International Journal of Applied Chemical and Biological Sciences (2024), 5(3), 1-12
ISSN- 2582-788X

RESEARCH PAPER

Article Identifier: https://identifier.visnav.in/1.0001/ijacbs-24g-15001/

Phylogenetic Analysis and Investigation of the Consensus Sequence of Alpha- Amylase-Producing *Bacillus* Species for Potential Amylase with Enhanced Activity

Jibril Mohammed [1]*, Emmanuel Oluwakorede Opadokun [2], Ayokanmi Joseph Aremu [3], Judah Opeyemi Ajibowo [1]

[1] Department of Microbiology, Faculty of Pure and Applied Sciences, Kwara State University, Malete, PMB 1530, Ilorin, Kwara State, Nigeria.
[2] Program in Biotechnology, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand.
[3] Program in Bioinformatics and Computational Biology, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand.

* For correspondence: jibrilmohammed2001@gmail.com

Received on: 15 July 2024 *Published on: 3 September 2024*

ABSTRACT

Alpha-amylase enzymes from *Bacillus* species are integral to various industrial processes due to their efficient starch hydrolysis capabilities. This study aimed to conduct a phylogenetic analysis to understand the evolutionary relationships between alpha-amylase-producing *Bacillus* species and investigate the consensus sequence derived from the multiple sequence alignment for its potential for alpha-amylase with enhanced activity. Sixteen (16) alpha-amylase protein sequences from *Bacillus* species were retrieved from the NCBI database and imported to Expasy server to calculate the physicochemical features. The Geneious Prime software was employed to perform multiple sequence alignment and construct a phylogenetic tree. The consensus sequence obtained from the MSA was analysed by determining its physicochemical features, predicting the functional domains, and predicting the 3D secondary structure. The physicochemical features of the sixteen sequences revealed a notable range of features, and the multiple sequence alignment revealed several conserved regions with consecutive identical residues, including highly conserved region sequence such as NGTLMQYFEW, conserved region sequences such as ENG, NHD, and less conserved region sequences such as DG, PL, LA, and YG. The consensus sequence analysis revealed physicochemical features, including 534 amino acids, molecular weight of 60,553.30 Da, theoretical pI of 5.47, negative residues (64), positive residues (49), instability index (24.92), aliphatic index (70.30), GRAVY (-0.537), and extinction coefficients (156650). The analysis also revealed

functional domains like alpha-amylase and glycosyl hydrolase, and a 3D secondary structure prediction showed a stable conformation suitable for industrial applications. This study enhanced our understanding of the evolutionary and functional relationships among alpha-amylase- producing Bacillus species and the potential of utilising consensus sequences in enzyme engineering.

Keywords: Alpha-amylase, *Bacillus* species, Consensus sequence, Phylogenetic analysis, Protein

1. INTRODUCTION

Alpha-amylase is a calcium-dependent metalloenzyme that breaks down starch molecules by randomly cleaving α-1,4 glycosidic bonds, producing maltotriose and maltose from amylose and maltose, glucose, and "limit dextrin" from amylopectin [1]. Alpha-amylase is a vital enzyme in various industries, including food, detergent, and biofuel production. Alpha-amylase can be derived from various sources, such as plants, animals, and microorganisms. However, microorganisms, particularly bacteria, are widely used for enzyme production due to their high productivity and ease of manipulation [2]. Among bacteria, *Bacillus* species have been reported to possess the highest amylolytic ability [3]. The most commonly utilised thermostable amylase in the starch industries is produced by *Bacillus* species such as *Bacillus subtilis, Bacillus mesentericus, Bacillus cereus, Bacillus polmyxa, Bacillus amyloliquefaciens, Bacillus coagulans, Bacillus licheniformis, Bacillus steriothermophilus, and Bacillus megaterium* [4-5].

Despite the extensive study of alpha-amylase, there is still a pressing need for enzymes with improved activity to optimise industrial processes. Enhanced enzyme activity can lead to increased efficiency and reduced costs in various applications. To address this need, bioinformatics approaches can be employed to identify potential alpha-amylases with enhanced activity. Phylogenetic analysis, a key bioinformatics tool that enables the study of evolutionary relationships between organisms and their protein, providing insights into the conservation and variation of enzyme sequences, can be used to uncover patterns of evolutionary adaptation and identify potential hotspots for enhanced enzyme activity.

