IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 6, June 2016

Primary, Secondary and Tertiary Structural Analysis of Disease Resistance Protein RGA4 of ZEA Maize Using Bioinformatics Tools

Payal Patel[1], Hetalkumar Panchal[2]

Under Graduate Student, ASPEE SHAKILAM Agricultural Biotechnology Institute (ASABI),

Navsari Agricultural University, Surat (Gujarat), India[1]

Associate Professor, Dept of Biochemistry and Bioinformatics, ASPEE SHAKILAM Agricultural Biotechnology

Institute (ASABI), Navsari Agricultural University, Surat (Gujarat), India[2]

Abstract: Objective: Disease resistance protein RGA4 of Zea Maize is mitochondrial carrier-like protein having UniProtKB ID: B6STS5. B6STS5 has been identified as an important protein involved in ADP binding. Therefore B6STR5 is considered as a significant protein for various diseases. The experimental 3D structure of B6STS5 is not available. Therefore, present study aims for analysis of primary, secondary and tertiary structure of disease resistance protein RGA4 of Zea Maize using bioinformatics tools and proposing the best 3D model after evaluating various parameters. Methods: Primary structure analysis was done by ProtParam, Secondary structure analysis was done by SOPMA and Jpred4 and Tertiary structure analysis was done by using two different softwares namely SPDBV and I-Tasser. All the predicted 3-D models were analyzed and validated by PROCHECK, ProQ and SolvX. Results: Model2 predicted by I-Tasser showed top results with 67.10% of the residues in the most favorable region, 22.00% in the allowed region, 7.90% in the generously allowed region and 3.00% in the disallowed region. The RMSD between the modeled and the template structure was found to be 1.96 Å. Quality of predicted Model2 developed by I-Tasser had checked by ProQ and found the best LGscore of 2.485 and MaxSub of 0.061 which indicates that the model is very good. PROCHECK and SolvX also confirmed the same. Conclusion: In this study, homology model was developed for B6STS5 using SPDBV and I-Tasser. The models developed were validated using PROCHECK, ProQ and SolvX. These analyses validated the homology model2 produced by I-Tasser is best, robust as well as reliable enough to be used for future study.

Keywords: B6STS5, Homology modeling, SPDBV, I-TASSER, PROCHECK, ProQ, SolvX, ProtParam, SOPMA.

I. INTRODUCTION

Zea Maize L. (Corn) is a staple cereal for human food in Central and South America, and many parts of Africa. Most sweet corns used for human consumption are yellow; high in vitamin A. Grain corn used for animal feed is more often white and sometimes called horse-corn. On an average 9% of production loss worldwide is due to disease [1]. This varied significantly by region with estimates of 4% in northern Europe and 14% in West Africa and South Asia (http://www.cabicompendium.org/cpc/economic.asp). A great number of diseases and pests attack on corn at various stages of growth. Plant diseases can considerably decline not only the net crop yields but also the crop quality by releasing toxins that affect human health, as the outcome of disease outbreak is getting severe across the globe.

The nature has blessed the crop plants with an inherent mechanism to defend themselves from the invasion of pathogens, termed resistance, which restricts further incursion and proliferation of potential pathogens. The complex network of inherent defense system in plants is comprised of three steps that include pathogen detection, signal transduction, and defense response initiation [2–4]. Induction of defense response involves recognition of specific pathogen effectors by specialized host genes, called resistance (R) genes. The host plant then initiates transcription of the defense response (DR) gene, including the pathogenesis-related (PR) gene that confers local or systemic resistance [5, 6].

Because of selective pressure from multitude of pathogens, plants have evolved post invasion mechanisms, which are controlled by dominant resistance genes that detects specific pathogen effectors molecules (for example, Avirulence molecule (Avr)) through direct or indirect means and initiates active defense response. The R-gene mediated resistance is fundamentally racespecific which is only effective against pathogen strains expressing the cognate effector recognised by the R protein. This mechanism is frequently associated with hypersensitive response (HR), resulting in death of the infected cells, also known as gene-for-gene (R-Avr) interaction.

