12291-013-0408-y

Wang, H., & Griffiths, M. W. (2009). Mg2+-free buffer elevates the transformation efficiency of Vibrio parahaemolyticus by electroporation. Letters in Applied Microbiology, 48(3), 349-354. doi.org/10.1111/j.1472-765X.2008.02531.x

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., & Schwede, T. (2018). SWISS-MODEL: Homology modeling of protein structures and complexes. Nucleic Acids Research. 46(W1), W296-W303. doi.org/10.1093/nar/gky427

Wende, A., Johansson, P., Vollrath, R., Dyall-Smith, M., Oesterhelt, D., & Grininger, M. (2010). Structural and biochemical characterization of a halophilic archaeal alkaline phosphatase. Journal of Molecular Biology, 400(1), 52-62. doi.org/10.1016/j.jmb.2010.04.057

Wojciechowski, C. L., Cardia, J. P., & Kantrowitz, E. R. (2002). Alkaline phosphatase from the hyperthermophilic bacterium T. Maritima requires cobalt for activity. Protein Science, 11(4), 903-911. doi.org/10.1110/ps.4260102

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: Protein structure and function prediction. Nature Methods, 12(1), 7-8. doi.org/10.1038/nmeth.3213