Furthermore, the consensus sequence, derived from the multiple sequence alignment of related enzymes, represents a hypothetical ancestor sequence that can reveal conserved motifs and functional residues crucial for enzyme activity. By investigating the consensus sequence of alpha-amylase-producing *Bacillus* species for potential alpha-amylase with enhanced activity, we can identify potential alpha-amylases with enhanced activity, laying the foundation for further experimental validation. This research aims to conduct a phylogenetic analysis to understand the evolutionary relationships between the alpha-amylase-producing *Bacillus* species and investigate the consensus sequence derived from the multiple sequence alignment for its potential for alpha-amylase with enhanced activity. By doing so, we seek to contribute to the development of more efficient industrial processes, increased productivity, and reduced costs.

2. METHODOLOGY

2.1. Tools

The National Centre for Biotechnology Information (NCBI) database was used for retrieving alpha-amylase protein sequences. The Expasy (ProtParam) server was utilised for physicochemical property analysis. Geneious Prime software (version 2020) was employed for

protein sequence alignment and phylogenetic tree construction. InterProscan was used to identify the domains within the consensus sequence, and SWISS-MODEL was used to predict the secondary structure.

2.2. Retrieval of Alpha-Amylase Protein Sequences from the Database

Alpha-amylase protein sequences from 16 different *Bacillus* species were retrieved from the NCBI database using the keywords "alpha-amylase" and "*Bacillus*". Sequences were obtained in FASTA format, and those with close range length (511–516) of amino acids were selected to ensure improved alignment quality, reduced bias, and enhanced identification of conserved regions and functional motifs. Sequence identity and functionality were verified using NCBI annotations.

2.3. Characterization of Physicochemical Properties of Proteins

The retrieved protein sequences were submitted to the Expasy server to calculate physicochemical properties, including molecular weight, number of amino acids, theoretical pI, number of negatively charged residues, number of positively charged residues, extinction coefficients, instability index, aliphatic index, and grand average of hydropathicity (GRAVY). The results were recorded and analysed to identify potential functional implications, providing insights into the protein's structure and function.

2.4. Multiple Sequence Alignment

Protein sequences were imported into Geneious Prime software, a comprehensive platform for sequence analysis. The Geneious aligner tool was employed to generate a high-quality alignment using default parameters (alignment type: global alignment with free end gaps, cost matrix: BLOSUM62, gap open penalty: 12, gap extension penalty: 3, refinement iterations: 2). The alignment was then visualised using Geneious Prime's document viewer, allowing for a detailed examination of consensus sequence, conserved regions, and variation.

2.5. Phylogenetic Tree Construction

The aligned protein sequences were used to construct a phylogenetic tree to infer evolutionary relationships among the alpha-amylase-producing *Bacillus* species. The Geneious Tree Builder tool was selected for tree construction, using Jukes-Cantor as the genetic distance model and the neighbour-joining method as the algorithm, with a boot rap value of 1000 and a replicate value of 100 to assess the robustness of the tree [6]. The resulting phylogenetic tree was visualised and interpreted using Geneious Prime's document viewer, providing insights into the evolutionary history and relationships among the alpha-amylase-producing *Bacillus* species. The tree was examined for well-supported nodes, branch lengths, and clustering patterns, shedding light on the diversification and adaptation of these enzymes across different *Bacillus* species.

2.6. Consensus Sequence Investigation for Potential Amylase with Enhanced Activity

The consensus sequence derived from the multiple sequence alignment was manually edited to remove all ambiguous residues (X and Z values). The modified consensus sequence was then analysed using Expasy to determine its physicochemical properties, including molecular weight, number of amino acids, theoretical pI, number of negatively charged residues, number of positively charged residues, extinction coefficients, instability index, aliphatic index, and grand average of hydropathicity (GRAVY). Additionally, InterProScan within the Geneious Prime was employed to identify

and annotate potential domains, and the secondary structure of the modified consensus sequence was predicted using SWISS-MODEL, and the 3D structure was visualized. The results of these analyses were then examined to determine the potential for enhanced alpha-amylase activity.

3. RESULTS AND DISCUSSION

3.1. Physicochemical Characterization

The physicochemical properties of the studied *Bacillus* species, as presented in Table 1, exhibit notable diversity across several parameters. The molecular weights range from 58,060.80 to 60,088.46 Da, and

Table 1. Physicochemical Features of Alpha-Amylase Protein Sequences Observed in *Bacillus* Species