Genetic resistance in plants is often divided into two major classes; Qualitative, or major-gene, resistance, is based on single major-effect resistance genes (R genes) and generally provides race-specific, high-level resistance. Quantitative

resistance typically has a multi-genic basis and generally provides non-race-specific intermediate levels of resistance.

More than 50 major R genes have been cloned in plants [7]. With some exceptions, most of these genes are dominant and share certain conserved domains such as a nucleotide binding site (NBS) and a leucine-rich repeat (LRR) region [8, 12] . A set of genes known as "R gene analogs" (RGAs) have been defined that, while they have no demonstrated function in disease resistance, share these domains. By analyzing the publicly available genomic sequences, 585 RGAs have been defined in the rice cultivar Nipponbare [13] and 149 in Arabidopsis [14]. In maize, 228 RGAs have been identified [15] using partial sequence data derived from several different maize lines. Once the complete genome sequence of the standard maize line B73 is available, a more complete analysis will be possible.

In all these species, RGAs were found to be located all over the genome, often clustering with groups of three or more RGAs mapping to the same locus, mirroring the clustering of plant R genes such as at the Rp1 locus described above [16]. Plant RGAs are both highly divergent and rapidly evolving [17], this fact, together with the high level of genetic diversity found within maize [18], suggests that a huge diversity of RGAs, and therefore a huge array of recognitional specificities, is likely available within maize germplasm.

The experimental 3D structure of Disease resistance protein RGA4 of Zea Maize is not available therefore there is need for the creation of the homology model. Computational approaches can provide homology modeling, which can be further used in molecular dynamic simulations, and automatic docking in order to demonstrate the function of proteins and to illustrate the mode of substrate binding. These types of methods can be used successfully in enzyme–substrate systems and can provide useful information for future studies. Main objective of the present work is to predict a three-dimensional (3D) model of B6STS5 using different software's namely I-Tasser [19-21] and SPDBV [22], along with the primary and secondary structural information. After comparing the results of various software's we have proposed the best 3D model on the basis of various structure evaluating parameters for future studies.

II. MATERIALS AND METHODS

Primary Structural Analysis
Primary Structural Analysis refers to compute various parameters like molecular weight, amino acid composition, extinction coefficient, estimated half-life, theoretical pI, and grand average of hydropathicity (GRAVY), aliphatic index and instability index. The primary sequence of the Disease resistance protein RGA4 of Zea Maize was obtained from UniProtKB database with a sequence Accession Number B6STS5, Entry Name B6STS5_MAIZE, sequence length 411 aa [23]. The complete primary sequence analysis was done by using the ProtParam tool [24], available through the ExPASy server at SIB bioinformatics resource portal.

Secondary Structural Analysis
Secondary Structural Analysis refers to assigning various regions of protein sequence as likely to fold in alpha helices, beta strands or turns. The complete secondary sequence analysis was done by using the Self-Optimized Prediction method With Alignment (SOPMA) tool [25], available through NPS@ interactive Web server dedicated to protein sequence analysis and available for the biologist community. NPS@ is the "protein part" of the "Pôle Bio-Informatique Lyonnais" (PBIL) located at the Institute of Biology and Chemistry of Proteins [26]. Another tool used for secondary structure analysis is JPred: Protein Secondary Structure Prediction server it has been in operation since approximately 1998 [27]. JPred incorporates the Jnet algorithm in order to make more accurate predictions. In addition to protein secondary structure JPred also makes predictions on Solvent Accessibility and Coiled-coil regions by Lupas method.

Tertiary Structural Analysis
Tertiary Structural Analysis refers to prediction of the arrangement of secondary structures as well as their side-chains into three-dimension space. The biological function of a protein is often intimately dependent upon its tertiary structure. From the available experimental data, it has been observed that proteins with similar amino acid sequences usually adopt similar structures. Therefore, the easiest and also the most accurate way to predict the protein tertiary structure is to build the structure based on sequence relatives that have high sequence similarities to the target protein according to the sequence alignment results. Such an approach is called comparative modelling. In most cases those sequence relatives and the target proteins belong to the same functional family in biology, i.e. they are homologues of each other. Thus, traditionally, comparative modelling is also called homology modelling. Homology modeling refers to constructing an atomic-resolution model of the query (Target) protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein called template protein. The query protein is aligned with the template and the secondary structure is predicted between the two and the model is developed.

