

Characterization of type I L-asparaginase encoding gene from the thermohalophilic bacterium *Bacillus subtilis* CAT3.4 from Wawolesea Hot Spring, Southeast Sulawesi, Indonesia

MUZUNI^{1,*}, MUHAMAD ILHAM², ARDIANSYAH², SURIANA², YUTRIANA PUTRI¹, KARTINA¹

¹Department of Biotechnology, Faculty of Mathematics and Natural Sciences, Universitas Halu Oleo. Kampus Bumi Tridharma, Anduonohu, Kendari 93232, Southeast Sulawesi, Indonesia. Tel.: +62-401-3190403, Fax.: +62-401-3190006, *email: muzuni_fmipa@uho.ac.id

²Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Halu Oleo. Kampus Bumi Tridharma, Anduonohu, Kendari 93232, Southeast Sulawesi, Indonesia

Manuscript received: 30 May 2023. Revision accepted: 30 September 2023.

Abstract. Muzuni, Ilham M, Ardiansyah, Suriana, Putri Y, Kartina. 2023. Characterization of type I L-asparaginase encoding gene from the thermohalophilic bacterium *Bacillus subtilis* CAT3.4 from Wawolesea Hot Spring, Southeast Sulawesi, Indonesia. *Biodiversitas* 24: 5167-5178. This study aims to determine the molecular characteristics of the gene encoding type I L-asparaginase from CAT3.4 thermohalophilic bacterial isolates from the hot springs of Wawolesea, North Konawe, Southeast Sulawesi. This enzyme can be used as a cancer therapy agent and prevents the formation of acrylamide in food products. It is a kind of exploratory research. The characterization was done by amplifying the *ansA* gene sequence encoding L-asparaginase from CAT3.4 isolate by polymerase chain reactions (PCR) technique using primers AsnBac1-F1 (5'-ACGCGATATTTCTTTTGGCCGG-3') and AsnBac1-R1 (5'-CAGTGAAGAGGTGCATGGTATG-3'). The amplified PCR product was used as a template for sequencing by the Sanger method. The amino acid coding regions (CDS) obtained were bioinformatically characterized using the NCBI website for BLASTn analysis, restriction site and hydrophobicity profiles were analyzed using the BioEdit program, phylogenetic tree was analyzed using the MEGA X program, and type I L-asparaginase amino acid sequence was analyzed using the ExPASy translate program. The characterization results showed that the target gene has a high similarity to the *ansA* gene sequence of 20 *Bacillus subtilis* strains, i.e., 99-100%, is closely related to the *ansA* gene of *Bacillus subtilis* strain SRCM103629 and *Bacillus subtilis* strain GOT9, can be identified using restriction enzymes *Mlu*I and *Bst*I to differentiate the species of organisms they produce, has CDS encoding 329 amino acids with a dominant composition of polar amino acids (54.1%) and has a hydrophobicity profile of amino acids dominated by hydrophilic regions. All these characteristics have confirmed that the characterized gene is the *ansA* gene encoding type I L-asparaginase from *Bacillus subtilis*. The finding of the *ansA* gene from *Bacillus subtilis* in a thermohalophilic region in Southeast Sulawesi, Indonesia is a novelty of the research.

Keywords: *ansA* gene, DNA sequencing, gene characterization, PCR, specific primer

INTRODUCTION

L-asparagine is a non-essential amino acid in humans that is synthesized from L-aspartic acid. L-asparagine plays various important roles in the biosynthesis of different types of glycoproteins and other types of proteins (Eichler and Imperiali 2018). L-asparagine is known to induce uncontrolled proliferation of cancer cells in tumor tissue under certain conditions (Krall et al. 2016). The metabolism of this molecule is controlled by 2 key enzymes, L-asparagine synthase (biosynthesis) (Lomelino et al. 2017) and L-asparagine enzymes (degradation) (Vimal and Kumar 2022).

L-asparaginase is an enzyme that functions to degrade the optical isomer of the amino acid asparagine (Batool et al. 2016). L-asparaginase is widely used in various industrial sectors related to the presence of L-asparagine molecules, including the biomedical industry (cancer therapeutic agent) (Sengupta et al. 2021) and the food industry (to prevent the formation of acrylamide in food products) (Shakambari et al. 2019). L-asparaginase is divided into two types based on its location in the cells of

the microorganism that produces it: type I (located in the cytoplasm) and type II (periplasm). Type I and II L-asparaginase have the same primary function, catalyzing the conversion of L-asparagine to L-aspartate, but differ in efficiency and affinity for substrate molecules (Loch and Jaskolski 2021).

The bacterial species known to have the L-asparaginase encoding gene in their genome are *Erwinia carotovora* (Kishore et al. 2015), *Escherichia coli* (Mihooliya et al. 2020), *Thermus thermophilus* (Dumina and Zhgun 2023), *Pseudomonas aeruginosa* (Kuwabara et al. 2015), *Acinetobacter glutaminasificans* (Cachumba et al. 2016), *Proteus vulgaris* (Ren and Zhenhua 2011), *Streptomyces gulbargensis* (Al Yousef 2022), and *Bacillus subtilis* (Jiao et al. 2020).

Researchers have been actively exploring the ability of microbes to produce L-asparaginase since this enzyme was identified for its physiological ability in lymphoblastic leukemia (ALL) and Hodgkin's lymphoma cancer cells (Van Trimont et al. 2022). These explorations are generally based on obtaining bacterial isolates that encode the L-asparaginase gene in their genome and have

physiological properties of enzymes adapted to their habitat. Jamaluddin et al. (2018) succeeded in isolating and culturing several thermohalophilic bacterial cultures from the Wawolesea hot springs that could produce L-asparaginase enzymes, one of which was encoded by CAT3.4, which was further identified by Muzuni et al. (2022b) as phenotypically similar to *B. subtilis* subs *subtilis* BS-2. This bacterium grows optimally at a temperature of 30°C to 65°C, a pH of 5 to 8, NaCl content of 2 to 10%. This strengthens the notion that these bacteria have L-asparaginase, which structurally has a protein adapted to high temperature and high salt environments.

The structural and physiological properties of a protein are greatly influenced by the sequence of its constituent amino acids, which is determined by the sequence of the DNA encoding the protein in the producing organism (Vasudevan and Vaidyanathan 2017). This indirectly links assumptions about how an organism can perform specific metabolic processes in extreme environments to the composition or sequence of the DNA encoding the enzymes involved in these metabolic processes. Thus, this study determined the molecular characteristics of the gene encoding type 1 L-asparaginase from CAT3.4 thermohalophilic bacterial isolates from the hot springs of Wawolesea, North Konawe, Southeast Sulawesi.

Gene characterization can also involve several types of analysis, including BLAST (Basic Local Alignment Search Tool) to determine inter-sequence similarity (Pearson 2013), hydrophobicity analysis to determine the location/position of proteins in cells, and restriction site analysis to determine where various restriction enzymes cut in the gene (Maurya 2019). Sequencing results can also be analyzed to see the evolutionary relationships of genes encoding the same protein in different types of organisms (Al-Atiyat and Aljumaah 2014). The characterization results can be used as a reference when the *ansA* gene, which encodes type 1 L-asparaginase, is applied in the field of genetic engineering to optimize its function and efficiency.

MATERIALS AND METHODS

This study was conducted from December 2022 to May 2023 at the Laboratory of the Microbiology Unit, Department of Biology and the Laboratory of Molecular and Environmental Biology, Faculty of Mathematics and Natural Sciences, Halu Oleo University, Kendari, Southeast Sulawesi Province, Indonesia.

Procedures

Media preparation and equipment sterilization

Nutrient agar (NA) media was prepared by mixing 8 g of nutrient broth (NB) and 20 g of agar in 1000 mL of sterile distilled water. The media solution was then heated using a hot plate and homogenized using a magnetic stirrer. The media was then sterilized using an autoclave at 1 atm 121°C for 15 min. Equipment, such as Petri dishes, test tubes, and other equipment, sterilized using an autoclave at 1 atm 121°C for 15 min.

Rejuvenation and preparation of liquid cultures of CAT3.4 thermohalophilic bacterial isolate

Rejuvenation of CAT3.4 thermohalophilic bacterial isolates was performed by inoculating the bacterial isolates onto slant NA (Nutrient Agar) media by the scratch method using inoculating loops, which were then incubated in an incubator at 50°C for 24 h. The liquid culture of CAT3.4 thermohalophilic bacterial isolates was prepared for DNA isolation. The liquid culture of CAT3.4 thermohalophilic bacterial isolates was prepared by picking the bacterial colonies from the NA medium and aseptically inoculating them into 10 mL of liquid NB media. The culture was then incubated at 50°C for 24 h (Muzuni et al. 2021).

