REVIEW OF RESEARCH
IMPACT FACTOR : 5.7631(UIF) UGC APPROVED JOURNAL NO. 48514 ISSN: 2249-894X
VOLUME - 8 | ISSUE - 4 | JANUARY - 2019

BIOINFORMATICS BASED STRUCTURAL AND FUNCTIONAL ANALYSIS OF AMYLASES FROM HALOALKALIPHILIC ACTINOMYCETES

Ashish B. Kalasava
UGC-CAS Department of Biosciences, Saurashtra University, Rajkot, Gujarat (India).

ABSTRACT :

During the last two decades, omics have developed with various approaches such as metagenomics, proteomics and Genomics to explore the residing microorganisms in habitats. Actinomycetes are gram-positive microorganisms belonging to domain Bacteria and family Actinomycetaceae. The coastal region, marine sediment, salt lakes, saline soil, brines are among the common habitats for the exploration of the potential of halophiles and salt-tolerant and halophiles actinomycetes. Amylases (endo-1, 4-d-glucanohydrolases, EC 3.2.1.1) are enzymes that hydrolyze alpha 1, 4-d glucosidic linkages in polysaccharides to produce different products; such as glucose and maltose and cover up to 30% of the total enzyme market in the world. Amylases producing Halophilic actinomycetes were characterized at taxonomical level from domain to species by using NCBI taxonomy. Comparative studies based on the sequences of amylase obtained from NCBI protein data for Halophilic actinomycetes, Nocardiopsis alba and Nocardiopsis sp. were analyzed using various tools / software such ProtParam, SwissPort or TrEMBL SWISS-MODEL, GMQE, QMEAN Z-score, MolProbity and Ramachandran plot and many more. The results revealed the uniqueness of the position of amino acids. The analysis of the physical and chemical parameters of the protein sequences has significantly proved the diversity and adaption strategies of the Halophilic actinomycetes.

KEYWORDS : *Haloalkaliphilic actinomycetes, Comparative Studies.*

1. INTRODUCTION

Till date, more than millions of actinomycetes have been discovered from the order *Actinomycetales*. Diversity of actinomycetes species clearly indicates its significance and biotechnological applications. Large number of the habitats is reported with the presence of actinomycetes. The extreme habitats such Antarctic soil, hot springs, highly moderate and saline areas are reported for the residing actinomycetes. The microbes showing the ability to grow and survive at high salt concentration are termed as Halophilic and Halo-tolerant microorganisms respectively. From 1 to 20% NaCl is required to grow Halophilic and halo-tolerant microorganisms. These microorganisms produce a wide array of economically important products including amylases. Amylase (EC 3.2.1.1) catalyzes the hydrolysis of internal 1, 4-glucosidic bonds in starch and related poly- and oligosaccharides. α, β, γ types of amylases are determined based on the end products produced by them. Every domain of life shows the production of amylolytic enzymes displaying starch-hydrolyzing activity amylase. *Nocardiopsis baichengensis* and *Nocardiopsis chromatogenes* (Li *et al.,* 2006*); Nocardiopsis halophila* (Al-Tai and Ruan, 1994)*; Nocardiopsis kunsanensis* (Chun *et al*., 2000) are known to produce amylases.

Comparison of the amino acids prolife of the halophilic amylases strongly indicates abundance of acidic amino acids as compared to others along with lower value of pI. Majority of the acidic amino acids act as stabilizers for enzyme stability and activity in the presence of salt.

Bioinformatics based study covers data mining of *Nocardiopsis sp.* reported for amylase production in terms of its homologous features, uniqueness, amino acids profiles and 3D structure characterization. The

selected protein sequences of *Nocardiopsis sp.* amylase available in NCBI were retrieved and bioinformatics analysis by ProtParam and protein 3D structure prediction analysis performed for its protein characterization. A distinctive feature in primary structure, functional features, conserved motifs, and hydropathy profile of *Nocardiopsis sp.* amylase have been evidently illuminated to grow in presences of high salt and pH. It is used to capture the applications in biotechnological, food and pharmaceutical industries where enzymes with unique features in polyextreme conditions of salt, solvent, and high temperature are highly acceptable.

2. MATERIALS AND METHODS

2.1. RETRIEVAL OF AMYLASE SEQUENCING

Amylases produced by Halophilic actinomycetes were retrieved from Protein (https://www.ncbi.nlm.nih.gov/protein/) databases of NCBI in FASTA format. The NCBI Protein database is a collection of sequences from several sources, including translations from annotated coding regions in Genbank and SwissPort, database is a freely accessible database of protein sequence and functional information (Bethesda, 2016; Consortium, 2014). All the sequences were significantly different from each other. Comparative analysis of halophilic actinomycetes was carried.

2.2. CLASSIFICATION OF ORGANISMS

NCBI Taxonomy browser (https://www.ncbi.nlm.nih.gov/taxonomy), which provides the information of organism's classification ranks was used to classify the organisms based on the taxonomic domain, family and genus.