The accuracy of the homology model is related to the degree of sequence identity and similarity between template and target. The selection of a suitable template and an optimal sequence alignment is essential to the success of homology modeling. BLASTp [28] was performed to find a template structure of a known protein from Protein Data Bank (PDB). Unfortunately we didn't get suitable hit so we have selected various PDB files extracted by I-Tasser server to model the structure of target protein by using offline SPDBV tool (Table 1). The crystal structure of full-length apoptotic protease-activating factor 1 (Apaf-1) (PDB ID:

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 6, June 2016

3SFZA) identified as best template with 22 percentage sequence identity of the whole template chains with query sequence having normalized Z Score 4.21, therefore this structures was used as templates to generate the model in SPDBV. The energy of modeled structure was minimized using the energy minimization facility available in SPDBV. The other software used to generate the homology model was I-TASSER (Iterative Threading ASSEmbly Refinement) server which automatically generates high-quality 3D structure and biological function of protein molecules from their amino acid sequences. I-TASSER implements multiple threading algorithms and iterative structure assembly simulations to find optimal sub-fragments within a database structures or within a user-specified structure.

Table 1: Top 10 threading templates used by I-TASSER server for B6STS5

PDB Hit	Iden 1	Iden 2	Coverage	Norm. Z Score
1vt4A	0.12	0.18	0.80	2.79
1z6tB	0.14	0.20	0.82	1.99
1vt4I	0.13	0.18	0.79	2.05
3sfzA	0.18	0.22	0.59	4.21
1z6tA	0.11	0.20	0.83	2.23
2a5yB	0.12	0.17	0.86	3.06

Assessment of homology model

The validation of structure model obtained from SPDBV and I- Tasser was performed by inspecting the backbone conformation of the modeled structure was calculated by analyzing the phi (φ) and psi (ψ) torsion angles using PROCHECK [29], as determined by Ramachandran plot. The ProQ web server [30] (available at Stockholm Bioinformatics Center) was also used. With ProQ different ranges are given for a model as LGscore>1.5 fairly good model, >2.5 very good model, >4 extremly good model, MaxSub>0.1 fairly good model, >0.5 very good model, >0.8 extremly good model. ERRAT is a protein structure verification algorithm that is especially well-suited for evaluating the progress of crystallographic model building and refinement. The program works by analyzing the statistics of non-bonded interactions between different atom types. This is extremely useful in making decisions about reliability. Verify 3D will provide you with a visual analysis of the quality of a putative crystal structure for a protein and analyzes the compatibility of an atomic model of the protein with its amino acid sequence. Prove Calculates the volumes of atoms in macromolecules.

Structure Validation, Solvation Preference analysis was performed by SolvX server: This server computes the solvation profile for a protein structure. [31] This is useful when assessing the quality of a homology model or an existing PDB structure used in homology modelling. Solvation preference is a measure of solvent accessibility for each residue within a protein; a well-packed structure should have an overall solvation preference below zero. Generally, the more negative this figure is, the better the model. RMSD analysis and other related analysis is done by I-Tasser.

