WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Bioinformatics Analysis of α-Amylase Three-Dimensional Structure in *Aspergillus oryzae*

***Harshita Budhadev*[1*]**

[1]Department of Bioscience and Biotechnology, Banasthali University, Rajasthan India
Email: harshita.budhadev0809@gmail.com
DOI: https://doi.org/10.55248/gengpi.4.923.52600

ABSTRACT:

The primary objective of this study is to provide insight into bimolecular structures from a bioinformatics standpoint and offer information about the relevant bioinformatics software tools in this domain. It is well-established that each protein possesses a distinct chemical or structural function, underscoring the uniqueness of their three-dimensional structures. Within the scope of this research, we conducted an investigation into α-amylases in *Aspergillus oryzae*. Among enzymes responsible for starch conversion, α-amylases, specifically classified as endo-1, 4-α-D-glucan glucanohydrolase [E.C.3.2.1.1], hold particular significance and have been extensively studied. Our research relied on the utilization of the PDB and NCBI databases, along with software tools such as Chimera, Predict Protein, and Multalign. Through the application of these software tools, we performed a range of analyses, encompassing the determination of residue composition, identification of secondary structures, detection of conserved regions, and characterization of ligand binding sites. Starch-hydrolyzing enzymes, including amylases, pullulanases, and glucoamylases, play pivotal roles in various industries such as food, chemicals, and pharmaceuticals. In specific applications like the detergent and bakery industries, meeting the requisite properties of α-amylase presents a considerable challenge. The utilization of bioinformatics software tools has significantly enhanced our comprehension of protein three-dimensional structures and their functional attributes.

Keywords: Amylase; 3-dimensional Structures; Starch Conversion; Bioinformatics; Enzymes.

Introduction:

The most well-known amylolytic enzymes include α-amylase (EC 3.2.1.1), β-amylase (EC 3.2.1.2), and glucoamylase (EC 3.2.1.3), which exhibit significant distinctions from one another (Xia et al. 2021). These differences extend beyond their primary and tertiary structures, encompassing their catalytic mechanisms and reaction pathways. Consequently, these enzymes have been categorized into distinct glycoside hydrolase (GH) families: GH13 for α-amylases, GH14 for β-amylases, and GH15 for glucoamylases (Lovell et al. 2003).

Among the enzymes responsible for starch conversion, α-amylases (endo-1, 4-α-D-glucan glucanohydrolase [E.C.3.2.1.1]) hold particular significance and have been the subject of extensive study (Ly et al. 1999). Starch-hydrolyzing enzymes, including amylases, pullulanases, and glucoamylases, play crucial roles in the food, chemical, and pharmaceutical industries. Notably, in specific applications such as the detergent and bakery industries, the desired properties of α-amylase present notable challenges (Patil et al. 2021).

Fungal amylases have been instrumental in the production of starches, starch derivatives, and starch saccharification products (Tateno et al. 2007). The primary fungi employed for α-amylase production include *Aspergillus oryzae*, *A. niger*, and *Rhizopus oryzae*. Figure 1 illustrates the three-dimensional structures of amylases.

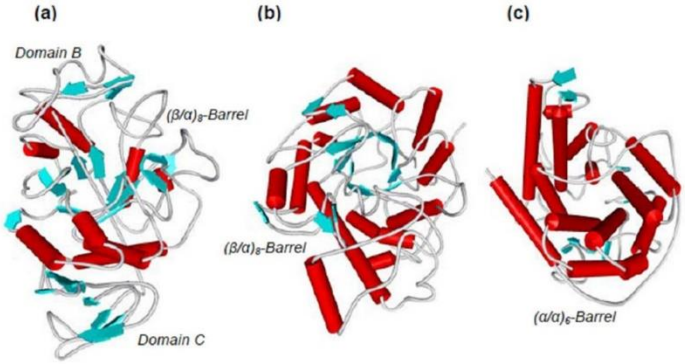

Figure 1. Three-Dimensional Structures of Amylases (a) Glycoside Hydrolase 13 (GH13) α-amylase from *Aspergillus oryzae* (PDB code: 2TAA). (b) Glycoside Hydrolase 14 (GH14) β-amylase from soybean (PDB code: 1BYA). (c) Glycoside Hydrolase 15 (GH15) glucoamylase from *Aspergillus awamori*. These images provide visual representations of the three-dimensional structures of these amylases, highlighting their structural diversity within the enzyme family.

