Isolation and characterization of α-amylase encoding gene in *Bacillus amylo liquefaciens* PAS

Achmad Rodiansyah, Sitoresmi Prabaningtyas, Mastika M. Ulfah, Ainul F. Mahmudah, Uun Rohmawati

Laboratory of Microbiology, Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Negeri Malang, Semarang No.5 Malang 65145, Indonesia

*Corresponding author: sitoresmi.prabaningtyas.fmipa@um.ac.id

ABSTRACT Amylolytic bacteria are a source of amylase, which is an essential enzyme to support microalgae growth in the bioreactor for microalgae culture. In a previous study, the highest bacterial isolate to hydrolyze amylum (namely PAS) was successfully isolated from Ranu Pani, Indonesia, and it was identified as *Bacillus amylo liquefaciens*. That bacterial isolate (*B. amylo liquefaciens* PAS) also has been proven to accelerate *Chlorella vulgaris* growth in the mini bioreactor. This study aims to detect, isolate, and characterize the PAS’s α-amylase encoding gene. This study was conducted with DNA extraction, amplification of α-amylase gene with polymerase chain reaction (PCR) method with the specific primers, DNA sequencing, phylogenetic tree construction, and protein modeling. The result showed that α-amylase was successfully detected in PAS bacterial isolate. The α-amylase DNA fragment was obtained 1,468 bp and that translated sequence has an identity of about 98.3% compared to the *B. amylo liquefaciens* α-amylase 3BH4 in the Protein Data Bank (PDB). The predicted 3D protein model of the PAS’s α-amylase encoding gene has amino acid variations that predicted affect the protein’s structure in the small region. This research will be useful for further research to produce recombinant α-amylase.

KEYWORDS α-amylase; *Bacillus amylo liquefaciens*; homology modeling; Ranu Pani

1. Introduction

Amylases, including α-amylase, β-amylase, and glucoamylase, are the most known as amylolytic enzymes, which they can be found in amylolytic bacteria (Gopinath et al. 2017). The α-amylase is a general enzyme with a dominant application in starch-related industries, such as textiles, papers, and pharmaceuticals (Mehta and Satyarnayana 2016). This enzyme can hydrolyze the α-1,4 glycosidic linkages of polysaccharides for resulting simpler molecules like monosaccharides (Abd-Elhalem et al. 2015). The α-amylases produced by bacteria are often used in industry because the microbial strains are easy to culture under designed conditions correlated with high production of α-amylase (Gopinath et al. 2017). The genetic engineering to produce recombinant α-amylase could improve their stability in the extreme conditions (Far et al. 2020).

The exploration of bacteria communities in several East Java lakes, Indonesia, had found many potential bacteria types, including amylolytic bacteria (Prabaningtyas and Witjoro 2017; Prabaningtyas et al. 2018; Nafi’ah et al. 2021). Ranu Pani, one of a lake located in Lumajang with an altitude of 2,200 meters above sea level (masl), contains about 18.18% organic substrates that allow the decomposition of organic matter by microorganisms, especially amylolytic bacteria (Gazali et al. 2015). The highest activity from amylolytic bacteria (isolate code: PAS) isolated from Ranu Pani based on the 16S rRNA gene barcode is identified as *Bacillus amylo liquefaciens*, which the sequence similarity is homogenous within other species in the *Bacillus subtilis* group reaching > 99% identity (Rodiansyah et al. 2021). Moreover, the biochemical characterization of that bacterial isolate also showed the similar characteristics with *B. subtilis*, reaching 66% identity. Bacterial isolate PAS could reduce the complex sugar about of 27,391 ppm with the enzyme activity of 0.01 units/mL (Nisa et al. 2021) and the amylum hydrolysis index of that isolate was about of 5.9 (Nafi’ah et al. 2021). Based on that result, the amylolytic activity of PAS bacterial isolate was relatively high.

The α-amylase discovery produced by microorganisms from different environments could provide novel amylases suitable for many applications in related industries (Gupta et al. 2014). The α-amylase sequence characterization from this bacterial isolate is important for further study, especially for their application to enhance the microalgae biomass in the mini bioreactor (Fuentes et al. 2016; Han et al. 2016) and to produce α-amylase recombining...
nant (Niu et al. 2009). In our study, the culture of Chlorella vulgaris with the co-culture method in the mini bioreactor containing modified-Gusrina medium combined with PAS bacterial culture had proved to improve microalgae growth and biomass production (Nafi’ah et al. 2021).

In this study, we successfully isolate and characterize the α-amylase encoding gene from B. amyloliquefaciens isolated from Ranu Pani, Indonesia. This result is essential to confirm the amylolytic activity from PAS bacterial isolate based on its encoding DNA sequence and to build recombinant DNA for α-amylase expression and enzyme engineering study.