III. RESULTS AND DISCUSSION

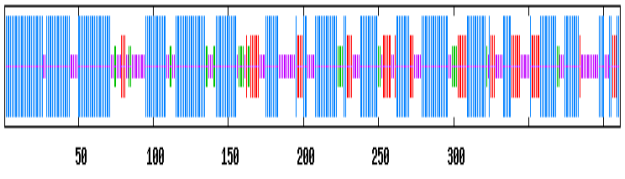
Primary Structural Analysis

Primary Structural Analysis of B6STS5 was done by ProtParam server and the results obtained are like this number of amino acids: 411, Molecular weight: 45323.6, Theoretical pI: 5.67, Amino acid composition: Ala (A) 47 (11.4%), Arg (R) 31 (7.5%), Asn (N) 10 (2.4%), Asp (D) 26 (6.3%), Cys (C) 8 (1.9%), Gln (Q) 16 (3.9%), Glu (E) 30 (7.3%), Gly (G) 28 (6.8%), His (H) 7 (1.7%), Ile (I) 14 (3.4%), Leu (L) 38 (9.2%), Lys (K) 18 (4.4%), Met (M) 14 (3.4%), Phe (F) 18 (4.4%), Pro (P) 13 (3.2%), Ser (S) 33 (8.0%), Thr (T) 17 (4.1%), Trp (W) 8 (1.9%), Tyr (Y) 3 (0.7%), Val (V) 32 (7.8%), Pyl (O) 0 (0.0%), Sec (U) 0 (0.0%), (B) 0 (0.0%), (Z) 0 (0.0%), (X) 0 (0.0%), Total number of negatively charged residues (Asp + Glu): 56, Total number of positively charged residues (Arg + Lys): 49, Atomic composition: Carbon C: 1982, Hydrogen H: 3156, Nitrogen N: 570, Oxygen O: 603, Sulfur S: 22, Formula: C1982H3156N570O603S22, Total number of atoms: 6333, Extinction coefficients: 48970, Abs 0.1% (=1 g/l) 1.080, assuming all pairs of Cys residues form cystines, Extinction coefficient 48470 Abs 0.1% (=1 g/l) 1.069, assuming all Cys residues are reduced, Estimated half-life: The N-terminal of the sequence considered is M (Met). The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro), >20 hours (yeast, in vivo), >10 hours (Escherichia coli, in vivo), The instability index (II) is 53.40 This classifies the protein as unstable, Aliphatic index: 83.36, Grand average of hydropathicity (GRAVY): -0.188.

Secondary Structural Analysis

Secondary Structural Analysis was done by SPOMA and found various secondary structure region in the protein having the different amount of amino acids like in Alpha helix (Hh) : 232 is 56.45%, 310Helix (Gg) : 0 is 0.00%, Pi Helix (Ii): 0 is 0.00% all in blue vertical lines, Beta bridge (Bb) : 0 is 0.00%, Extended strand (Ee) : 53 is 12.90% all in red vertical lines, Beta turn (Tt) : 29 is 7.06% in green vertical lines, Bend region (Ss) : 0 is 0.00%, Random coil (Cc): 97 is 23.60%, Ambiguous states (?) : 0 is 0.00%, Other states: 0 is 0.00%. Another tool used for secondary analysis is Jpred4: Protein Secondary Structure Prediction Server and results are showed in figure.

(A)



DOI 10.17148/IJARCCE.2016.5602

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 6, June 2016

(B)

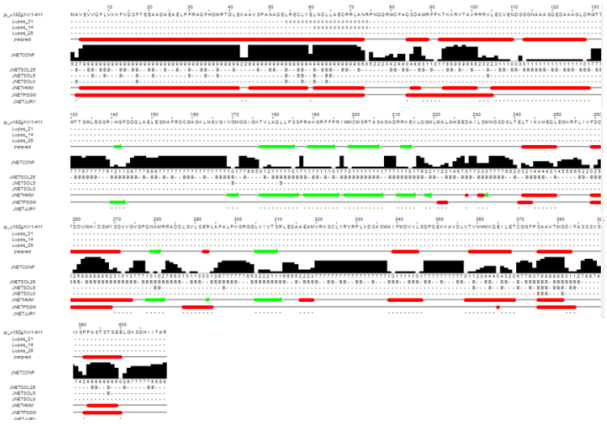
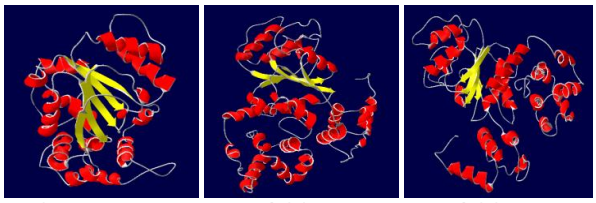


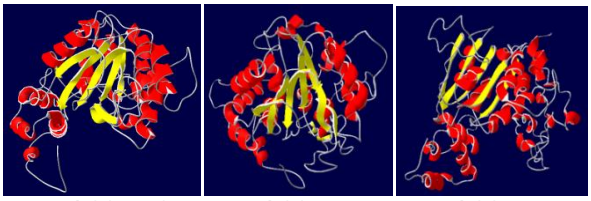
Fig. 1: Showing results of SOPMA (A) and Jpred4 (B).