2.3. ANALYSIS OF AMINO ACID FREQUENCY AMONG THE SELECTED ACTINOMYCETES

Total twelve Sequences were retrieved from the NCBI database. Based on the origin, all sequences were clustered into three groups. Each group comprised four sequences. The twelve amylase sequences from amylase producing actinomycetes were compared to understand the amino acid composition and its role in the structural stability of the protein. The ProtParam (https://web.expasy.org/protparam/) tool was used to predict the amino acid frequency of the proteins. It is a tool which allows the computation of various physical and chemical parameters for a given protein sequence including the amino acid composition (Vaidya *et al.,* 2018).

2.4. PROTEIN 3D STRUCTURE PREDICTION AND ANALYSIS

The structures were gauged throughout homology modelling by SWISS-MODEL (https://swissmodel.expasy.org/interactive) based on the GMQE (Global Model Quality Estimation), QMEAN Z-score, MolProbity and Ramachandran plot. It makes authentic protein models and has simple access to modelling results, their visualization and elucidation (Waterhouse *et al.*, 2018).

3. RESULT AND DISCUSSION

3.1. RETRIEVAL OF ORGANISMS

Twelve protein sequences were considered for the analysis of the amylases. The sequences were retrieved from NCBI using protein database (Table 1). Eight out of twelve sequences belonged to *Nocardiopsis sp.* while the remaining four were of *Actinopolyspora erythraea*.

	Organisms	NCBI AC_No.
Amylase Producing Halophilic Actinomycetes	Alpha-glucosidase [*Actinopolyspora erythraea*]	KGI81297.1
	Glycosidase [*Actinopolyspora erythraea*]	KGI81207.1
	Hypothetical protein IL38_20700 [*Actinopolyspora erythraea*]	KGI79873.1
	Hypothetical protein IL38_20695 [*Actinopolyspora erythraea*]	KGI79872.1
Amylase from *Nocardiopsis alba*	Alpha-amylase [*Nocardiopsis alba*]	WP_014910485.1
	Alpha amylase [*Nocardiopsis alba* ATCC BAA-2163]	AFR09166.1
	Alpha amylase [*Nocardiopsis alba* ATCC BAA-2164]	AFR06474.1
	Alpha amylase [*Nocardiopsis alba* ATCC BAA-2165]	AFR08628.1
Amylase from *Nocardiopsis sp.*	Alpha-amylase [*Nocardiopsis sp.* CNT312]	WP_028649691.1
	Alpha-amylase [*Nocardiopsis sp*. SBT366]	WP_049578795.1
	Alpha-amylase [*Nocardiopsis sp.* RV163]	WP_047867596.1
	Alpha-amylase [*Nocardiopsis sp.* JB363]	SIO90115.1

Table 1: Retrieved organisms from NCBI

3.2.CLASSIFICATION OF ORGANISMS

The amino acid sequences of the twelve amylase producing actinomycetes were analyzed by ProtParam (https://web.expasy.org/protparam/) for the amino acid composition and its role in the structural stability of the protein (Table 2). A tool was used to predict the amino acid frequency of the protein

	Organisms	No. of AA	Molecular Weight	Theoretical pI	Atomic Composition					Instability Index	Aliphatic Index
					C	H	N	O	S		
Amylase Producing Halophilic Actinomycetes	Alpha-glucosidase [*Actinopolyspora erythraea*]	551	62304.8	4.98	2784	4165	797	826	9	35.34	71.03
	Glycosidase [*Actinopolyspora erythraea*]	464	50610.44	4.51	2221	3337	613	721	14	30.22	69.01
	Hypothetical protein IL38_20700 [*Actinopolyspora erythraea*]	433	47021.50	4.49	2060	3235	589	653	10	43.50	87.92
	Hypothetical protein IL38_20695 [*Actinopolyspora erythraea*]	409	44473.06	4.89	1969	3077	565	595	9	32.69	87.31
Amylase from *Nocardiopsis alba*	Alpha-amylase [*Nocardiopsis alba*]	601	63905.70	4.24	2796	4200	774	921	16	19.90	66.57
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2163]	555	61748.54	4.64	2762	4168	768	837	7	41.99	78.76
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2164]	541	60875.75	4.71	2729	4116	772	809	6	41.58	79.48
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2165]	620	69562.15	9.57	3070	4740	976	866	12	52.78	65.50
Amylase from *Nocardiopsis sp.*	Alpha-amylase [*Nocardiopsis sp.* CNT312]	708	74576.09	4.38	3277	4869	905	1070	16	28.60	61.24
	Alpha-amylase [Nocardiopsis sp. SBT366]	606	63798.25	4.13	2789	4155	763	930	17	24.68	62.48
	Alpha-amylase [*Nocardiopsis sp.* RV163]	592	63059.90	4.44	2769	4137	771	895	16	23.59	64.31
	Alpha-amylase [*Nocardiopsis sp.* JB363]	655	73128.35	5.84	3273	5030	934	955	12	31.91	81.18

*C - Carbon, H - Hydrogen, N - Nitrogen, O - Oxygen, S - Sulfur.