2. Materials and Methods

2.1. Isolate and media

A single bacterial isolate with the highest potency to hydrolyze amylose named B. amyloliquefaciens PAS was obtained from the previous study. This isolate stored in nutrient agar was inoculated into 5 mL nutrient broth (NB). The NB medium was prepared with 5 g peptone (Merck KGaA, Darmstadt, Germany) and 3 g beef extract (Merck KGaA, Darmstadt, Germany) homogenized in aquadest for 1,000 rpm, 37 °C, overnight. The bacterial cells were harvested from the medium with serial centrifugation at 7,500 rpm, 37 °C, overnight. The bacterial cells were harvested from the medium with serial centrifugation at 7,500 rpm for 5 min, and hold at 4 °C. Next, the amplification products were checked on 1 percent agarose gel electrophoresis from SeaKem® LE (Lonza, Basel, Switzerland) with 1 kb DNA marker (Geneaid, New Taipei City, Taiwan). The gel was run in a Mupid-exU system (Takara, California, USA) with voltages at 50V for 1 h. Then, the gel was visualized on the UV-transilluminator.

The PCR product was next used for DNA sequencing. Pair-read sequencing was carried out by the Sanger sequencing method through the 1st Base Malaysia DNA sequencing service. The AM-PAS forward and reverse primers also were used for DNA sequencing. The sequencing results were checked with FinchTV version 1.5.0 (available at: www.digitalworldbiology.com/FinchTV) (Geospiza 2004), and the contig sequence was built with a DNA baser (available at: www.dnabaser.com) (SRL 2014).

2.3. Multiple alignment and phylogenetic analysis

The PAS’s α-amylase DNA sequence consensus from the sequence contig, namely AM-PAS, was aligned using the BlastX program from NCBI (available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al. 1990), which had set to protein data bank (PDB) database. The AM-PAS DNA contig sequence was converted into the protein sequence with MegaX software (available at: www.megasoftware.net) (Kumar et al. 2018) and ORF finder (available at: https://www.ncbi.nlm.nih.gov/orffinder/) (Wheeler et al. 2003).

The AM-PAS protein sequence was re-aligned in the global database using the protein blast (BlastP) program (Altschul et al. 1997). The protein sequences from BlastP were downloaded, and those sequences were used for multiple alignments with ClustalX software (available at: www.clustal.org) (Larkin et al. 2007). The aligned sequences from ClustalX were analyzed and compared the diversity of its amino acids composition using a graphic view tool in BioEdit software (available at: https://bioedit.software.informer.com/) (Hall 1999). Moreover, that alignment file was used for phylogenetic tree construction using MegaX software. The phylogenetic tree was constructed with the UPGMA method (Sneath et al. 1973) and calculated with the Dayhoff method (Dayhoff et al. 1978), including the bootstrap test with 1000 replicates (Felsenstein 1985).

2.4. Homology modeling protein

The 3D structural protein model was constructed through to the homology modeling method using the SWISS-MODEL web server (Waterhouse et al. 2018). This server is accessible and automatic tools to predict 3D protein structure based on the homolog amino acids composition. The PyMol software version 1.8 was used to evaluate the 3D protein structure (available at: https://pymol.org/2/) (Schrodinger 2010). I-Tasser and COFACTOR webserver was used to determine the secondary structures and pre-
dict the functional insights of the protein, including ligand binding sites and gene ontology (available at: https://zhanglab.ccmb.med.umich.edu/) (Roy et al. 2010, 2012).

3. Results and Discussion

3.1. PCR product and DNA sequencing

The amplicons were obtained with a correct band with a length of about 1,500 bp in 1% gel electrophoresis with a 1 kb DNA ladder (Figure 1). Based on that result, the expected target DNA with a length of about 1545 bp targeted by AM-PAS primers could be amplified during the PCR. However, the unspecific amplicon also presents in the gel with a length of about 750 bp. Therefore, before taking for DNA sequencing, the target band was purified first with gel extraction method included in DNA pre-treatment from 1st Base DNA sequencing service to eliminate unspecific amplicon.

The sequence assembly from pair-read sequencing has a length of 1,468 bp (Figure 2). The contig sequence does not carry the start codon because we have trimmed the ambiguities reads from DNA sequencing results with FinchTV software. This contig sequence was used for the following analysis.

3.2. α-amylase sequence alignment and the phylogenetic analysis

BlastX program with PDB database showed that the AM-PAS nucleotide sequence has the highest similarity and query cover up to 99% with accession domain PRK09441 described as α-amylase (Figure 3). This result could confirm that the AM-PAS nucleotide sequence was the α-amylase encoding gene.