Materials and Methods:

In this study, we conducted a comprehensive analysis of the three-dimensional structure of alpha-amylase enzymes derived from *Aspergillus oryzae*, exploring various facets of their molecular properties. The research commenced with the retrieval of the amino acid sequences corresponding to *Aspergillus oryzae* alpha-amylase, also known as Taka-Amylase A (PDB ID: 2TAA), from the Protein Data Bank website (Long et al. 1987).

To augment our investigation, we gathered amino acid sequences of this enzyme from multiple strains available in the NCBI databases. These sequences were subsequently aligned using the Multalign software tool to discern any variations or conserved regions among them (Gouet et al. 2003).

The three-dimensional structural analysis of the enzyme was performed with the assistance of Chimera UCSF software, allowing us to visualize and examine its spatial arrangement in detail. Additionally, Predict Protein software was instrumental in predicting the amino acid composition of the enzyme and elucidating its secondary structure characteristics.

Through these analytical steps, we aimed to gain a comprehensive understanding of the alpha-amylase enzymes in *Aspergillus oryzae*, shedding light on their sequence variations, structural attributes, and potential functional properties.

Results and Discussion:

Protein Structure Analysis:

Taka-Amylase A (EC=3.2.1.1) is composed of three distinct chains designated as A, B, and C. Each of these chains consists of 479 amino acids within its monomeric structure, inclusive of a signal domain. The molecular weight of Taka-Amylase A is measured at 54,810 Da.

A detailed amino acid sequence analysis for each of the three chains of Taka-Amylase A can be found in Table 1, sourced from www.uniprot.org.

Furthermore, the identification of conserved regions within the enzyme's sequence was accomplished through the utilization of Multalign software. A graphical representation of these conserved regions is provided in Figure 2, offering valuable insights into the structural and functional attributes of the enzyme.

Table 1. Analysis of each chain of the Amino acid sequence of Taka-amylase (www.uniprot.org).

Serial No.	Feature Key	Position(s)	Length	Description
1	Single peptide	1 – 21	29	
2	Chain	21 – 499	479	Alpha-amylase Atype-1/2
3	Active Site	227	1	Nucleophile
4	Active Site	251	1	Proton Donor
5	Binding Site	56, 104, 143 225, 255, 317 365	1,1,1 1,1,1 1	Substrate
6	Site	318	1	Transition State Stabilizer

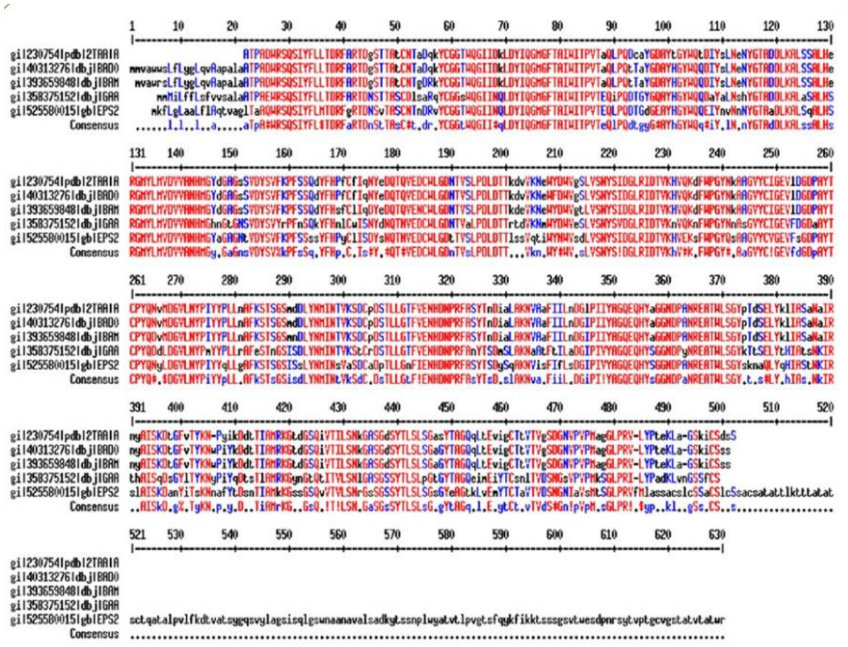

Figure 2. Multiple sequence alignment was conducted using the Multalign software, and the accession numbers for the sequences included in the alignment are as follows:

Taka-Amylase A (2TAA|A) Alpha-amylase [*Aspergillus awamori*] (BAD06002.1) Alpha-amylase [*Aspergillus sojae*] (BAM28635.1) Alpha-amylase [*Aspergillus kawachii* IFO 4308] (GAA91738.1) Alpha-amylase Amy13A [*Penicillium oxalicum* 114-2] (EPS26265.1). This alignment was carried out to identify and highlight conserved regions within these amylase enzyme sequences, aiding in our understanding of shared structural and functional elements across different amylase variants.

Forecasting the secondary structural elements and estimating solvent accessibility for the specific protein chain.

The PROFsec method employs evolutionary information gathered from multiple sequence alignments and a multi-level system to predict secondary structural elements and solvent accessibility. It classifies secondary structure into three states: helix (H, encompassing alpha-, pi-, and 3_10-helix), strand (E, representing extended strands in beta-sheet conformation with at least two residues in length), and loop (Gubbi et al. 2006).

This prediction system relies on a network of neural networks, with an anticipated average accuracy surpassing 72%. An evaluation of enzyme molecules reveals that the active form of the enzyme comprises three chains arranged in the same sequence, and the secondary structure of the enzyme corresponds to the structure depicted in the accompanying diagram (Rost et al. 1994). Specifically, the enzyme's secondary structure consists of 59% loop, 20.9% helix, and 20.09% strand. Notably, the bulk structure of the enzyme predominantly consists of loops, with helical and stranded regions being roughly equal in proportion.

The figures presented in Figures 3 and 4 illustrate the amino acid composition and corresponding secondary structure, respectively. Additionally, the enzyme's structure has been examined in terms of solvent accessibility, as illustrated in Figure 5.

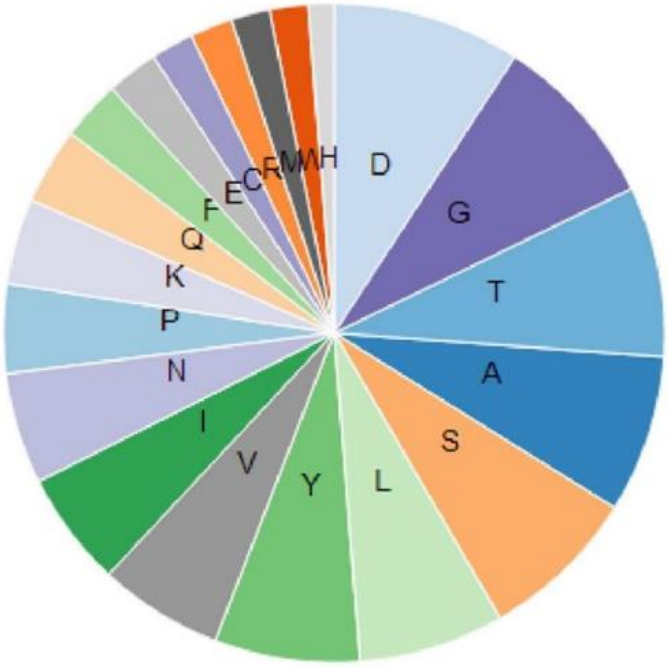

Figure 3. Amino acid composition.

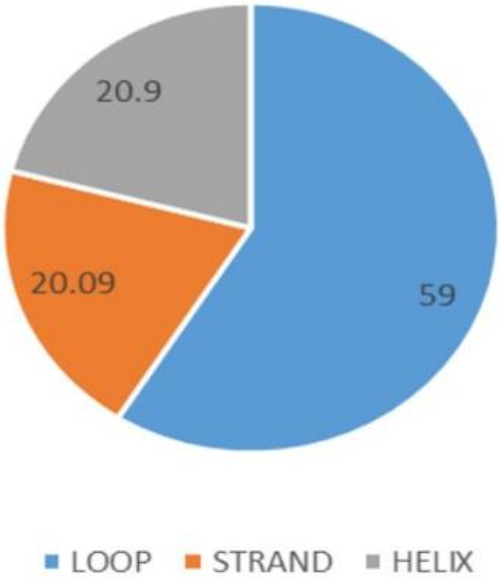

Figure 4. Composition of the secondary structure.

