BIOINFORMATICS STUDY OF PHY GENE ENCODING PHYTASE FROM VARIOUS OF MICROORGANIMS

N. Mayasari[1*], L. Triratna[2], A. Azhar[1], H. Maulana[1], Dyin[1], Y. Widyastuti[2], N. Herlina[2], Abun[1], R. Ridwan[2], Rohmatussolihat[2], A.S. Hasbi Al Islahi[3]

[1]Faculty of Animal Husbandry, Universitas Padjadjaran,
Jl. Ir. Soekarto KM. 21 Jatinangor, Sumedang Jawa Barat Indonesia
[2]Research Center for Applied Zoology, BRIN. KST Soeakrno Cibinong
Jl. Raya Bogor KM 46 Cibinong Bogor Jawa Barat Indonesia
[3]PT. Berdikari Persero. Jl. Medan Merdeka Barat No 1 Jakarta Indonesia

Abstract

Phytase is an important enzyme for livestock especially for monogastric because this enzyme is able to hydrolyze phytic acid. The addition of phytase in ration has been known ncreased digestibility of phosphate, reduced soil pollution, and increased feed digestibility. In Indonesia, the avaibility of phytase is limited. The sources and technology to produce phytases is needed. Thus, various studies including bioinformatic study is important to fullfill the requirement of phytase in Indonesia. This study aimed to study phy gene encoding phytase from different microorganisme.. Data regarding a phy gene sequences encoding phytase from different microrganisms, amino acid sequences of phytase, 3D structure with active site and binding site of phytase, the phylogenetic tree of phy gene and phytase were evaluated. The results showed that coding sequence (CDS) size of the phy gene varies from 1104 to 1959 base pairs with the amino acid sequence size from 368 to 652 AA. The active site of phytase from Aspergillus niger are the amino acid Histidine (His59) and Aspartic acid (Asp339) where this phytase is dominated by the alpha helix structure. While, phytase from Bacillus amyloliquefaciens has 6 calcium binding sites, namely GLU15, PRO29, VAL73, ASP280, ILE 312, dan ASP313, where beta sheet and loop are very dominant in compiling this enzyme. Further study on bioinformatics, gene expressuon, enzyme activities and biological evaluation is needed to support the development of phytase production.

Key words: *bacteria, bioinformatic, feed additive, fungi and gene phy*

INTRODUCTION

Phytase is an enzyme that capable of hydrolyzing phytate (myo-inositol hexaphosphate) into myo-inositol and inorganic phosphate [1]. Phytat is a phosphate compound that is dominantly stored in grains, especially cereals and legumes [2]. Phytate in cereals and legumes is one of the main compositions in monogastric animal feed as a source of phosphate [3]. However, these animal are not able to digest or utilize phosphate in the form of phytic acid [4,5]. This is due to the low activity of phytase in the digestive tract. Thus, it is necessary to add additional sources of phosphate in inorganic phophate form to monogastric animal feed which becomes ineffective and uneconomical. Phytate is showing anti-nutritional effects whrereas they have an complex interaction with mineral such as iron, zinc, calcium and also with proteins [6,7]. In addition, the effect of undigested phytic acid causes environmental pollution, because phytic acid also cannot be decomposed in the soil [8].

Phytase is produced by various microorganisms, both prokaryot and eukaryotic, such as bacteria, yeast and fungi [5]. Some of the phytase-producing fungi are *Aspergillus, Myceliophthora, Mucor, Penicillium, Rhizopus* and *Trichoderma* [9,10,11,12]. While, some phytase-

[*]Corresponding author: novi.mayasari@unpad.ac.id
The manuscript was received: 11.09.2023
Accepted for publication: 15.02.2024

Article licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (http://creativecommons.org/licenses/by-nc-sa/4.0/)

producing bacteria are *Lactobacillus, Bacillus*, and *Escherichia coli.*

Phytase is encoded by certain genes present in the genome of phytase-producing microorganisms. The gene encoding phytase is called the phy gene. PhyA and phyB ara usually genes from *Aspergillus*, phyC gene derived from *Bacillus*, while AppA gene is derived from *E.coli* [13]. Based on gene data bank (The National Center for Biotechnology Information or NCBI), the size of the phy gene varies depending on the phytase source microbe. Like wise, the size of the length of the amino acid that compose phytase are varies [14].

