The structure characteristic of IAA n-acetyl-transferase enzyme produced by two species of bacteria (Bacillus subtilis and Bacillus amyloliquefaciens)

R F W Putrie1,2, I N P Aryantha1, Iriawati1 and S Antonius3

1 Doctoral Programmes, School of Life Sciences and Technology, Bandung Institute of Technology. JL Ganesa 10, Bandung, 4013, West Java-Indonesia.
2Research Center for Biotechnology, Indonesian Institute of Sciences. JL. Raya Bogor KM. 46 Cibinong 16911, West Java-Indonesia.
3Research Center for Biology, Indonesian Institute of Sciences. JL. Raya Bogor KM. 46 Cibinong 16911, West Java-Indonesia.

Email: rahayufwputrie@gmail.com

Abstract. Exogenic indole-3-acetic acid (IAA) hormone has been known to be produced by plant-associated bacteria for regulating plant growth and development. The genus of Bacillus as the most common colonizer of the plant has the capability to produce this hormone. IAA n-acetyltransferase is an enzyme that plays role in the production of tryptophan-dependent IAA hormone on bacteria. Generally, enzymes as proteins have certain characteristics according to their function. The aim of this study was to compare and analyze the protein structure characteristics of IAA n-acetyltransferase enzyme produced by two species of Bacillus, such as B. subtilis and B. amyloliquefaciens. The analytical modeling based on NCBI database showed that protein structure characteristics produced by these species are similar to 3D protein models and the types of amino acids that build up the enzyme. However, the amount of α-helix, β-sheet and the number of amino acids that make up it remains different. In addition, another similarity was also found that the enzymes of the two species do not have transmembrane proteins. These results can contribute to theoretical knowledge related to the characteristics of structural proteins from enzymes involved in IAA hormone production.

1. Introduction
Indole 3 acetic acid (IAA) hormone is the main signal to regulate many aspects of vascular differentiation in plants. This hormone plays an important role in regulating the growth of plants such as cell elongation, vascular tissue development, and apical dominance [1]. In addition, IAA also acts as an indicator of drought-resistant hormones in plants [2]. This plant hormone is known to be produced by symbiotic bacteria to support the growth of the plant host. Production of exogenous IAA is one of the natural mechanisms secreted by plant growth-promoting bacteria [3].

Bacillus is a genus of bacteria that is closely associated with plants, both as rhizosphere bacteria and endophytic bacteria [4]. In previous studies, the data has shown that Bacilli become the dominant symbiont genus in pineapple plant which is able to grow in semi-arid ecosystem [5]. Their genus also has been known for the ability to produce IAA hormone in the absence and in the presence of L-tryptophan [6]. This is interesting, but the genus of Bacillus has many members, very large and diverse. This species is divided into three main groups based on the phenotype of sporangia, adult spores,
biochemical tests and growth properties [7]. However, two Bacillus species are known to have the ability to produce exogenous IAA, namely *B. subtilis* and *B. amyloliquefaciens* [8]. *B. subtilis* (Mt3b) have been known for producing exogenous IAA and this ability could be improved by creating optimal growth conditions [3]. *B. amyloliquefaciens* FZB 42 has also been shown to produce IAA dependent tryptophan [9]. In term of the ability to synthesize IAA, tryptophan has been identified as the main precursor for IAA biosynthesis pathways in bacteria [10,11].

IAA n acetyltransferase enzyme encoded by the ysnE gene is known to be involved in tryptophan-dependent IAA hormone production [9,10]. That is important enzyme in the process of forming the IAA hormone via indole-3-pyruvate (IPyA) pathway, precisely in the process of transferring acyl groups. IAA n-acetyltransferase function to catalysis of the transfer of an acetyl group to a nitrogen atom on the acceptor molecule. If no acyl groups are transferred, the IAA hormone cannot be formed. The position of its function is in the first stage of the dependent tryptophan pathway, namely at the conversion stage from tryptophan to indole-3-pyruvit acid [12].