S/N	Organism	Accession No.	No. of amino acids	Molecular Weight (Da)	Theoretical pI	No. of Negatively charged residue	No. of Positively charged residue	Instability index	Aliphatic index	Grand average of hydropathicity (GRAVY)	Extinction coefficients
1	*Bacillus albus*	WTM77626.1	513	58120.26	5.43	61	45	23.35	68.97	-0.602	151150
2	*Bacillus amyloliquefaciens*	QGW08575.1	514	58402.92	5.74	67	56	33.24	66.40	-0.635	136710
3	*Bacillus anthracis*	XAS94459.1	513	58445.58	5.38	63	46	21.39	67.45	-0.638	151150
4	*Bacillus cereus*	RCL15329.1	513	58307.47	5.34	63	47	22.26	66.88	-0.619	152640
5	*Bacillus fungorum*	WP_099685063.1	513	58110.22	5.37	62	45	24.14	69.16	-0.591	151150
6	*Bacillus halotolerans*	WAT23122.1	516	58822.24	5.89	62	49	37.41	64.44	-0.658	152180
7	*Bacillus licheniformis*	ACT63870.1	512	58411.01	6.28	61	53	28.67	68.79	-0.599	141180
8	*Bacillus mojavensis*	QQF63604.1	516	58828.30	5.95	63	51	38.36	65.00	-0.657	150690
9	*Bacillus sonorensis*	WPP37429.1	511	58060.80	5.82	61	50	30.71	69.82	-0.518	142795
10	*Bacillus thuringiensis*	ALQ69061.1	513	58358.44	5.32	63	46	22.72	67.84	-0.637	151150
11	*Bacillus tropicus*	UOK44626.1	513	58253.38	5.37	62	62	22.99	66.51	-0.622	152640
12	*Bacillus velezensis*	WJN56765.1	513	58267.67	5.69	67	56	32.94	65.01	-0.661	136710
13	*Bacillus wiedmannii*	SDC57199.1	513	58111.25	5.37	64	46	20.06	69.34	-0.608	151150
14	*Cytobacillus firmus*	USK40237.1	514	60088.46	4.72	91	58	30.48	78.89	-0.570	124220
15	*Cytobacillus oceanisediminis*	USK45710.1	512	60001.63	4.86	92	62	30.77	75.80	-0.597	122730
16	*Geobacillus stearothermophilus*	QOR84891.1	511	59916.46	5.76	76	67	28.40	83.93	-0.468	121700

Table 2. Conserved Regions of Alpha-Amylase Alignment with Consecutive Identical Residues

Region	Sequence	Length	Category
36 – 45	NGTLMQYFEW	10	Highly Conserved
226 – 228	ENG	3	Conserved
366 – 368	NHD	3	Conserved
263 – 264	DG	2	Less Conserved
324 – 325	PL	2	Less Conserved
389 – 390	LA	2	Less Conserved
406 - 408	YG	2	Less Conserved

sequence lengths span from 511 to 516 amino acids.

The isoelectric points (pI) vary between 4.72 and 6.28, indicating differences in the proteins' net charge at physiological pH. The negative residue counts range from 61 to 92, while positive residues vary from 45 to 67. The instability index values (20.06–38.36) suggest that all the proteins are stable. Additionally, the aliphatic index ranges from 64.44 to 83.93, reflecting variations in the content of aliphatic amino acids, which can influence protein thermostability. This diversity in physicochemical properties suggests that each *Bacillus* species has adapted to distinct environmental niches, leading to variations in their protein properties.

The range of molecular weights may indicate differences in protein structure and potential functional roles. For instance, proteins with higher molecular weights might have additional domains or subunits that contribute to specific functions. The variability in isoelectric points suggests that the proteins may exhibit different substrate specificities or binding properties. Proteins with lower pI values are more acidic and might interact differently with substrates compared to those with higher pI values. This can affect enzyme activity and stability under various pH conditions, which is critical for industrial applications where pH stability is a key factor. The instability index values, all below 40, indicate that these proteins are generally stable, which is advantageous for industrial processes requiring prolonged enzyme activity. The aliphatic index, indicative of the volume occupied by aliphatic side chains, further supports the potential stability of these

Table 3. Physicochemical Features Observed in Modified Consensus Sequence

Properties	Values
No. of amino acids	534
Molecular Weight	60553.30
Theoretical pI	5.47
No. of Negatively charged residue	64
No. of Positively charged residue	49
Instability index	24.92
Aliphatic index	70.30
Grand average of hydropathicity (GRAVY)	-0.537
Extinction coefficients	156650

