International Journal of Genomics and Proteomics
ISSN: 0976-4887 & E-ISSN: 0976-4895, Volume 4, Issue 1, 2013, pp.-64-71.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000227

MOLECULAR MODELING OF α-AMYLASE FROM GERMINATED SOYBEAN (*Glycine max*) AND ITS FUNCTIONAL DIVERSITY

KUMARI A.[1], SINGH V.K.[2] AND KAYASTHA A.M.[1]*

[1]School of Biotechnology, Faculty of Science, Banaras Hindu University, Varanasi-221 005, UP, India.
[2]Centre for Bioinformatics, School of Biotechnology, Faculty of Science, Banaras Hindu University, Varanasi-221 005, UP, India.
*Corresponding Author: Email- kayasthabhu@gmail.com

Received: October 31, 2011; Accepted: April 15, 2013

Abstract- Starch hydrolyzing α-amylase was purified from germinating soybean seeds. Amino acid sequence of soybean α-amylase (Accession No. Gm0237X0071) was taken from protein databases (http://soybeangenome.siu.edu/; http://soybase.org/; http://soybeangenome.org/) and used for identification of full length gene with available clone data at NCBI (http://www.ncbi.nlm.nih.gov/). *Glycine max* strain Williams 82 clone GM_WBb0115J10 (AC235387.1) was used for similarity search and annotation of full-length gene. The present *in-silico* investigation deals with full length gene (TPA BK007878) identification; and *cis*-acting elements study; identified the important promoter's i.e. TATA, CAAT, GATABOX, DOFCOREZM, -300ELEMENT, WBOX, MYBST1, and EBOX for multifarious uses. A template structure (3AMK chain a) from *Oryza sativa* branching enzyme was selected for comparative modeling using an automated approach. Homology model was constructed using software DS Modeler and the quality of refined model was investigated using PDBSum, ERRAT and other bioinformatics softwares. The modeled structure showed acceptable Ramachandran statistics and remarkable active site residues. Structural analysis of the predicted model of α-amylase from soybean also gives an idea about potential sites inferring the region of catalytic active site responsible for inhibitory action; and opens the new opportunities for further investigations.

Keywords- α-Amylase, *Glycine max,* Homology modeling, Ramachandran plot, DS Modeler

Citation: Kumari A., Singh V.K. and Kayastha A.M. (2013) Molecular Modeling of α-Amylase from Germinated Soybean (*Glycine max*) and Its Functional Diversity. International Journal of Genomics and Proteomics, ISSN: 0976-4887 & E-ISSN: 0976-4895, Volume 4, Issue 1, pp.- 64-71.

Introduction

α-Amylase (α-1,4 glucan-4-glucanohydrolase, EC 3.2.1.1) categorized as family 13th of glycosyl hydrolase, functions the endohydrolysis of α,1→4 glycosidic linkages present in various polysaccharides [1-3]. The end product of its action constitutes glucose, maltose and oligosaccharides with varying length with an α-configuration [4]. They are ubiquitous in nature occupy presence in all sections of life viz. plants, animals and microbes [5].

α-Amylases are composed of three structural domains. The largest is Domain A which forms a central eight-stranded $(\beta/\alpha)_8$-barrel, to one end of which are located the active site residues [4, 6-7]. Domain B serves to form a calcium binding site against the wall of the $(\beta/\alpha)_8$-barrel of Domain A. Domain B is probably responsible for the differences in substrate specificity and stability among the α-amylases [8]. Domain C is made up of antiparallel β-structure and is only loosely associated with Domains A and B [9-10]. Residues equivalent to Asp-206, Asp-297, Glu-230 (catalytic site residues) and His-122, His-296 (part of substrate binding site) are conserved in all α-amylases [7,11-13]. Calcium ion, which is located at the interface between the A and B domains, is conserved in all α-amylases with known three dimensional structures [14-17].

Knowledge of the three-dimensional structure of protein complexes provides a valuable understanding of the function of molecular systems [18]. Protein structure determination is growing rapidly as can be seen from the large structural data available in Protein Data Bank (PDB) [19]. However, determination of the structure of protein complexes still remains a difficult task because of the experimental hurdles. Thus, computer algorithms and web server are used to predict the structure of the proteins that have not all been solved structurally. Homology modeling can be used to predict model structure of unknown protein (target) from its amino acid sequences based on the structure of related protein of known structure (template) [20-23]. Four steps are used to create a model (i) template selection (ii) alignment of template sequence with the target (iii) model construction (iv) evaluation of generated model [24].

Recently, Starch hydrolyzing α-amylase from germinated soybeans seeds (*Glycine max*) has been purified 400-fold to electrophoretic homogeneity with a final specific activity of 384 units/mg [25]. Peptide map obtained by MALDI along with its molecular mass information was used for the confirmation of the soybean α-amylase. In the present study, homology modeling of α-amylase from soybean was done using DSMODELER [26]. This software constructs three-dimensional structure of the protein to show the functional and binding site with their specific domains. Predicted model was evalu-

ated with the lowest value of PROCHECK statistics (http://nihserver.mbi.ucla.edu/SAVES_3/saves.php) and quantitative measures of the predicted model building was carried out using Qmean (http://swissmodel.expasy.org/qmean/cgi/index.cgi) and Vadar (http://redpoll.pharmacy.ualberta.ca/vadar/) servers.

Methodology

Mass Spectrometry, Database Searching and Sequence Alignment

Soybean α-amylase was purified according to Kumari et al. [25] and analyzed by matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS; Ultraflex III, Brüker-Daltonik, Bremen, Germany). Single band on SDS-PAGE was excised and was digested overnight with MALDI grade trypsin (Sigma). The digested mixture was analyzed by mass spectrometry to produce a peptide mass spectrum from which molecular masses of all of the proteolytic fragments can be read. The MASCOT software [27] was used to compare the peptide mass patterns obtained with those of all proteins from the theoretical soybean α-amylase proteome. The molecular weight search (MOWSE) scoring scheme [28] was used for unequivocal identification of proteins. Since α-amylase from soybean is already available in protein database (http://Glycine max genome.siu.edu/; http://soybase.org/); so the peptide masses as obtained with mass spectrometry was used for the confirmation of the protein.

Sequence of soybean α-amylase was aligned with the known sequence in database using the ClustalW [29]. These known sequences were taken from http://www.expasy.org/enzyme/3.2.1.1. It gives clear information about the best match for the given sequence and aligned them accordingly. There is no similar protein sequence available in NCBI public database but tblastn study gives an idea about the availability of α-amylase gene at genomic level so we did comparative study based on existing α-amylase sequence obtained from (http://www.matrixscience.com) with accession no. Gm0237X0071. It showed similarity with *Glycine max* strain Williams 82 clone GM_WBb0115J10 (AC235387.1). Full length gene prediction was done for obtained strain using Fgenesh (http://sun1.softberry.com/berry.phtml) and -1000 upstream region was retrieved for promoter study using PLACE [30].