Specific primer designing

Primer was designed by collecting type 1 L-asparaginase gene sequences from different species of *Bacillus subtilis* bacteria obtained from GeneBank databases at the website <http://www.ncbi.nlm.gov>. The sequences were aligned using the ClustalW Alignment BioEdit program to determine the forward and reverse primer sequences based on the conserved sequences at the 3' and 5' ends. The results of the specific primary design are AsnBac1-F1 (5'-ACGCGATATTTCTTTTGGCCGG-3') and AsnBac1-R1 (5'-CAGTGAAGAGGTGCATGGTATG-3') (Muzuni et al. 2022a).

Isolation of genomic DNA

The genomic DNA of thermohalophilic bacteria CAT3.4 was isolated using the modified alkaline lysis method of Muzuni et al. (2021). The modification made was in the form of giving the vortex procedure after the addition of solution II. The first step in DNA isolation was to place a total of 1.5 mL of bacteria cultured in NB media for 24 h at 50°C into a microcentrifuge tube and centrifuge at 5000 rpm for 2 min. Pellets were collected and 100 µL of solution I (50 mM glucose, 25 mM Tris-Cl, 10 mM EDTA) was added and incubated on ice (0°C) for 5 min. Then 200 µL of solution II (0.2 N NaOH, 1% SDS) was added to the sample and inverted 5 times. After that, the mixture was vortexed and incubated in a water bath for 30 min, vortexing every 10 min. The sample was then mixed with 150 µL of solution III (5 mM sodium acetate, 11.5 mL of glacial acetic acid, 28.5 mL H₂O) and centrifuged at 10000 rpm for 10 min at 8°C.

The supernatant was collected and 1x volume phenol-chloroform (PC) was added. It was then centrifuged at 10000 rpm, 4°C for 10 min. The supernatant was taken and mixed with 3 M sodium acetate pH 5.2 to 0.1x volume of the supernatant then diluted with 1x volume of absolute ethanol, incubated at -20°C for 30 min and centrifuged at 10000 rpm, 4°C for 20 min. The DNA pellet was then washed with 500 µL of 70% alcohol, dried and dissolved in 30-50 µL of H₂O.

Type 1 L-asparaginase gene amplification

The type I L-asparaginase gene from thermohalophilic bacteria CAT3.4 was amplified by Polymerase Chain Reaction (PCR) technique using the composition of the PCR mix solution (Table 1).

Table 1. The composition of the type 1 L-asparaginase gene amplification mix solution

Solution	Final concentration
Genomic DNA of thermohalophilic bacteria CAT3.4	10.0 ng
AsnBac1-F1 Primer	0.5 mM
AsnBac1-R1Primer	0.5 mM
2x master mix	1.0 x
dH ₂ O	up to 10.0 µL
Total volume	10 µL

The PCR process consisted of initial denaturation at 94°C for 5 min; a PCR process of 35 cycles, which included denaturation at 94°C for 1 min, annealing at 55°C for 1 min, extension at 72°C for 90 seconds; and final extension at 72°C for 5 min. The PCR products, which were about 1000 bp in size, were then confirmed by the electrophoresis method.

Characterization of *L*-asparaginase gene and amino acid sequences

Purification and sequencing of PCR products were done in 1st Base Singapore. The sequencing results were characterized using a program according to the type of analysis needed, namely BLASTn NCBI, Bioedit, MEGA, and Expasy. The analysis to determine sequence similarity was performed using BLASTn on the NCBI site, the restriction site was mapped and the hydrophobicity profiles of amino acids was calculated using the BioEdit application, and the analysis of amino acid sequences was performed using the ExPASy translate program and the MEGA X application with the maximum likelihood 1000x replication algorithm for phylogenetic construction based on nucleotide and amino acid sequences (Muzuni et al. 2022a).

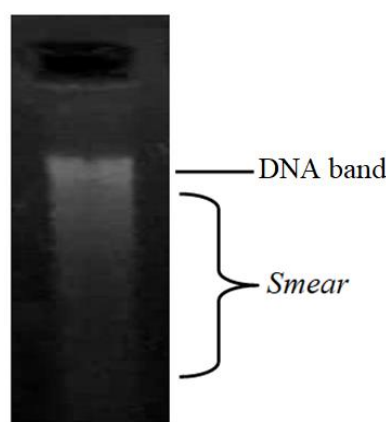
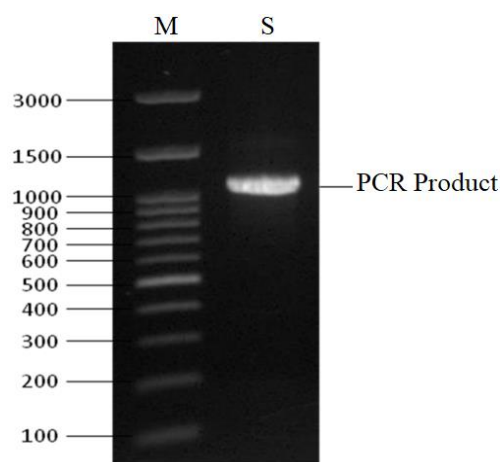
RESULTS AND DISCUSSION

Genomic DNA isolation

The results of genomic DNA isolation were visualized using the electrophoresis method, which separates charged particles under the influence of an electric field (Sambrook and Russel 2001). The electrophoresis results show intact DNA bands concentrated under the wells of the agarose gel (Figure 1). Below the intact DNA was a smear pattern, most likely DNA that was degraded during the isolation process. The isolated DNA was ready to be used as a template for amplification of the *ansA* gene from the CAT3.4 isolate.

Amplification of the *ansA* gene encoding type 1 L-asparaginase

Amplification of the gene coding for type 1 L-asparaginase was performed using AsnBac1-F1 and AsnBac1-R1 primers. AsnBac1-F1 primers are specifically designed to attach to the upstream region of the start codon, and AsnBac1-R1 primers downstream of the stop codon so that the gene is fully amplified from start to stop codon. The visualization results by electrophoresis showed that the target gene was successfully amplified, as indicated by the acquisition of firm and thick bands (Iqbal et al. 2016). Based on the comparison with the marker, the amplicon size was approximately 1000-1100 bp (Figure 2), according to the size prediction using the placement of AsnBac1-F1 and AsnBac1-R1 primers in the design template sequence. Further confirmation of the actual size was obtained from the sequencing results. A single DNA band (homologous in size) that is firm and thick is one of the markers of a successful PCR process (Hermono et al. 2017).

**Figure 1.** Visualization of the results of DNA isolation using the electrophoretic method on 1% agarose**Figure 2.** Results of *ansA* gene amplification using AsnBac1-F1 and AsnBac1-R1 primers. S: Sample, M: Marker

```

1  CAGTGAAGAGGTGCATGGTATGAAAAAATTATTGATGTTGACAACTGGGGGAACGATTGC 60
61  TTCAGTTGAAGGGGAAAATGGGCTGGCTCCCGGAGTCAAGGCTGATGAATTATTAAGTTA 120
121 CGTATCAAAACTTGATAACGATTACACAATGGAACTCAGTCGCTTATGAATATAGACAG 180
181 CACCAATATGCAGCCTGAATACTGGGTGGAAATAGCGGAAGCCGTTAAGGAAAATTATGA 240
241 TGCCTATGACGGGTTTGTATTACTACGGTACAGATACAATGGCCTATACATCTGCCGC 300
301 ACTATCGTATATGCTGCAGCATGCCAAAAGCCGATGTGCATCACCGGCTCGCAGATTCC 360
361 GATCACGTTCCAAAAACCGATGTAAAAAAAATATTACAGATGCCGTTTCGATTTCCTG 420
421 TGAAGGCGTGGGCGCGTTTATGTTGTATTTGACGGCAGAGTCATTAGGGAACGCGTGC 480
481 GATCAAATTAAGAACGAAAAGCTACGACGCATTTGAAAGCATCAATTACCCATATATCGC 540
541 TTTTATCAATGAAGACGGGATCGAATACAACAACAGTAACGGAACCTGAGAACGACAC 600
601 CTTCACAGTTGACACCTCACTATGCACAGATGTATGTCTGCTGAAGCTGCATCCAGGCTT 660
661 AAAGCCCGAAATGTTTGTATGCCCTGAAAAGCATGTACAAAGGAATTGTCATTGAGAGTTA 720
721 TGGCAGCGGAGGGGTGCCGTTTGAAGGCAGAGACATTTGTCAAAAGTGAATGAGAGTTA 780
781 CGAAAGCGGCATTGTCGTGGTCATTACGACTCAATGTCTTGAAGAAGGCGAAGACATGAG 840
841 CATTACGAAGTTGGCCGCAGAGTCAACCAAGACTTAATTATCCGATCAAGAAATATGAA 900
901 CACAGAAGCGATTGTGCCAAAATTGATGTGGGCACTAGGTCAGTCTTCGGATCTTCCTGT 960
961 CGTCAAGAGAATTATGGAACGCCGATAGCTGATGACGTTGTCCTGTAAATGTAAAAAC 1020
1021 GATAACCAGCAAAGAGAAAGAAGTTAAAGATTATGTAAACGCCCAAAAAGAATATCGC 1080
1081 GT 1082