Table 2: Classification of Atomic Composition

3.3. COMPARATIVE ANALYSIS OF AMINO ACID FREQUENCY

Looking into the detailed analysis of the amino acid composition of the sequences, higher numbers of the acidic amino acids as compared to basic amino acids were detected. The catalytic domain sequences

of the amylases from different actinomycetes were analyzed through ProtParam to understand the lineage of α-amylases (Table 3-6 and Figure 1-4). Lower values of pI were observed in the eleven sequences out of twelve, indicating its relevance in the adaptation of proteins in saline habitats.

Table 3: Comparisons of Halophilic Actinomycetes amino acids frequency

Amylase Producing Halophilic Actinomycetes	Aliphatic	Sulphur Containing	Hydroxy	Aromatic	Heterocyclic	Basic	Acidic
Alpha-glucosidase [*Actinopolyspora erythraea*]	35.2	1.6	9.3	10.5	8.5	14.1	20.7
Glycosidase [*Actinopolyspora erythraea*]	36.9	3	14.2	10.5	3.9	9.6	21.9
Hypothetical protein IL38_20700 [*Actinopolyspora erythraea*]	42.1	2.3	12	6.7	5.8	9.9	21.2
Hypothetical protein IL38_20695 [*Actinopolyspora erythraea*]	41.1	2.2	13.4	7.1	7.8	11.5	17

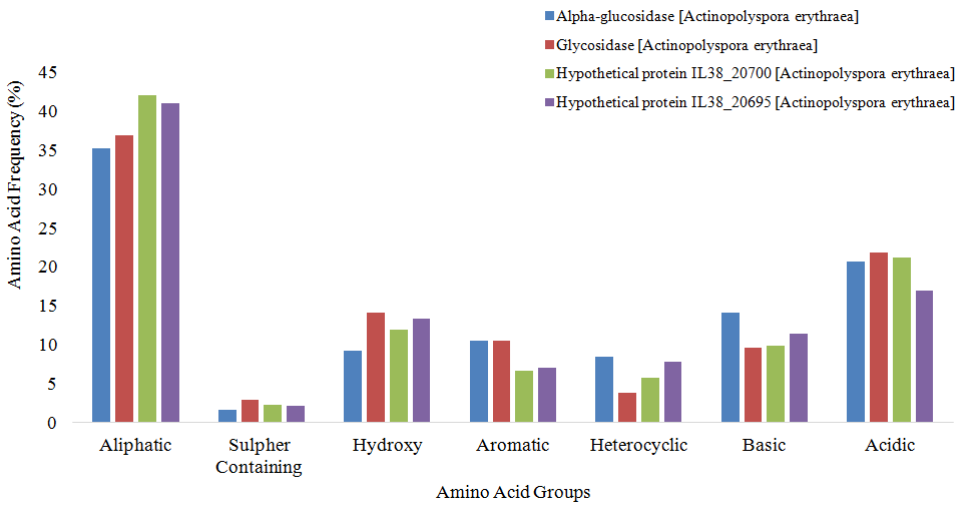


Figure 1: Comparisons of Halophilic Actinomycetes amino acids frequency

Amylase from *Nocardiopsis alba*	Aliphatic	Sulphur Containing	Hydroxy	Aromatic	Heterocyclic	Basic	Acidic
Alpha-amylase [*Nocardiopsis alba*]	39.2	2.6	12.2	8.8	7	7.7	22
Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2163]	38.1	1.3	9.7	9.9	7.7	12.3	20.9
Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2164]	37.1	1.1	10	10.6	8.5	12.4	20.3
Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2165]	35.5	1.9	6.5	8.1	11.1	19.4	17.5