Sequence Analysis, Template Search and Model Generation

Sequence of α-amylase from soybean was aligned with diverse species using ClustalW and phylogenetic tree was constructed using UPGMA method [31,32]. The conserved motifs present in these sequences were analyzed using MEME (Multiple EM for Elicitation) [33]. Parameters have been set with number of different motifs: 30, minimum motif width: 100 and maximum motif width: 300, manually. The result obtained from PDB advance search showing their structural similarity with branching enzyme from *Oryza sativa* (PDB ID: 3AMK, chain a) at a resolution of 2.30 Å [34].The three dimensional structure of α-amylase from soybean has been predicted using DSMODELER [26].

Refinements and Evaluation

Energy minimization of the generated model was done using the Conjugate Gradient technique to remove the bad contacts between protein atoms. Model was evaluated according to the lowest energy value of the objective function and the backbone conformation of the model using DSMODELER [26].

Phi/Psi Ramachandran plot obtained with stereo chemical properties were assessed using PROCHECK server (http://nihserver.mbi.ucla.edu/SAVES_3/saves.php) [35-36]. Structural comparison between template and target structure were calculated using PROSA server (https://prosa.services.came.sbg.ac.at/prosa.php) and (ii) UCSF Chimera 1.5.1. UCSF Chimera 1.5.1. (http://www.cgl.ucsf.edu/chimera/) is a structural comparison matchmaker tool based on Needle-Wunsch Alignment algorithm BLOSUM -62 matrix [37]. QMEAN server (http://swissmodel.expasy.org/qmean/cgi/index.cgi) model quality estimation was used to analyzed QMEAN score / QMEANclust score, Residue error, Energy profiles and plot and Volume area dihedral angle for fractional accessible surface area, residue volume, 3D profile and stereo/packing quality index were done with VADAR (http://vadar.wishartlab.com/). QMEAN and VADAR were specially designed for quantitatively and qualitatively assessing protein structures determined by X-ray crystallography, NMR spectroscopy, 3D-threading or Homology modeling.

Active Site Prediction

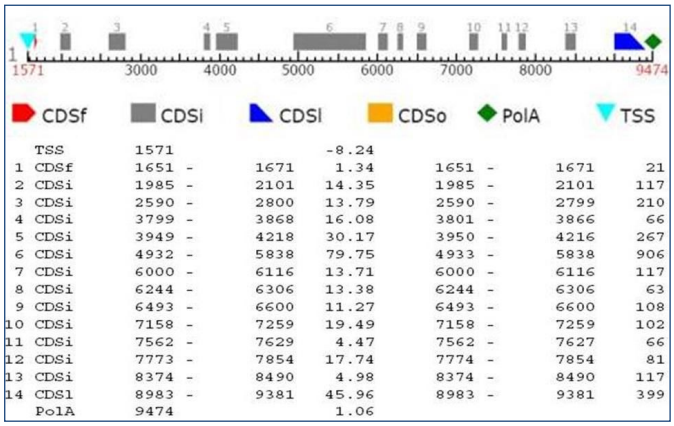
Possible binding sites of the final obtained model were searched using Q-site Finder [38]. Q-site Finder determines the potential sites for ligand binding in docking calculations. Obtained binding sites were compared to the active site of the template for the determination of residues forming the binding pocket and to find their role in biological functions.

Results

Mass Spectrometry and Database Searching and Sequence Alignment

Peptide mass spectrum obtained from the digestion of purified soybean α-amylase by trypsin provides fingerprint of great specificity. This peptide masses significantly revealed 9 tryptic peptides masses corresponding to α-amylase from soybean with accession no. Gm0237x00071 with 17% sequence coverage [25]. The obtained full length peptide sequence of soybean α-amylase was used for sequential and structural alignment. After sequence alignment using tblastn study it was observed that soybean α-amylase showed similarity with soybean (*Glycine max*) strain Williams 82 clone GM_WBb0115J10 (AC235387.1) and matches at position 104382 to 109759. Complete full length gene was fetched out and submitted to NCBI and obtained TPA Accession: TPA BK007878. Complete hypothetical mRNA showed the combination of gene with 14 CDS, TSS site at 921 position and Poly A site at 8824 position [Fig-1A]. *cis*-acting element study showed the gene containing regulatory promoters, as listed in [Table-1]. Seed germination specific promoter has been prominently identified in upstream region. Important promoters investigated were TATA, CAAT, GATABOX, DOFCOREZM, -300ELEMENT, WBOX, MYBST1 and EBOX. The length of α-amylase for soybean is 741 amino acids with a molecular mass of 84 kDa for (Accession No. Gm0237X0071) whereas, predicted α-amylase gene (i.e. TPA BK007878) containing 883 amino acids residues showed a molecular mass of 100 kDa. It was observed that 157 amino acids were missing in purified soybean α-amylase sequence having an accession no. Gm0237X0071. The sequence was:

MAESLTIIVRSKQYLATQKPVNLALGYRNPHGYGFSFGSRRSI-
HERVSSHFKGIAVMTDDKSTMSSTEEDLENIGIFHIDPSLKPYKDH
FKYRLKRYVDQKKLIEEYEGGLEEFSQGYLKFGFNREEGGIVYCE
WAPAAQEAQIIGDFNGWDGSNHQ



	TSS	1571		-8.24			
1	CDSf	1651	-	1671	1.34	1651 - 1671	21
2	CDSi	1985	-	2101	14.35	1985 - 2101	117
3	CDSi	2590	-	2800	13.79	2590 - 2799	210
4	CDSi	3799	-	3868	16.08	3801 - 3866	66
5	CDSi	3949	-	4218	30.17	3950 - 4216	267
6	CDSi	4932	-	5838	79.75	4933 - 5838	906
7	CDSi	6000	-	6116	13.71	6000 - 6116	117
8	CDSi	6244	-	6306	13.38	6244 - 6306	63
9	CDSi	6493	-	6600	11.27	6493 - 6600	108
10	CDSi	7158	-	7259	19.49	7158 - 7259	102
11	CDSi	7562	-	7629	4.47	7562 - 7627	66
12	CDSi	7773	-	7854	17.74	7774 - 7854	81
13	CDSi	8374	-	8490	4.98	8374 - 8490	117
14	CDSi	8983	-	9381	45.96	8983 - 9381	399
	PolA	9474			1.06		

Fig. 1A- Full length gene prediction of α-amylase.