According to Chelius and Wodzinski, phytase of the genus *Aspergillus* is the most widely used in industry and has high activity [15, 16]. Meanwhile, based on the results of several studies related to phytase from genus *Bacillus*, on average it has thermostable properties [17, 18, 19]. For that reason, in this bioinformatics study, the focus will be on phytase from *Bacillus* and *Aspergillus*.

MATERIAL AND METHOD

Phy Gene Analysis

The phy gene sequences of the gene source microorganisms were collected from the gene data bak (NCBI). The source of microorganism for the gene ara fungi of the genus Aspergillus and bacteria of the genus Bacillus. The collected sequences are whole gene or coding sequences (CDS). Partial sequences are not selected. The results of the sequence collection were then aligned usingbioedit v.7 and analyzed the phylogenetic tree at the gene level.

Protein Analysis

The amino acid sequences that make up phytase from the same microorganisme as the gene source were collected in notepad file. The sequences were aligned using bioedit v.7 and their phylogenetic tree was analyzed at the amino-acid level.

Bioinformatics analysis

Bioinformatics study used data from Gene bank and Protein Data Bank (PDB). The data used from PDB is the 3D phytase structures of *Aspergillus niger* (ID: 3K4P), *Aspergillus fumigatus* (ID: 1QWO), *Klebsiella sp*. ASR1 (ID: 2WNI), *Selonomonas ruminantium* (ID: 1U24), *Bacillus Sp*. (ID: 1POO), *Bacillus amyloliquefaciens* (ID: 2POO) and *E. Coli* (ID: 7Z1J). The analysis uses a combination of bioedit v.7 and spdbv v.7.

RESULTS AND DISCUSSIONS

Gene *phy* Analysis

There are 38273 data of phytase nucleotide in the NCBI. From all these data, there are those that show only in the form of partial sequences, whole gene or coding sequence (CDS) and some have known the whole genome of the microbial gen source. Microbial gene sources of genes in the data bank are very varied, ranging from plants, fungi and bacteria.

There are 71 CDS and whole gene phy from Bacillus. This CDS means that the phy gene sequence can encode or express phytase in its entirely from the start codon to before stop codon. CDS differs from partial sequences in that when the protein or enzymes are expressed it is most likely inactive. The size of the phy gene sequences of Bacillus varies from 1146 to 1724 base pairs (bp). The species of this *Bacillus* is varies, namely *B. Subtilis, B. Licheniformis, Bacillus sp, B. Cereus, B. Glycinifermentas, B. Amyloliquefaciens, B. Amilosiamensis,* and *B. Arthopaeus*, (Table 1).

There are 113 coding sequences (CDS) of the genus Aspergillus. The size of the phy gene from Aspergillus tend to be more varied and longer than those Bacillus. The sizes of phy gene from Aspergillus are 858, 1392, 1395 (2), 1398 (3), 1401 (7), 1406 (6), 1440 (3), 1455 (4), 1482, 1488, 1491 (5), 1494(4), 1506 (4), 1512, 1515 (8), 1525, 1528, 1551, 1553, 1554 (8), 1569 (5), 1570, 1584 (6), 1590 (2), 1596 (8), 1611, 1638, 1770, 1812, 1860, 1861, 1887, 1934, 1959 (5), 1962, 1971, 2071, 2116, 2319, 2327, 2379, 2232, and 2665 amino acid. Species of Aspergillus as a a source of phy genes are *A.flavus, A.niger, A.fischeri, A.fumigatus, A.japonicus, A.orizae, A.awamori, A.terreus, A.crisstatus, A.sclerotialis, A.eucalypticola, dan A.tubingensis*. The phy gene sequence of *Aspergillus* tends to be longer due to the

Article licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (http://creativecommons.org/licenses/by-nc-sa/4.0/)

addition of the signal peptide encodSing sequence in the up stream part of the gene. This signal peptide will carry phytase out of the cell to become active. So that phytase from *Aspergillus* is an enzyme that is secreted out of cells.