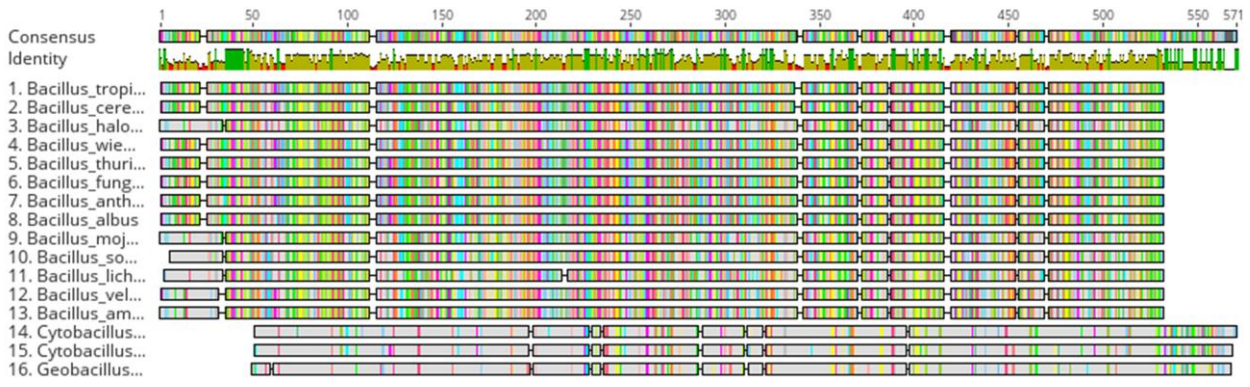

Figure 1. Multiple Sequence Alignment of Alpha-Amylase Sequences

proteins, especially at higher temperatures. The variation in physicochemical properties underscores the importance of considering these factors when exploring the functional capabilities and potential industrial applications of alpha-amylases from different *Bacillus* species. Understanding these properties can guide the selection and engineering of enzymes for specific industrial processes, aiming to enhance activity, stability, and substrate specificity.

3.2. Sequence Analysis

The multiple sequence alignment (Figure 1) reveals a fascinating pattern of conserved residues and regions, along with a consensus sequence (Figure 2), despite the notable variation in amino acid sequences among the different *Bacillus* species. While the alignment exhibits

