Chapter 5: *In Silico* Study on Molecular Adaptation of acidophilic α-amylase: with Special Reference to *Aspergillus niger*

Introduction

Acidophilic amylase produced from different microorganisms like bacteria, fungi, archaebacteria etc. Among the bacteria *Bacillus* sp. and fungi *Aspergillus* sp. are mostly used in the production of this amylase. It has immense application in different industries such as glucose and fructose syrup producing industry, bakery industry, fruit juice producing industry and in different digestive syrup producing industry (Souza and Maghalhaes, 2010). This α-amylase enzyme is widely distributed in plants and soils therefore; it plays a significant role in breakdown of starch (Mehta and Satyanarayana, 2016). There are many properties of enzymes like enzyme kinetics, downstream processing and yield of enzyme can be refined by this protein engineering process (Garg et al., 2014). Subsequently in different industries remodelled enzymes will be used.

In silico study is a computational methodology, which is used in molecular docking analysis of different drugs, proteins etc. *In silico* study not only reduces the time of discovery but also reduces the expensive laboratory work. *In silico* study is applied in whole cell analysis of prokaryotes and eukaryotes, optimization of different product yield, drug designing of different medicine like cancer, HIV etc (Bortalato et al., 2011).

Some *In silico* analysis of different α-amylases from different organisms have already been reported. *In Silico* Inhibition study of α-amylase by diosmetin and galangin was reported by Madeswaran et al. (2014), inhibition by Butein and Tricetin was reported by Madeswaran et al. (2014), inhibition of porcine pancreatic α -amylase by anthrocyanin were reported by Sui et al. (2016), *in vitro* and *in Silico* interaction study of α-amylase and molecule from *Stevia rebaudiana* was reported by Singla et al. (2017). *In silico* and *in vitro* interaction study of α-amylase with seed extract of Viciafaba and their antidiabetic activity was reported by Choudhary and Mishra (2017). Comparative i*n silico* analysis of different α-amylase family GH57 was reported by Martinovičová and Janeček (2018). An important review work on bacterial and archaeal α-amylase diversity was reported by Mehta and

Satyanarayana (2016). The (β/α)8 Scaffold based α-amylase evolution was reported by Pujadas and Palau (1997). Comparative study among barley, rice and wheat α-amylase was reported by Zhang and Li (2017). Sequence diversity and molecular phylogenetics study of wheat α-amylase was analyzed by Pandey et al. (2016). Gene sequence analysis of Indian rock oyster α -amylase was performed by Thongsaiklaing et al. (2014). But comparative study among different α -amylase, produced by an industrially important species *Aspergillus niger* have not been done yet. In present study, comparative study between normal and acid α -amylase was performed to understand the adaptive changes among them. This study will help to remodel normal enzyme to acidophilus α -amylase.

Materials and Methods

5.2.1. Sequence and structure retrieval

Multiple protein sequences of *Aspergillus niger* α -amylase were retrieved from different protein sequence databases, like EBI (https://www.ebi.ac.uk/), IMG JGI (https://img.jgi.doe.gov/), NCBI (https://www.ncbi.nlm.nih.gov/) and PDB (https://www.rcsb.org/). Among the retrieved sequences total 36 sequences were selected for analysis, where total 5 sequences were reported as acid α -amylase. Three X-ray crystallographic structures were also retrieved from PDB for structure based comparative analysis.

5.2.2. Multiple Sequence Alignment

Selected sequences were initially subjected to multiple sequence alignment (MSA) using Clustal X2 Madeswaran et al. (2014). The result of MSA was represented by CLC-Bio sequence viewer [http://resources.qiagenbioinformatics.com/manuals/clcsequenceviewer/current/User_Manual.pdf].

5.2.3. Phylogenetic tree construction

The result of MSA was used for phylogenetic tree construction using Phylip 3.69 Madeswaran et al. (2014). The phylogenetic tree was represented through Dendroscope (Huson et al., 2007).

5.2.4. Structural comparison and active site analysis

Secondary structure comparison and active site analysis of three X-ray crystallographic structures were performed by Profunc server [http://www.ebi.ac.uk/thornton-srv/databases/profunc/] whereas tertiary structure alignment was performed by Pymol (Yuan et al., 2017).