```

Figure 3. Nucleoside sequence based on electropherogram data; note: the gray background shows CDS of type 1 L-Asparaginase

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
✓	Bacillus subtilis strain SEM-9 chromosome .complete genome	1999	1999	100%	0.0	100.00%	CP021123.1
✓	Bacillus subtilis strain SRCM103629 chromosome .complete genome	1999	1999	100%	0.0	100.00%	CP035413.1
✓	Bacillus subtilis strain GOT9 chromosome .complete genome	1993	1993	100%	0.0	99.91%	CP047325.1
✓	Bacillus subtilis strain TO-A JPC .complete genome	1993	1993	100%	0.0	99.91%	CP011882.1
✓	Bacillus subtilis TOA .complete genome	1993	1993	100%	0.0	99.91%	CP005997.1
✓	Bacillus subtilis strain CV16 chromosome .complete genome	1982	1982	100%	0.0	99.72%	CP062497.1
✓	Bacillus subtilis strain SRCM103689 chromosome .complete genome	1977	1977	100%	0.0	99.63%	CP035391.1
✓	Bacillus sp. wens01 chromosome .complete genome	1973	1973	100%	0.0	99.63%	CP024921.1
✓	Bacillus subtilis subsp. subtilis strain 2K1.1 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP032872.1
✓	Bacillus subtilis subsp. subtilis strain N3-1 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP032865.1
✓	Bacillus subtilis subsp. subtilis strain N2-2 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP032863.1
✓	Bacillus subtilis subsp. subtilis strain N1-1 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP032861.1
✓	Bacillus subtilis subsp. subtilis strain GFR-12 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP032852.1
✓	Bacillus subtilis subsp. subtilis strain 2RL2-3 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP032857.1
✓	Bacillus subtilis subsp. subtilis strain N4-2 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP032867.1
✓	Bacillus subtilis subsp. subtilis strain MH-1 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP032853.1
✓	Bacillus subtilis strain SRCM103886 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP035162.1
✓	Bacillus subtilis strain ATCC 21228 chromosome .complete genome	1971	1971	100%	0.0	99.54%	CP020023.1
✓	Bacillus subtilis strain DKU_NT_03 .complete genome	1971	1971	100%	0.0	99.54%	CP022891.1
✓	Bacillus subtilis strain DKU_NT_02 .complete genome	1971	1971	100%	0.0	99.54%	CP022890.1

Figure 4. BLASTn analysis results. The identity of the organism in GeneBank that has the highest nucleotide sequence similarity to the nucleotide sequence of the *ansA* gene from CAT 3.4 isolate

Sequencing of the *ansA* gene encoding type 1 L-asparaginase

Sequencing is intended to obtain entire nucleotide sequences from the electropherogram data of PCR products. Sanger sequencing uses dideoxynucleoside triphosphate (ddNTP), which contains a ribose sugar with a hydroxyl group (-OH) that has lost an O atom (deoxygenated) at the C2 and C3 atoms. When ddNTP is added to single-stranded DNA (ssDNA) by the polymerase, a forced termination of chain elongation occurs (Negi et al. 2014). With this method, it is possible to obtain all ssDNA

strands of different sizes complementary to the template. The entire ssDNA strand contains ddNTP labeled with a fluorescent dye, so the detector interprets it as peak data with a unique color corresponding to the ddNTP it represents (ddTTP: red; ddATP: green; ddCTP: blue; ddGTP: black) (Bisht and Panda 2014). Based on the sequencing results, the nucleotide sequence was obtained with a size of 1082 bp, and the CDS (coding sequence) of the *ansA* gene encoding type 1 L-asparaginase was found to be intact (from the start codon/ATG to the stop codon/TAA) with a gene size of 990 bp (Figure 3).

Sequence characterization of the *ansA* gene encoding type 1 L-asparaginase

BLAST analysis

The BLASTn (Basic Local Alignment Sequence Tools) analysis aims to obtain information on the types of organisms likely to encode the *ansA* gene (as well as the identity of the target organism) and the similarity of the *ansA* gene sequence with other sequences. BLASTn uses an algorithm that functions by matching input nucleotide sequences to all sequences in GeneBank and then calculating their statistical similarity (Wheeler and Bhagwat 2007).

The BLASTn results show the top 20 microorganisms with the highest sequence similarity to the *ansA* gene sequence of isolate CA T3.4, all of which are *B. subtilis* with multiple strains. *Bacillus subtilis* strain SEM-9 and *B. subtilis* strain SRCM103629 had the highest similarity with 100% identity percentage (Figure 4). The percent identity value indicates how likely it is that the nucleotide sequence of the *ansA* gene is the same as the sequence in the organism database. Thus, the *ansA* gene of the CAT 3.4 isolate has the same order and sequence size as the *ansA* gene sequence of *B. subtilis* strain SEM-9 and *B. subtilis* strain SRCM103629.

The maximum score in Figure 4 is the total score obtained from the alignment between the *ansA* gene sequences of the CAT 3.4 isolate and all sequences in the GeneBank databases. In contrast, the total score was obtained from all alignments with parallel sequences. The results of the BLASTn analysis show that all organisms shown had the same maximum score and total score; for example, *B. subtilis* strain SRCM103629 had a maximum score of 1999 and a total score of 1999 (Figure 4). This shows that the alignment between the gene sequences and CAT3.4 isolate with the genome of *B. Subtilis* strain SRCM103629 occurs only once, and no other sequences exist in the genome of *B. subtilis* strain SRCM103629 has similarities to the gene sequence and CAT3.4 isolate.

The alignment results with the other 19 top bacterial strains have the same maximum score and total score, meaning that the sequence alignment occurred only once

for each organism. This eliminates the possibility of random alignment, which could cast doubts on the BLASTn results obtained. These results are also confirmed by the query coverage values, which were all 100% with an error value of 0. The query coverage percentage represents the value of how long the *ansA* gene sequence is aligned with the database sequence. An error value of 0 confirms the exact alignment result of a randomly probable alignment. Wahyuni and Sepianto (2018) stated that a nucleotide sequence is confirmed to be from the same species if the query coverage and percentage identity are close to 100% with an E value close to 0. In general, it can be concluded that the input sequence (*ansA* gene) is a gene sequence that is also shared by 20 strains of *B. subtilis* (Figure 4).

Analysis of phylogenetic trees using the MEGA X program

The gene encoding type 1 L-asparaginase (*ansA* gene) was found in several species of bacteria of the genus *Bacillus* and in several other types of bacteria of different genera. From an evolutionary point of view, it is almost certain that these bacteria originally had the same ancestral species as the *ansA* gene in their genome. Over time, the evolution of derived species into different species also adds genetic variation to the genome (Weller and Wu 2015).

The results of the phylogenetic tree construction show the formation of 5 clades (groups) (Figure 5). Clade I consisted of CAT 3.4 isolates, *B. subtilis* strain SRCM103629, *B. subtilis* strain GOT9, *B. subtilis* strain CV16, *B. subtilis* strain ATCC21228, *Bacillus tequilensis* strain PV9W, *Bacillus rugosus* strain SPB7 and *Bacillus valismortis* strain Bac111. Clade II consisted of *Bacillus halotolerans* strain F41-3 and *Bacillus mojavensis* strain UCMB5075. Clade III comprised *Bacillus atrophaeus* strain SRCM101359, *Bacillus endophyticus* strain FH5, and *Bacillus sonorensis* strain HGS2.10A. Clade IV comprised *Bacillus nakamurai* strain NRRL B-41091, *Bacillus amyloliquefaciens* Y2 and *Bacillus velezensis* strain UCMB5044. Clade V consisted of *Bacillus licheniformis* strain 14ADL4.