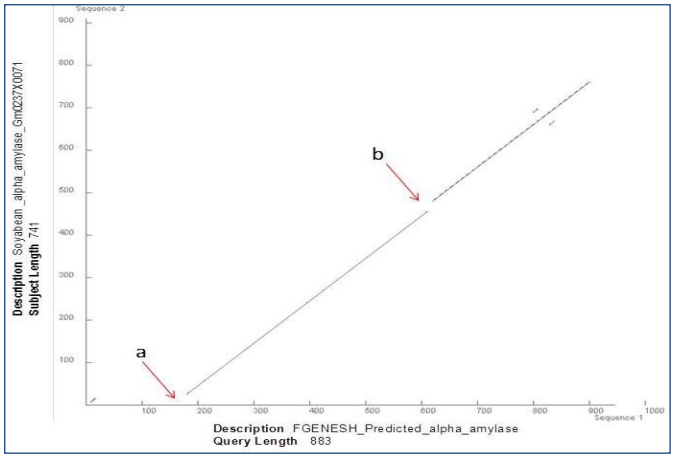
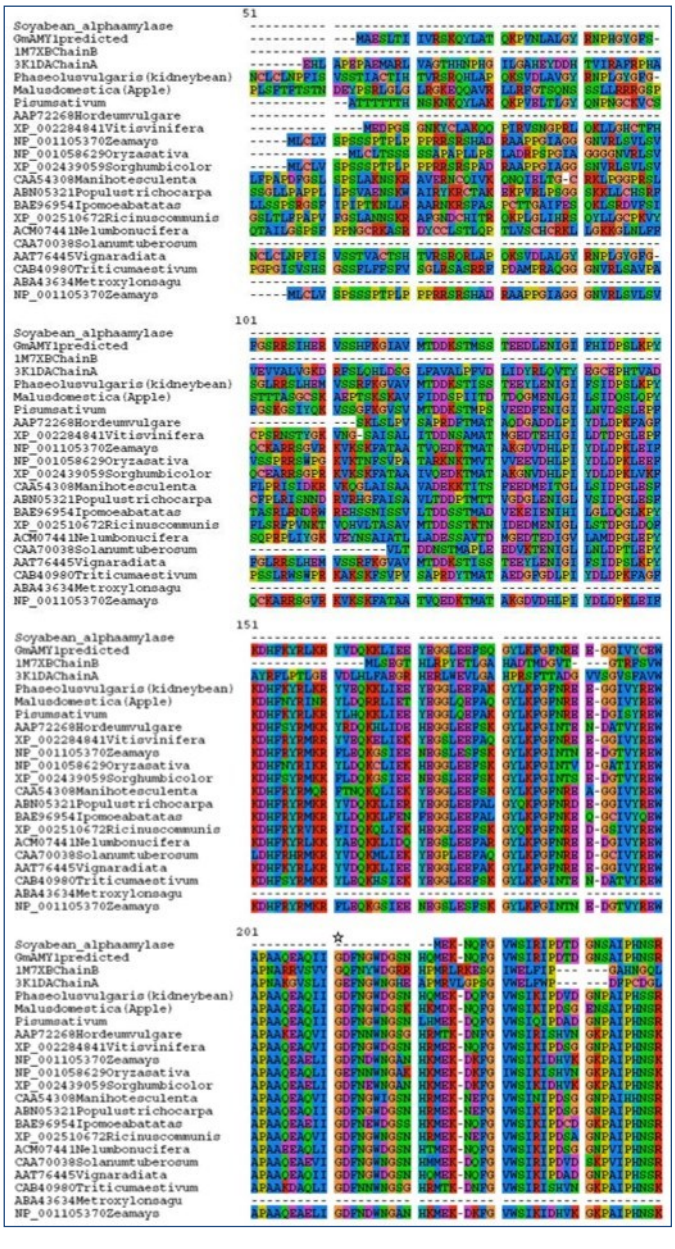


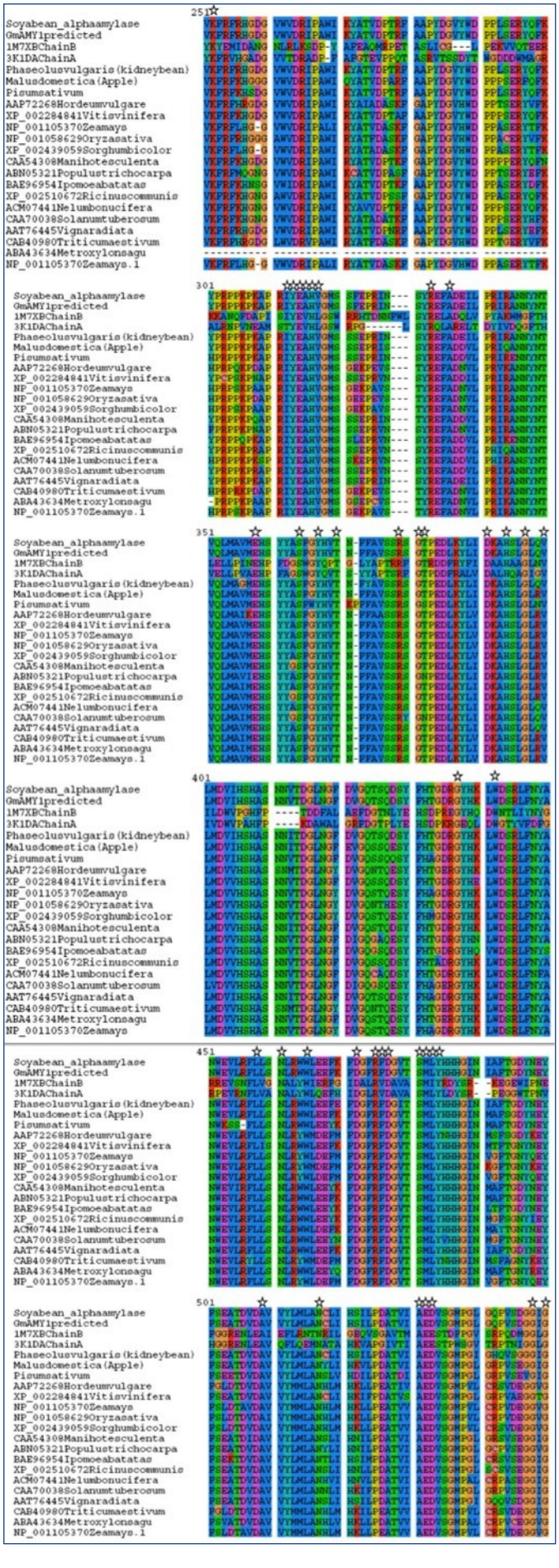
Fig. 1B- Dot plot analysis for soybean α-amylase and predicted amylase (Gm0237X0071, TPA BK007878). Arrow (a) indicates the missing fragment; arrow (b) indicates difference in residues