5.3. Results and Discussion

Amino acid sequences of 36 different α-amylase from *Aspergillus niger* were retrieved from different biological databases (Table 5.1). Among them 7 sequences were from EBI database with 1acid-stable α-amylase (A0A124BXE9), 6 sequences from one organism *Aspergillus niger* CBS 513.88 from IMG JGI database. Among which one sequence was acid α-amylase (640407924 XP_001394335). 20 sequences retrieved from NCBI with 2 acid-stable α-amylase (Gnd AQ42198.1, ADX42122.1) and 1 acid α-amylase (CAK48325.1). Another 3 sequences of three X-ray crystallographic structure; 2AAA, 2GUY and 2GVYwere retrieved from PDB structure database.Among the three structures, 2AAA: A was found as the acid amylase. So, along with neutral pH stable α- amylase, acid α- amylase protein sequences were also retrieved from all the databases. The sequences were selected on the basis of at least one amino acid sequence diversity.

Table 5.1 Protein sequence retrieval of α-amylase and acidophilus α-amylase from different biological sequence and structure databases.

EBI (https://www.ebi.ac.uk/)		
A0A124BXE9	Acid-stable α-amylase	*Aspergillus niger*
P56271	Acid α-amylase	
A0A124BVF5	α-amylase	
A0A117DYY7		
A0A100ILZ9		
A0A117E185		
A0A100IUD4		
IMG JGI (https://img.jgi.doe.gov/)		

640407924 XP_001394335	acid α-amylase	*Aspergillus niger* CBS 513.88 clone An11: NT_166526
640411740 XP_001397301	α-amylase	*Aspergillus niger* CBS 513.88 clone An15: NT_166530
640414456 XP_001402054		*Aspergillus niger* CBS 513.88 clone An04: NT_166539
640400989 XP_001389762		*Aspergillus niger* CBS 513.88 clone An01: NT_166518
640407205 XP_001393627		*Aspergillus niger* CBS 513.88 clone An09: NT_166525
640409402 XP_001395749	extracellular α-amylaseamyA/amyB	*Aspergillus niger* CBS 513.88 clone An12: NT_166527
NCBI (https://www.ncbi.nlm.nih.gov/)		
GAQ40175.1	α-amylase	*Aspergillus niger*
GAQ43652.1		
GAQ47563.1		
GAQ35931.1		
GAQ43653.1		
EHA25810.1		*Aspergillus niger* ATCC 1015
EHA27086.1		
XP_001395328.2		*Aspergillus niger* CBS 513.88
XP_003188777.1		
XP_001397301.1		
XP_001393626.1		
EHA28336.1		*Aspergillus niger* ATCC 1015
EHA27488.1		
EHA24955.1		
EHA23512.1		
XP_001393627.2		*Aspergillus niger* CBS 513.88
XP_001389762.2		
GAQ42198.1	acid-stable α-amylase	*Aspergillus niger*
ADX42122.1		
CAK48325.1	acid α-amylase	
PDB (https://www.rcsb.org/)		
2AAA:A	X-ray diffraction study of Calcium binding in α-amylase	*Aspergillus niger*
2GUY:A	Orthorhombic crystal structure (space group P21212) of α-amylase	
2GVY:A	Monoclinic crystal form of α-amylase in complex with maltose	

Functionally similar protein may or may not have sequential similarity. Those protein who shares homologous relationships, have maximum percentage of identical amino acid compositions. The diversity among functionally similar protein sequences occurred due to interaction of protein with its immediate environment (Li et al., 2008). To become stabilize within the changed environment similar as well as dissimilar amino acid substitution, insertion and deletion occurs among the proteins. To understand the pattern of amino acid diversity, selected protein sequences of *Aspergillus niger* α-amylase were