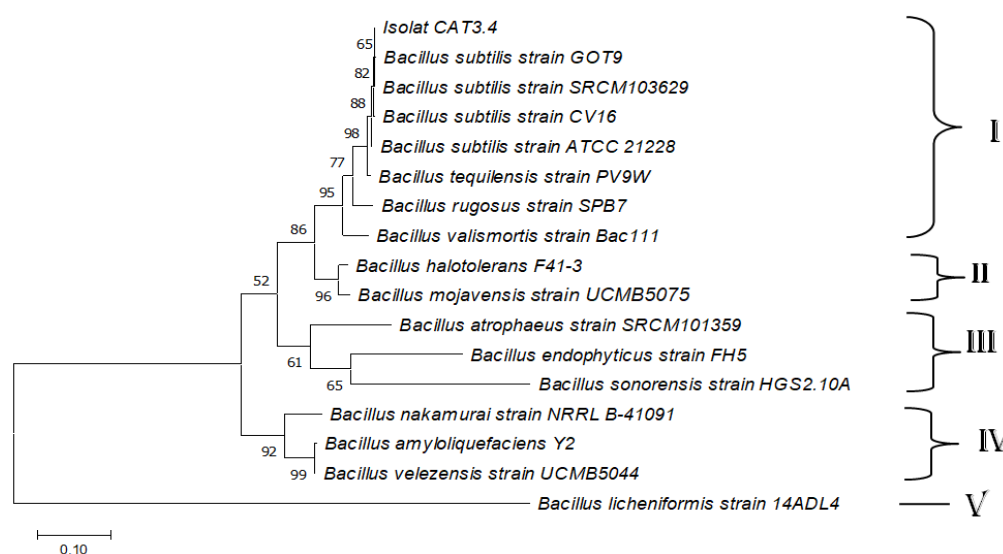


Figure 5. The phylogenetic tree was constructed using the *ansA* gene's nucleotide sequence from isolate CAT 3.4 and comparator bacterial species from the genus *Bacillus*. The numbers on the branches represent the bootstrap values

Clade I in the phylogenetic tree was divided into six subclades, with the first subclade consisting of CAT3.4 isolate, *B. subtilis* strain SRCM103629, and *B. subtilis* strain GOT9 with a bootstrap value of 65; the second subclade consisting of members of the first subclade and *B. subtilis* strain CV16 with a bootstrap value of 82; the third subclade consisting of members of the second subclade and *B. tequilensis* strain ATTC 21228 with a bootstrap value of 88; the fourth subclade consisting of members of the third subclade and *B. tequilensis* strain PV9W with a bootstrap value of 98; the fifth subclade consisting of members of the fourth subclade and *B. rugosus* strain SPB7 with a bootstrap value of 77; and the sixth subclade consisting of members of the fifth subclade and *B. valismortis* strain Bac111 with a bootstrap value of 95. Clade II consisted of *B. halotolerans* strain F41-3 and *B. mojavensis* strain UCMB5075 with a bootstrap value of 96. Clade III consisted of 2 subclades; the first consists of *B. endophyticus* strain SRCM101359 and *B. sonorensis* strain HGS2.10A with a bootstrap value of 65. Meanwhile, the second subclade comprised members of the first subclade and *B. atrophaeus* strain SRCM101359, with a bootstrap value 61. Clade IV also consists of 2 subclades, namely the first subclade consisting of *B. amyloliquefaciens* strain Y2 and *B. velezensis* strain UCMB5044 with a bootstrap value of 99 and the second subclade consisting of members of the first subclade and *B. Nakamura* strain NRLL B-41091 with a bootstrap value of 92. Clade V had no subclade because it had only one member, *B. licheniformis* strain 14ADL4.

The *ansA* gene from the CAT3.4 isolate was most closely related to the *ansA* gene from *B. subtilis* strain SRCM103629, forming the first subclade in Clade I with a bootstrap value of 65. It was closer to all members of Clade I than to members of Clade II, Clade III, and Clade IV. The *ansA* gene from CAT 3.4 isolate clustered in a clade with all members of the *B. subtilis* species, in agreement with Muzuni et al. (2022b) who found that CAT3.4 isolate is similar to *B. subtilis* subs. *subtilis* BS.2 based on its 16S RNAs. Although the *ansA* gene is not a species-determining molecular marker, genetic variation within individuals of a species is generally lower (more uniform) than that within individuals of higher taxa (Van Rossum et al. 2020) so that the *ansA* gene of CAT3.4 isolates will tend to cluster and be more closely related to the *ansA* gene of members of the species *B. subtilis* than to that of other members of the genus *Bacillus*, as shown in the clustering and branching patterns of the phylogenetic tree.

The pattern of grouping the *ansA* gene of CAT3.4 isolates with other members of Clade I is closely related to the form of adaptation to the environmental conditions in which it lives. All bacterial species in Clade I belong to the thermophilic bacteria except *B. subtilis* strain GOT9, which belongs to the halophilic bacteria. Microorganisms adapt to both regular and periodic environmental changes by adapting cellular systems, acquiring special abilities through horizontal gene transfer (HGT) or gene neofunctionalization, and reorganizing dynamic cellular systems to adapt to temporary environmental changes (Ho and Zhang 2018).

Adaptation of the cellular system occurs through simple mutations that occur every one or more generations that accumulate at a locus and can determine the survivability of genes (Jordan et al. 2022). This contrasts with HGT, which lasts for a short time but becomes integrated into the genome occurs over a long period of time (Mortier-Barrière et al. 2020). Environmental factors, such as ambient temperature, may be one of the factors that cause spontaneous nucleotide sequence alterations (Chu et al. 2018). Based on their ability to live within a specific temperature range, they are generally thermophilic bacteria, with *B. subtilis* having a growth temperature range of 45-55°C (Andriani et al. 2017), *B. tequilensis* strain PV9W having a growth temperature range of 25-50°C (Fachrial et al. 2020), *B. rugosus* strain SPB7 having a growth temperature range of 20-45°C (Bhattacharya et al. 2020) and *B. valismortis* strain Bac111 having a growth temperature range of 10-50°C (Roberts et al. 1996). All members of this Clade I have a relatively uniform maximum growth temperature, which is thought to be one of the reasons that the nucleotides of the *ansA* gene are closely related to each other, so that they are grouped in the same clade.

Based on the results of the phylogenetic tree construction, the *ansA* gene from the CAT3.4 isolate also has the most distant relationship to the *ansA* gene from *B. licheniformis* strain 14ADL4, which also provides an evolutionary clue to the ancestral genes of the two species. Bacteria can evolve in several ways, including the accumulation of mutations at specific loci in the genome over time and the many generations of individual bacterial lineages (Gibson and Eyre-Walker 2019). The branching length of a branch from the putative ancestor is defined as the genetic distance or the amount of evolutionary genetic change, which is a vector of time to the rate of evolution; sometimes, in phylogenetic construction, assumptions are made about the molecular clock, which is the constant rate of the evolutionary process, so that theoretically it can be predicted the start time changes in the organism of the individual ancestor (Shamir 2001). The long distance branching indicates that the separation of the *ansA* gene of the CAT3.4 isolate from the *ansA* gene of *B. licheniformis* strain 14ADL4 occurred much earlier than the other reference bacteria.

Analysis of restriction-enzyme cut sites

Restriction enzymes are a group of enzymes naturally found in bacteria and archaea that protect microbes from viral infections and foreign DNA molecules. Restriction enzymes bind to short nucleotide sequences in DNA and catalyze the opening of the double strand at the binding site, causing the DNA to be fragmented into smaller pieces. Because of these properties, restriction enzymes are generally used as the standard in molecular studies of sequence identification mechanisms (Di Felice et al. 2019). The analysis of the restriction sites of the *ansA* gene of CAT3.4 isolates aims to map the recognition sites of different restriction enzymes, which can later be used as a distinguishing feature of the *ansA* gene encoded by other bacterial species.

The restriction site analyses using the BioEdit programs includes all restriction enzymes with 4 bp and 6 bp recognition sites and are palindromic. The specificity of restriction enzyme cleavage can be used as a gene marker at the species level or at the *Bacillus* genus level. The difference in the location of the recognition site for the same restriction enzyme illustrates the genetic variation of the *ansA* gene within the species and the genus *Bacillus*.

Enzymes with broad cut points can potentially be used as molecular markers of the *ansA* gene at the *Bacillus* genus level, namely *BsrGI*, *PstI*, *SphI*, and *SspI*. The *BsrGI* enzyme has a cutting site in almost all analyzed *ansA* gene sequences (except for *B. valismortis* strain Bac111 and *B. sonorensis* strain HGS2.10A); this deficiency can be compensated by *EagI*, which has a cutting site in the *ansA* gene sequence of both bacterial species, so that the combination of enzymes *BsrGI* and *EagI* can be used as molecular markers of the *ansA* gene at the *Bacillus* genus level (Table 2).