Enzymes are proteins that have different characteristics according to their biological functions [13]. There are about 300 types of amino acids, but only 20 amino acids are involved and play a role in protein synthesis and biological function [14]. The enzyme-making proteins produced by species of the genus of Bacillus may have different characteristics from one another. Any indication that protein structure different of this enzyme effect the IAA formation. Therefore, in this study, we investigated of protein structure characteristic of IAA n-acetyltransferase enzyme producing by two species of Genus Bacillus, especially *B. subtilis* and *B. amyloliquefaciens*. The research was a preliminary study and contributed to theoretical knowledge of the protein biochemical analysis related to the characteristics of structural proteins from enzymes that involved in IAA hormone production, especially in Bacillus.

2. Materials and Methods
The microbial sequence data used in the study was obtained from NCBI sequence database and processed using bioinformatic tools. The EC number of targeted enzymes is EC:2.3.1.-. Our studies based on the research data obtained previously in Putrie *et al.* [5] combine with Mishra *et al.* [8] for the basic reference for selecting the species. The higher exogenous IAA producers of *Bacillus*, namely *B. subtilis* and *B. amyloliquefaciens*, were used in this study.

2.1. Protein structure analysis
Amino acid sequences of IAA n-acetyltransferase enzyme was derived from two species Bacillus at the NCBI site. *B. subtilis* with accession number NP_390711.1 and *B. amyloliquefaciens* with A2RH37. The structure protein analyzed were included the secondary and tertiary proteins structure of the enzyme. The secondary structure of proteins was predicted using online software https://open.predictprotein.org/ by entering each sequence database of each species. PredictProtein is a meta service for sequence analysis, including structural and functional protein features [15]. Next, the amino acid sequences that form *α*-helix and *β*-sheet (strands) was observed. Analysis of tertiary protein structure was carried out using online software provided on the website http://swissmodel.expasy.org/interactive [16]. In a model that has the same function with the highest level of similarity, it is further analyzed for the shape of the structure and similarities or differences of the two proteins.

2.2. Protein transmembrane analysis
The next step was to analyze the prediction of the structure of the transmembrane protein [15]. Prediction and visualization of each sequence were carried out by using the website http://www.enzim.hu/hmmtop/html/submit.html to determine the number and location of transmembrane alpha helices based on HMMTOP. Furthermore, the presence or absence of transmembrane helices, their number and location of sequences were observed. Visualization of each sequence was done by using the website http://tdmas.bioinfo.se/DAS/index.html to determine the number and location of the transmembrane alpha helix based on the DAS curve. The area of the watershed curve that exceeds the strict cut-off threshold and transmembrane area were analyzed.
2.3. Hydrophobicity and solubility protein analysis
The analysis referred to Hamdani et al. [16]. Each amino acid sequence was analyzed for hydrophobicity by Kyte-Dollyte Hydropathy Plot at http://fasta.bioch.virginia.edu/fasta_www2/fasta_www.cgi?rm=misc1. Furthermore, Protein-SOL solubility was predicted on the website: https://proteinsol.manchester.ac.uk/. The solubility of each of these proteins was observed, including predictions of hydrophobicity and protein solubility.

2.4. Isoelectric point (pI) dan molecular weight (MW) analysis
Analysis of the molecular weight and isoelectric point (Pi) of each protein sequence was estimated by Compute pI/Mw by using tools from http://web.expasy.org/compute_pi/ [16]. The isoelectric point for each protein sequence was observed. Each protein sequence then was entered to the http://protcalc.sourceforge.net/ page, in the "Charge" column, checking "Isoelectric point" and " List charges of pH range" contents by default. In the "Charge at pH" section, enter each PI obtained from the pI/MW analysis results. Protein charge at pH = pI, pH> pI, and pH <I. The pH suitable for the PAGE process of the two proteins above was analyzed. Isoelectric point prediction and protein molecular weight were analyzed to determine the relationship of pI and protein charge, as well as the analysis function for SDS-PAGE and protein PAGE.