Table 4: Comparisons of *Nocardiopsis alba* amino acids frequency

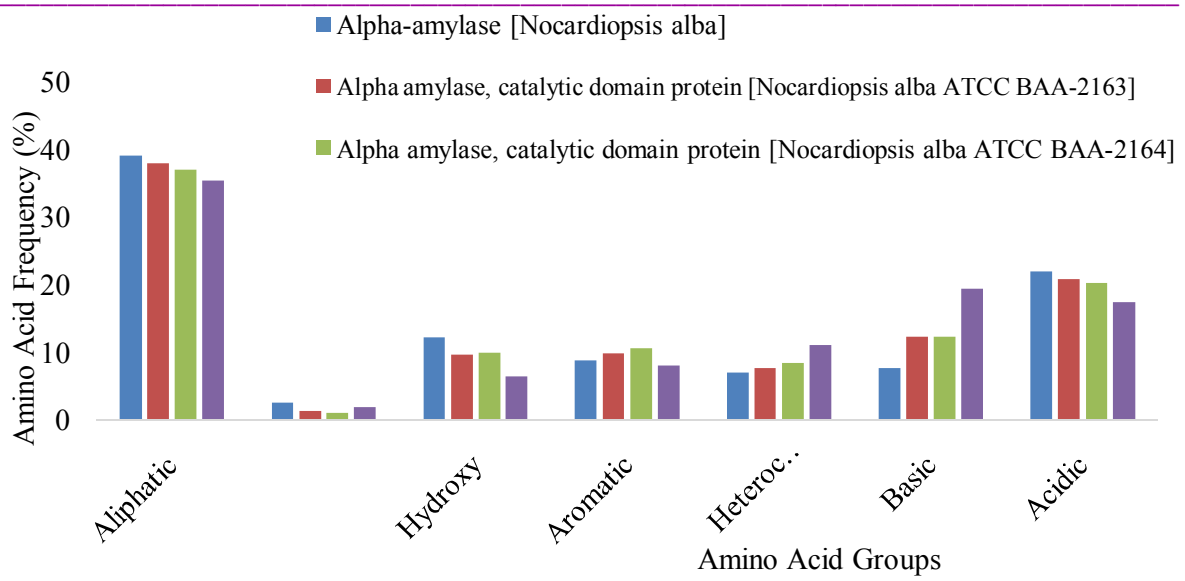


Figure 2: Comparisons of *Nocardiopsis alba* amino acids frequency

Amylase from *Nocardiopsis sp.*	Aliphatic	Sulphur Containing	Hydroxy	Aromatic	Heterocyclic	Basic	Acidic
Alpha-amylase [*Nocardiopsis sp.* CNT312]	39.8	2.2	13.8	9.9	6.4	8.2	19.7
Alpha-amylase [*Nocardiopsis sp.* SBT366]	39.2	2.8	13.7	9.4	6.9	6.5	21.6
Alpha-amylase [*Nocardiopsis sp.* RV163]	39.2	2.8	12.8	9.7	5.9	8.6	20.9
Alpha-amylase [*Nocardiopsis sp.* JB363]	28.1	1.8	11.1	8.7	7.8	15.6	17.9

Table 5: Comparisons of *Nocardiopsis sp.* amino acids frequency

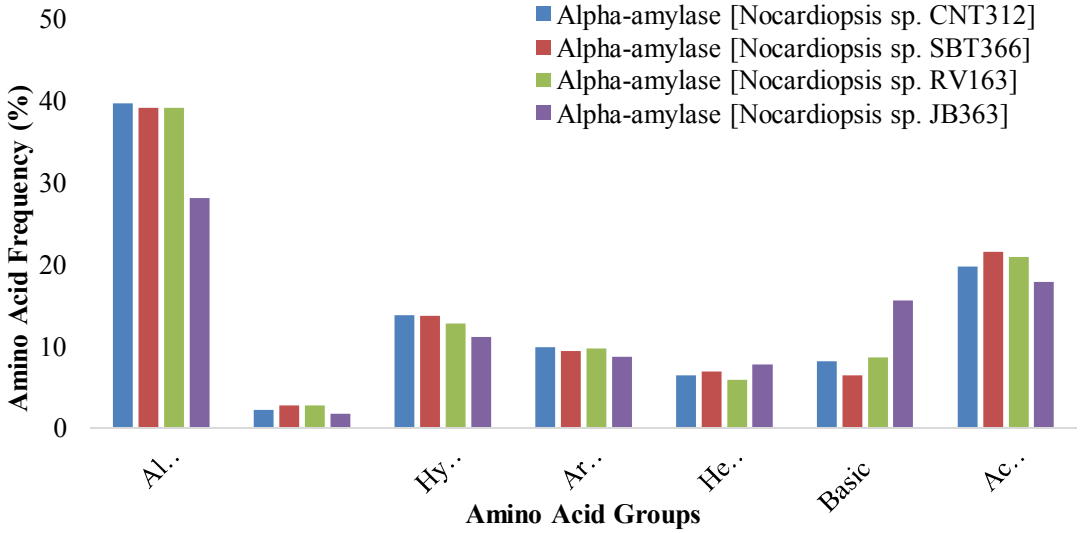


Figure 3: Comparisons of *Nocardiopsis sp.* amino acids frequency

	Organisms	Aliphatic	Sulphur Containing	Hydroxy	Aromatic	Heterocyclic	Basic	Acidic
Amylase Producing Halophilic Actinomycetes	Alpha-glucosidase [*Actinopolyspora erythraea*]	35.2	1.6	9.3	10.5	8.5	14.1	20.7
	Glycosidase [*Actinopolyspora erythraea*]	36.9	3	14.2	10.5	3.9	9.6	21.9
	Hypothetical protein IL38_20700 [*Actinopolyspora erythraea*]	42.1	2.3	12	6.7	5.8	9.9	21.2
	Hypothetical protein IL38_20695 [*Actinopolyspora erythraea*]	41.1	2.2	13.4	7.1	7.8	11.5	17
Amylase from *Nocardiopsis alba*	Alpha-amylase [*Nocardiopsis alba*]	39.2	2.6	12.2	8.8	7	7.7	22
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2165]	38.1	1.3	9.7	9.9	7.7	12.3	20.9
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2165]	37.1	1.1	10	10.6	8.5	12.4	20.3
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2165]	35.5	1.9	6.5	8.1	11.1	19.4	17.5
Amylase from *Nocardiopsis sp.*	Alpha-amylase [*Nocardiopsis sp.* CNT312]	39.8	2.2	13.8	9.9	8.5	8.2	19.7
	Alpha-amylase [*Nocardiopsis sp.* SBT366]	39.2	2.8	13.7	9.4	8.5	6.5	21.6
	Alpha-amylase [*Nocardiopsis sp.* RV163]	39.2	2.8	12.8	9.7	8.5	8.6	20.9
	Alpha-amylase [*Nocardiopsis sp.* JB363]	28.1	1.8	11.1	8.7	8.5	15.6	17.9