In addition to being used as a molecular marker for the *ansA* gene at the *B. subtilis* species level and the *Bacillus* genus level, the results of mapping the *ansA* gene restriction sites can also be used as molecular markers for the producing species. Typically, bacterial identity is determined by a PCR technique based on 16S rDNA using universal primers annealed to the conserved region of all bacterial species (Sune et al. 2020). The combination of the PCR technique with restriction enzymes allows the wide range of non-conserved gene sequences, including the *ansA* gene, to be used to determine the identity of the organism that possesses the gene. This technique is known as polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP).

PCR-RFLP involves the characterization of PCR products using multiple restriction enzymes to obtain DNA fragments of varying number and size that form a unique banding pattern when visualized by electrophoresis (Rohit et al. 2016). When the *ansA* gene is amplified using the primer pairs AsnBac1-F1 and AsnBac1-R1 which are then digested with *MluI*, 2 fragments of 455 bp and 535 bp sizes will be obtained.

This fragmentation pattern can be used as an indicator of the identity of *B. subtilis* when the *ansA* gene sequence is used as a barcode. However, such an identification will be difficult to maintain for two reasons: (i) the genetic variation at the genus level or above is very high, so that primary acquisition using conserved sequences in the same gene will be difficult, (ii) if mutations occur at the restriction recognition site and remain to produce a functional protein, individual bacteria will survive and pass on the gene on to the next generation, which may lead to identification errors after a period of time.

Amino acid sequence analysis of type 1 L-asparaginase

The translation results using Exspasy yielded 329 amino acids (Figure 6) with the following composition: alanine 6.08%, cysteine 1.22%, aspartic acid 6.69%, glutamic acid 7.6%, phenylalanine 2.74%, glycine 7.29%, histidine 0.91%, isoleucine 8.21%, lysine 6.08%, leucine 7.60%, methionine 3.95%, asparagine 4.26%, proline 3.65%, glutamine 3.04%, arginine 3.04%, serine 6.08%, threonine 7.60%, valine 8.81%, tryptophan 0.61%, and tyrosine 4.56%.

Amino acids are divided into two groups based on their polarity (Nanni et al. 2014), namely the nonpolar and polar amino acids. The polar amino acids that make up type 1 L-asparaginase from CAT3.4 isolates (valine, glycine, tryptophan, tyrosine, proline, methionine, isoleucine, alanine, and phenylalanine) have a percentage of 45.9% and the polar amino acids (cysteine, leucine, histidine, arginine, lysine, aspartic acid, glutamine, asparagine, serine, glutamic acid, and threonine) have a percentage of 54.1%.

Based on UniProt data for the *B. subtilis* type 1 L-asparaginase enzyme with the access code UniProtKB-P26900 (ASPG1_BACSU), this enzyme has an active site consisting of a catalytic site at the 12th amino acid sequence in the form of threonine (T) and a substrate binding site at the 54th amino acid sequence, namely serine (S) and 85-86 threonine (T)-aspartic acid (D). Type 1 L-asparaginase is a homodimeric protein, meaning that it is composed of two protein subunits with identical amino acid sequences (Mou et al. 2015).

```

1  M K K L L M L T T G G T I A S V E G E N G L A P G V K A D E 30
31 L L S Y V S K L D N D Y T M E T Q S L M N I D S T N M Q P E 60
61 Y W V E I A E A V K E N Y D A Y D G F V I T H G T D T M A Y 90
91 T S A A L S Y M L Q H A K K P I V I T G S Q I P I T F Q K T 12
121 D A K K N I T D A V R F A C E G V G G V Y V V F D G R V I Q 150
151 G T R A I K L R T K S Y D A F E S I N Y P Y I A F I N E D G 180
181 I E Y N K Q V T E P E N D T F T V D T S L C T D V C L L K L 210
211 H P G L K P E M F D A L K S M Y K G I V I E S Y G S G G V P 240
241 F E G R D I L S K V N E L I E S G I V V V I T T Q C L E E G 270
271 E D M S I Y E V G R R V N Q D L I I R S R N M N T E A I V P 300
301 K L M W A L G Q S S D L P V V K R I M E T P I A D D V V L 329

```

Figure 6. Type 1 L-asparaginase amino acid sequence from CAT3.4 isolates based on analysis using the Exspasy program. T: active site, S and TD: ligand-binding sites (based on Uniprot data, accessed 2021). Note: N: asparagine (polar), C: cysteine (polar), L: leucine (nonpolar), H: histidine (polar), R: arginine (polar), Y: tyrosine (nonpolar), K: lysine (polar), D: aspartic acid (polar), G: glycine (nonpolar), W: tryptophan (nonpolar), I: isoleucine (nonpolar), V: valine (nonpolar), M: methionine (nonpolar), P: proline (nonpolar), A: alanine (polar), F: phenylalanine (nonpolar), S: serine (polar), E: glutamic acid (polar), T: threonine (polar), Q: glutamine (polar)

Table 2. Cutting sites of various restriction enzymes in the *ansA* gene sequence using the BioEdit program

Restriction enzyme	Recognition sites	Position																
		Isolat CAT3.4	<i>B. subtilis</i> ATTC 21228	<i>B. subtilis</i> CV16	<i>B. subtilis</i> SRCM103629	<i>B. subtilis</i> GOT9	<i>B. Tequilensis</i> PV9W	<i>B. rugosus</i> SPB7	<i>B. Halotolerans</i> F41-3	<i>B. mojavensis</i> UCMB5075	<i>B. valismortis</i> Bac111	<i>B. nakamurai</i> NRL1 B-41091	<i>B. amyloliquefaciens</i> Y2	<i>B. velezensis</i> UCMB5044	<i>B. atrophaeus</i> SRCM101359	<i>B. endophyticus</i> SRCM101359	<i>B. sonorensis</i> HGS2.10A	<i>B. licheniformis</i> 14ADLA
<i>Bsr</i> GI	T'GTAC_A	675	675	675	675	675	675	675	675	675	-	679	696	696	672	602	-	130
<i>Mlu</i> I	A'CGCG_T	455	455	455	455	455	-	-	489	489	455, 489	-	-	-	-	-	439	217
<i>Pst</i> I	C_TGCA_G	300	300	300	300	300	300	300	300	300	300	-	-	-	300	280	-	507, 622
* <i>Sna</i> BI	TAC'GTA	103	103	103	103	103	103	103	103	103	103	103	124	124	-	-	-	-
<i>Sph</i> I	G_CATG'C	305	305	305	305	305	305	305	-	305	305	305, 566	326	326	-	305	-	395
* <i>Ssp</i> I	AAT'ATT	376	376	376	376	376	376	376	376	376	-	-	-	-	182	733	376	-
<i>Eag</i> I	C'GGCC_G	-	-	-	-	-	-	835	835	835	835	835	856	856	-	-	835	-
* <i>Sma</i> I	CCC'GGG	-	-	-	-	-	-	73	72	72	-	73	94, 657	94, 657	633	-	-	-
<i>Ase</i> I	AT'TA_AT	-	-	-	-	-	-	-	905	905	528, 905	905	926	926	902	504	-	-
<i>Bsr</i> I	CG'CG	457	457	457	457	457	-	368, 457	491	491	457, 491	459, 521	407, 480	407, 480	659	441	398, 441	219, 248
* <i>Bfr</i> BI	ATG'CAT	-	-	-	-	-	-	-	-	-	-	491	512	512	-	-	-	172, 374
<i>Apa</i> LI	G'TGCA_C	-	-	-	-	-	-	-	-	-	-	-	581	581	601	-	-	488
* <i>Eco</i> RV	GAT'ATC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	979	751
<i>Hind</i> III	A'AGCT_T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	587	759

Note: (*) Blunt end, (') Cutting sites, (_) Cutting sites is on another strand

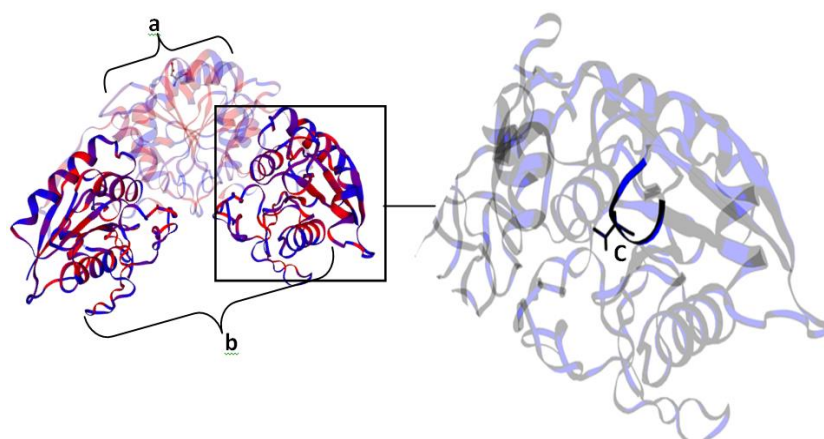


Figure 7. 3D structure of type 1 L-asparaginase, a: C-terminal region, b: N-terminal region, and c: active site/catalyst site. (red regions are hydrophobic, blue are hydrophilic) (UniProt/P26900 accessed 17 Maret 2023)