3. Results and Discussion
The results obtained include secondary and tertiary protein structure analysis. Secondary structure analysis to determine differences in location and number of α-helix and β-sheet (strands). Table 1 showed that the α-helix number of the two species was similar and the number of β-sheets (strands) was higher in B. amyloliquefaciens. The location and amount of α-helix and β-sheet (strands) of amino acids from the enzyme affected to its tertiary shape. Transitions usually occur in protein filaments with long alpha helical domains that can affect the stiffness, strength, and energy dissipation capacity of large deformations [17]. However, the α-helix fold resulted no significant difference in shape, but the β-sheet produced a more diverse shape. In addition, the folding shape such as the twist of the β-strand or the relationship between two β-sheets was influenced by the composition of the amino acids that build it [18].

Table 1. Differences location of α-helix and β-sheet (strand) structures of IAA n-acetyltransferase

| Structure               | B. subtilis          | B. amyloliquefaciens |
|-------------------------|----------------------|----------------------|
| α-helix (amino acid sequence number) |                      |                      |
| 10-25 (16 aa)           | 31-47 (17 aa)        |
| 30-36 (7 aa)            | 52-55 (4 aa)         |
| 38-41 (4 aa)            | 59-63 (5 aa)         |
| 80-82 (3 aa)            | 101-105 (5 aa)       |
| 87-101 (15 aa)          | 108-124 (17 aa)      |
| 117-126 (10 aa)         | 139-147 (9 aa)       |
| β-sheet (strand)        |                      |                      |
| (amino acid sequence number) |                    |                      |
| 2-5 (4 aa)              | 6 (1 aa)             |
| 46-53 (8 aa)            | 9-13 (5 aa)          |
| 56-65 (10 aa)           | 19-26 (8 aa)         |
| 70-78 (9 aa)            | 67-74 (8 aa)         |
| 106-111 (6 aa)          | 78-86 (9 aa)         |
| 130-132 (3 aa)          | 91-97 (7 aa)         |
| 144-150 (7 aa)          | 128-133 (6 aa)       |
|                         | 151-152 (2 aa)       |
|                         | 165-171 (21 aa)      |
The amino acid types of both species indicated that they have the same type of amino acid, but with different quantities as shown in Figure 1. In *B. subtilis* the highest quantity of amino acids is E whereas in *B. amyloliquefaciens* is L. The composition of amino acids from an enzyme affected to its activity specifically. The most reactive amino acids tend to be nucleophilic that contain sulfur or aromatic side chains [19].

![Amino acid composition of IAA n-acetyltransferase enzyme from (a) B. subtilis and (b) B. amyloliquefaciens](image1)

**Figure 1.** Amino acid composition of IAA n-acetyltransferase enzyme from (a) *B. subtilis* and (b) *B. amyloliquefaciens*

Analysis of tertiary protein structure is useful for finding models that have the same function and the highest similarity by using a Swiss model web server. The results showed that the 3D modelling structure of the enzyme from two species were similar, but there were differences at the three observation points as shown in Figure 2. Homology model construction is useful for identifying structural templates, alignment of target sequences and template structures, modeling and evaluating model quality. In the enzyme protein from species *B. subtilis* showed 100% identity sequence homology with *ysnE* protein, while in the enzyme protein from species *B. amyloliquefaciens* by 63.01%. In both of coverage results there were not receptor regions for ADP ligands because these areas were conserve. Enzyme activity is limited by its tertiary structure. This is due to the fact that many amino acids with high activity potential are buried in the nucleus of proteins, where they cannot be accessed by pro-oxidants [19]. For that reasons tertiary structure analysis is important.

![3D modelling protein structure of IAA n-acetyltransferase enzyme from (a) B. subtilis and (b) B. amyloliquefaciens](image2)

**Figure 2.** 3D modelling protein structure of IAA n-acetyltransferase enzyme from (a) *B. subtilis* and (b) *B. amyloliquefaciens*
The analysis results of the transmembrane protein structure prediction of IAA n-acetyltransferase from both species indicated that they did not have a transmembrane protein as shown in Figure 3. Threshold strict cut off can be interpreted as a limit point. There are 2 types of cut off, namely "strict cut off" and "loose" cut off. In general, it can be concluded that if there is a watershed curve area that exceeds the threshold strict cut-off is an area that belongs to the transmembrane while below the cut-off is not a transmembrane protein.