Homology modeling using SPDBV

The model was generated by SPDBV by using templates top identified structural analogs in PDB by I- Tasser to model protein were not having the good Ramachandran plot result. So we have filtered all the modeled structure by SPDBV for further analysis. The main feature of using SPDBV was we can do all the required analysis standalone software.



SPDBV I-TASSER 1 I-TASSER 2



I-TASSER 3 I-TASSER 4 I-TASSER 4

Fig. 2: Ribbon diagrams of the modeled B5STS6 with SPDBV and I-Tasser; αHelices, β-strands and loops are colored red, yellow and gray, respectively.

Homology modeling using I-TASSER

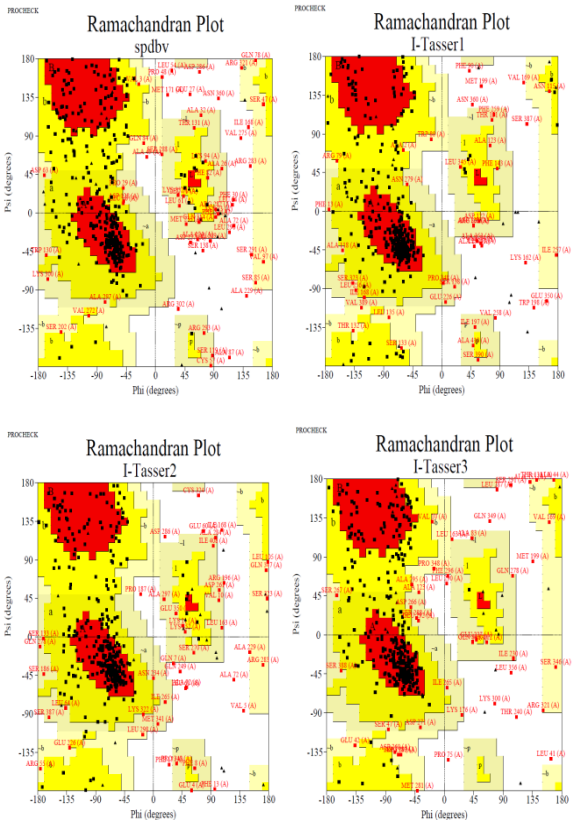
In this method the target sequences were first threaded using a representative PDB structure library to search for the possible folds by Profile- Profile Alignment (PPA), Hidden Markov Model, PSI- BLAST profiles, Needleman-Wunch and Smith-Waterman alignment algorithms. The top 10 alignments are from the following threading programs MUSTER, dPPAS, Neff-PPAS, PPAS, wdPPAS, SPARKS-X, SP3, HHSEARCH2, PROSPECT2, and FFAS03. The PDB ID: 3SFZ had the best Z-score using all the ten algorithms and was used for modeling B5STS6 structure (Table 1). I-TASSER server predicted 5 models from which the model with best C-Score of -1.56 was selected with estimated accuracy of 0.532 (TM-Score) and 3.10Å (RMSD). C-score is a confidence score for estimating the quality of predicted models by I-TASSER. It is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. C-score is typically in the range of [-5, 2], where a C-score of higher value signifies a model with a high confidence and vice-versa.

