Comparative structural analysis of OsCKX2 protein of Oryza sativa

Pranati Swain
Central Rice Rsearch Institute, Cuttack, Odisha, India

"Dedicated to my beloved parents & sister Mrs. Pritilata Nayak & Mr. Purnananda Swain, Miss Prativa Swain"

***Abstract-* OsCKX2 is an important protein of *Oryza sativa* which helps in high grain production. The objective is to find out the structure of protein and to represent the best 3D model by using insilico analysis. In *Oryza sativa* the OsCKX2 is homologous to maize (*zea mays)* cytokine oxidasedehydrogenase, which is responsible for plant growth and development. The following study was obtained by predicting the 3D structure of OsCKX2 in *Oryza sativa*u using homology modeling concept with the help of bioinformatic tools i.e, 3D-JIGSAW, SWISS-MODEL, MODELLER9.12 and PHYRE2. The 3D structures were further evaluated and steriochemical analysis was done by using RAMPAGE server for Ramachadran plot analysis. The best model was selected on the basis of Z-score and 3D-1D score.**

***Index Terms- oryza sativa,* OsCKX2 protein, 3D structure, structure prediction, homology**

I. INTRODUCTION

Rice is cultivated in more than 100 countries worldwide on approximately 150 million hectares of land where India ocuupies first position in production of rice. The *GN1A* gene of rice is found in japonica subspecies having Cytokinin dehydrogenase2**(**OsCKX2) protein. This protein is generally responsible for high grain production in rice(Shuyu Li et al.,[1], Ashikari et al.,[2]). The uniprot ID of the protein is Q4ADV8. The function of this protein is it catalyzes the oxidation of cytokinins, which is a family of N(6)-substituted adenine derivatives and theses are plant hormones helping for well growth and development. The substituent of protein is an isopentenyl group changes the number of reproductive organs by regulating the cytokinin accumulation in inflorescence meristems. The protein leads in high grain production and higher cytokinin levels in inflorescence meristems(Schmuelling et al.,[3]). The protein is mostly expressed in leaves, culms, inflorescence meristems, and flowers and in vascular tissues. The protein belongs to the oxygen-dependent FAD-linked oxidoreductase family having 1 FAD-binding PCMH-type domain. The protein interacts with FAD, flavin-adenine dinucleotide, the coenzyme or the prosthetic group of various flavoprotein oxidoreductase enzymes, in either the oxidized form, FAD, or the reduced form, FADH2. Cytokinin acts as a growth regulator in plants. The binding of cytokinin to the receptor leads to transcription. The enzyme catalyzes the oxidation of one compound with the reduction of another. Protein helps in the cytokinin signaling pathway. The phytohormones are synthesized in the root apical meristem and transported through the plant in the xylem sap. Cytokinins are involved in several physiological processes such as promoting cell division, chloroplast maturation, regulating cell growth , differentiation, monitoring nutrient uptake and senescence(Ashikari M et al[4]). Auxin also regulates the cell cycle and tissue morphogenesis.

II. MATERIALS AND METHODS

The genbank ID of the *GN1A* gene was retrived from RAP-DB rice database which showed the protein evidence against Uniprot database.

The protein sequence, name, origin, protein attributes, general annotation and entry information was retrieved from the National centre of Biotechnology information (NCBI) or from Uniprot database. The amino acid sequence is available in FASTA format.

The secondary structure prediction was done by using CFSSP tool(Jabeen et al.,[5], Isvoran A et al.,[6]). GOR V is another tool for secondary structure prediction of proteins(Taner et al.,[7], Kloczkowski[8]).

Comparative homology modeling of the protein was done by using PHYRE2 server(Kelly et al.,[9]), Modeller9.12(Madhusudhan et al.,[10]), 3DJIGSAW tool(Bates et al.,[11]), SWISS-MODEL tool. After predicting the 3D structures structural evaluation and stereo-chemical analysis was done by using various evaluation and validation tools like PROCHECK saves server and RAMPAGE program.

The backbone conformations like allowed regions, disallowed regions of Psi/Phi Ramachandran plot was predicted by using above evaluation tools(Lovel et al.,[12]).

Energy calculations on a protein chain were performed by atomic empirical mean force potential ANOLEA (Melo et al.,[13]).

A plot obtained from the Verify3D (Eisenberg D.,[14]) structure evaluation Server that represented the average3D-1D profile score. Verify3D analyzes the compatibility of an atomic model with the corresponding amino acid sequence.