subjected to multiple sequence alignment. Though all the sequences were α-amylase and selected from same species, they showed significant level of sequence diversity. An amino acid sequence insertion of 535at N-terminal and 87at C-terminal end was observed for the sequencesA0A124BXE9 (from EBI) and GAQ42198.1 (from NCBI). Both of them are acid stable in nature. Some locations were observed with conserved amino acids, such as position 613-628, 673-681, 811-821 and 951-965. Whereas some amino acid positions were 100% conserved like, 620-G, 628-P, 646-Y, 662-G, 636-T, 670-L, 684-D, 686-V, 791-L, 798-V, 817-G, 819-R, 821-D, 851-G. and some positions were found with similar substitution for one or two sequences like, 626-I/L(2), 667-L/V(2), 852-E/D(1), and 990-L/V(1).According to Van der Marrel et al. (2002)some amino acids, like Asp(D) at the β-sheet 2, Arg(R) and Asp(D) at the β-sheet 4, Glu(E) at the β-sheet 5 and His(H)-Asp(D) at the β-sheet 7, were 100% conserved among maximum α-amylase family proteins and directly acts as the catalytic amino acids. Here in present analysis position 684-D, 819-R, 821-D and 852-E/D (1) were representing the specific amino acids positions. Some amino acid diversity was observed for the His (H)-Asp (D) combination at 5β-sheet position (Van der Marrel *et al.*, 2002). At the alignment position 919, H was replaced by Q for 9 sequences, D and E by each one position. Whereas at the alignment position 920, D was replaced by I, G, S and A in 4 sequences. So, from the above result it has been identified that mainly catalytic amino acid residues remained unchanged due to their functional similarity and diversity occurred at other structural positions.

The result of multiple sequence alignment in phy format was used for Phylip3.69 to construct the phylogenetic tree. According to the phylogenetic tree represented as phylogram, five separate clades were found (Fig. 5.1). Most interestingly the seven sequentially different α-amylase from one whole genome sequence of *Aspergillus niger* CBS 513.88, downloaded from IMG-JGI, were found in seven different locations in the tree. Above result indicated that seven different sequences of *Aspergillus niger* CBS 513.88 may be the isoforms of α-amylases present in the same organism. The acid α-amylase sequence of *Aspergillus niger* CBS 513.88 (accession number 640407924 XP_001394335) was found at the bottom cluster. Most noticeably all the acid α-amylase retrieved from different biological databases were also found together in the bottom

cluster. The separate clustering of acid α- amylase indicated their drastic sequential adaptation in their changed environment. Phylogenetic distribution of seven isoform of α-amylase from one organism (*Aspergillus niger* CBS 513.88), indicated its increased capacity of degrading starch at neutral as well as acidic condition. Their expression may occurs according to the change in external environment.