Table 2: Top Identified structural analogs in PDB Used by I-Tasser to model protein

Rank	PDB Hit	TM-score	RMSD	Identity	Coverage
1	3iz8A	0.790	1.96	0.112	0.822
2	2a5yC	0.541	3.11	0.120	0.608
3	1z6tC	0.540	3.88	0.121	0.642
4	3sfzA	0.532	3.10	0.152	0.608
5	3izaA	0.508	3.47	0.129	0.601
6	4xguA	0.501	5.37	0.121	0.691
7	1ksfX	0.493	5.24	0.080	0.671
8	4xgcD	0.485	4.75	0.100	0.623
9	3cf2D	0.481	5.71	0.101	0.686

Model validation

Validation of the model including the geometric properties of the backbone conformations, were analyzed using various structure evaluation programs. Ramachandran plot calculations were calculated with PROCHECK program. Model2 predicted by I-Tasser indicated that 67.10% of the residues in the most favorable region, 22.00% in the allowed region, 7.90% in the generously allowed region and 3.00% in the disallowed region (Fig. 3).



DOI 10.17148/IJARCCE.2016.5602

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 6, June 2016

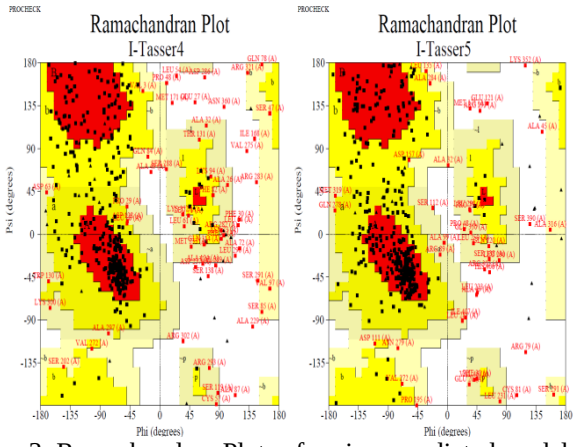


Fig. 3: Ramachandran Plots of various predicted models

These results revealed that the majority of the amino acids are in a phi-psi distribution that is consistent with a right-handed α-helix, and the model is reliable and of good quality. Whereas other models produced by I-Tasser and SPDBV did not have such best scores. Model2 developed by I-Tasser had ProQ LGscore of 2.485 and MaxSub of 0.061 indicated that the model developed by I-Tasser is was very good whereas other two model come in the criteria of fairly good model. RMSD between the template and model developed by SPDBV and I-Tasser are shown in table 2.

All these results suggest that the model2 developed by I-Tasser is comparatively robust and can be used in subsequent stages of analysis. Therefore, the PROCHECK (Table 3), SolvX, Verify_3D, results confirm the quality of predicted 3D structure as more reliable and within an acceptable range.

SolvX is a program that evaluates the atomic solvation preference of full-atom 3D protein models. Solvation preference is a measure of solvent accessibility for each residue within a protein; well-packed structures should have an overall solvation preference value less than zero. This program is particularly useful when evaluating the quality of a theoretical 3D model of a protein compared with experimentally resolved structures. Results of SolvX program is represented by various graphs produced by the server.

The thin red line on each graph shows the residue-by-residue solvation preference, and the thick, coloured line shows average solvation preference over a sliding window of 11 residues. Model 2 predicted by I-Tasser is having the -8.0 overall score so it is considered as best model. (Fig. 4).

SolvX discriminate between correct and incorrect three-dimensional structures for a given sequence, or to identify the correct sequence placement in a given structure. Backbone co-ordinates were taken from experimentally known structures or hypothetical models and side-chain conformations were optimized by an efficient Monte Carlo algorithm using simulated annealing and simple potential functions.

Table 3: Ramachandran Plots statistics of various predicted models

Model	AA in Core Region	AA in Additionally Allowed Region	AA in Generously Allowed Region	AA in Disallowed Region
SPDBV	65.20%	23.60%	7.60%	3.50%
I-Tasser1	63.30%	25.30%	7.90%	3.50%
I-Tasser2	67.10%	22.00%	7.90%	3.00%
I-Tasser3	56.50%	29.10%	8.70%	5.70%
I-Tasser4	62.80%	26.90%	7.10%	3.30%
I-Tasser5	58.50%	27.10%	7.70%	6.70%