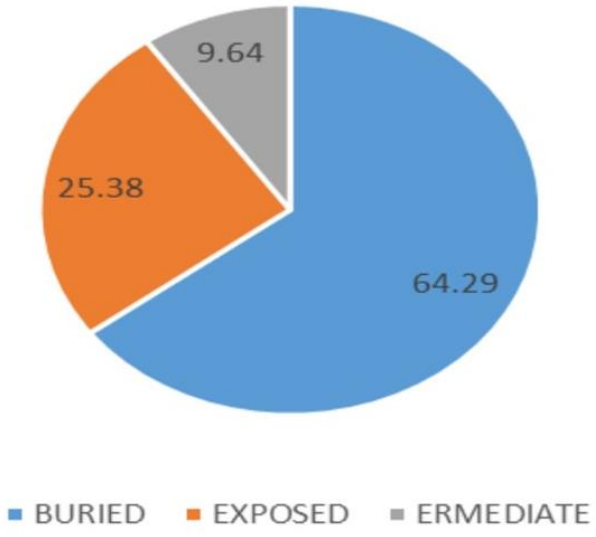

Figure 5. Accessibility of the solvent within the structure of the enzyme.

The NetPhos 2.0 server facilitates neural network predictions for serine, threonine, and tyrosine phosphorylation sites within eukaryotic proteins. This software was employed to identify phosphorylation sites on enzymes, as depicted in Figure 6, which illustrates the phosphorylation site on the enzyme structure.

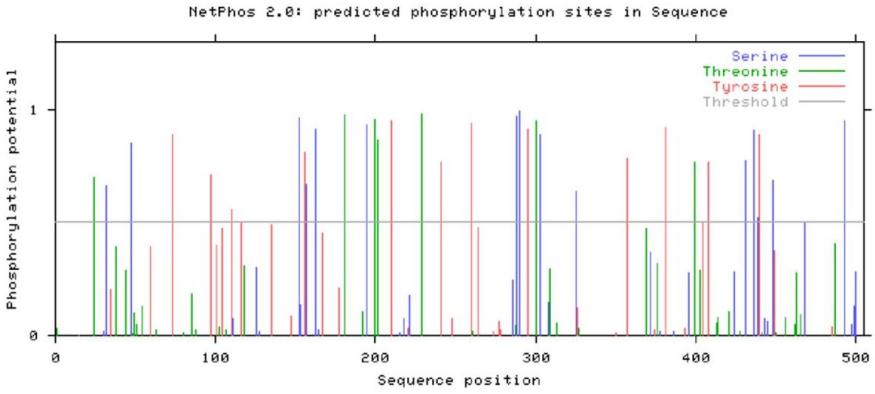

Figure 6. Sites of Phosphorylation within the enzymatic structure.

Remarkably, this enzyme possesses three calcium ions, with one in each of the three chains, positioned at residue 479, as visualized in Figure 7. It is conceivable that the essential calcium ion resides in proximity to both site C and the maltose residue, potentially contributing to the tightening of catalytic sites conducive to amylase binding or facilitating the supply of water molecules for the enzymatic reaction (Cotton 2011).