III. RESULTS AND DISCUSSION

The secondary structure which was predicted by SOPMA showed that it has 35.93% of alpha helix(Hh), 16.46% extended strand(Ee), 6.19% beta turn(Tt), 41.42% random coils(Cc). The following graph which obtained by using CFSSP software shows the helix, sheet, turn and coil respectively(fig.1).

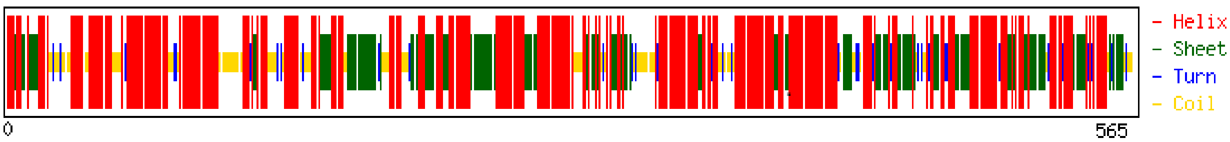

Fig.1 secondary structure predicted by CFFSP (chou & fasman secondary structure prediction server)

The main objective for predicting the structure of proteins is to predict the homolog's of known 3D structure. This process is called template-based homology modeling or fold recognition.

The homology modeling helps to understand the function, interaction as well as location of the protein. The OsCKX2 protein of *Oryza sativa* have *565* amino acids which is homolog to maize cytokine oxidase dehydrogenase with the PDB-ID 3S1D having A chain. Four 3D models were obtained by using various protein structure prediction tools(fig.1, fig.2, fig.3, fig.4).

The sequence similarity with homolog by using swiss model was 57.40%, by using phyre2 was 57%.00, by using modeller9.12 was 55.00% and by using 3DJIGSAW was 55.08%.

so the best model was generated by using SWISS-MODEL(fig.5).

The predicted models were visualized using various visualization tools i.e, Pymol, Rastop, Jmol, Discovery studio, Rasmol and the easy visualization tool was Pymol.