Table 1 Microorganisms has CDS of phy gene

No	Microorganisms	Total	Gene Size (bp)
1	*Bacillus sp.*	2	1200 dan 1724 base pare
2	*B. subtilis*	18	1152, 1209, 1210, 1247, 1363
3	*B. licheniformis*	18	1473
4	*B. cereus*	4	1544
5	*B. amyloliquefaciens*	11	1146
6	*B. Glycinifermentans*	2	1152
7	*B. arthopaeus*	5	1158
8	*B. inaquosorum*	2	1149
9	*B. halotoleran*	5	1149
10	*B. cabrialesii*	2	1149

Phylogenetik Analisis

Based on the phylogenetic tree analysis, it can be seen that at the nucleotide sequence level of the phy gene from *Bacillus*, there are 2 major groups. What's interesting is that even though they are from the same species, they can be in different groups, which means the level of evolution is different. Meanwhile, the phylogenetic tree which is composed of amino amino sequences of phytase from *Bacillus* are consist of 1 small group and 1 large group. Where this small group only consists of four *B. cereus*.

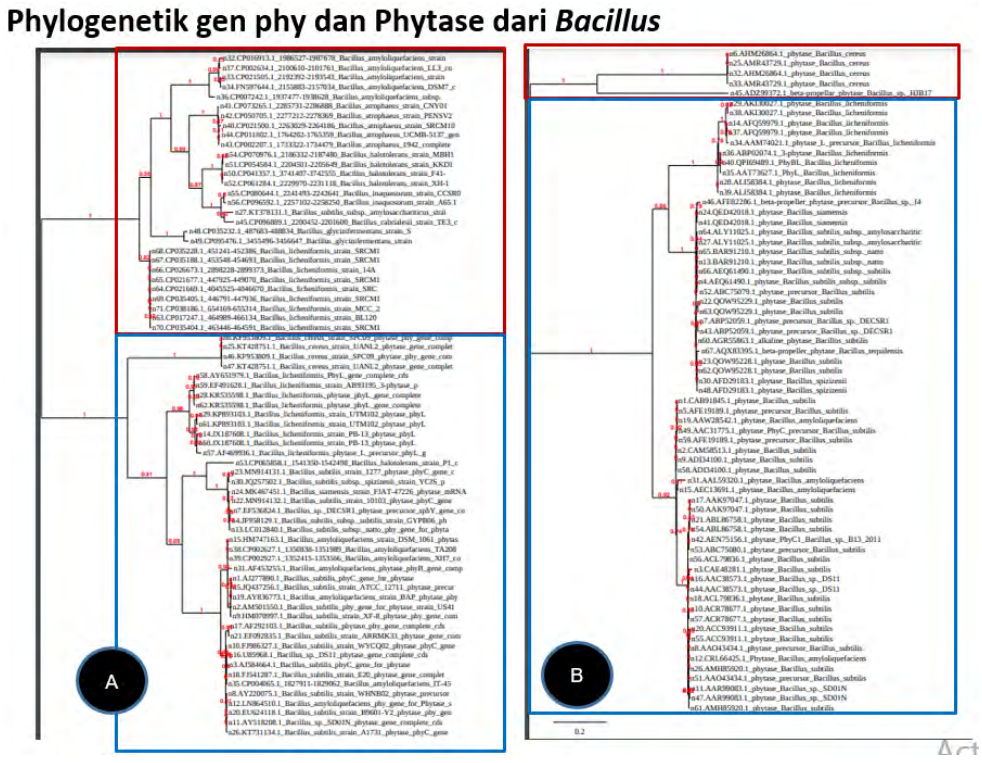

Fig. 1 Phylogenetic tree from *Bacillus*. (A) *phy* gene. (B) Phytase

Article licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (http://creativecommons.org/licenses/by-nc-sa/4.0/)

Fig. 2 Phylogenetic tree from *Aspergillus*

A molecular phylogenetic tree composed of the nucleotide sequences of the phy gene from Aspergillus. There are 2 medium groups and 1 large group. It can be seen that even though they are from the same species, their evolutionary level can be different, at the nucleotide level.