Figure 2. Protein Alignment Consensus Sequence

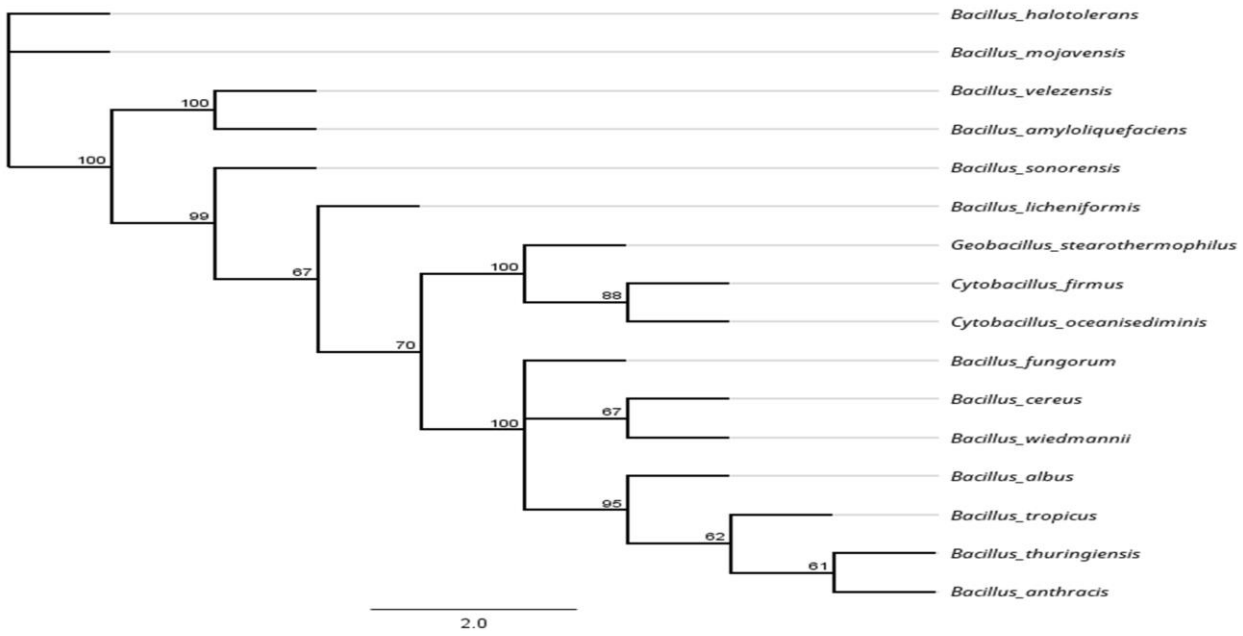

Figure 3. Phylogenetic Tree of Alpha-Amylase Producing *Bacillus* Species

some gaps, indicating the potential loss of ancestral amino acids during evolution, the conserved regions with consecutive identical residues (Table 2) suggest that functional sites essential for amylase activity have been preserved. The highly conserved region from positions 36 to 45 (NGTLMQYFEW) likely represents a critical segment of the enzyme, possibly involved in substrate binding or catalysis. The presence of such a conserved sequence across different species underscores its fundamental role in enzyme function and stability. Additionally, shorter conserved regions such as ENG (226-228), NHD (366-368), DG (263-264), PL (324-325), LA (389-390), and YG (406-408) indicate other functional or structural elements essential for the enzyme's activity.

These shorter motifs may play roles in maintaining the enzyme's three-dimensional structure, facilitating proper folding, or interacting with substrates or cofactors. The fact that these conserved sequences are maintained across diverse *Bacillus* species suggests strong evolutionary pressure to preserve these functional elements, despite variations in other parts of the protein. This conservation implies that these regions are critical for the enzyme's catalytic efficiency and overall stability. The less conserved regions, though showing some degree of conservation, may indicate areas where evolutionary adaptations have occurred, possibly to optimise the enzyme for different environmental conditions or substrates. These regions

	Bacillus_tro...	Bacillus_ce...	Bacillus_hal...	Bacillus_wi...	Bacillus_th...	Bacillus_fu...	Bacillus_an...	Bacillus_alb...	Bacillus_m...	Bacillus_so...	Bacillus_lic...	Bacillus_vel...	Bacillus_a...	Cytobacillu...	Cytobacillu...	Geobacillus...
Bacillus_tropicus		94.932%	69.246%	93.774%	97.276%	94.163%	96.693%	97.082%	70.019%	69.201%	69.574%	68.085%	68.085%	15.353%	15.975%	14.700%
Bacillus_cereus	94.932%		69.439%	95.914%	94.747%	94.553%	94.163%	93.774%	70.213%	69.006%	69.961%	68.472%	68.472%	15.560%	15.768%	14.700%
Bacillus_halotolerans	69.246%	69.439%		68.605%	69.186%	68.992%	68.217%	69.186%	98.643%	83.953%	79.377%	83.689%	83.915%	15.768%	15.768%	14.286%
Bacillus_wiedmannii	93.774%	95.914%	68.605%		94.152%	95.127%	94.737%	94.737%	69.380%	68.359%	69.126%	68.217%	68.217%	15.975%	15.975%	14.907%
Bacillus_thuringiensis	97.276%	94.747%	69.186%	94.152%		94.932%	97.466%	96.881%	69.961%	68.945%	69.320%	68.023%	68.023%	15.353%	16.183%	14.493%
Bacillus_fungorum	94.163%	94.553%	68.992%	95.127%	94.932%		94.542%	95.322%	69.767%	69.141%	69.709%	67.442%	67.442%	15.560%	16.183%	14.700%
Bacillus_anthracis	96.693%	94.163%	68.217%	94.737%	97.466%	94.542%		96.881%	68.992%	68.555%	68.738%	67.248%	67.248%	15.353%	15.975%	14.493%
Bacillus_albus	97.082%	93.774%	69.186%	94.737%	96.881%	95.322%	96.881%		69.961%	69.336%	69.126%	67.829%	67.829%	15.353%	15.975%	14.907%
Bacillus_mojavensis	70.019%	70.213%	98.643%	69.380%	69.961%	69.767%	68.992%	69.961%		84.736%	79.767%	84.272%	84.496%	15.768%	15.975%	14.493%
Bacillus_sonorensis	69.201%	69.006%	83.953%	68.359%	68.945%	69.141%	68.555%	69.336%	84.736%		80.431%	80.235%	80.431%	16.805%	16.805%	15.735%
Bacillus_licheniformis	69.574%	69.961%	79.377%	69.126%	69.320%	69.709%	68.738%	69.126%	79.767%	80.431%		76.459%	76.654%	15.353%	15.353%	15.321%
Bacillus_velezensis	68.085%	68.472%	83.689%	68.217%	68.023%	67.442%	67.248%	67.829%	84.272%	80.235%	76.459%		99.415%	14.938%	15.145%	14.493%
Bacillus_amyloliquefaci...	68.085%	68.472%	83.915%	68.217%	68.023%	67.442%	67.248%	67.829%	84.496%	80.431%	76.654%	99.415%		14.938%	14.938%	14.286%
Cytobacillus_firmus	15.353%	15.560%	15.768%	15.975%	15.353%	15.560%	15.353%	15.353%	15.768%	16.805%	15.353%	14.938%	14.938%		79.688%	62.622%
Cytobacillus_oceanisedi...	15.975%	15.768%	15.768%	15.975%	16.183%	16.183%	15.975%	15.975%	15.975%	16.805%	15.353%	15.145%	14.938%	79.688%		63.601%
Geobacillus_stearother...	14.700%	14.700%	14.286%	14.907%	14.493%	14.700%	14.493%	14.907%	14.493%	15.735%	15.321%	14.493%	14.286%	62.622%	63.601%	