After the conversion from nucleotide to protein sequence, the AM-PAS has a length of about 489 amino acids. The protein blast result showed that the AM-PAS amino acid sequence has a slight variation compared to α-amylases in B. amyloliquifaciens (PDB id. 3BH4). However, it was highly variable compared to α-amylase in other species in genera Bacilli (Table 1).

To show the AM-PAS’s amino acids variation with other α-amylase B.amyloliquefaciens, the graphic view of protein sequences from AM-PAS with 3BH4 is already in Figure 4. Based on Figure 4, it is showed that the AM-PAS protein sequence contains about half of the signaling peptide in the early region, which consists of amino acid LLFVSLPTKTSA, and several amino acids gaps in the end region. Overall, the amino acids composition of AM-PAS was quite identical to α-amylase 3BH4.

The feature of well-characterized α-amylase consists of signaling peptides in the amino acid position 1-31, and the description of α-amylase functional started from position 32-514 (Bateman 2019). The AM-PAS protein sequence was identical with amino acid in the functional sites with 3BH4 protein sequence, such as metal binding and active sites. In contrast, amino acid variations were present in the non-functional region of the protein. The black square in Figure 4 shows the feature of key for metal-binding that they are located at amino acids 190, 214, 225, 231 described as Ca\(^{2+}\) and Na\(^{+}\) cofactor metal-binding while in the other positions (133, 212, 233, 235, 266, 331, 438, 461) just specific for Ca\(^{2+}\) metal binding.

<table>
<thead>
<tr>
<th>Sequence information</th>
<th>Distance</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>3BH4 α-amylase (B. amyloliquifaciens)</td>
<td>0.019</td>
<td>98%</td>
</tr>
<tr>
<td>1E3X α-amylase (B. amyloliquifaciens)</td>
<td>0.079</td>
<td>92%</td>
</tr>
<tr>
<td>1VJS α-amylase (Bacillus licheniformis)</td>
<td>0.220</td>
<td>78%</td>
</tr>
<tr>
<td>1O80 α-amylase (B. licheniformis)</td>
<td>0.225</td>
<td>77%</td>
</tr>
<tr>
<td>1BLI α-amylase (B. licheniformis)</td>
<td>0.228</td>
<td>77%</td>
</tr>
<tr>
<td>1W9X α-amylase (Bacillus halmapalus)</td>
<td>0.365</td>
<td>63%</td>
</tr>
<tr>
<td>2GJP α-amylase (B. halmapalus)</td>
<td>0.373</td>
<td>62%</td>
</tr>
<tr>
<td>2DIE α-amylase Bacillus sp. Ksm-1378 (Bacillus sp.)</td>
<td>0.398</td>
<td>60%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acid positions</th>
<th>AM-PAS</th>
<th>3BH4</th>
</tr>
</thead>
<tbody>
<tr>
<td>78</td>
<td>T(Threonine)</td>
<td>L(Leucine)</td>
</tr>
<tr>
<td>83</td>
<td>I(Isoleucine)</td>
<td>N(Asparagine)</td>
</tr>
<tr>
<td>160</td>
<td>G(Glycine)</td>
<td>E(Glutamic acid)</td>
</tr>
<tr>
<td>346</td>
<td>E(Glutamic acid)</td>
<td>R(Arginine)</td>
</tr>
<tr>
<td>414</td>
<td>K(Lysine)</td>
<td>N(Asparagine)</td>
</tr>
</tbody>
</table>
FIGURE 2 Graphic view of AM-PAS nucleotide sequence contig from pair-reads sequencing.

FIGURE 3 BlastX result of AM-PAS nucleotide sequence. It shows that the AM-PAS nucleotide sequence has a specific hit with PRK09441 domain α-amylase.

FIGURE 4 The graphic view of AM-PAS protein sequence alignment compared with 3BH4 protein sequence.
Active sites were remarked with a red square, with features key description for nucleophile and proton donor located in amino acids 262 and 292, respectively. The transition state stabilizer located in amino acid 359, remarked with a blue square (Alikhajeh et al. 2010; Bateman 2019).

The phylogenetic tree with the UPGMA method is shown in Figure 5. The construction of this tree according to α-amylase protein sequences from B. amylo liquefaciens and other related Bacillus spp α-amylase protein sequences showed that AM-PAS sample located in one clade with 3BH4 and 1E3X (Red square in Figure 5). AM-PAS clade has a high confidence value with a bootstrap score of about 100. The phylogenetic tree confidence can be interpreted using bootstrap value, the high bootstrap score indicating that the tree can be trusted (Gregory 2008). Phylogenetic based on protein sequence can detect the functional protein and maybe inherited during evolution (Rao et al. 2014).