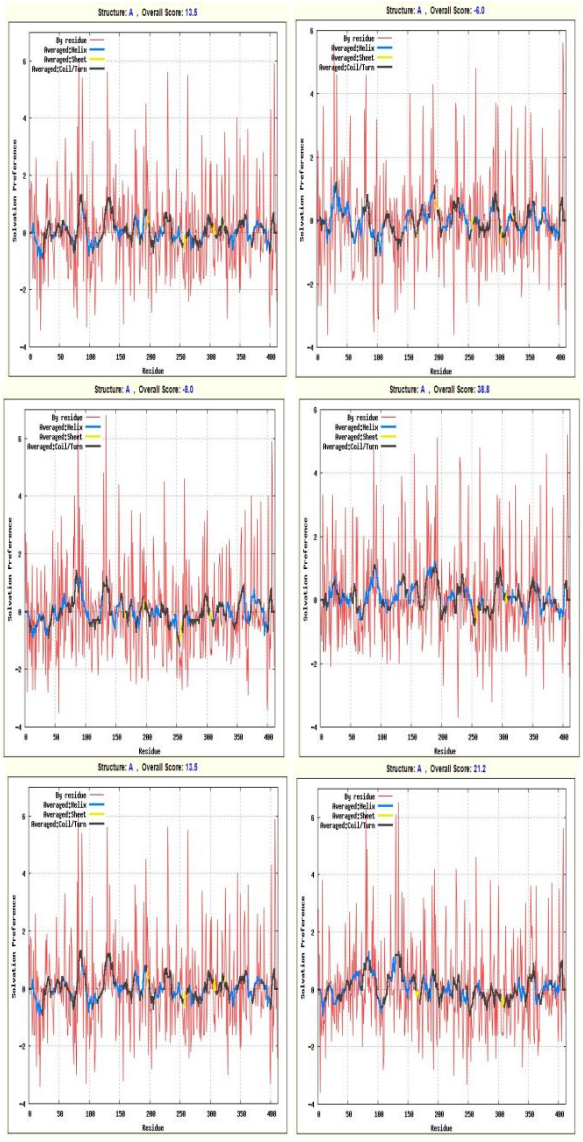


Fig. 4: The thin red line on each graph shows the residue-by-residue solvation preference, and the thick, coloured line shows average solvation preference over a sliding window of 11 residues.

DOI 10.17148/IJARCCE.2016.5602

IV. CONCLUSION

Disease resistance protein RGA4 of Zea Maize is mitochondrial carrier-like protein having UniProtKB ID: B6STS5. B6STS5 has been identified as an important protein involved in ADP binding. Therefore B6STR5 is considered as a significant protein for various diseases. The experimental 3D structure of B6STS5 is not available. Therefore, present study aims for analysis of primary, secondary and tertiary structure of disease resistance protein RGA4 of Zea Maize using bioinformatics tools and proposing the best 3D model after evaluating various parameters.

The models produced by all software's were further assessed by Procheck, SolvX and Verify_3D. Looking on the various results produced by all the analysis tools it can be suggested that, model produced by I-Tasser server is reliable for this type of protein. On the basis of various structure validation tools only one structure would be the best among five. We have to evaluate structure on various parameters and then decide the most appropriate structure for further analysis. The protein model suggested in present research could be used for further analysis in the area functional analysis and drug discovery.

ACKNOWLEDGEMENT

We sincerely thank to Principal of ASPEE SHAKILAM Agricultural Biotechnology Institute, Navsari Agricultural University, Surat for allowing us to utilize the research facilities.