Figure 4. Distance Matrix Score Heatmap Showing the Sequence Similarity Between the Studied *Bacillus* Species

```
MMFKRITIVG LSVVLFLPSI YGGSKYADTV NNGTLMQYFE WYAPNDGNHW RLQNDAENLA  60
QKGITVWIPP AYKGTTQNDV GYGAYDLYDL GEFNQKGTVR TKYGTKAQLK SAIEALHKQN 120
IDVYGDVVLN HKGGADYTET VTAVEVDPNN RNEVSGDYEI KAWTGFNFPG RGDTYSDFKW 180
KWYHFDGTDW DEGRKLNRIY KFRGIGKAWD WEVSSENGNY DYLMYADLDF DHPDVANEMK 240
NWGTWYANEL NLDGFRLDAV KHIHSFLRDW VNHVRQQTGK EMFTVAEYWQ NDIGTLNNYL 300
AKTNYNQSVF DAPLHYNFHY ASGNGYDMRN ILNGTVVSNH PALAVTFVEN HDTQPGQSLE 360
SVSPWFKPLA YAFILTRAEG YPSVFYGDYY GTKGNSSEIP ALKDKIEPIL TARKNFAYGT 420
QRDYLDHPDV IGWTREGDSV HANSGLATLI SDGPGGSKWM DVGKNNAGEV WDITGNQTDT 480
VTINKDGWGF HVSGGSVSIY VQQEKTGLNI PIAAIAAVYA FMIFIYLWKR SKKE       534
```

Figure 5. Modified Consensus Sequence

might be involved in fine-tuning the enzyme's activity, stability, or interaction with other molecules.

3.3. Evolutionary Analysis

The phylogenetic tree (Figure 3) constructed from the aligned sequences reveals a clear clustering pattern, with the *Bacillus* species grouping into distinct clades that correlate with their sequence similarity and evolutionary relationships. The distance matrix score heatmap (Figure 4) provides a visual representation of the closeness of the study organisms, with higher percentages indicating closer relationships. This phylogenetic tree offers valuable insights into the