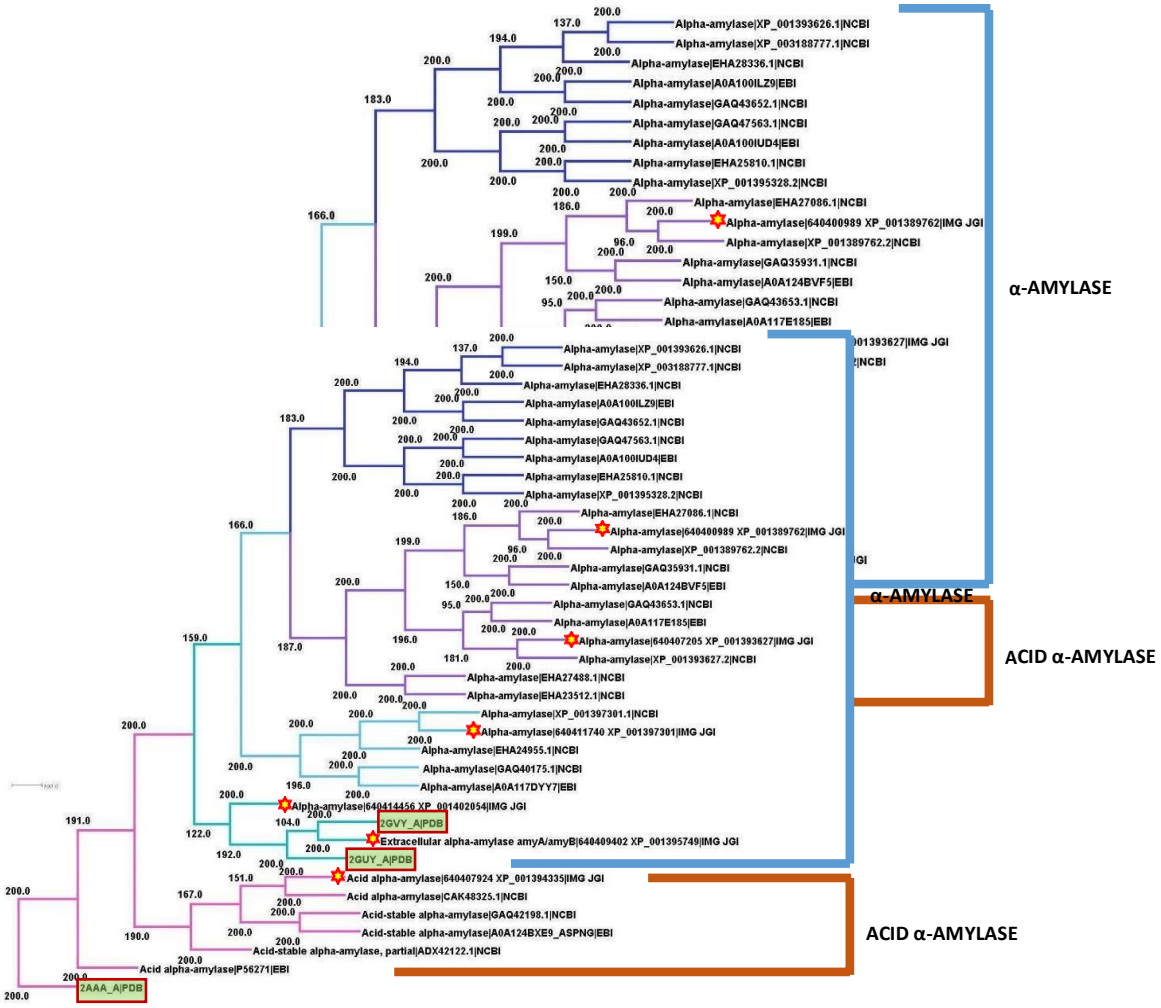

Fig 5.1. Phylogenetic tree of 36 α-amylase protein sequences from different *Aspergillus niger* strains, constructed with 100 bootstrap values.

Among the three X-ray crystallographic structures (2AAA, 2GUY and 2GVY), 2AAA was the acid α-amylase. So, to understand the differences between α-amylase and acid α-amylase, three structures were compared at their secondary and tertiary structure level. At secondary conformation two α-amylase 2GUY and 2GVY showed near about similar type of alpha helix and beta sheet conformations. But in comparison with them some positions

were not found in 2AAA acid α-amylase, like position 65 with short beta sheet, 150 and 270 with short alpha helix (Fig. 5.2).