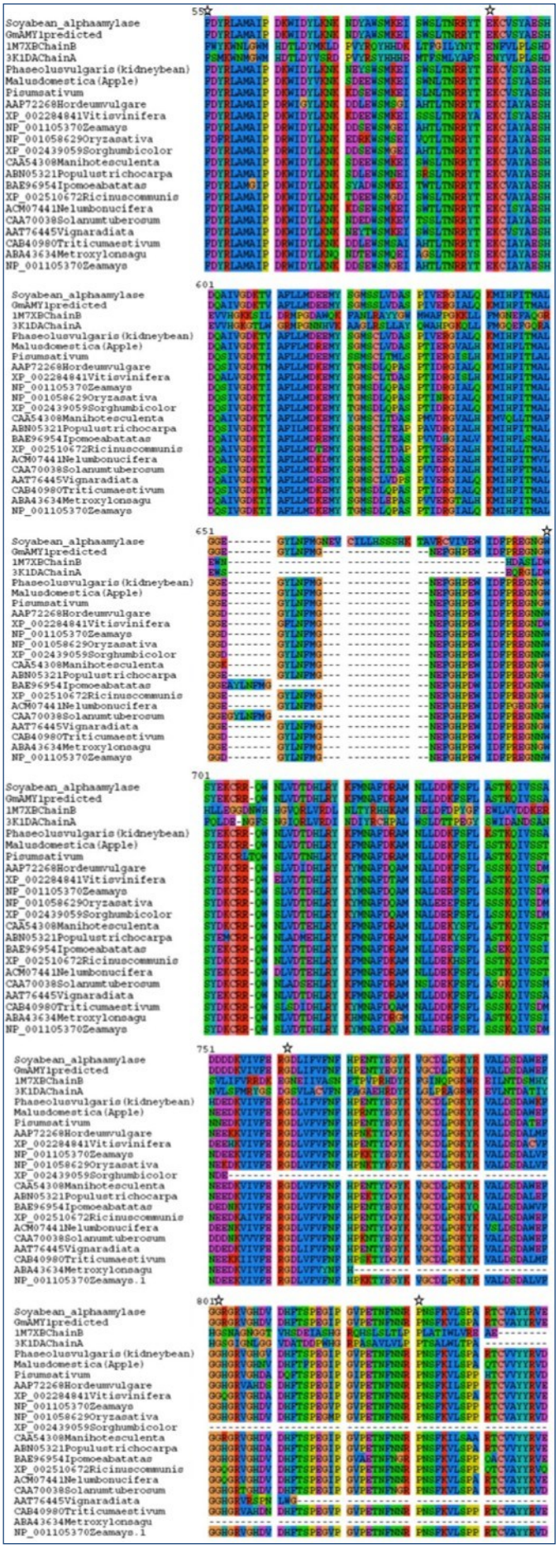
Table 1- List of identified regulatory elements

Promoter	Sequence	Function	PMID
GATBOX	GATA	Light regulation	2535536
CAATBOX 1	CAAT	Tissue specific promoter activity	2710102
DOFCOREZM	AAAG	DNA binding proteins	10074718
-300ELEMENT	TGHAAARK	Identification of an enhancer element for the endosperm-specific expression	2152160
WBOXHVISO1	TGACT	Sugar-responsive elements	12953112
MYBST1	GGATA	Transcriptional activator	7957104
EBOXBNNAPA	CANNTG	Light-responsive and tissue-specific	15821875
AAR1AT	NGATT	Response regulators operate as transcriptional activators.	11135105
TATABOX5	TTATTT	cis elements and trans-acting factors	7630938
TBOXATGAPB	ACTTTG	Promoter analysis of the nuclear gene	11442054
NODCON2GM	CTCTT	Nodule specific genes from soybean.	11442054
PYRIMIDINE BOXOSRAMY1A	CCTTTT	Functional dissection of a sugar-repressed α-amylase gene RT (Ramy1A) promoter in rice embryos	9506846

There was also difference in fragment CILLHSSSHKTAVRCVIV in case of Gm0237X0071 [Fig-1B]. Based on sequence similarity we find 21 similar sequences from different species, retrieved sequences were used for multiple sequence alignment for checking the diversity among predicted α-amylase and other diverse species. By alignment study it was observed that there were unique consensus regions at various places as shown in [Fig-2]. Residues equivalent to Met-1; Val-28; Ile-89, Tyr-90; Glu-91, Ala-92, His-93; Val-94; Gly-95; Arg-107, Ala-110; Glu-132; Ser-138; Try-141; Val-143; Thr-144; Phe-152; Asp-161; His-164; Gly-167; Val-170; Gly-207; Trp-212; Leu-228; Asn-231; Trp-235; Asp-242; Ser-251; Met-252; Lys-253; Try-254; Ala-279; Asn-287; Ala-301; Glu-302; Asp-303; Gly-318; Gly-320: Phe-321; Glu-361; Trp-462; Gly-524; Gly-564 and Pro-593 are conserved in soybean amylase (Gm0237X0071), whereas residue equivalent to Met-158; Val-185; Ile-246; Tyr-247; Glu-248; Ala-249; His-250; Val-251; Gly-252; Arg-264; Ala-267; Glu-289; Ser-295; Tyr-298; Val-300; Thr-301; Arg-309; Ser-310; Gly-311; Thr-312; Asp-32; His-324; Gly-327; Val-330; Gly-367; Trp-372; Leu-388; Asn-391; Trp-395; Asp-402; Ser-411; Met-412; Lys-413; Try-414; Ala-439; Asn-447; Ala-461; Glu-462; Asp-463; Gly-478; Gly-480; Phe-481; Glu-521; Trp-608; Gly-669; Gly-709 and Pro-738 are well-conserved in predicted soybean amylase gene (TPA BK007878). Two major clusters were obtained from phylogenic tree based on α-amylase sequences from different species. Both the α-amylase (i.e. Gm0237X0071 and TPA BK007878) are closer to cluster A species i.e. *Phaseolus vulgaris, Vigna radiata*, 3AMK_chain A of *Oryza sativa,* respectively [Fig-3]. 11 unique motifs have been identified after MEME study [Fig-4A], [Fig-4B]. Motif diversity showed their evolutionary significance, which are inferred by phylogenetic study.