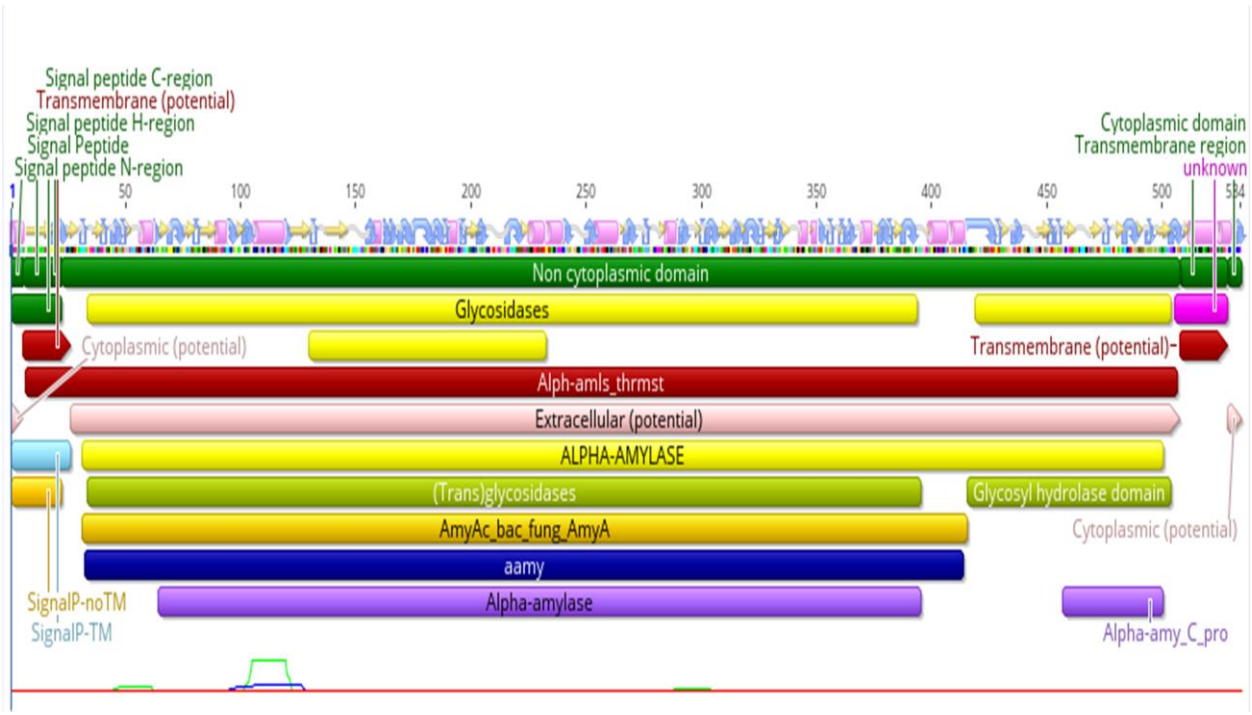

Figure 6. Domain Prediction from Modified Consensus Sequence of Alpha-Amylase

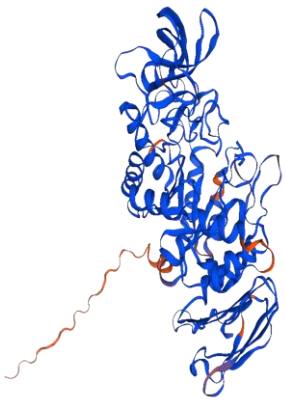

Figure 7. Secondary Structure Prediction of Modified Consensus Sequence of Alpha Amylase

evolutionary history of these *Bacillus* species, suggesting that they share a common ancestor but have since diverged and evolved distinct characteristics. The tree topology reveals that species with higher sequence identities are more closely related, indicating a recent common ancestor. The sequence variations observed among the species may be attributed to genetic drift, gene duplication, or horizontal gene transfer events. The conserved regions identified in the sequence analysis may represent ancestral traits maintained across the genus *Bacillus*, while the variations could have evolved to accommodate specific functional requirements. For instance, *Bacillus velezensis* and *Bacillus amyloliquefaciens* form a well-supported clade with a bootstrap value of 100 and an identity score of 99.41%, suggesting these species share a recent common ancestor. This close relationship may reflect similar functional characteristics in their alpha-amylase genes, which could be advantageous for industrial applications requiring consistent enzyme performance. Species like *Bacillus mojavensis* and *Bacillus halotolerans* occupy separate branches, indicating significant evolutionary divergence. This divergence could suggest adaptation to different environmental conditions or functional specialisations in their alpha-amylase production, which might be exploited for specific industrial processes. Certain species, such as *Bacillus cereus* and *Bacillus wiedmannii*, are closely related with a bootstrap value of 67, indicating a more recent divergence. This close relationship implies potential similarities in their alpha-amylase production capabilities, making them suitable candidates for comparative studies on enzyme efficiency and application. *Bacillus thuringiensis* and *Bacillus anthracis* exhibit very high identity scores of 97.47%, indicating highly conserved alpha-amylase sequences. This conservation suggests these enzymes may perform similar functions and could be interchangeably used in industrial applications, offering consistency and reliability in enzyme performance. The moderately high identity scores (around 80-85%) between species like *Bacillus mojavensis* and *Bacillus velezensis* indicate functional similarities while also hinting at potential variations. These variations can be exploited for specific biotechnological processes where slight differences in enzyme activity or stability may be advantageous. The tree also places Cyto*Bacillus firmus* and Cyto*Bacillus oceanisediminis* in the same clade with high bootstrap values, emphasizing their similar evolutionary paths and the potential novelty of their alpha-amylase enzymes. These species, together with *GeoBacillus stearothermophilus*, display lower identity scores (14–17%) with other *Bacillus* species, reflecting greater evolutionary divergence and potentially different functional adaptations of their alpha-amylase enzymes. The phylogenetic tree not only infers the evolutionary relationship within the *Bacillus* genus but also highlights the diverse functional potential of alpha-amylase enzymes across different species. This diversity in evolutionary paths and enzyme characteristics suggests that different *Bacillus* species could be targeted for specific industrial applications, depending on the desired enzyme properties.