Table 6: Comparisons of amino acids frequency between three groups

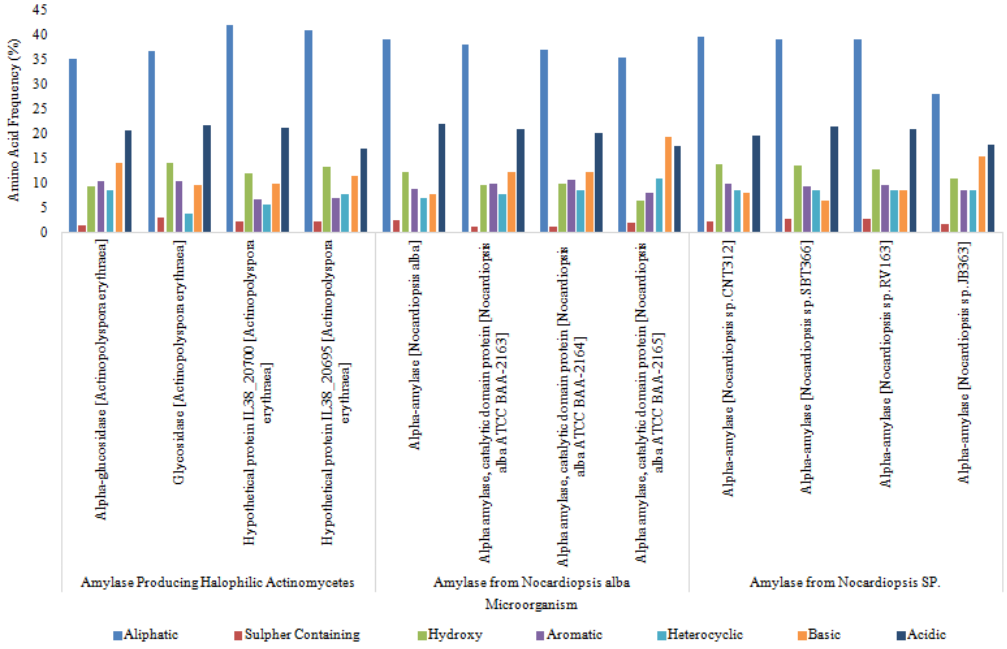


Figure 4: Comparisons of amino acids frequency between three group

3.4. THE PREDICTION AND ANALYSIS OF THE PROTEIN 3D STRUCTURE

The selected proteins were homology modeled by SWISS-MODEL and evaluated for the predicted structures through the structural assessment server of Swiss model workspace. The details of the structure and the evaluation data are given in the table 6 (A & B). The predicted structure shows that the alpha amylase enzymes do not require any metal ion for their optimum activity. The scores of GMQE, QMEAN,

MolProbity, and Ramachandran favored the quality of the predicted structures that are significantly better Table 7 (A & B) and Figure 5.

	Organisms	NCBI AC_No.	Swiss Model Template	Swiss Identity %	Query Coverage	GMQE	QMEAN	MolProbity	Ramachandran Favoured
Amylase Producing Halophilic Actinomycetes	Alpha-glucosidase [*Actinopolyspora erythraea*]	KGI81297.1	3wy1.1.A	39.10 %	9-549	0.68	-3.08	2.13	88.31%
	Glycosidase [*Actinopolyspora erythraea*]	KGI81207.1	1kxh.1.A	50.35 %	34-461	0.78	-1.81	1.55	94.37%
	Hypothetical protein IL38_20700 [*Actinopolyspora erythraea*]	KGI79873.1	6aav.1.A	28.82 %	17-431	0.61	-3.66	2.31	89.83%
	Hypothetical protein IL38_20695 [*Actinopolyspora erythraea*]	KGI79872.1	6aav.1.A	27.82 %	9-409	0.67	-2.59	2.06	93.98%
Amylase from *Nocardiopsis alba*	Alpha-amylase [*Nocardiopsis alba*]	WP_014910485.1	1xh2.1.A	39.29 %	40-497	0.55	-3.89	1.92	89.69%
	Alpha amylase [*Nocardiopsis alba* ATCC BAA-2163]	AFR09166.1	6aav.1.A	40.57 %	4-504	0.68	-1.41	1.81	93.59%
	Alpha amylase [*Nocardiopsis alba* ATCC BAA-2164]	AFR06474.1	3wy1.1.A	40.43 %	3-552	0.64	-4.01	1.76	90.99%
	Alpha amylase [*Nocardiopsis alba* ATCC BAA-2165]	AFR08628.1	5h2t.2.B	61.86 %	2-536	0.69	-2.71	1.54	91.02%
Amylase from *Nocardiopsis sp.*	Alpha-amylase [*Nocardiopsis sp.* CNT312]	WP_028649691.1	3bmv.1.A	19.93 %	36-438	0.51	-3.04	2.02	86.17%
	Alpha-amylase [*Nocardiopsis sp.* SBT366]	WP_049578795.1	1ciu.1.A	26.52 %	49-708	0.47	-5.18	2.12	90.61%
	Alpha-amylase [*Nocardiopsis sp.* RV163]	WP_047867596.1	1clv.1.A	42.89 %	26-485	0.58	-4.11	1.78	89.30%
	Alpha-amylase [*Nocardiopsis sp.* JB363]	SIO90115.1	7mgy.1.A	64.43 %	1-653	0.87	-0.77	1.31	95.39%