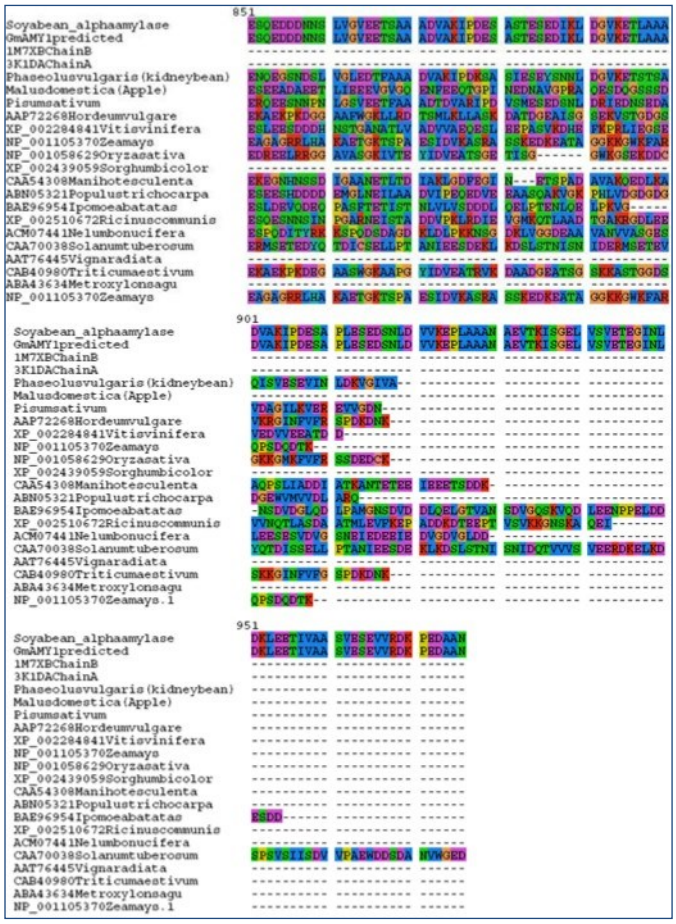


Fig. 2- Sequence alignment of soybean α-amylase with predicted amylase and with 21 diverse amylases from different species. Asterisk denotes the consensuses amino acid residue

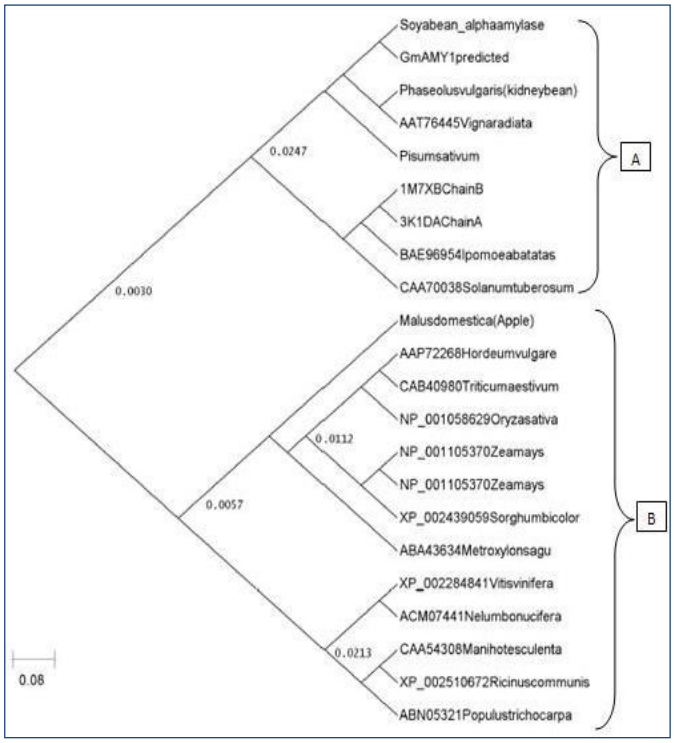


Fig. 3- Phylogenic tree based on α-amylase sequences from different species.

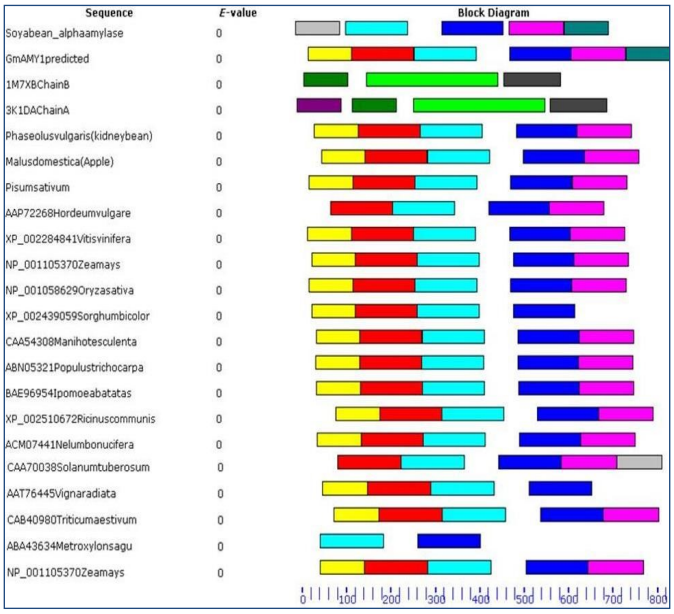


Fig. 4a- Schematic distribution of respective conserved motifs; identified by means of MEME software

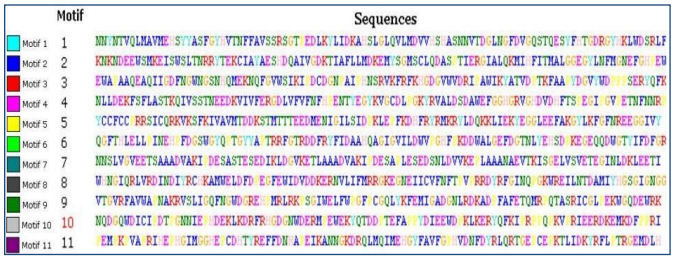


Fig. 4b-Multilevel consensus sequences for the MEME defined motifs; among different species

Homology Modeling of α-Amylase from Soybean

To the best of our knowledge three dimensional structure for α-amylase from soybean is not known till date. The 3D structure of soybean α-amylase can be used to predict its molecular functions, active sites, and its interactions with suitable inhibitors. Three models were built by DSMODELER based on CHARMm force field using Conjugate Gradient (CONJ) [26] method that exhibits better convergence than the steepest descent method. Three distinct domains organization is fairly-well conserved in the generated model [Fig-5]. The predicted model of α-amylase from soybean consists of 14 α-helices and 17 β-sheets [Fig-6].

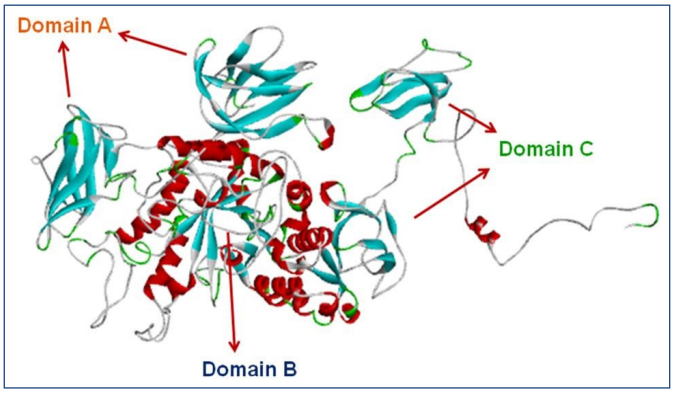


Fig. 5- The final 3D structure of α-amylase from soybean with three distinct domains A, B,and C as obtained after energy minimization.