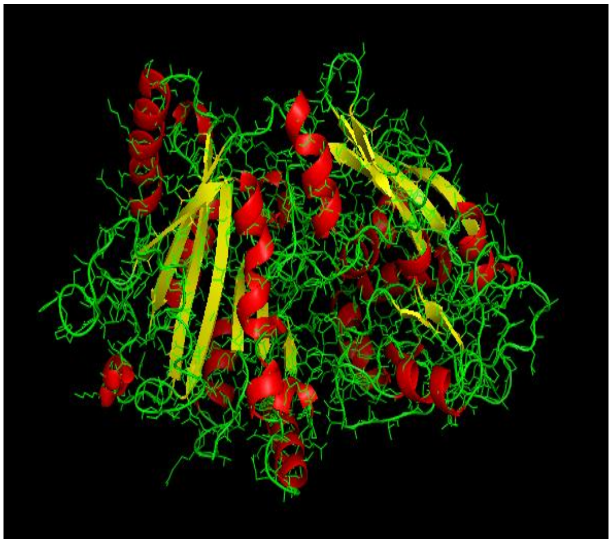

Fig.2 3D model of OsCKX2 protein

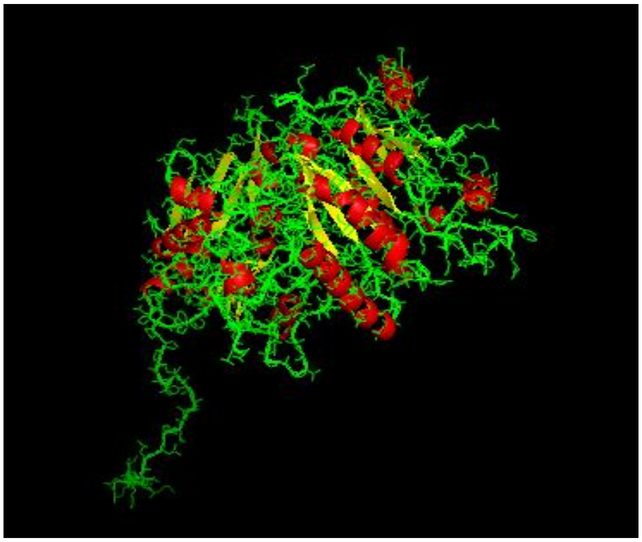

Fig.3 3D model of OsCKX2 protein by modeller9.12 using 3DJIGSAW

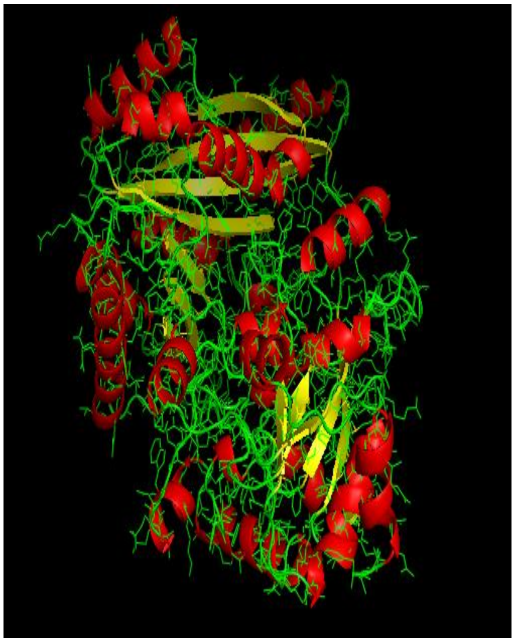

Fig.4 3D model of OsCKX2 protein using PHYRE2

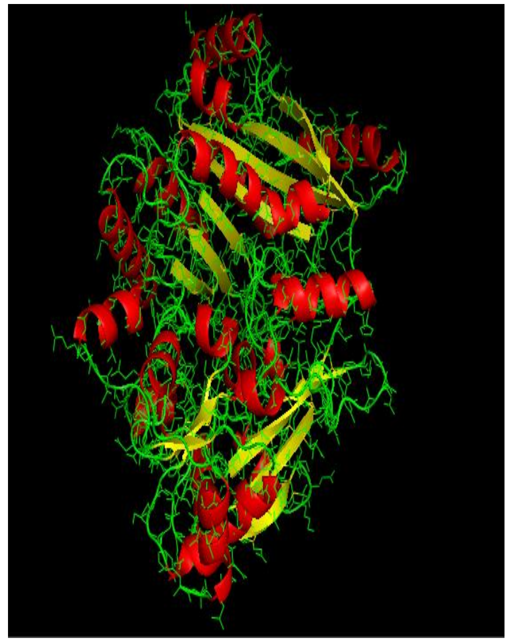

Fig.5 3D model of OsCKX2 protein using swiss model

The objective of using four different model building tool was to predict the perfect alignment to a 3D model. SWISS-MODEL tool is used most frequently to generate the 3D models from a small number of rigid bodies, obtained from the core of the aligned regions (Blundell et al.,[15]). The confirmation of final three-dimensional (3D) structure was predicted by the Phyre2 server, which is an automatic fold recognition server for predicting the structure and the function of proteins based on homology modeling.

The predicted models might posses error due to energy of the model or from the resemblance of a given characteristic of the model to real structures. So the evaluation of the model was carried out by using ANOLEA and Verify3D tools

Table 1 Model evaluation of proteins by using ANOLEA and Verfy3D tools

Protein	**Modeling program**	**Evaluation program**	**Torsion angle energy**	**Z-score**	**3D-1D profile**
	3DJIGSAW	ANOLEA	-38.19	-1.65	
		Verify3D	-	-	0.73
	MODELLER9.12	ANOLEA	-72.67	-0.57	
OsCKX2		Verify3D	-	-	0.77
	PHYRE2	ANOLEA	-90.42	-1.52	
		Verify3D	-	-	0.77
	SWISSMODEL	ANOLEA	-72.67	-0.57	
		Verify3D	-	-	0.77