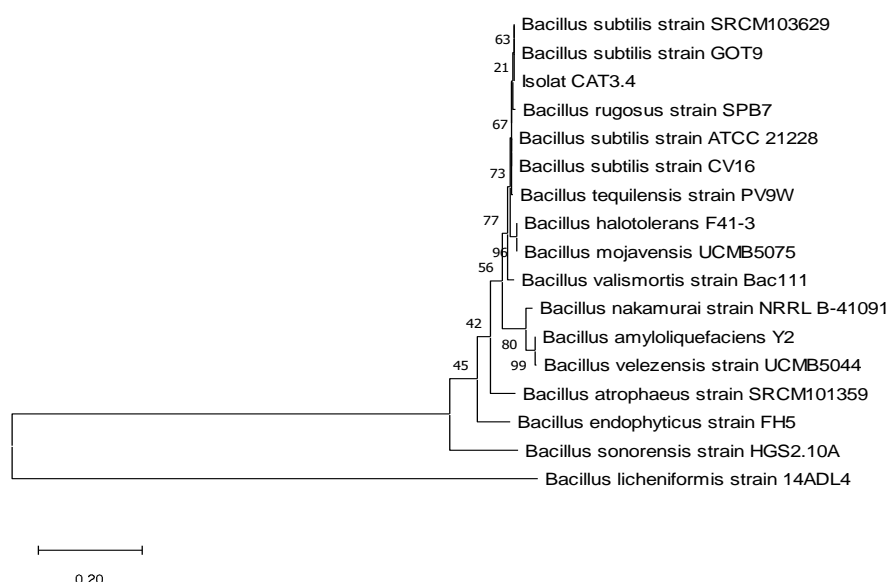


Figure 8. A phylogenetic tree constructed using the amino acid sequence of type 1 L-asparaginase

The amino acids that make up type 1 L-asparaginase are folded so that the amino acids at sequences 12, 54, and 85-86, the amino acids that make up the active site, are close together (Figure 7). The type 1 L-asparaginase dimer has three central regions: two N-terminals (3-187 in subunits I and II) and a C-terminus (205-319 in subunits I and II). The active site of L-asparaginase is located at both the N-terminal subunits I and II. Type 1 L-asparaginase is an enzyme that is active in the cytoplasm with a low affinity compared to L-asparaginase type 2 (Sharafi et al. 2017). This may be closely related to the nature of type 2 L-asparaginase, a homotetramer with four active sites in one molecule.

Construction of a phylogenetic tree based on amino acid sequence

The results of amino acid sequence analysis are input data for phylogenetic tree construction. This construction is based on multiple codons that have different nucleotide arrangements but are translated as the same type of amino group (synonymous codons), which are the result of

mutations in the gene sequence without altering the encoded amino acids or the resulting protein functionalities (DeRisi et al. 2019). Based on this, the phylogenetic relationship between the CAT3.4 isolate and other reference bacteria based on the amino acid sequence of the type 1 L-asparaginase gene has the possibility of forming a different branching pattern when it is presented based on the nucleotide sequence of the encoding gene.

The results of the phylogenetic tree construction show that the enzyme type 1 L-asparaginase isolate CAT3.4 has a close relationship with type 1 L-asparaginase from all members of *B. subtilis* along with *B. rugosus* strain SPB7 and *B. tequilensis* strain PV9W (Figure 8). This is also supported by the very short branch length (well below the scale), indicating that the genetic variation at the amino acid level for type 1 L-asparaginase from both CAT3.4 isolates and all the comparator bacterial species is less than the genetic variation at the nucleotide level (consistent with the grouping of nucleotide sequences into the same clade but with different branch lengths (Figure 5).

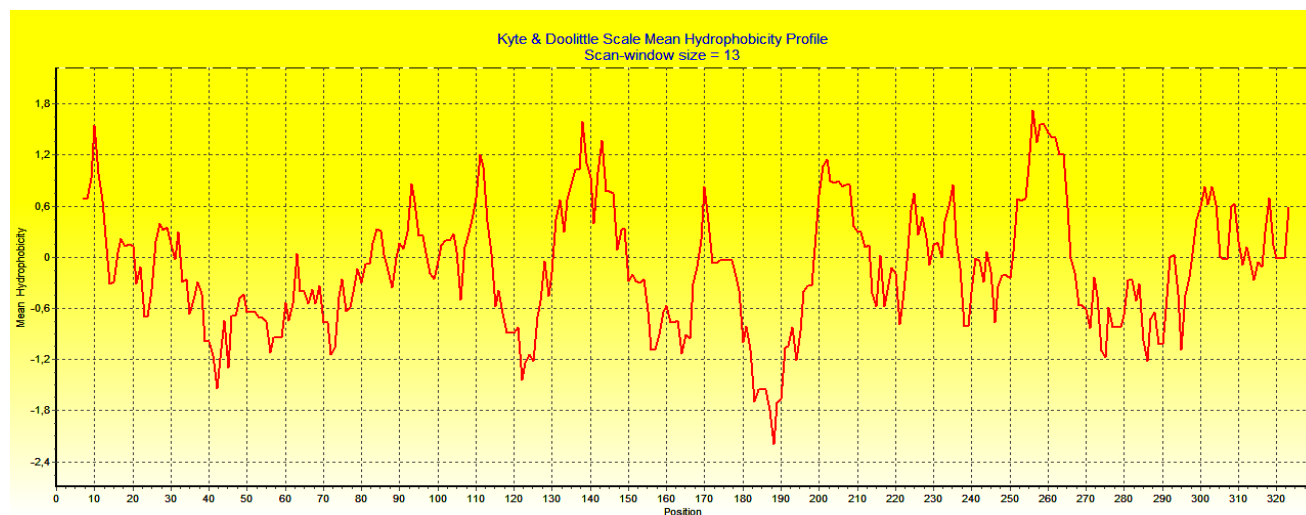


Figure 9. Hydrophobicity profile of the amino acid type 1 L-asparaginase isolate CAT3.4

Based on the branching pattern of the constructed phylogenetic tree, it is known that *B. Subtilis* strain GOT9 clusters on the same branch with CAT3.4 isolates and *B. subtilis* strain SRCM103629. This clustering pattern is different from the phylogenetic tree clustering pattern obtained using nucleotide sequences. This is because the nucleotide sequence between CAT3.4 isolates and *B. subtilis* strain GOT9 has an identity percentage of 99.91% based on BLASTn (Figure 4), but based on BLASTp results (data not shown), it is known that the amino acid similarity between the three species is 100%, confirming the presence of silent mutations, type of substitution at the 366th nucleotide.

Analysis of the hydrophobicity profile of type 1 L-asparaginase

The hydrophobicity profile of the amino acid sequences that make up type 1 L-asparaginase from CAT3.4 isolates was analyzed using the BioEdit program with the Kyte and Doolittle index model (also known as the hydropathy index model) to obtain the hydrophobicity profile of the sequence. This model works by assigning a value to each amino acid unit that also represents the average hydropathy property of the sequence. The higher the hydropathy value of an amino acid is, the more hydrophobic the amino acid will be (Higgs et al. 2017).

The analysis results of the hydrophobicity profile of the amino acid sequence of type 1 L-asparaginase are presented in graphical form with two axes (X and Y). The Y axis is the average hydrophobicity value, and the X axis is the sequence of amino acids (Figure 9).

The hydrophobicity profile describes the ability of each amino acid unit or the entire straight-chain amino acid sequence to interact with the external environment, which is determined by the polarity of the side chain (R) of each amino acid residue (Zhu et al. 2016). In a polar environment (e.g., cytoplasm), hydrophilic or water-soluble polar amino acids will be located on the surface of the

protein. Conversely, hydrophobic nonpolar amino acids will fold inward and form hydrophobic interactions in the protein core to maintain overall protein stability (Sahoo et al. 2023).

Polar amino acids such as glutamic acid and lysine (composing 13.68% of the total amino acid constituents) also help to strengthen the hydrogen bonds between amino acids on the surface by interacting with the side chains of nonpolar amino acids so that the protein structure has more stable hydrogen bonds (Kuroda and Gray 2016). Uniprot data (accessed in January 2021) shows that the type 1 L-asparaginase enzyme is active in the cell cytoplasm (a polar environment). Type 1 L-asparaginase from the CAT3.4 isolate has a hydrophobicity profile dominated by the region's hydrophilic region, which is supported by polarity analysis of the amino acid sequence (54.1%) (Figure 9). Type 1 L-asparaginase from CAT3.4 isolates contains 45.9% nonpolar amino acids distributed in groups along the sequence, groups of hydrophobic amino acids that enhance the stability of the proteins they compose.