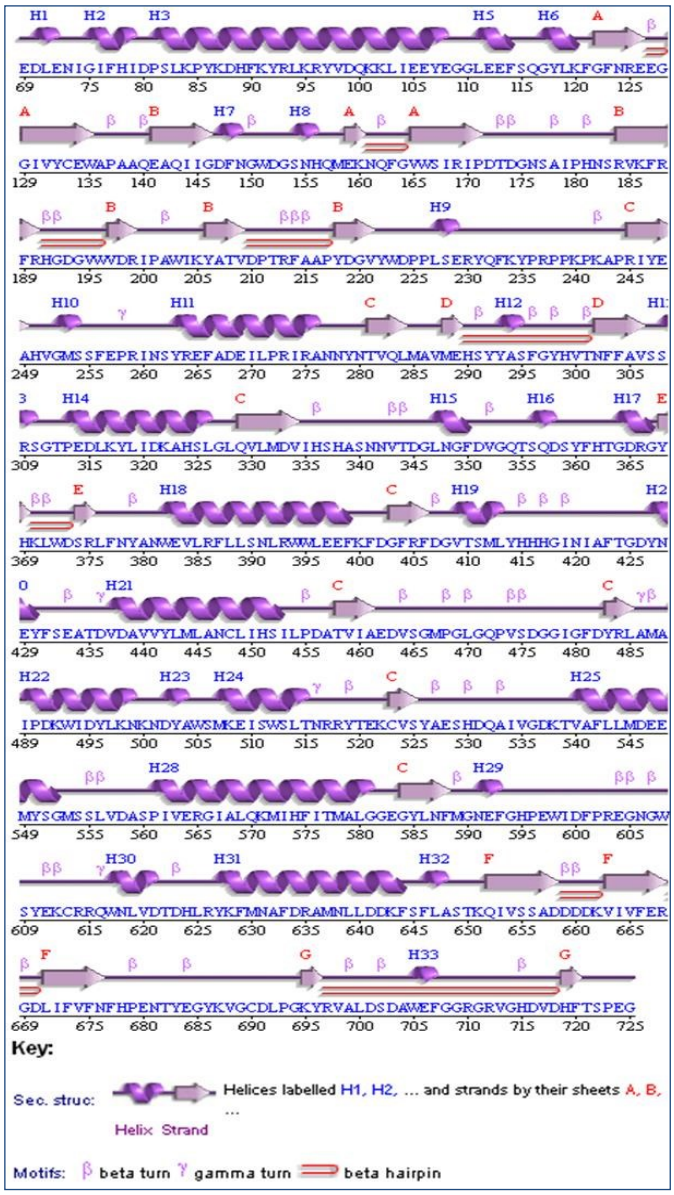


Fig. 6- Secondary structure prediction for generated model

Validation of the Predicted Structure

The stereo chemical quality of the predicted structure was assessed after energy minimization. Energy was-12828.653 after energy minimization. The total numbers of residues in selected model were observed as: 94.4% residues in most favored region, 4.7% in additional allowed region, 0.9% in generously allowed region and no residues in disallowed region in Ramachandran plot statistics [Fig-7A] [Table-2]; which indicates that generated model quality is good. QMEAN score / QMEANclust score of the whole model reflecting the predicted model reliability ranging from 0 to 1. In this predicted model QMEAN score 0.716 with global scores estimated absolute quality Z-score: -0.52 result show that model is reliable. Fractional accessible surface area volumes of all residues close to 1.0±0.1, Statistics of hydrogen bonds of predicted model show equal to expected mean hbond distance score 2.2 sd=0.4, mean hbond energy observed -1.7 sd=1.0 (74% residues) closest to expected -2.0 sd=0.8 (expected 75% residues). Dihedral Angles were observed closest to expected, Total Accessible Surface Area score 27365.1 Angs. with expected score 22651.6 Angs., Total volume (packing) score observed 93847.6 Angs. with expected 92249.8 Angs., Stereo/Packing, 3D quality index results shows that less than 1% of error residues in predicted model. [Fig-7B], [Fig-7C]. Over all quality score was 82.866 for predicted models. PROSA score for Template and Target were -10.48 and -10.89, respectively [Fig-8A]; whereas root mean square deviation of 0.160Å between 651 atom pairs was observed UCSF Chimera 1.5.1. [Fig-8B]. Structural comparison between template and target structures show less deviation at atomic level. Validated and refined α-amylase model from soybean was successfully submitted to Protein Model Database (http://mi.caspur.it/PMDB/) with PMDBID PM0078685.

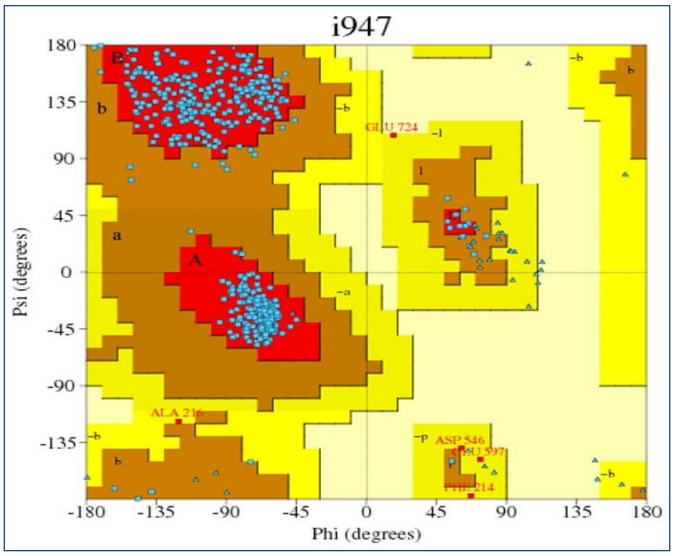


Fig. 7A- Ramachandran plot, based on most favoured, allowed generous and disallowed residues. Quantitative measures of the model protein

Table 2- Ramachandran Plot statistics after energy minimization

S. No.	Description	% Value
1	Residues in most favoured regions	94.40
2	Residues in additional allowed regions	4.70
3	Residues in generously allowed regions	0.90
4	Residues in disallowed regions	0.00