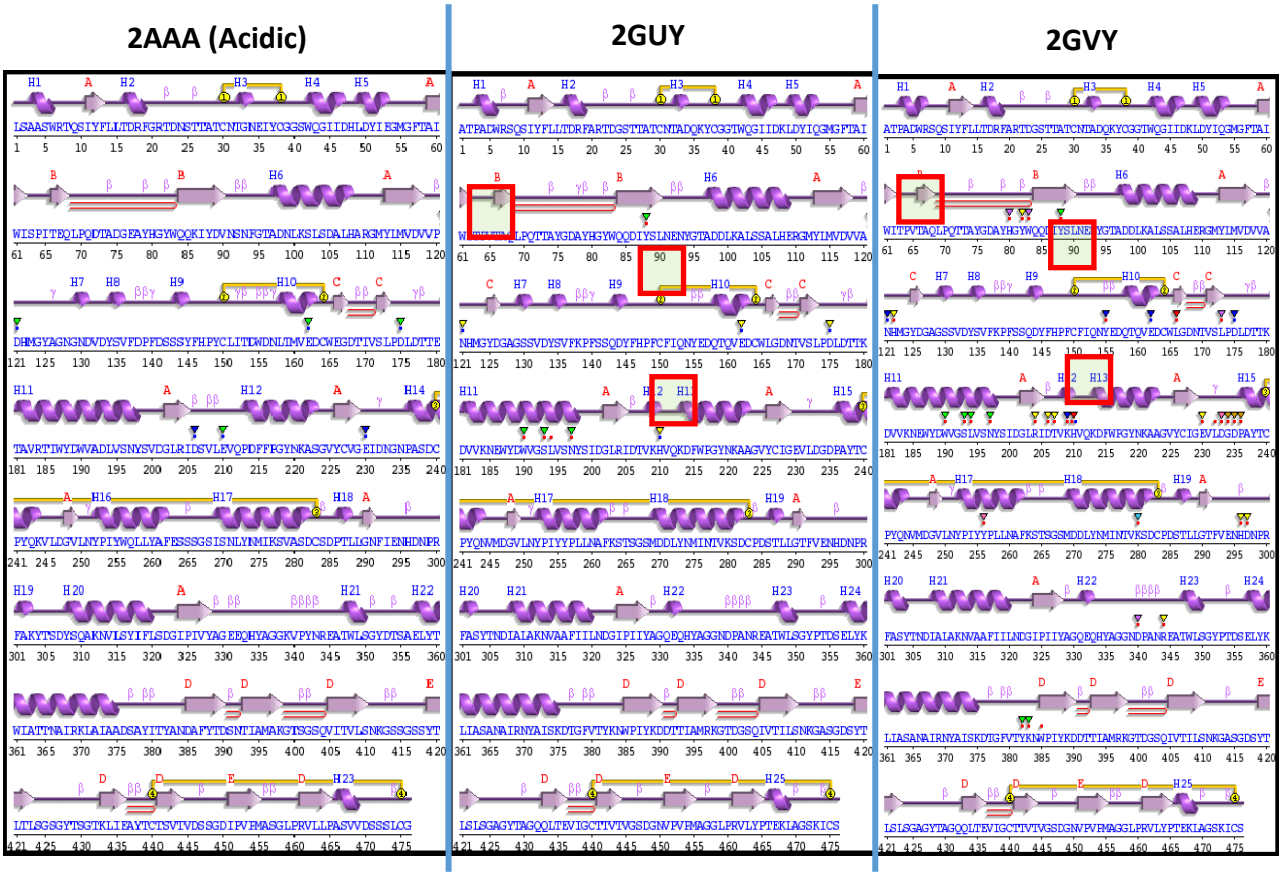

Fig 5.2. Phylogenetic tree of 36 α-amylase protein sequences from different *Aspergillus niger* strains, constructed with 100 bootstrap values