3.3. 3D protein model from AM-PAS sequence

The homology modelling method was recently used to identify key amino acids in various organisms, including bacteria; this technique is beneficial for comparing and reproducing complex protein structural based on amino acid sequences (Ali and Shafiq 2015; Pramanik et al. 2017; Waterhouse et al. 2018). Generally, 30% of amino acid sequence similarity is considered a threshold for homology modelling accuracy (Xiang 2006). The compatible and satisfied template was used for the α-amylase 3D model obtained from B. amylo liquefaciens with PDB accession id 3BH4 (Alikhajeh et al. 2010). This model has the highest similarity of the amino acid sequence about of 98.3% and coverage about >95%, which has local quality verification (Q-mean) about 0.8-1 (Figure 6a). The amino acid residues from the Ramachandran plot are mainly located in the favored region (Figure 6b). This Q-mean score and Ramachandran plot provide for scoring the model’s quality and estimation of the per-residue model quality with statistical calculation (Benkert et al. 2008, 2009). AM-PAS sample compared with the template model (3BH4) has a variation at amino acids position 79, 83, 160, 346, 414 (Listed in Table 2).

The tertiary protein model from the AM-PAS sequence contains the helix, sheet, and secondary coil structures (Figure 7). The genetic variation on the AM-PAS protein sequence formerly predicted with homology modelling impacts in the beta-sheet that has shorter compared with α-amylase 3BH4, specifically at position 347 (Figure 8d). The altered secondary protein structure by mutations in the codons can contribute to hydrogen bonds, disulfide bonds, and hydrophobic interaction that directly changes the secondary and tertiary structures of the proteins (Bunz 2008).

The α-amylase from B. amylo liquefaciens is classified in the family enzyme glycoside hydrolase (GH)13 (Janeček et al. 2015; Bateman 2019). Most α-amylases in this family have three domains for ligands binding. The A domain starts from amino acids number 3 to 103 and 207 to 396, forming 8- secondary structures (beta/alpha bar-

FIGURE 8 Structure 3D protein model from AM‐PAS (green) compared with 3BH4 protein (yellow). The pictures (a), (b), (c), (d), and (e) were zoomed in from protein complex pictures specified in amino acids 160, 78, 83, 346, 414, respectively. The 3D structure of AM‐PAS remarked with red, while 3BH4 remarked with cyan. The different structure appears only in picture d.

rel). The B domain starts from amino acids number 104 to 206 located between domain A specified at the third beta-strand and the third alpha-helix. The C domain starts from amino acids number 397 to 482, which have eight-stranded beta-barrel (Hwang et al. 1997). Domain A and domain B probably played a major role in the substrate-binding catalytic process and stability. The ligand still can be catalyzed without the C domain (Janeček et al. 2003; Janeček and Kuchtová 2012), but domain C is essential in recombinant enzyme properties (Montor-Antonio et al. 2017). Therefore, all three domains are still needed in protein engineering to produce stable protein.

Naturally, the member of polysaccharides cannot directly utilize by organisms. Amylolytic bacteria can convert the polysaccharides into monosaccharides which are essential for microorganisms. Monosaccharides like glucose as the product from hydrolysis reaction by amylase are beneficial for microalgae in the co-culture. Bacterial isolate PAS can produce amylase of about 0.01 units/mL (Nisa et al. 2021). Another report mentioned that B. amyloliquefaciens with optimizing culture condition could produce amylase up to 54.93 units/mL (Deb et al. 2013). Most species in genera Bacilli have also benefits as growth-promoting and protecting agents from pathogens (Srivastava et al. 2016; Lopes et al. 2018). B. amyloliquefaciens could be considered a beneficial bacterium for the microalgae culture because they had proved to stimulate microalgae growth and improve biomass (Kang et al. 2021; Nafi’ah et al. 2021). This study was the first report of the exploration of α-amylase-producing bacteria isolated from Ranu Pani, Indonesia.

4. Conclusions

The α-amylase encoding sequence was successfully detected in B. amyloliquefaciens PAS isolated from Ranu Pani, Indonesia. The AM-PAS nucleotide sequence was showing a genetic variation that impacts the variation of amino acids. However, based on the protein modelling method, the predicted 3D protein model has a similar structure with α-amylase in the database with PDB.id 3BH4, reaching more than 98% identity. Therefore, this research will be helpful for further research to produce recombinant α-amylase.

Acknowledgments

The authors are grateful to the Laboratory of Microbiology, Department Biology, Universitas Negeri Malang that provides the laboratory facilities for this study. This research and publication were funded by Penerimaan Negara Bukan Pajak (PNPB) Universitas Negeri Malang with contract No. 20.3.252/UN32.14.1/LT/2019.

Authors’ contributions

SP, AR designed the study. MMU designed primers in this study, AR, MMU, AFM, UR carried out the laboratory work. AR analyzed the data and wrote the manuscript. All authors read and approved the final version of the manuscript.
Competing interests

The author declare that they have no competing interest.

References


SRL HB. 2014. DNA Baser Sequence Assembler.