Bioinformatics analysis

The seven phytases with known 3-dimensional structure from PDB used as bencmark for data processing can be seen in table 2. Each 3 –dimensional structure can be seen in figure 1.

Table 2 Phytase from Protein Data Bank

No	Microorganims	ID	Size	E.C number	Active site	Binding site
1	*A.niger*	3K4P	444 AA	3.1.3.8	His59 and Asp339	-
2	*A.fumigatus*	1QWO	442 AA	3.1.3.8	NEP58 (loop) &ASP377 (brthasheet)	-
3	*Klebsiella sp ASR1*	2WNI	418AA	3.1.3.8	His358(loop)	-
4	*Selenomonas ruminantium*	1U24	337AA	3.1.3.72		-
5	*Bacillus sp.*	1POO	355	3.1.3.3.1.3.8		GLU15, PRO29, VAL73, ASP280, ILE 312, ASP313.
6	*Bacillus amuloliquefaciens*	2POO	355	3.1.3.8		
7	*E. coli*	7Z1J	412 AA	3.1.3.26		

Article licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (http://creativecommons.org/licenses/by-nc-sa/4.0/)

From table 2 it can be seen that different microorganisms as source of phytase can produce the different size lengths of amino acid (AA), different E.C number and different reaction when hydrolyzing phytic acid.

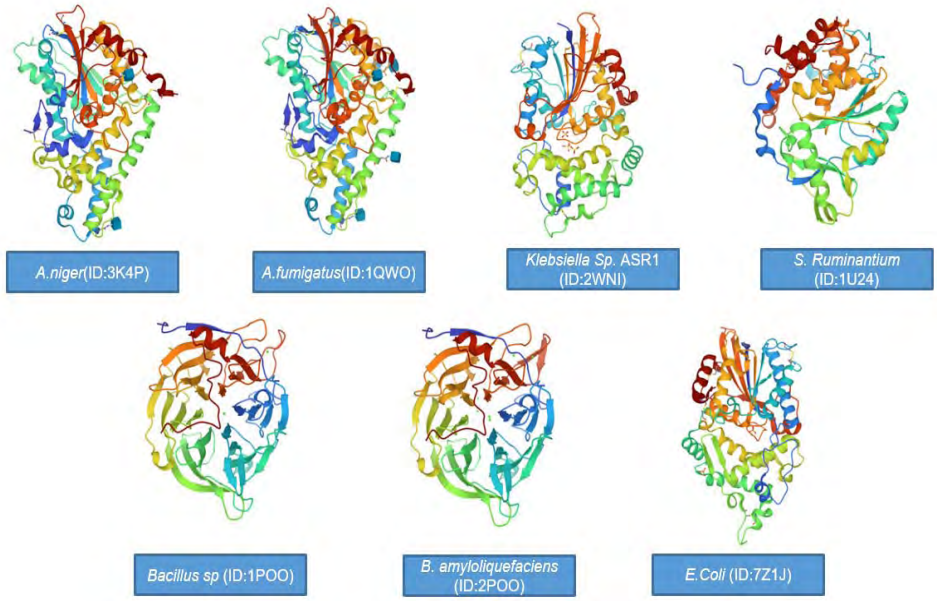

Fig. 3 3D Structure of Phytase from different microorganism

The phytase structure from *A.niger*, *A.fumigatus*, *Klebsia sp ASR1*, *S.ruminantium* and *E. Coli* was dominated by the alpha helix. Mean while, phytase from *Bacillus sp* and *B. amyloliquefaciens* was dominated by beta sheet structure. In *A.niger*, it is known that 2 amino acids play role in the active site, namely His59 and Asp339.