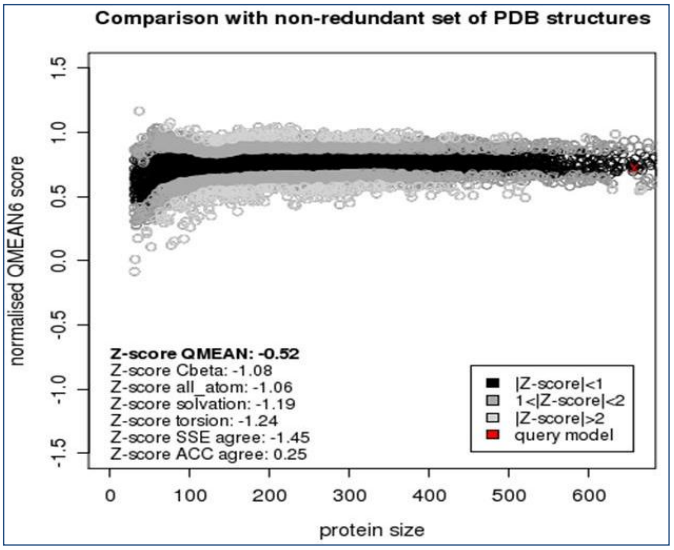


Fig. 7B- QMEAN score and Global scores estimated absolute quality Z-score

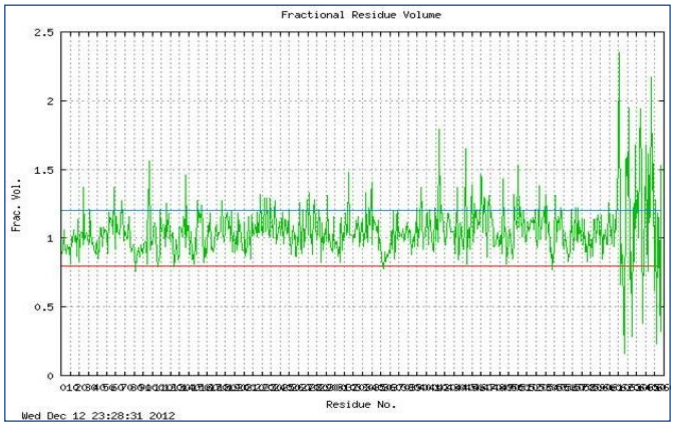


Fig. 7C- Volume area dihedral angle reporter results from VADAR.

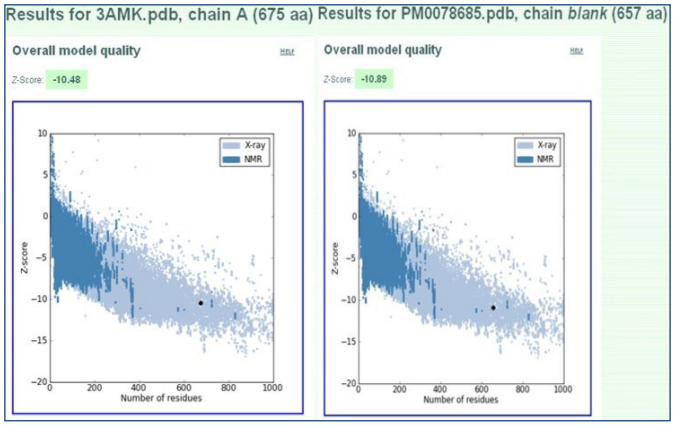


Fig. 8A- Superimposition of modeled structures of α-amylases from soybean with the template 1M7X using the combinatorial extension (CE) method (Template structure in red colour and modeled structure in green colour).

Fig. 8B-Superimposition of modeled structures of α-amylases from Glycine max with the templet 1M7X using UCSF Chimera 1.5.1.

Active Site Identification

Active site was calculated by using active site finder tool of Q-sitefinder. A total of 10 active sites have been predicted [Fig-9] and after sequence comparison it has been determined that showed that the best possible binding sites Gm0237X0071 were-His- 91,Val -92, Gly-93, Met-94, Ser-95, Ser-96, Phe-97, Glu-98, Asn-102, Ser-103, Tyr-104, Phe-107, Tyr-134, Ser-148, Arg-150, Ser-151, Gly-152, Lys-468, where as in case of TPA BK007878 structure; the best possible sites showing the possible binding regions were Phe-114, Tyr-118, Glu-134, Trp-135, Ala-136, Phe-189, Arg-199, Thr-208, Met-288, Val-300, Thr-301, Asn-302, Phe-303, Phe-304, Ala-305,Val-334, Ile-335, His-336, His-338, Ser-340, Asp-345, Gly-346, Asn-348,GLY349, Phe-350, Asp-351, His-369, Lys-370, Phe-377, Val-384, Phe-387, Leu-388, Leu-389, Asn-391, Leu-392 and Tyr-442. With comparison with template protein structure (3AMK.pdb) Lys-101, Lys-234, Arg-237, Lys-322, Ser-325, His-369, Leu-371, Trp-372, Ser-411, His-417, Phe-423, Tyr-430, Val-464, Ser-465, Lys-498, Asn-499, Lys-500, Asp-502, Pro-561, Lys-628, Asn-631, Ala-632, Pro-679 and Arg-748 residues were found to be best active binding sites.

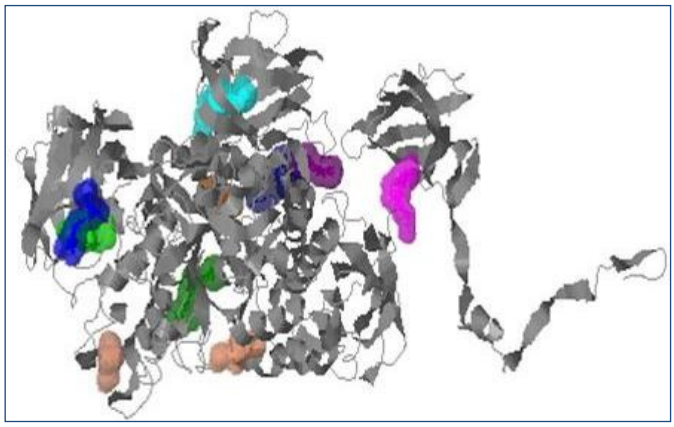


Fig. 9- Active site identification in predicted structure

Discussion

Predicted gene of α-amylase contains 13 introns and 14 exons. *Cis*-acting element study was carried out by retrieving promoter region and number of regulatory elements were observed closely related with 3 main physiological phenomenon i.e. endosperm specificity, stress response and hormone response. Earlier studies revealed that despite considerable structural diversity, all α-amylase from different sources shared a similar topology and fold. α-Amylases from soybean have similar domain topology and conserved regions as predicted by sequence alignment and homology modeling. Homology modeling can be used to generate 3D model of unknown protein from its homologous protein. Predicted 3D structure of the protein gives important information which is based on the quality of the model generated. This can be useful for drug discovery processes i.e. for the selection of target protein for which drug has to develop, for designing mutants and to identify active and binding site residues [39-40]. α-Amylase from soybean showed higher similarity with branching enzyme from *Oryza sativa* (PDB ID: 3AMK, chain a). The template and the target have 76% identity with 89% positive identity with 0 E-value and score was 1169.07bits (3023). The template structure consists of three major domains, an NH_2-terminal seven-stranded β-sandwich domain, a COOH-terminal domain, and a central (β/α) barrel domain containing the enzyme active site. The central domain is similar to that of all the other amylase family enzymes [34].