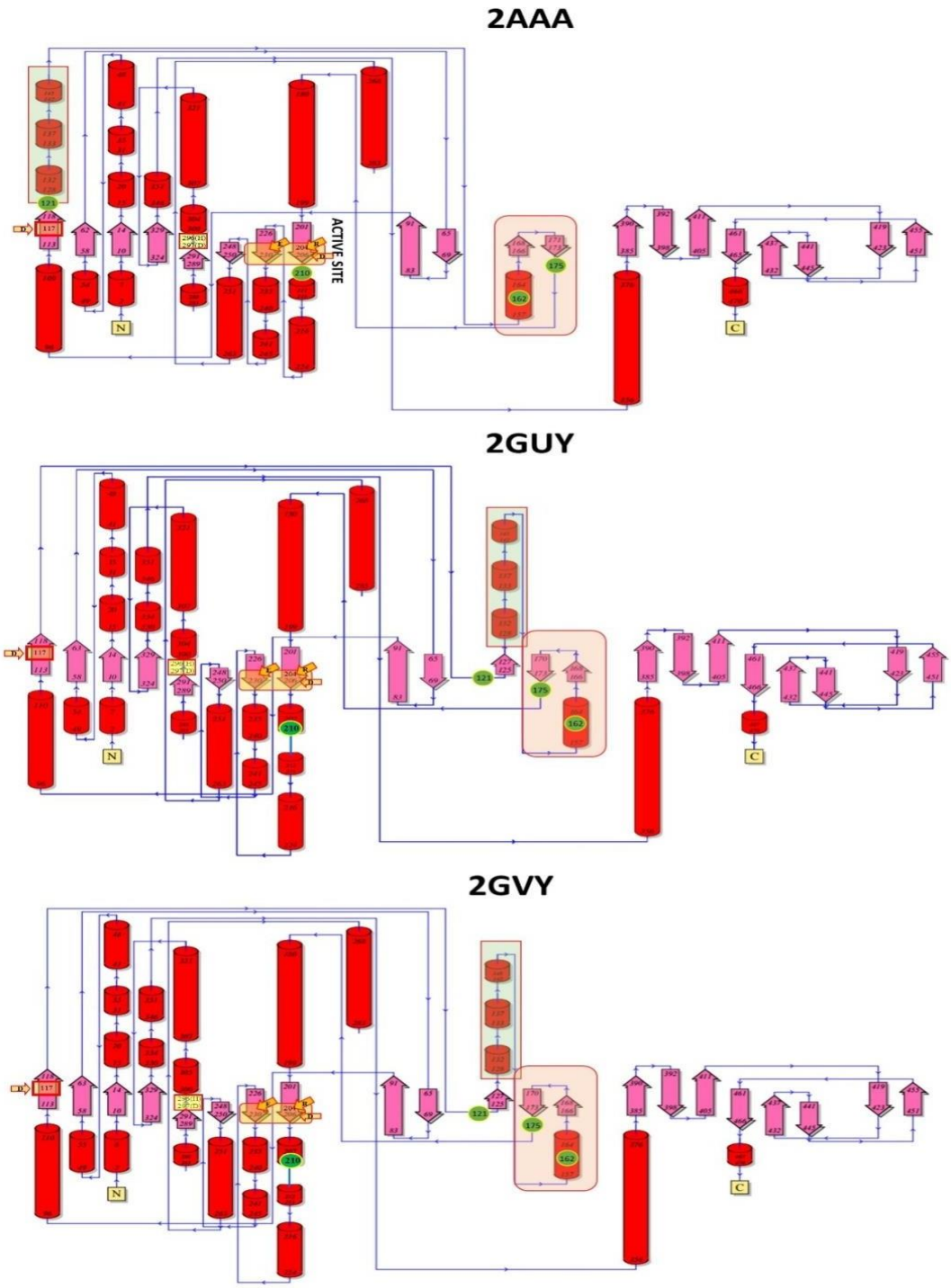

Fig 5.3. Topology structure of three x-ray crystallographic structures. Arrow indicating the active site residues and green bullet indicated the calcium (Ca^{2+}) binding residues.

showed diversity within acid α-amylase2AAA (Figure 5.3). In case of 2AAA, the alpha helix domain present in between the position 121-145 was found just before the beta sheet ended at the amino acid position 118. Whereas in both the case of normal α- amylase (2GUY and 2GVY) the above said alpha helix domain was found in the interior position of the structures. The orientation of alpha-beta barrel position 157-175 was also found in altered direction between acid α-amylase and α-amylase. Except all the diverged position, a conserved parallel and antiparallel beta-sheet structure was observed among all the structures, and according to the X-ray crystallographic structure analysis all the protein have their conserved catalytic amino acid residues like 117-D, 204-R, 206-D, 230-E, 296-H and 297-D (Fig. 5.3). Among these sites, first four positions were found within the β5, β8 and β9 positions of the respective structures. But 296-H and 297-D were found in the 11β-α loop position. However, change in topology position of calcium (Ca^{2+}) binding residue 121-N was observed for acid α-amylase (2AAA) in comparison with others (Fig. 6.3).So, from the above results it could be concluded that though acid α-amylases have sequential diversity in comparison to normal α-amylase, the active site and related structural conformations remain alike among all of them.

According to the literature Boel et al. (1990) proper folding of active site was dependent upon some conserved amino acids located within different positions of the protein. As change in topology position of Ca^{2+}binding amino acid 121-N was observed for acid α-amylase (2AAA), the site was analyzed further. Comparative study of those conserved Ca^{2+}binding amino acids among selected X-ray crystallographic structures revealed that the common amino acid His for the position 210 in 2GUY and 2GVY was replaced by Glu210 in acid α-amylase 2AAA (Fig. 5.4. A, B). Most interestingly, it was found that above replacement of His by Glu was remarkably present in case of all the acid α-amylases represented in figure 4B.

According to Suzuki et al. (1989) and Conrad et al. (1995) Ca^{2+} binding to the α-amylase increase the enzyme stability at changed pH. They also reported that modification or mutation at some position of the enzyme helps to increase the Ca^{2+} binding capacity as well as its pH stability at a range of 8-10.5. So, Ca^{2+}binding site play an important role in

stabilization of the enzyme in different pH. Among 20 amino acids, His is a basic amino acid, but Glu is known as the acidic amino acid. According to Li et al. (2008) more diversity of amino acids within functionally similar protein is the result of more interaction of proteins with different external environments. Here, change in pH influenced the adaptation of α-amylase in acidic medium by replacing the basic amino acid with an acidic one to maintain the exact structure of Ca^{2+}binding site. Similar type of amino acid alteration for all the acid α-amylase again signifies the pH mediated change at Ca^{2+}binding site.