The objective of using Verify3D method was to test the compatibility of an atomic model with respective to it's amino acid sequence. Verify3D generally uses the 3D structure profile for analyzing each residue, it's location, alpha, beta, loop, polar, apolar etc. -1 is considered as bad score whereas +1 is considered as good score(Eisenberg, 1997[16]). The energy constituent of protein may not fully 100% correct in the predictive manner. So an energy minimization tool named ANOLEA was performed on a protein chain, evaluating the "Non- Local Environment" (NLE) of each heavy atom in the molecule (Melo etal.,[17]). ANOLEA result generally shows high energy (%) and pseudo Z-score. From the Z-score, high energy region(HEz) of the protein profile the stability of protein can be predicted. Lower Z-score means higher reliability and vice versa. Two same qualities of models were formed by using Phyre2 and SWISS-MODEL server where the Z-score is -0.57 while the models generated by using 3DJIGSAW server gave the Z-score of -38.19 and the models generated by using MODELLER9.12 tool gave the Z-score of -72.67. The correct and specific alignment method obtained by combining, weighting and screening the results of several multiple alignment programs help to find the best Z-score and the final three-dimensional structure was built using the modeling package MODELLER. The 3D-1D profile scores, obtained from Verify3D, of OsCKX2 models is 0.77 obtained from Phyre2 and SWISS-MODEL servers. MODELLER9.12 and 3DJIGSAW homology modeller resulted in -72.67 and -38.19 respectively. The 3D structures which were generated in SWISS-MODEL and Phyre2 servers are identical having the same number of α helixes and β sheets. Z-score of SWISS-MODEL, MODELLER and Phyre2 modeler is much smaller compared to 3DJIGSAW server. Furthermore, the 3D-1D score of 0.77 indicates a good quality 3D model, compared to the *Zea mays* model. Ramachandran plots analysis was generated by using Procheck saves server tool which showed the allowed and disallowed region respectively((fig.6, fig.7, fig.8, fig.9). The RAMPAGE program (Lovel et al., 2002[18]) for all four 3D models, also shows all residues in the most favorable regions and disallowed regions(fig.10, fig.11, fig.12, fig.13). The number of favored region in the 3DJIGSAW OsCKX2 model is 78.4%, for MODELLER is 86.6%, for Phyre2 is 88.1% and for SWISS MODEL is 88.1% respectively. The models formed by MODELLER, SWISS-MODEL and Phyre2 tool showed that most of the region of protein lie in the favored region which states it is a stable protein.