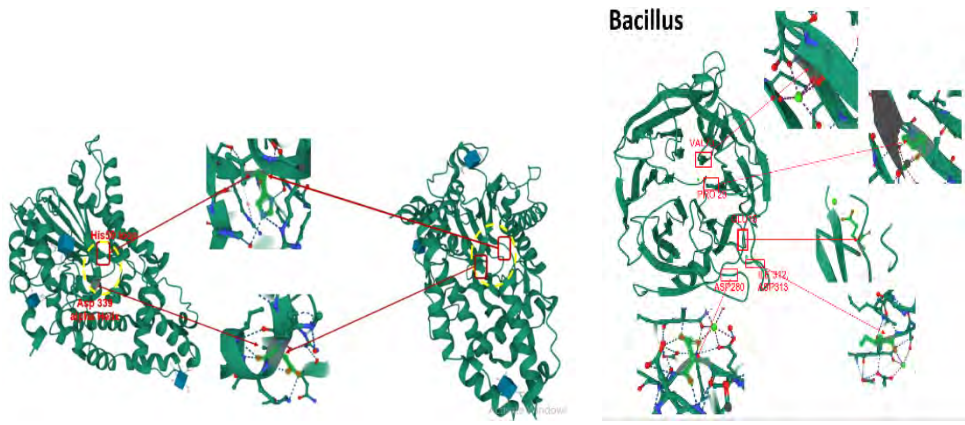

Fig. 4 Active site of Phytase from *A. niger* (A) and Binding site of Phytase from *Bacillus sp* (B)

Article licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (http://creativecommons.org/licenses/by-nc-sa/4.0/)

In *Aspergillus* only 2 amino acids are involved in the active site. While in *Bacillus* therea 6 amino acids involved in the binding site, namely GLU15, PRO29, VAL73, ASP280, ILE 312, ASP313.

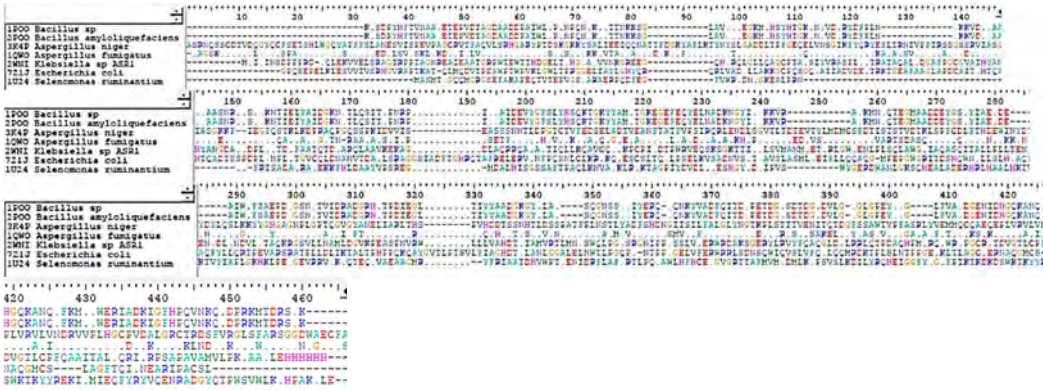

Fig. 5 Alignment of 7 phytases from different microorganisms

The alignment results showed that the level of similarity between the 7 phytases from different microorganisms wasvery low. However, the phytase of *Bacillus sp*. and *Bacillus amylolyquefacient* had 100% identical homology. Meanwhile, the level of similarity of *A.niger* with *A.fumigatus* was 66%. In addition, the level of homology berween one another is not more than 11%. *Bacillus* with *Aspergillus* is 11%. *Bacillus with Klebsiella* is 6%. *Aspergillus niger* with *Klebsiella* is 9%. *Klebsiella* with *S. Ruminatium* is 8%.

CONCLUSIONS

There are 38273 nucleotide data that make up the phytase gene. These nucleotides are sourced from various organisms. There are 71 complete genes in the form of CDS sourced from *Bacillus* and 113 genes from *Aspergillus*. Bioinformatics studi from *Aspergillus* showed very diverse sequenses of gene and protein constituents. This diversity is due to the difference species of microbial gene sources. However, when analyzing the kinship through phylogenetic tree, it can be seen that even though they come from the same species, at the nucleotide level the phy genes can be in different group. Then, at the level of amino acids will produce different levels of kinship. The active site of phytase from *Aspergillus* are His59 and Asp339. While the binding site of phytase from *Bacillus* are GLU15, PRO29, VAL73, ASP80, ILE312, ASP13.

ACKNOWLEDGMENTS

The authors are greatful to Badan Riset dan Inovasi Nasional (BRIN) for providing financial support to carry out this research.