The structurally conserved regions were determined by multiple sequence alignment, based on the Needleman and Wunsch Algorithm [37]. Homology modeling provides high quality structure alignment for structure prediction. Phi and Psi torsion angles of predicted models were checked by PROCHECK. The RMSD of the equivalent C_α atoms of the modeled structure was 0.161 Å. In Ramachandran plot analysis, no residues were found in disallowed region. The Ramachandran plot for the model showed 100% of the residues in

the allowed regions and contains non-glycine and non-proline residues. Structural alignment was done using Swiss-PDB viewer (http://spdbv.vital-it.ch/**)** to calculate the structure deviation between template (3AMK) and selected model of α-amylase from soybean. Active site prediction showed that residues Lys-101, Lys-234, Arg-237, Lys-322, Ser-325, His-369, Leu-371, Trp-372, Ser-411, His-417, Phe-423, Tyr-430, Val-464, Ser-465, Lys-498, Asn-499, Lys-500, Asp-502, Pro-561, Lys-628, Asn-631, Ala-632, Pro-679 and Arg-748 were well-conserved and containing active binding site for TPABK007878. This model can be helpful for generating hypotheses and to explore and design new potent inhibitors of soybean α-amylase.

Acknowledgments

AK would like to thank University Grants Commission (UGC, India) for a fellowship. SUB-DIC, Centre for Bioinformatics, School of Biotechnology is thankfully acknowledged.

References

[1] Robyt J.F. and French D. (1970) *Journal of Biological Chemistry*, 245, 3917-3927.

[2] Thoma J.A., Spradlin J.E. and Dygert S. (1971) *The Enzymes*, 5, 115-199.

[3] Larson S.B., Greenwood A., Cascio D., Day J. and McPherson A. (1994) *Journal of Molecular Biology*, 235, 1560-1584.

[4] Van der Maarel M.J.E.C., Van der Veen B., Uitdehaag J.C.M., Leemhuis H., Dijkhuizen L. (2002) *Journal of Biotechnology*, 94, 137-155.

[5] Hagenimana V., Vezina L.P. and Simard R.E. (1992) *Journal of Agriculture Food Chemistry*, 40, 1777-1783.

[6] Svensson B. and Sogaard M. (1991) *Society Transactions*, 20, 30-42.

[7] Brayer G.D., Luo Y. and Withers S.G. (1995) *Protein Science*, 4, 1730-1742.

[8] Svensson B. (1994) *Plant Molecular Biology*, 25, 141-157.

[9] Nielsen P.K., Bonsager B.C., Fukuda K., Svensson B. (2004) *Biochimica et Biophysica Acta*, 1696, 157-164.

[10] Buisson G., Duée E., Haser R., Payan F. (1987) *EMBO Journal*, 6, 3909-3916.

[11] Rogers J.C. (1985) *Biochemical and Biophysical Research Communications*, 128, 470-476.

[12] Swift H.J., Brady L., Derewenda Z.S., Dodson E.J., Dodson G.G., Turkenburg J.P., Wilkinson A.J. (1991) *Acta Crystallography*, 47, 535-544.

[13] Kadziola A., Abe J., Svensson B., Haser R. (1994) *Journal of Molecular Biology*, 231, 104-121.

[14] Kumar S.S.R., Singh S.A. and Appu-Rao A.G. (2009) *Biochimie*, 91, 548-557.

[15] Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G., Jensen H., Thim L. and Woldike H.F. (1990) *Biochemistry*, 29, 6244-6249.

[16] Machius M., Declerck N., Huber R., Wiegand G. (1998) *Structure*, 6, 281-292.

[17] Machius M., Wiegand G. and Huber R. (1995) *Journal of Molecular Biology*, 246, 545-559.

[18] Gabb H.A., Jackson R.M. and Sternberg M.J.E. (1997) *Journal of Molecular Biology*, 272, 106-120.

[19] Bernstein F.C., Koetzle T.F., Williams G.J.B., Meyer E.F., Brice M.D., Rodgers J.R., Kennard O., Shimanouchi T., Tasumi M. (1977) *Journal of Molecular Biology*, 112, 535-542.

[20] Blundell T.I., Sibanda B.L., Sternberg M.J.E., Thornton J.M. (1987) *Nature*, 326, 347-352.

[21] Fiser A., Feig M., Brooks C.L.III, Sali A. (2002) *Accounts of Chemical Research*, 35, 413-421.

[22] Johnson M.S., Srinivasan N., Sowdhamini R. and Blundell T.L. (1994) *Critical Reviews in Biochemistry and Molecular Biology*, 29, 1-68.

[23] Greer J. (1981) *Journal of Molecular Biology*, 153, 1027-1042.

[24] MartiRenoma M.A., Stuart A.C., Fiser A., Sanchez R., Melo F., Sali A. (2000) *Annual Review of Biophysics & Biomolecular Structure*. 29, 291-325.

[25] Kumari A., Singh V.K., Fitter J., Polen T., Kayastha A. (2010) *Phytochemistry*, 71, 1657-1666.

[26] Nayeem A., Sitkoff D. and Krystek S. (2006) *Protein Science*, 15, 808-824.

[27] Perkins D.N., Pappin D.J., Creasy D.M., Cottrell J.S. (1999) *Electrophoresis*, 20, 3551-3567.

[28] Pappin D.J.C., Hojrup P. and Bleasby A.J. (1993) *Current Biology*, 3, 327-332.

[29] Thompason J.D., Higgins D.G. and Gibson T.J. (1994) *Nucleic Acids Research*, 22, 4673-4680.

[30] Higo K., Ugawa Y., Iwamoto M., Korenaga T. (1999) *Nucleic Acids Research*, 27, 297-300.

[31] Sharma R.D., Goswami N., Lynn A.M., Sharma R.D., Goswami N., Lynn A.M., Rajnee, Sharma P.K., Jawaid S. (2009), *Bioinformation*, 4, 242-244.

[32] Kushwaha H., Gupta N., Singh V.K., Kumar A., Yadav D. (2008) *Online Journal of Bioinformatics*, 9, 130-145.

[33] Bailey T.L, Williams N., Misleh C., Li W.W. (2006) *Nucleic Acids Research*, 34, 369-373.

[34] Abad M.C., Binderup K., Rios-Steiner J., Arni R.K., Preiss J. and Geiger J.H. (2002) *Journal of Biological Chemistry*, 277, 42164-42170.

[35] Ramachandran G.N. and Ramakrishnan C. (1963) *Journal of Molecular Biology*, 7, 95-99.

[36] Vidya N., Vadivukkarasi B., Manivannan G., et al. (2008) *In Silico Biology*, 8, 471-483.

[37] Needleman S.B. and Wunsch C.D. (1970) *Journal of Molecular Biology*, 48, 443-453.

[38] Laurie A.T. and Jackson R.M. (2005) *Bioinformatics*, 21, 1908-1916.

[39] Aparoy P., Leela T., Reddy R.N., Reddanna P. (2009) *Journal of Molecular Graphic & Modeling*, 27, 744-750.

[40] Aparoy P., Reddy R.N., Guruprasad L., Reddy M.R., Reddanna P. (2008) *Journal of Computer Aided Molecular Design*, 22, 611 -619.