Table 7 (A): Information of predicted three dimensional structures and its accuracy

	Organisms	Seq. Similarity	Coverage	Range	Cβ	All atom	Solavation	Torsion	QMEANDisCo Global
Amylase Producing Halophilic Actinomycetes	Alpha-glucosidase [*Actinopolyspora erythraea*]	0.39	0.93	9-549	-2.81	-2.54	-0.86	-2.18	0.69 ± 0.05
	Glycosidase [*Actinopolyspora erythraea*]	0.45	0.92	34-461	-2.98	-1.34	-1.46	-0.86	0.81 ± 0.05
	Hypothetical protein IL38_20700 [*Actinopolyspora erythraea*]	0.34	0.92	17-431	-1.33	-2.39	-1.65	-3.08	0.63±0.05
	Hypothetical protein IL38_20695 [*Actinopolyspora erythraea*]	0.34	0.98	9-409	-2.33	-2.35	-1.91	-1.61	0.66±0.05
Amylase from *Nocardiopsis alba*	Alpha-amylase [*Nocardiopsis alba*]	0.39	0.75	40-497	-3.05	-2.33	-2.38	-2.66	0.68±0.05
		0.31	0.86	48-593	-2.86	-3.22	-2.47	-2.95	0.54±0.05
		0.34	0.88	51-596	-3.61	-3.58	-4.42	-3.78	0.35±0.05
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2163]	0.39	0.88	4-504	-0.99	-1.73	-1.15	-0.81	0.73±0.05
		0.35	0.93	3-552	-1.75	-2.75	-1.35	-3.05	0.64±0.05
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2164]	0.39	0.94	2-536	-1.57	-2.30	-1.31	-1.89	0.69±0.05
	Alpha amylase, catalytic domain protein [*Nocardiopsis alba* ATCC BAA-2165]	0.48	0.66	36-438	-3.12	-2.63	-0.79	-2.30	0.73±0.05
Amylase from *Nocardiopsis sp.*	Alpha-amylase [*Nocardiopsis sp.* CNT312]	0.31	0.84	53-708	-2.81	-2.66	-3.10	-2.91	0.52±0.05
		0.35	0.81	60-707	-3.95	-3.51	-3.67	-3.86	0.41±0.05
		0.33	0.83	60-707	-4.56	-3.31	-3.58	-.3.83	0.41±0.05
		0.34	0.83	60-707	-3.61	-3.19	-3.48	-3.97	0.41±0.05
		0.35	0.82	60-707	-3.23	-3.74	-3.49	-4.38	0.39±0.05
		0.34	0.77	100-707	-2.46	-3.27	-3.34	-3.97	0.39±0.05
		0.34	0.73	60-596	-3.68	-4.10	-4.13	-4.50	0.38±0.05
	Alpha-amylase [*Nocardiopsis sp.* SBT366]	0.38	0.74	39-498	-3.54	-2.07	-2.71	-1.93	0.67±0.05
		0.31	0.85	49-597	-3.00	-2.92	-3.30	-2.91	0.54±0.05
	Alpha-amylase [*Nocardiopsis sp.* RV163]	0.41	0.75	26-485	-3.02	-2.35	-1.93	-3.02	00.71±0.05
		0.30	0.87	42-585	-4.21	-3.26	-3.34	-3.85	0.53±0.05
		0.34	0.89	42-588	-4.05	-3.68	-3.88	-4.62	0.39±0.05
	Alpha-amylase [*Nocardiopsis sp.* JB363]	0.50	0.98	1-653	-1.18	-0.11	-0.17	-0.52	0.85±0.05

Table 7 (B): Information of predicted three dimensional structures and its accuracy