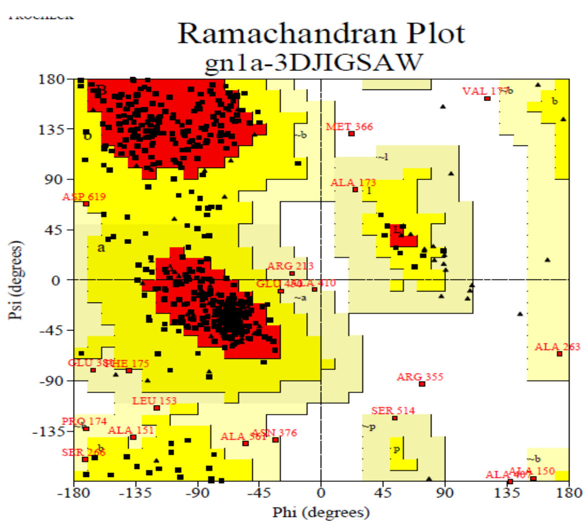

Fig.6 ramachandran plot of 3DJIGSAW model

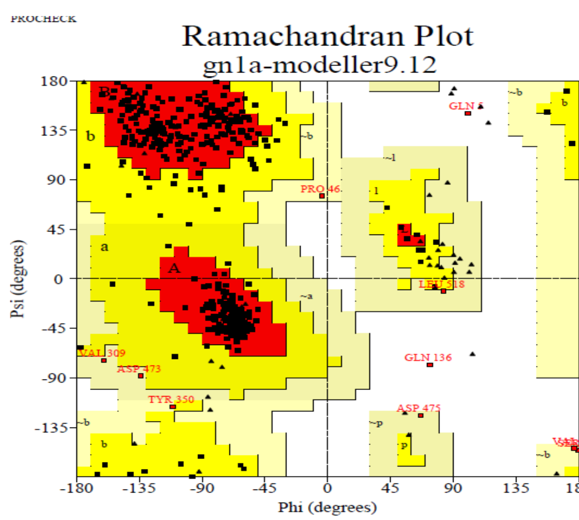

Fig.7 ramachandran plot of modeller9.12 model

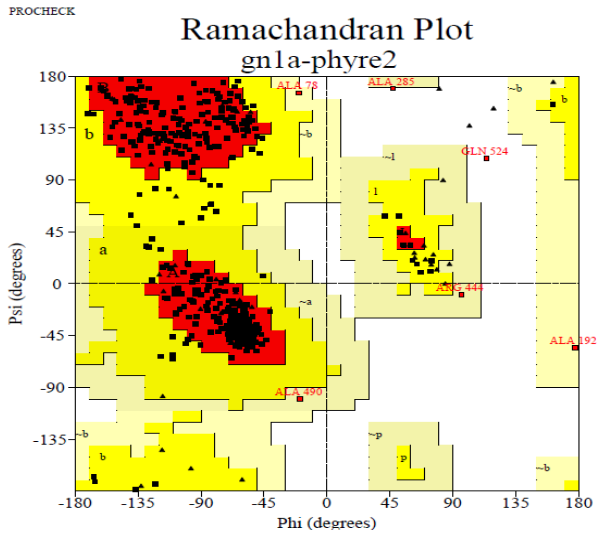

Fig.8 Ramachandran plot of PHYRE2 model

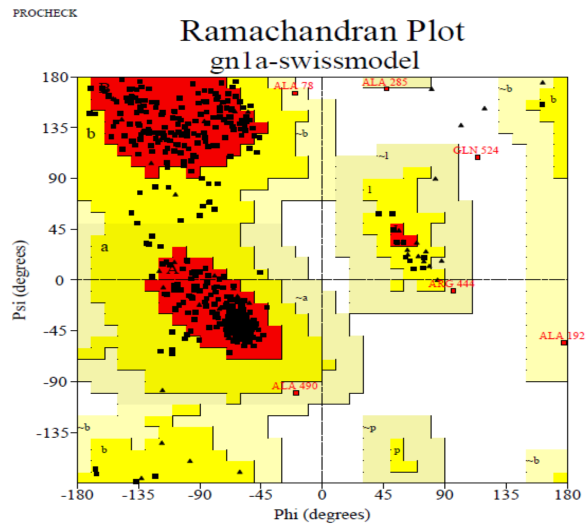

Fig.9 ramachandran plot of swiss-model model

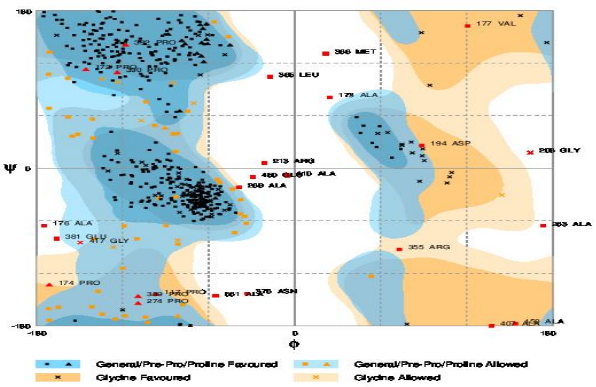

Fig.10 Ramachandran plot of 3DJIGSAW model

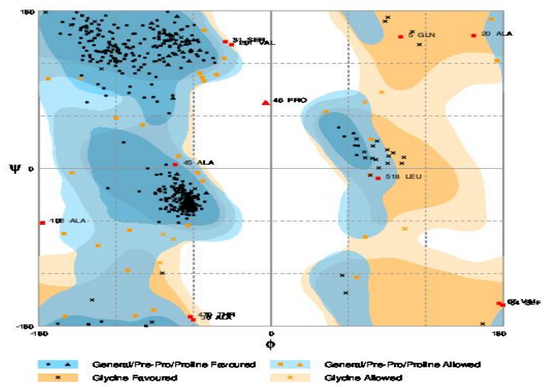

Fig.11 Ramachandran plot of modeller9.12

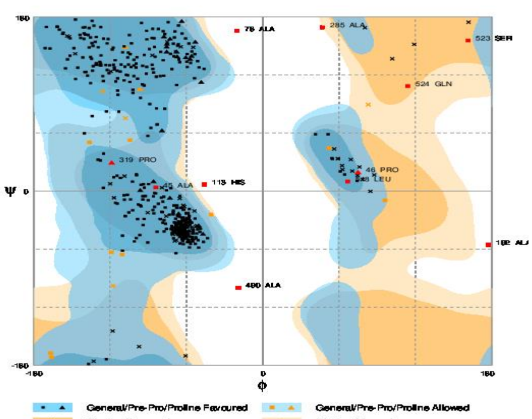

Fig.12 Ramachandran plot of PHYRE2 model

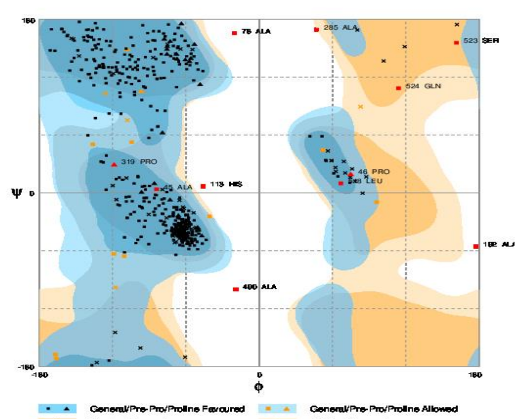

Fig.13 Ramachandran plot of swiss-model model

IV. CONCLUSION

The structure of protein was predicted successfully by using both secondary and tertiary structure prediction method. Homology modeling was performed by using various bioinformatic tools. The phi-psi bond along with the allowed and disallowed regions in Ramachandran plot was analysed. The best model was selected depending upon using Z-score and 3D-1D score. The rice protein which was analysed showed homologous to the Zea mays protein(cytokine oxidasedehydrogenase). This protein generally helps for high grain production in cereal and crops.