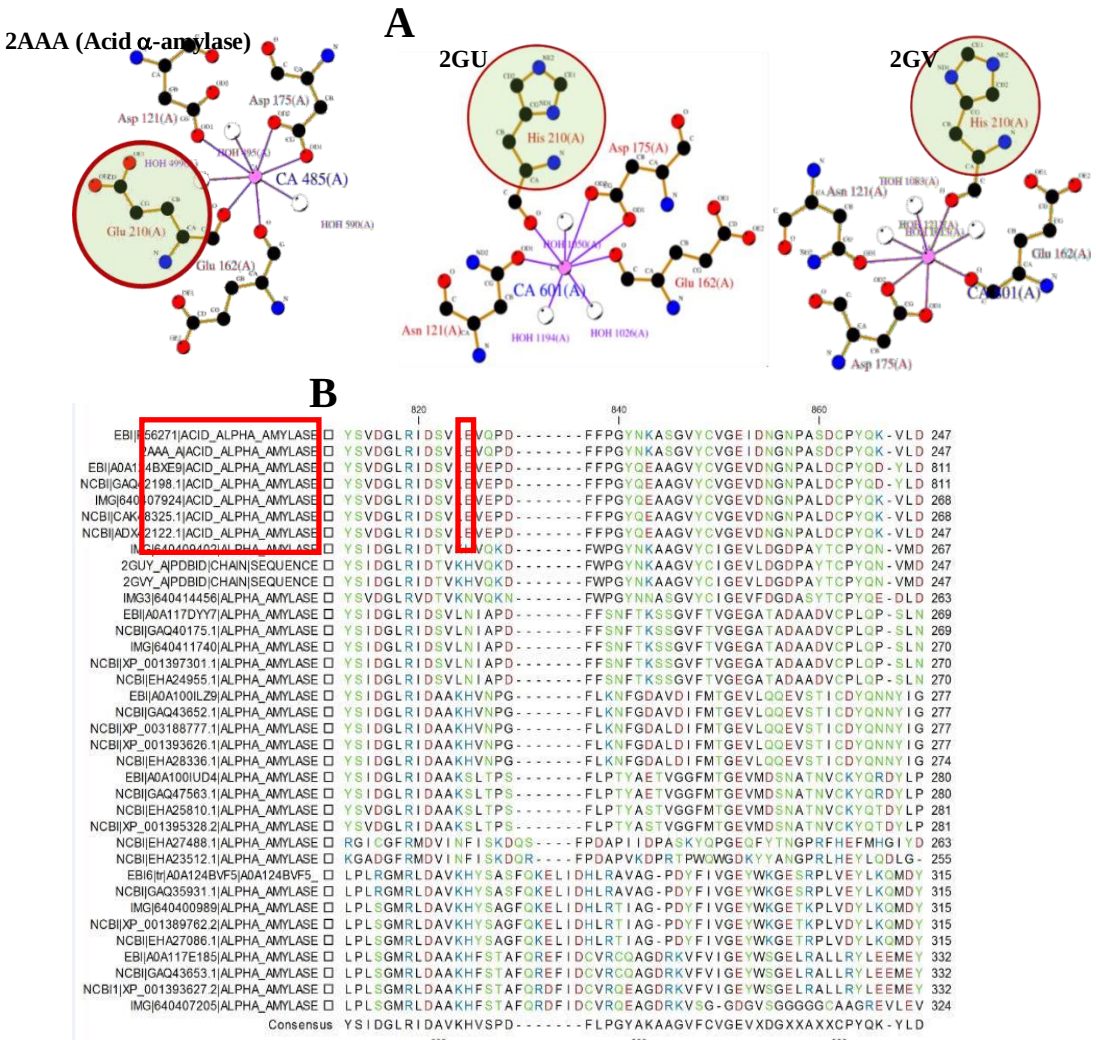

Fig 5.4. (A) Residues responsible for catalytic site formation on three X-ray crystallographic structures of α-amylase. (B) Multiple sequence alignment showing substitution of His (H) by Glu (E) at the alignment position 825 for all the acid α-amylase.

5.4. Conclusion

The molecular make up of cells from different genus are different. So, interaction of functionally similar protein produced by different genus are also different. The sequential diversity of functionally similar proteins from different genus, reflects the genetic adaptation. Whereas sequential diversity of functionally similar proteins from same genus as well as species reflects protein's interaction with their diversified external environment. In present study it was clearly observed that acidic environment governed the required changes among protein sequences of acid α-amylases from different strains of *Aspergills niger*. Beyond strain specificity environmental adaptation played an important role in α-amylase protein sequence diversity. The specific change of His to Glu at the Ca^{2+}binding amino acid residue within all acid α-amylase is an important clue to make genetically engineered potent α-amylase, which may functionally active in low pH condition.