REFERENCES

1. G. Estepa, R.M. Guerra-Hernández & B.G. Villanova,"Phytic acid content in milled cereal products and breads", *Food Research International*, **1999**, *32*, 217–221.
2. G.M. Lolas and P. Markakis,"Phytase of navy beans", *J. Food Sci*, **1997**, *42*, 1094-1097,
3. P.H. Selle and V. Ravindran,"Phytate-degrading enzymes in pig nutrition", *Livest.sci*, **2008**, *113*, 99-122,.
4. H. Brinch-Pedersen, Hatzak et al, "Heat Stable phytases in transgenic wheat (Triticum aestivum L):deposition pattern, thermostability and phytate hydrolysis", *J.Agric.Food Chem*, **2006**, *54*, 4624-4632,
5. A. Vohra and T. Satyanarayana,"Phytase: microbial sources, production, purification, and potential biotechnological application", *Crit.rev biotechnol*. **2003**, *23*, 29-60.

Article licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (http://creativecommons.org/licenses/by-nc-sa/4.0/)

6. C.B. sharma, M. Goel and M. Irshad, "Myo-inositol hexaphosphate as a potential inhibitor of α-amylase of different origin", *Phytochemistry*, **1987**, *17*, 201-204.
7. E. Graf (Ed), "Phytic acid, chemistry and Applications", *Pillsbury, Pillatus Press, Minneapolis, MN*, **1986**.
8. B.L. Lim, P. Yeung, C. Cheng, and J.E. Hill, "Distribution and diversity of phytate-mineralizing bacteria", *The ISME Journal*, **2007,** *1*, 4, 321-330.
9. Y.H. Tseng, T.J. Fang, and S.M. Tseng, "Isolation andcharacterization of a novel phytase from Penicillium simplicissimum", *Folia Microbiol,* **2000,** *45*, 121-127.
10. A. Sabu, S. Sarita, A. Pandey et al., "Solid – state fermentation for production of phytase by Rhizopus oligosporus", *Appl. Biochem. Biotech*, **2002,** *102*, 251-260,.
11. K. Roopesh, S. Ramachandran, K Nampoothiri et al., "Comparison of phytase production on wheat bran and oilcakes in solid-state fermentation, Recent advancement in white biotechnology through fungal", *fungal biology eds A*, **2006,** 65-99,
12. D. Dailin, S.Z. Elsayedlsayed, E. Sukmawati et al., "Fungal phytases: biotechnological applications in food and feed industries, in recent advancement in white biotechnology through fungi", *fungal biology, eds A,* **2019,** 65-99.
13. Li.K. Zou, H.N. Wang, X. Pan et al., "Expression, purification and characterization of a phyA^{m}-phyCs fusion phytase", *Journal of Zhejiang University SCIENCE B,* **2008***, 9*, 7, 536-545.
14. https://www.ncbi.nlm.nih.gov/
15. A. Casey anf G. Walsh, "Identification and characterization of a phytasae of potential commercial interest", *Journal of Biotechnology*, **2004**, *110*, 3, 313-322.
16. Chelius, M.K. and Wodzinski, R.J, "Strain improvementof Asepergillus niger for phytase production", *Appl. Microbial Biotechnolgy*, **1994,** *41*, 79-83.
17. Y.O. Kim, J.K Lee, H.K Kim, et al., "Cloning of the thermostable phytase gene (phy) from Bacillus sp. DS11 and its overexpression in Escherichia coli", *FEMS Microbiology Letters*, **1998**, *162*, 185-192.
18. K.R. Salimi, M. Hashemi, M. Safari, and M. Mousivand, "A novel phytase characterized by thermostability and high pH tolerance from rice phyllosphere isolated Bacullus subtilis B.S.46", *Journal of Advanced Research*, **2016**, *7*, 381-390.
19. Z. Zhang, J. Yang, P. Xie et al., "Characterization of a thermostable phytase from Bacillus licheniformis WHU and further stabilization of teh enzyme through disulfide bond engineering", *Enzyme and Microbiology Technology*, **2020,** *142*.

Article licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (http://creativecommons.org/licenses/by-nc-sa/4.0/)