Figure 3. DAS curve of transmembrane protein of IAA n-acetyltransferase enzyme from (a) B. subtilis and (b) B. amyloliquefaciens

Characteristics of transmembrane structures is the discovery of the presence of transmembrane domains (transmembrane helices) in proteins that are domains embedded in cell membranes so that this region has a helical structure that contains many hydrophobic amino acid residues. In previous analysis, the enzyme from both species showed that they did not have a transmembrane protein, which assumed that this enzyme was hydrophilic. Futhermore, the results of the hydrophobicity analysis of the two species also indicated that their enzymes were hydrophilic as shown in Figure 4. IAA n-acetyltransferase enzyme from B. subtilis and B. amyloliquefaciens had almost the same hydrophobicity tendencies.

Figure 4. Hydrofobicity characteristic of IAA n-acetyltransferase enzyme from (a) B. subtilis and (b) B. amyloliquefaciens

The result was confirmed again by solubility analysis. Protein Sol is a web server for calculating or predicting solubility of protein. Solubility is one of the independent characteristics of proteins that can be determined by the amino acids that build up proteins under special experimental conditions [20]. The enzyme of B. subtilis species has a solubility value (QuerySol) of 0.581 and B. amyloliquefaciens 0.323
as shown in Figure 5. The average population value for the experimental dataset (PopAvrSol) is 0.45, and hence if QuerySol is greater than 0.45 is estimated to have higher solubility and vice versa if lower scale solubility values are predicted to be less soluble [21]. Based on Figures 4 and 5, the trend match results were obtained. Proteins of both Bacillus species tend to be hydrophilic and water soluble due to lack of transmembrane protein.

![Solubility](image)

**Figure 5.** Solubility characteristics of IAA n-acetyltransferase enzyme from (a) B. subtilis and (b) B. amyloliquefaciens

Moreover, enzyme surfaces have a net charge that depends on the amount and identity of the charged amino acids. At certain pH, positive and negative charges on the amino acid balance. In the balance condition, the net charge of the molecule is zero and electrically neutral, thus the condition is called the isoelectric point (pI). The table of enzyme characteristics included the value of pI, molecular weight (MW) and estimated pH as shown in Table 2. Based on this table, we can see that the range of protein isoelectric points of the enzyme *B. amyloliquifaciens* was greater than *B. subtilis*. The value of pI and protein load are related, protein load at pH = pI = 0, protein load at pH > pI = negative, and protein load at pH < pI = positive. The pI /MW ratio can be used to estimate the presence of protein/enzyme in 2D gel. In addition, the application of the analysis function can be used to determine the most optimal pH value for enzyme analysis with SDS-PAGE and PAGE protein. The pH adjustment affects the net charge protein needed for SDS-PAGE analysis, to do SDS-PAGE the protein needs to be negatively charged.

**Table 2.** Characteristics of IAA n-acetyltransferase enzyme from *B. subtilis* and *B. amyloliquefaciens*: isoelectric point (pI), molecular weight (MW) and estimated pH.

| Character Type          | B. subtilis | **Species** |
|-------------------------|-------------|-------------|
| Isoelectric point (pI)  | 5.80        | 8.79        |
| Molecular weight (MW)   | 17052.42 gram/mol | 19855.69 gram/mol |
| Estimated pH            | 6.20        | 8.59        |
| Protein charges :       |             |             |
| pH = pI                 | 0           | 0           |
| pH > pI                 | -1.4        | -1.5        |
| pH < pI                 | 0.9         | 0.4         |
| Isoelectric range       | 5.80-6.20   | 8.79-8.59   |
Activity of the enzyme is influenced by enzyme-making proteins. This research was considered to conduct a preliminary study of the protein biochemical analysis via bioinformatics tools. Protein structure characterization in this study was not included binding site. The further study needs to be carried out regarding the research of the significant impacts and the dissimilarities in IAA production more deeply.