REFERENCES

[1] Li S., Zhao B., Yuan D., Duan M., Qian Q., Tang L., Wang B., Liu X., Zhang J., Wang J., Sun
J., Liu Z., Feng Y.Q., Yuan L., Li C. (2012) Rice zinc finger protein DST enhances grain production through controlling Gn1a/OsCKX2 expression www.pnas.org/cgi/doi/10.1073/pnas.1300359110

[2] Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles E.R, Qian Q, Kitano H, and Matsuoka M(2005) Cytokinin oxidase regulates rice grain production. Science 309(5735) 741-745.

[3] Schmuelling T., Werner T., Riefler M., Krupkova E., Bartrina y Manns I.(2003) Structure and function of cytokinin oxidase/dehydrogenase genes of maize, rice, Arabidopsis and other species J. Plant Res. 116:241-252(2003) [PubMed] [Europe PMC] [Abstract].

[4] Ashikari M et al.(2005) Cytokinin Oxidase Regulates Rice Grain Production. Science 309, 741 (2005); DOI: 10.1126/science.1113373.

[5] Jabeen U, Zubair A, Ahmed W, Nazir B, Sherwani K S(2013) Bioinformatics Study of the Sequence Analysis and Secondary Structure Prediction of Avicennia marina glyoxalase I Protein european academic research, vol. I, issue 6/ sepember 2013 ISSN 2286-4822, www.euacademic.org.

[6] Isvoran A, craciun D, ciorsac A, perrot N, beswick V, nedellec P, Sanson A, jamin N (2014) a bioinformatics study concerning the structural and functional properties of human caveolin proteins. J. Serb. Chem. Soc. 79 (2) 133–150.

[7] Taner Z. Sen, Robert L. Jernigan, Jean Garnier and Andrzej Kloczkowski(2005) GOR V server for protein secondary structure prediction. Structural bioinformatics Vol. 21 no. 11 2005, pages 2787–2788.

[8] Kloczkowski,1 K.-L. Ting,1 R.L. Jernigan,1† and J. Garnier(2002) Combining the GOR V Algorithm With Evolutionary Information for Protein Secondary Structure Prediction From Amino Acid Sequence PROTEINS: Structure, Function, and Genetics 49:154–166 . Lovell SC, Davis IW, Arendall III WB, et al. 2002. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins: Structure, Function and Genetics, 50: 437-450.

[9] Kelley L A & Sternberg M J E(2009) Protein structure prediction on the Web: a case study using the Phyre server. NATURE PROTOCOLS | VOL.4 NO.3

[10] Madhusudhan M S, Marc A, Renom M, Eswar N, John B, Pieper U, Karchin R, Shen M Y, and Sali A. Comparative Protein Structure Modeling. The Proteomics Protocols Handbook.

[11] Bates PA, Kelley LA, MacCallum RM, et al. 2001. Enhancement of Protein Modelling by Human Intervention in Applying the Automatic Programs 3D-JIGSAW and 3D-PSSM. Proteins: Structure, Function and Genetics, Suppl(5): 39-46.

[12] Lovell SC, Davis IW, Arendall III WB, et al. 2002. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins: Structure, Function and Genetics, 50: 437-450

[13] Melo F, Feytmans E. 1998. Assessing protein structures with a non-local atomic interaction energy. Journal of Molecular Biology, 2775: 1141-1152.

[14] Eisenberg D, Luthy R, et al. 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods in Enzymology, 277: 396-404.

[15] Blundell TL, Sibanda BL, Sternberg MJ, et al. 1987. Knowledge-based prediction of protein structures and the design of novel molecules. Nature, 326: 347-352.

[16] Eisenberg D, Luthy R, et al. 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods in Enzymology, 277: 396-404.

[17] Melo F, Feytmans E. 1998. Assessing protein structures with a non-local atomic interaction energy. Journal of Molecular Biology, 2775: 1141-1152.

[18] Lovell SC, Davis IW, Arendall III WB, et al. 2002. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins: Structure, Function and Genetics, 50: 437-450.