The results of the hydrophobicity analysis showed that there are several regions in the amino acid sequence of type 1 L-asparaginase that are strongly hydrophobic, namely amino acid sequences 130-150 (RFACEGVGGVYVFDGRVIQ), 200-210 (LCTDVCLLKL), and 250-265 (NELIESGIVVVITTQ). If the region matches the 3D data of the type 1 L-asparaginase protein, it is far away, surrounded by hydrophilic regions in the protein.

The *ansA* gene encoding type 1 L-asparaginase from isolate CAT3.4 (1) has a high similarity to the *ansA* gene sequence of 20 *B. subtilis* strains (99-100%) based on BLASTn results, (2) is closely related to the *ansA* gene of *B. subtilis* strain SRCM103629 and *B. subtilis* strain GOT9 based on their nucleotide phylogenetics, (3) can be identified using the restriction enzymes *MluI* and *BstI* as a differentiator for the species of organism producing it, (4) encodes 329 amino acids with a dominant composition of

polar amino acids (54.1%), and (5) has a hydrophobicity profile of amino acids dominated by the hydrophilic region. All these characteristics confirm that the gene to be characterized is an *ansA* gene active in the cytoplasm of the *B. subtilis* species.

The target gene has a high similarity to the sequence of the *ansA* gene of *B. subtilis* strains 99-100%, is closely related to the *ansA* gene of *B. subtilis* strain SRCM103629 and *B. subtilis* strain GOT9, can be identified using restriction enzymes *Mlu*I and *Bst*I, has a CDS encodes 329 amino acids with a dominant composition of polar amino acids and has a hydrophobicity profile of amino acids dominated by hydrophilic regions. All these characteristics confirmed the *ansA* gene encoding type 1 L-asparaginase from *B. subtilis*. The enzymes has play an important role as anti-cancer and prevent the formation of acrylamide in food. As a follow-up to this research, it is necessary to clone the gene into an expression vector to produce the enzyme on a large scale. Furthermore, the enzymes were tested on cancer cells and food ingredients so that they could reduce cancer cell division and acrylamide production.

ACKNOWLEDGEMENTS

The authors are grateful to the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia for the full research funding under contract number 42/UN29.20/PG/2023. Our gratitude is extended to the Microbiology Laboratory and the Molecular Biology and Environment Laboratory, Faculty of Mathematics and Natural Sciences, Halu Oleo University, Kendari, Indonesia for their assistance in sample analysis.

REFERENCES

- Al-Atiyat R, Aljumaah RS. 2014. Genetic distances and phylogenetic trees of different Awassi sheep populations based on DNA sequencing. *J Genet Mol Res* 13 (3): 6557-6568. DOI: 10.4238/2014.August.26.6.
- Al Yousef SA. 2022. *Fusarium* sp. L-asparaginases: purification, characterization, and potential assessment as an antileukemic chemotherapeutic agent. *Environ Sci Pollut Res* 29: 11243-11254. DOI: 10.1007/s11356-021-16175-5.
- Andriani Y, Safitri R, Rochima E, Fakhrudin SD. 2017. Characterization of *Bacillus subtilis* and *B. licheniformis* potentials as probiotic bacteria in Vanamei Shrimp Feed (*Litopenaeus vannamei* Boone, 1931). *Nusantara Biosci* 9 (2): 188-193. DOI: 10.13057/nusbiosci/n090214.
- Batool T, Makky EA, Jalal M, Yusoff MM. 2016. A comprehensive review on L-asparaginase and its applications. *Appl Biochem Biotechnol* 178 (5): 900-23. DOI: 10.1007/s12010-015-1917-3.
- Bhattacharya D, Villalobos SDLS, Ruiz VV, Selvin J, Mukherjee J. 2020. *Bacillus rugosus* sp. nov. producer of a diketopiperazine antimicrobial, isolated from marine sponge *Spongia officinalis* L. *Antonie Leeuwenhoek* 113: 1675-1687. DOI: 10.1007/s10482-020-01472-9.
- Bisht SS, Panda K. 2014. Advance in Biotechnology: DNA Sequencing Method and Application. Springer, India. DOI: 10.1007/978-81-322-1554-7_2.
- Krall AS, Xu S, Graeber TG, Braas D, Christofk HR. 2016. Asparagine promotes cancer cell proliferation through use as an amino acid exchange factor. *Nat Commun* 7: 11457. DOI: 10.1038/ncomms11457.
- Cachumba JJM, Antunes FAF, Peres GFD, Brumano LP, Santos JC, Da Silva SS. 2016. Current applications and different approaches for microbial L-asparaginase production. *Braz J Microbiol* 47: 77-85. DOI: 10.1016/j.bjm.2016.10.004.
- Chu XL, Zhang BW, Zhang QG, Zhu BR, Lin K, Zhang DY. 2018. Temperature responses of mutation rate and mutational spectrum in an *Escherichia coli* strain and the correlation with metabolic rate. *BMC Evol Biol* 18 (1): 126. DOI: 10.1186/s12862-018-1252-8.
- DeRisi JL, Huber G, Kistler A, Retallack H, Wilkinson M, Yllanes D. 2019. An exploration of ambigrammatic sequences in namaviruses. *Sci Rep* 9 (1): 17982. DOI: 10.1038/s41598-019-54181-3.
- Di Felice F, Micheli G, Camilloni G. 2019. Restriction enzymes and their use in molecular biology: An overview. *J Biosci* 44: 38. DOI: 10.1007/s12038-019-9856-8.
- Dumina M, Zhgun A. 2023. Thermo-L-Asparaginases: from the role in the viability of thermophiles and hyperthermophiles at high temperatures to a molecular understanding of their thermoactivity and thermostability. *Intl J Mol Sci* 24 (3): 2674. DOI: 10.3390/ijms24032674.
- Eichler J, Imperiali B. 2018. Biogenesis of asparagine-linked glycoproteins across domains of life-similarities and differences. *ACS Chem Biol* 13 (4): 833-837. DOI: 10.1021/acscchembio.8b00163.
- Fachrial E, Putri RRJS, Lister INE, Anggraini S, Harmileni, Nugroho TT, Saryono. 2020. Molecular identification of cellulase and protease producing *Bacillus tequilensis* UTMSA14 isolated from the geothermal hot spring in Lau Sidebuk Debuk, North Sumatra, Indonesia. *Biodiversitas* 21 (10): 4719-4725. DOI: 10.13057/biodiv/d211035.
- Gibson B, Eyre-Walker A. 2019. Investigating evolutionary rate variation in bacteria. *J Mol Evol* 87 (9): 317-326. DOI: 10.1007/s00239-019-09912-5.
- Hermono BAS, Bintari SH, Mustikaningtyas D. 2017. Identification of *Salmonella* sp in fruit juice snacks in Gunungpati District, Semarang by PCR. *Jurnal MIPA* 40 (2): 68-73.
- Higgs T, Stantic B, Hoque T, Sattar A. 2017. Hydrophobic-hydrophilic forces and their effects on protein structural similarity. In *Suppl Conf Proc*: 1-12. DOI: 10.4225/03/5a13709f243b5.
- Ho WC, Zhang J. 2018. Evolutionary adaptations to new environments generally reverse plastic phenotypic changes. *Nat Commun* 9: 350. DOI: 10.1038/s41467-017-02724-5.
- Iqbal M, Buwono ID, Kurniawati N. 2016. Comparative analysis of DNA isolation methods for detection of White Spot Syndrome Virus (WSSV) in Vaname Shrimp (*Litopenaeus vannamei*). *J Fish Mar* 7 (1): 54-65.
- Jamaluddin J, Alfin A, Muzuni M, Yanti NA. 2018. Exploration of thermophilic bacteria producing L-asparaginase in Wawolesea Hot Springs, Southeast Sulawesi. *Biowallacea: J Biol Res* 5 (1): 1-9.
- Jiao L, Chi H, Lu Z, Zhang C, Chia SR, Show PL, Tao Y, Lu F. 2020. Characterization of a novel type I L-asparaginase from *Acinetobacter soli* and its ability to inhibit acrylamide formation in potato chips. *J Biosci Bioeng* 129 (6): 672-678. DOI: 10.1016/j.jbiosc.2020.01.007.
- Jordan JA, Lenski RE, Card KJ. 2022. Idiosyncratic fitness costs of ampicillin-resistant mutants derived from a long-term experiment with *Escherichia coli*. *Antibiotics* 11 (3): 347. DOI: 10.3390/antibiotics11030347.
- Kishore V, Nishita KP, Manonmani HK. 2015. Cloning, expression and characterization of L-asparaginase from *Pseudomonas fluorescens* for large scale production in *E. coli* BL21. *3 Biotech* 5: 975-981. DOI: 10.1007/s13205-015-0300-y.
- Kuroda D, Gray JJ. 2016. Shape complementarity and hydrogen bond preferences in protein-protein interfaces: implications for antibody modeling and protein-protein docking. *Bioinformatics* 32 (16): 2451-2456. DOI: 10.1093/bioinformatics/btw197.
- Kuwabara T, Prihanto AA, Wakayama M, Takagi K. 2015. Purification and characterization of *Pseudomonas aeruginosa* PAO1 Asparaginase. *Procedia Environ Sci* 28: 72-77. DOI:10.1016/j.proenv.2015.07.011.
- Loch JJ, Jaskolski M. 2021. Structural and biophysical aspects of L-asparaginases: A growing family with amazing diversity. *IUCrJ* 8 (4): 514-531. DOI: 10.1107/S2052525221006011.
- Lomelino CL, Andring JT, McKenna R, Kilberg MS. 2017. Asparagine synthetase: Function, structure, and role in disease. *J Biol Chem* 292 (49): 19952-19958. DOI: 10.1074/jbc.R117.819060.
- Maurya KG. 2019. Restriction Enzyme. *Encyclopedia of Animal Cognition and Behavior*. Springer Nature, Switzerland. DOI: 10.1007/978-3-319-47829-6_210-1.