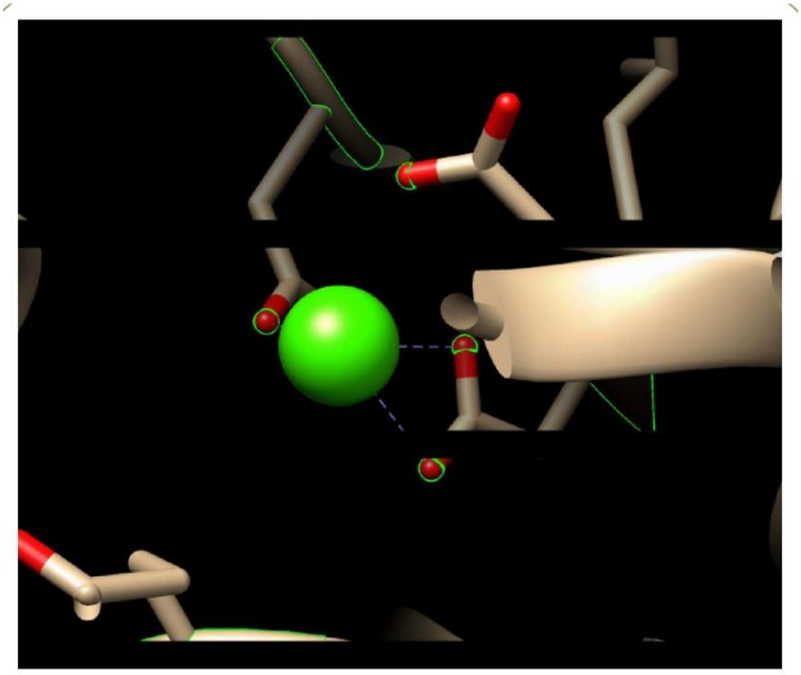

Figure 7. Calcium ion within the structure of the enzyme.

Conclusion:

The depiction of the entire chain folding reveals several distinctive features of the enzyme's structure, which align with established structural characteristics in general. The molecule comprises two discernible domains: the primary domain and the C-terminal domain, interconnected by a single polypeptide chain. Notably, a molecular model of Taka-amylase A (α-amylase) from *A. oryzae*, suggests that both histidine (His) and aspartic acid (Asp) residues in these homologous regions function as active sites. Furthermore, it is postulated that Asp and lysine (Lys) in Region 2, along with His in Region 4, may serve as substrate-binding sites. Additionally, another active site involving glutamic acid (Glu) and adjacent substrate-binding sites composed of valine (Val), leucine (Leu), and Asp between regions 2 and 4 have been proposed for consideration.

Understanding the three-dimensional structure of proteins is indispensable for comprehending their functions. The fusion of bioinformatics and biochemistry plays a pivotal role in enhancing our insight into biochemical models. In the contemporary era characterized by exponential growth in information, proficiency in utilizing bioinformatics software is paramount to swiftly and accurately achieve desired research outcomes.

References:

Cotton, Simon. Every molecule tells a story. CRC Press, 2011.

Gouet, P., Robert, X., & Courcelle, E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic acids research, 31(13), 3320-3323.

Gubbi, J., Lai, D. T., Palaniswami, M., & Parker, M. (2006). Protein secondary structure prediction using support vector machines and a new feature representation. International Journal of Computational Intelligence and Applications, 6(04), 551-567.

Long, C. M., Virolle, M. J., Chang, S. Y., Chang, S., & Bibb, M. J. (1987). alpha-Amylase gene of Streptomyces limosus: nucleotide sequence, expression motifs, and amino acid sequence homology to mammalian and invertebrate alpha-amylases. Journal of Bacteriology, 169(12), 5745-5754.

Lovell, T., Himo, F., Han, W. G., & Noodleman, L. (2003). Density functional methods applied to metalloenzymes. Coordination chemistry reviews, 238, 211-232.

Ly, H. D., & Withers, S. G. (1999). Mutagenesis of glycosidases. Annual review of biochemistry, 68(1), 487-522.

Patil, A. G., Khan, K., Aishwarya, S., Padyana, S., Huchegowda, R., Reddy, K. R., ... & Zameer, F. (2021). Fungal amylases and their industrial applications. Industrially Important Fungi for Sustainable Development: Volume 2: Bioprospecting for Biomolecules, 407-434.

Rost, B., & Sander, C. (1994). Combining evolutionary information and neural networks to predict protein secondary structure. Proteins: Structure, Function, and Bioinformatics, 19(1), 55-72.

Tateno, T., Fukuda, H., & Kondo, A. (2007). Production of L-Lysine from starch by Corynebacterium glutamicum displaying α-amylase on its cell surface. Applied microbiology and biotechnology, 74, 1213-1220.

Xia, W., Zhang, K., Su, L., & Wu, J. (2021). Microbial starch debranching enzymes: Developments and applications. Biotechnology Advances, 50, 107786.