3.4. Consensus Sequence Analysis

The physicochemical features of the modified consensus sequence are presented in Table 3. The modified consensus sequence is composed of 534 amino acids with a molecular weight of 60,553.30 Da and a theoretical isoelectric point (pI) of 5.47. The sequence has 64 negatively charged residues and 49 positively charged residues, which may influence the enzyme's interaction with substrates and stability under various pH conditions. The instability index of 24.92 indicates that the protein is stable. The aliphatic index of 70.30

suggests moderate thermal stability, which is beneficial for industrial applications. The grand average of hydropathicity (GRAVY) is -0.537, indicating that the protein is hydrophilic and likely soluble in aqueous environments, facilitating its use in various industrial processes. The high extinction coefficient (156650) is indicative of the protein's strong absorbance at 280 nm, which can be useful for protein quantification. The domain analysis (Figure 6) reveals the presence of several critical functional regions. The alpha-amylase domain is prominent, confirming the protein's enzyme role. The presence of glycosyl hydrolase domains indicates the enzyme's capability to break down glycosidic bonds, a key feature for amylase activity. Additionally, the extracellular domain suggests that the enzyme can function outside the bacterial cell, which is advantageous for industrial enzyme applications. The presence of both cytoplasmic and transmembrane regions indicates that the enzyme might interact with cellular membranes, potentially aiding in substrate binding and processing. The 3D structure (Figure 7) provides further insights into the enzyme's functionality. The structure is characterised by a well-defined catalytic domain with several alpha-helices and beta-sheets, typical of alpha-amylase enzymes. The catalytic site is likely located in the groove formed by these secondary structural elements, allowing for effective substrate binding and catalysis. These structural features suggest that the consensus sequence is well-equipped for efficient starch breakdown, which is essential for industrial applications. The combination of physicochemical properties, domain architecture, and structural characteristics suggests that the modified consensus sequence has several potential enhancements. In conclusion, the consensus sequence of alpha-amylase from *Bacillus* species indicates a robust potential for enhanced activity. The sequence's physicochemical stability, domain structure, and 3D conformation suggest that it is well-suited for industrial applications requiring efficient starch hydrolysis. Further experimental validation, including activity assays under various conditions, would be essential to confirm these predictions and fully harness the enzyme's capabilities for industrial use.

4. CONCLUSION

The phylogenetic analysis revealed distinct genetic clusters among alpha-amylase-producing *Bacillus* species, highlighting both evolutionary divergence and shared ancestry. Conserved regions identified through multiple sequence alignment underscored the preservation of enzymatic efficiency across species, providing insights into potential targets for industrial optimization. The consensus sequence analysis suggested its potential for enhanced alpha-amylase activity, indicating suitability for industrial applications. These findings lay the groundwork for further research aimed at optimising enzyme performance through biotechnological interventions.

5. RECOMMENDATION

To fully harness the enzymatic capabilities of the consensus sequence, further experimental validation, including activity assays under various conditions, is essential. Additionally, comparative studies on other microbial alpha-amylases could contribute to the identification of enzymes with unique performance characteristics. Such research could provide valuable insights into the development of more efficient industrial processes and enhanced enzyme applications.

6. ACKNOWLEDGEMENT

NA

7. CONFLICT OF INTEREST

The authors have declared that there is no conflict of interest.

8. SOURCE/S OF FUNDING

NA

9. REFERENCES

1. Tiwari SP, Srivastava R, Singh CS, Shukla K, Singh RK, Singh P, et al. (2015), Amylases: An Overview with Special Reference to Alpha Amylase, Journal of Global Biosciences; 4(1): 1886-1901.
2. Bhattacharjee I, Mazumdar D, Saha, PS (2019), Microbial Amylases and their Potential Application in Industries: A Review, The Pharma Innovation Journal; 8(6): 162-170.
3. Adedayo MR, Mohammed J (2024), Isolation, Identification, and Screening of Bacteria from Yam Waste for Amylase Production Ability, International Journal of Applied Chemical and Biological Sciences, 5(1), 12-21.
4. Ekka A, Namdeo N (2018), Screening, Isolation and Characterization of Amylase Producing Bacteria and Optimization for Production of Amylase, IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB); 4(2): 50-56.
5. Padhiar AR, Kommu S (2016), Isolation, Characterization and Optimization of Bacteria Producing Amylase, International Journal of Advanced Research in Biological Sciences. 3(7): 1-7.
6. Saha GC, Islam R, Olawale OA, Khan HI, Agboola OO, Pereira F, et al. (2023), Phylogenetic Analysis of Blast Resistance RMG7 (Resistance to Magnaporthe Grisea 7) Gene in Cereals (Wheat and Rice), International Journal on Recent and Innovation Trends in Computing and Communication, 11(10): 2358-2370.