- Mihooliya KN, Nandal J, Kumari A, Nanda S, Verma H, Sahoo DK. 2020. Studies on efficient production of a novel L-asparaginase by a newly isolated *Pseudomonas resinovorans* IGS-131 and its heterologous expression in *Escherichia coli*. 3 Biotech 10 (4): 148. DOI: 10.1007/s13205-020-2135-4.
- Mortier-Barrière I, Polard P, Campo N. 2020. Direct visualization of horizontal gene transfer by transformation in live *Pneumococcal* cells using microfluidics. Genes (Basel) 11 (6): 675. DOI: 10.3390/genes11060675.
- Mou Y, Huang PS, Hsu FC, Mayo SL. 2015. Computational design and experimental verification of a symmetric protein homodimer. PNAS 112 (34): 10714-10719. DOI: 10.1073/pnas.1505072112.
- Muzuni, Yanti NA, Prasetya WM. 2021. Characterization of the gene encoding chitinase enzyme from *Bacillus* isolates insulated from some locations in Southeast Sulawesi. J Phys Conf Ser 1899: 012017. DOI: 10.1088/1742-6596/1899/1/012017.
- Muzuni, Ningsih R, Yanti NA, Asniah. 2022a. Molecular identification of *Phytophthora* sp. from Indonesian cocoa using phylogenetic analysis. Pak J Biol Sci 25 (3): 245-253. DOI: 10.3923/pjbs.2022.245.253.
- Muzuni, Suriana, Yanti NA, Ardiansyah. 2022b. Phenotypic characterization and identification of potential L-sparaginase-producing thermophilic bacteria from Wawolesea Hot Spring, North Konawe, Southeast Sulawesi, Indonesia. Pak J Biol Sci 25 (11): 1021-1032. DOI: 10.3923/pjbs.2022.1021.1032.
- Nanni L, Lumini A, Brahmam S. 2014. An empirical study of different approaches for protein classification. Sci World J 2014: 1-17. DOI: 10.1155/2014/236717.
- Negi DS, Shrivastava P, Das SP. 2014. DNA sequencing by polymer synthesis with variable ratio of deoxynucleoside triphosphate and fluorescent dideoxynucleotide triphosphate. Asian J Biomed Pharm Sci 4 (32): 32-38. DOI: 10.15272/ajbps.v4i32.495.
- Pearson WR. 2013. An introduction to sequence similarity ("homology") searching. Curr Protoc Bioinformatics 42 (1): 3-1. DOI: 10.1002/0471250953.bi0301s42.
- Ren J, Zhenhua L. 2011. A PPY/PT for assaying L-asparaginase activity in *Proteus vulgaris* the construction and application of the piezoelectric transducer. Intl Conf New Technol Agric 2011: 845-848. DOI: 10.1109/ICAE.2011.5943923.
- Roberts MS, Nakamura LK, Cohan FM. 1996. *Bacillus vallismortis* sp. Nov., a close relative of *Bacillus subtilis*, isolated from soil in Death Valley, California. Intl J Syst Bacteriol 46 (2): 470-475. DOI: 10.1099/00207713-46-2-470.
- Rohit A, Maiti B, Shenoy S, Karunasagar I. 2016. Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) for rapid diagnosis of neonatal sepsis. Indian J Med Res 143 (1): 72-78. DOI: 10.4103/0971-5916.178613.
- Sahoo S, Shivani K, Padhy AA, Kumari V, Mishra P. 2023. Principles, methods, and applications of protein folding inside cells. In: Saudagar P, Tripathi T (eds.) Protein Folding Dynamics and Stability. Springer, Singapore. DOI: 10.1007/978-981-99-2079-2_13.
- Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual. 3rd Edition, Vol. 1. Cold Spring Harbor Laboratory Press, New York.
- Sengupta S, Biswas M, Gandhi K, Gota V, Sonawane A. 2021. Pre-clinical evaluation of novel L-asparaginase mutants for the treatment of acute lymphoblastic leukemia. Blood 138 (Supplement 1): 4919. DOI: 10.1182/blood-2021-150857.
- Shakambari G, Ashokkumar B, Varalakshmi P. 2019. L-Asparaginase: a promising biocatalyst for industrial and clinical applications. Biocatal Agric Biotechnol 17: 213-224. DOI: 10.1016/j.bcab.2018.11.018.
- Shamir R. 2001. Lecture 8: Algorithms for molecular biology, Tel Aviv University, Israel.
- Sharafi Z, Barati M, Khoshayand MR, Adrangi S. 2017. Screening for type II L-asparaginases: Lessons from the genus *Halomonas*. Iran J Pharm Res 16 (4): 1565-1573.
- Sune D, Rydberg H, Augustinsson ÅN, Serrander L, Jungeström MB. 2020. Optimization of 16S rRNA gene analysis for use in the diagnostic clinical microbiology service. J Microbiol Methods 170: 105854. DOI: 10.1016/j.mimet.2020.105854.
- Van Rossum T, Ferretti P, Maistrenko OM, Bork P. 2020. Diversity within species: interpreting strains in microbiomes. Nat Rev Microbiol 18 (9): 491-506. DOI: 10.1038/s41579-020-0368-1.
- Van Trimpont M, Peeters E, De Visser Y, Schalk AM, Mondelaers V, De Moerloose B, Lavie A, Lammens T, Goossens S, Van Vlierberghe P. 2022. Novel insights on the use of L-asparaginase as an efficient and safe anti-cancer therapy. Cancers 14 (4): 902. DOI: 10.3390/cancers14040902.
- Vasudevan D, Vaidyanathan K. 2017. Textbook of Biochemistry for Medical Students, Chapter-04. Proteins: Structure and Function, Amrita Institute of Medical Science and Research Center. DOI: 10.5005/jp/books/13014_5.
- Vimal A, Kumar A. 2022. L-asparaginase: Need for an expedition from an enzymatic molecule to antimicrobial drug. Intl J Pept Res Ther 28 (1): 9. DOI: 10.1007/s10989-021-10312-x.
- Wahyuni FD, Seprianto. 2018. *Candida antarctica* lipase B synthetic gene: A bioinformatics analysis. Bioscience 2 (2): 20-29. DOI: 10.24036/0201822100216-0-00.
- Weller C, Wu M. 2015. A generation-time effect on the rate of molecular evolution in bacteria. Evolution 69 (3): 643-652. DOI: 10.1111/evo.12597.
- Wheeler D, Bhagwat M. 2007. BLAST Quickstart: Example-driven web-based BLAST tutorial. Methods Mol Biol 395: 149-176. DOI: 10.1007/978-1-59745-514-5_9.
- Zhu C, Gao Y, Li H, Meng S, Li L, Francisco JS, Zeng XC. 2016. Characterizing hydrophobicity of amino acid side chains in a protein environment via measuring contact angle of a water nanodroplet on planar peptide network. Proc Natl Acad Sci 113 (46): 12946-12951. DOI: 10.1073/pnas.1616138113.