REFERENCES

[1] Oerke EC (2005) Crop losses to pests. J Agr Sci 144:31–43.doi: 10.1017/S0021859605005708
[2] Nimchuk Z, Eulgem T, Holt BF, Dangl JL (2003) Recognition and response in the plant immune system. Annual Reviews of Genetics 37: 579–609.
[3] Tiffin P, Moeller DA (2006) Molecular evolution of plant immune system genes. Trends in Genetics 22: 662–670.
[4] Chisholm ST, Coaker G, Day B, Staskawicz BJ (2006) Host-microbe interactions: shaping the evolution of the plant immune response. Cell 124: 803–814.
[5] Bishop JG, Dean AM, Mitchell-Olds T (2000) Rapid evolution in plant chitinases: molecular targets of selection in plant–pathogen coevolution. Proceedings of the National Academy of Sciences 97: 5322–5327.
[6] Dangl JL, Jones JDG (2001) Plant pathogens and integrated defence responses to infection. Nature 411: 826–833.
[7] Martin , G.B. , A.J. Bogdanove , and G. Sessa . 2003. Understanding the functions of plant disease resistance proteins . Annual Review of Plant Biology 54 : 23 – 61 .
[8] Johal , G.S. , and S.P. Briggs . 1992 . Reductase activity encoded by the HM1 disease resistance gene in maize . Science 258 : 985 – 7 .
[9] Multani , D.S. , R.B. Meeley , A.H. Paterson , J. Gray , S.P. Briggs , and G.S. Johal . 1998. Plantpathogen microevolution: Molecular basis for the origin of a fungal disease in maize . Proceedings of the National Academy of Sciences of the United States of America 95: 1686 – 91.
[10] Romer , P. , S. Hahn , T. Jordan , T. Strauss , U. Bonas , and T. Lahaye . 2007. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance Gene . Science 318 : 645 – 648 .
[11] Xiao, S. , S. Ellwood , O. Calis , E. Patrick , T. Li , M. Coleman , and J.G. Turner . 2001. Broad-spectrum mildew resistance in Arabidopsis thaliana mediated by RPW8. Science 291: 118 – 120.
[12] McDowell, J.M., and S.A. Simon. 2006. Recent insights into R gene evolution. Molecular Plant Pathology 7: 437 – 448.
[13] Monosi , B. , R.-J. Wisser , L. Pennill , and S.-H. Hulbert. 2004. Full-genome analysis of resistance gene homologues in rice. Theoretical and Applied Genetics 109: 1434 – 1447.
[14] Meyers , B.C. , A. Kozik , A. Griego , H. Kuang , and R.W. Michelmore . 2003. Genome-wide analysis of NBS-LRR-encoding genes in arabidopsis . Plant Cell 15: 809 – 834.
[15] Xiao, W. , J. Zhao , S. Fan , L. Li , J. Dai , and M. Xu . 2007 . Mapping of genome-wide resistance gene analogs (RGAs) in maize (Zea mays L.). TAG Theoretical and Applied Genetics 115: 501 – 508.
[16] Smith, S.M. , A.J. Pryor , and S.H. Hulbert . 2004. Allelic and haplotypic diversity at the rp1 rust resistance locus of maize. Genetics 167: 1939 – 1947.
[17] Ellis, J. , P. Dodds , and T. Pryor . 2000. Structure, function and evolution of plant disease resistance genes. Current Opinion in Plant Biology 3: 278 – 284.
[18] Buckler, E.S., B.S. Gaut , and M.D. McMullen . 2006. Molecular and functional diversity of maize. Current Opinion in Plant Biology 9: 172 – 176.
[19] Yang Zhang.: I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008; 9: 40.1-8.
[20] Ambrish Roy, Alper Kucukural, Yang Zhang.: I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols. 2010; 5: 725-738.
[21] Ambrish Roy, Jianyi Yang, Yang Zhang.: COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Res. 2012; 40: 471-477.
[22] Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18, 2714-2723.
[23] http://www.uniprot.org/
[24] Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607
[25] Geourjon C, Deleage G, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 1995 Dec;11(6):681-684
[26] https://prabi.ibcp.fr/htm/index.php
[27] Drozdetskiy A, Cole C, Procter J & Barton GJ. JPred4: protein secondary structure prediction server, Nucleic Acids Res. Web Server issue [doi:10.1093/nar/gkv332]
[28] Altschul SF, Madden TL, Schäffer AA, et al.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25:3389-402.
[29] Laskowski R A, MacArthur M W, Moss D S, Thornton J M (1993). PROCHECK - a program to check the stereochemical quality of protein structures. J. App. Cryst., 26, 283-291.
[30] Wallner B. and Elofsson A.: Can correct protein models be identified? Protein Science. 2003; 12: 1073-1086.
[31] Holm L, Sander C (1992) Evaluation of protein models by atomic solvation preference. J Mol Biol 225(1): 93-105.