4. Conclusion
The structure characteristics of IAA n-acetyltransferase enzyme from B. subtilis and B. amyloliquefaciens showed that there was a similarity on the type of amino acids that build up the enzyme. In addition, the two proteins that build up these enzymes were not included in the transmembrane protein group and tend to be hydrophilic. Three point at 3D modelling, the quantity of amino acid as well as the quantity of α-helix and β-sheet (strand) and their location were a slight difference. Isoelectric range and molecular weight of the enzyme from B. amyloliquifasciens were greater than B. subtilis.

Acknowledgement
The authors would like to thank Ministry of Research and Technology of the Republic of Indonesia for funding research through the National Competitive Research Programme.

References
[1] Fahad S, Hussain S, Bano A, Saud S, Hassan S, Shan D, Khan F A, Khan F, Chen Y, Wu C, Tabassum M A, Chun M X, Afzal M, Jan A, Jan M T and Huang J 2015 Environ. Sci. Pollut. Res. 22 4907–4921
[2] Pavlovic I, Petrík I, Tarkowska D, Lepedus H, Bok V V, Brkanac S R, Novak O and Salopek-Sondi B 2018 Int. J. Mol. Sci. 19 2866
[3] Wagi S and Ahmed A 2019 PeerJ 7 7258
[4] Wiyada M 2012 J. Microbiol. Biotechnol. 22 1597–1604
[5] Putrie R F W, Aryantha I N P, Iriawati and Antonius S 2020 Biodiversitas 21 3084-3093
[6] Junior A F C, Oliveira A G, Oliveira L A, Santos G R, Chagas L F B, Lopes da Silva A L and Costa J L. 2015 Bulg. J Agric Sci. 21 282–287
[7] Zeigler D R and Perkins J B 2009 The Genus Bacillus: practical handbook of microbiology second edition CRC Press Taylor & Francis Group
[8] Mishra V K and Kumar A 2015 Afr. J. Microbiol. Res. 9 1139-1149
[9] Idris E E, Iglesias D J, Talon M and Borris R 2011 MPMI 20 619–626
[10] Shao J, Li S, Zhang N, Cui X, Zhou X, Zhang G, Shen Q and Zhang R 2015 Microb Cell Fact. 14 130 1-13
[11] Li M, Guo R, Yu F, Chen X, Zhao H, Li H and Wu J 2018 Int. J. Mol. Sci. 19 443
[12] Danapriatna N 2014 Jurnal Ilmiah Solusi 1 82-88
[13] Deeni Y and Sojimade N 2015 Current App. Biotech 14 225-237
[14] Akram M, Asif H M, Uzair M, Akhtar N, Madni A, Shah S M A, Hassan Z and Ullah A 2011 e J. Med. Plant. Res. 5 3997-4000
[15] Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Honigschmid P, Schafferhans A, Roos M, Bernhofer M, Michael, Richter L, Ashkenazy H, Punt A, Schlessinger A, Bromberg Y, Schneider R, Vriend G, Sander C, Ben-Tal N and Rost B 2014 Nucleic Acids Res 42 W337–W343
[16] Hamdani A D, Madihah, Suryohastari R B, Virgianti D P and Putrie R F W 2018 Nusantara Biosci. 10 203-209
[17] Qin Z and Buehler M J 2010 PRL 104 198304
[18] Fujiwara K, Toda H and Ikeguchi M 2012 BMC Structural Biology 12 1-15
[19] Sbroggio M F, Montilha M S, Garcia de Figueiredo V R, Georgetti S R and Kurozawa L E 2016 Food Sci. Technol. 36 375-381
[20] Rasool N, Hussain W and Mahmood S 2017 J Proteomics Bioinform 10 324-328
[21] Niwa T, Ying B W, Saito K, Jin W, Takada S, Ueda T and Taguchi H 2009 *PNAS* **106** 4201–4206