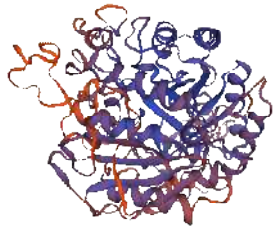
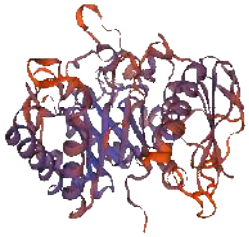
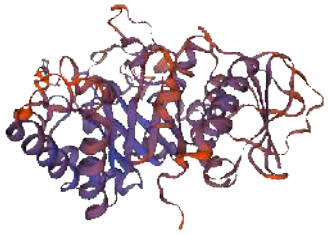
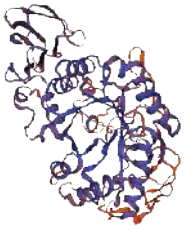
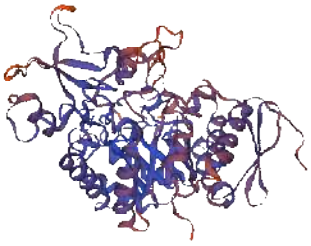
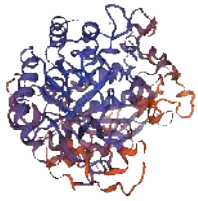
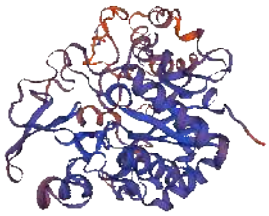
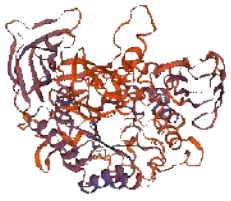
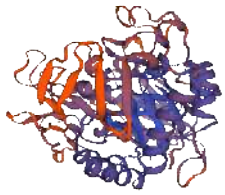
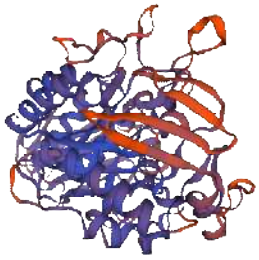
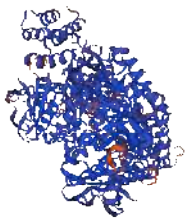
Alpha-glucosidase [*Actinopolyspora erythraea*]	Glycosidase [*Actinopolyspora erythraea*]	Hypothetical protein IL38_20700 [*Actinopolyspora erythraea*]
		
Hypothetical protein IL38_20695 [*Actinopolyspora erythraea*]	Alpha-amylase [*Nocardiopsis alba*]	Alpha amylase [*Nocardiopsis alba* ATCC BAA-2163]
		
Alpha amylase [*Nocardiopsis alba* ATCC BAA-2164]	Alpha amylase [*Nocardiopsis alba* ATCC BAA-2165]	Alpha-amylase [*Nocardiopsis sp.* CNT312]
		
Alpha-amylase [*Nocardiopsis sp*. SBT366]	Alpha-amylase [*Nocardiopsis sp.* RV163]	Alpha-amylase [*Nocardiopsis sp.* JB363]
		

Figure 5: Protein 3D structure prediction.

4. DISCUSSION

Twelve protein sequences considered for producing the amylases were retrieved from NCBI using a protein database. Eight of twelve sequences were from Nocardiopsis *sp.* the remaining four were from *Actinopolyspora erythraea*. Further, molecular and structural characterizations were carried out to view that amylases with similar/different biochemical properties will have similar/different genes/protein sequences and vice versa. Halophilic amylase from *Marinobacter algicola* was found to contain 571 amino acid residues (Sumit *et al.,* 2016). While our sequences were found to contain 409 to 708 amino acids.

Looking at the amino acid composition of the sequences, it was apparent that amylases contained higher numbers of the acidic amino acids as compared to the basic amino acids. This is in confirmation to the fact that the majority of proteins obtained from halophiles and analyzed for the amino acids compositions showed a higher abundance of acidic amino acids (Vaidya *et al.,* 2018).

The sequences of the amylase catalytic domains from different actinomycetes were analyzed through Protopram to understand the lineage of α-amylase. *Marinobacter algicola* α-amylase has a low pI value of 4.87, similarly a lower value of pI was also observed in eleven sequences out of twelve (Sumit *et al.,* 2016). It clearly showed the adaption of protein in saline habitats.

Marinobacter algicola α-amylase contained more negatively charged residues (48) than positively charged residues (30) in its structure. The sequences also showed the higher numbers of negatively charged residues (57- 96) compared to positively charged residues (27 - 91) (Sumit *et al.,* 2016).

The presence of acidic amino acids and lower pI values are also observed in halophilic α-amylases from *Halomonas meridiana* and *Kocuria varians Natronococcus* sp. strain Ah-36 (Coronado *et al.*, 2000 and Kobayashi *et al.*, 1992). The results are in agreement with above mentioned reports as each sequence contained a higher number of acidic amino acids as compared to the basic amino acids. Acidic residues enhance the binding of hydrated ions and help to maintain water on the protein surface. While comparing all amino acids composition, each halophilic sequence showed the higher percentage of aliphatic amino acids.

Lack of three-dimensional structures restricts our understanding about the structure–function association of the halophilic α-amylases. Various structural features present in halophilic proteins make them adapted to the saline conditions.

In order to correlate the salt stability of halophilic α-amylase, the three-dimensional structure of twelve protein sequences considered for producing the amylase were modeled using the online SWISS MODEL for protein 3D structure prediction. Figure No. 1 shows the model with tertiary structure topology. The Ramachandran plot is a way to visualize energetically favored regions for the backbone dihedral angles against amino acid residues in protein structure. As per SWISS MODEL guide, the score of 98% of the Ramachandran is considered as ideal. In each structure studied, 86.17% to 95.39% score was observed. Thereby, we can strongly validate our structure along with the stability.

GMQE is a quality estimate which combines the properties from the target template alignment and the template structure. As per the SWISS MODEL guide, the value of GMQE must be between 0 to 1. The values were in the range of 0.28 to 0.87 in our studies. It showed the reliability of the quality estimation with reference to alignment and structure.

Bioinformatics was used to understand the α-amylase properties with its structures and amino acids composition. It showed a high inclination of the negatively charged amino acids attributing salt stability. Present investigation has led to better understanding of the structural features responsible for the adaptation of the Halophilic α-amylases in salt.

5. CONCLUSION

Twelve protein sequences of the amylases retrieved from NCBI using a protein database were considered for the structural analysis. Eight sequences were related with the *Nocardiopsis sp.* While four were identified with the *Actinopolyspora erythraea.* Various approaches and bioinformatics tools; Ramachandran plot, ProtParam and 3D structure formation were employed to analyze the structure and

function relationship of the amylases. Bioinformatics analysis revealed the structural properties of the α-amylases and its association with the amino acids composition.

ACKNOWLEDGEMENT

UGC-CAS Department of Biosciences, Saurashtra University, Rajkot (Gujarat) to provide computational resources to carry out studies.

CONFLICT OF INTEREST

No conflict of interest related to the study

REFERENCES

1. Al-Tai, Amira Mahmoud, and Ji-Sheng Ruan. "Nocardiopsis halophila sp. nov., a new halophilic actinomycete isolated from soil." International Journal of Systematic and Evolutionary Microbiology 44, no. 3 (1994): 474-478.
2. Bethesda (MD). Entrez Help. 2016 [Available from: https://www.ncbi.nlm.nih.gov/books/NBK3837/].
3. Chun, Jongsik, Kyung Sook Bae, Eun Young Moon, Sang-Oun Jung, Hong Kum Lee, and Sang-Jin Kim. "Nocardiopsis kunsanensis sp. nov., a moderately halophilic actinomycete isolated from a saltern." International Journal of Systematic and Evolutionary Microbiology 50, no. 5 (2000): 1909-1913.
4. Chun, Jongsik, Kyung Sook Bae, Eun Young Moon, Sang-Oun Jung, Hong Kum Lee, and Sang-Jin Kim. "Nocardiopsis kunsanensis sp. nov., a moderately halophilic actinomycete isolated from a saltern." International Journal of Systematic and Evolutionary Microbiology 50, no. 5 (2000): 1909-1913.
5. Coronado, María-José, Carmen Vargas, Jürgen Hofemeister, Antonio Ventosa, and Joaquín J. Nieto. "Production and biochemical characterization of an α-amylase from the moderate halophile Halomonas meridiana." FEMS microbiology letters 183, no. 1 (2000): 67-71.
6. Kobayashi, Tetsuo, H. A. R. U. H. I. K. O. Kanai, T. A. K. A. Y. A. Hayashi, T. E. R. U. H. I. K. O. Akiba, R. Akaboshi, and K. Horikoshi. "Haloalkaliphilic maltotriose-forming alpha-amylase from the archaebacterium Natronococcus sp. strain Ah-36." Journal of bacteriology 174, no. 11 (1992): 3439-3444.
7. Kumar, Sumit, Rizwan Hasan Khan, and S. K. Khare. "Structural elucidation and molecular characterization of Marinobacter sp. α-amylase." Preparative Biochemistry and Biotechnology 46, no. 3 (2016): 238-246.
8. Li, Wen-Jun, Reiner M. Kroppenstedt, Dong Wang, Shu-Kun Tang, Jae-Chan Lee, Dong-Jin Park, Chang-Jin Kim, Li-Hua Xu, and Cheng-Lin Jiang. "Five novel species of the genus Nocardiopsis isolated from hypersaline soils and emended description of Nocardiopsis salina Li et al. 2004." International journal of systematic and evolutionary microbiology 56, no. 5 (2006): 1089-1096.
9. UniProt Consortium. "UniProt: a hub for protein information." Nucleic acids research 43, no. D1 (2015): D204-D212.
10. Vaidya, Atman, Varun S. Nair, John J. Georrge, and Singh Sp. "Comparative analysis of thermophilic proteases." J. Life Sci. Bioinform. Pharm. Chem. Sci 4 (2018): 65-91.
11. Waterhouse, Andrew, Martino Bertoni, Stefan Bienert, Gabriel Studer, Gerardo Tauriello, Rafal Gumienny, Florian T. Heer et al. "SWISS-MODEL: homology modelling of protein structures and complexes." Nucleic acids research 46, no. W1 (2018): W296-W303.