2012 International Conference on Biological and Life Sciences
IPCBEE vol.40 (2012) © (2012) IACSIT Press, Singapore

Structural Modeling of Human Desmocollin-2 Using I-TASSER Methods

Rohini K [1+], Srikumar P S [2] and Anbarasu K [3]

[1]Unit of Biochemistry, Faculty of Medicine, AIMST University, Bedong, Kedah, Malaysia

[2]Unit of Psychiatry, Faculty of Medicine, AIMST University, Bedong, Kedah, Malaysia

[3]School of Biosciences and Technology, VIT University, Vellore, Tamilnadu, India

Abstract. Desmosomes are well defined contact structures involved in cell-cell adhesion of humans. The proteins of cadherin class are desmoglein-2 and desmocollin-2 which are found to be critical in the disease mechanism of arrhythmogenic right ventricular cardiomyopathy (ARVC) in humans. The desmocolin-2 mutations area major contributory factor. With the advance of computational methods, there is a need for research on structural biology of desmocollin-2 which could lead to annotation of protein function and drug design. Analysis of desmocollin-2 domain was carried out by PFAM and secondary structure prediction using GOR IV. The 3D structural model was generated using I- TASSER based on threading algorithm. The assessment of structural models wasdone with C-score, TM score and RMSD. Model 1 was found to be the best and was validated using PROCHECK.

Keywords: Desmocollin, I-TASSER, C- score, RMSD.

1. Introduction

Arrhythmogenic right ventricular cardiomyopathy (ARVC), a heart muscle disorderis associated with heart failure, ventricular arrhythmias and sudden death[1]. Familial disease occurs in 30–50% of the cases with usually autosomal-dominant inheritance[2]. The desmosomes are specialized structures involving cell–cell contact within the intercalated disc (ID), which allows the heart to withstand mechanical strain during contractile cycles. In these structures, the desmosomal cadherins desmoglein-2 (DSG2) and desmocollin-2 (DSC2) link neighboring cells via interactions with their extracellular cadherin domains[3].

In cardiac desmosomes, two desmosomal cadherins are expressed: desmocollin-2 (DSC2) and desmoglein-2 (DSG2). Both have overlapping functions in binding plakoglobin (PG) and plakophilin-2 (PKP2), believed to be mediated by a conserved region in their cytoplasmic portion, the intracellular cadherin segment (ICS)[5]. Desmocollins are cadherin-like adhesion molecules of desmosomes, characterized by two splice variants (a form and b form) with different cytoplasmic portions. The DSC2a protein consists of the full ICS domain, whereas this module is significantly truncated in the DSC2b isoform[4].

Mutations in both DSG2 and DSC2cardiac desmosomal cadherins, also contributes to the disease. The crucial cardiac function of DSC2 was demonstrated by knock-down experiments in the zebrafish model. DSC2 is expressed as two different splice variants with different carboxy-termini[4].

The two mutations in DSC2 missense mutations p.E102K and p.I345T gene associated with ARVC, a genetically determined heart muscle disease characterized by structural, electrical, and pathological abnormalities of the right ventricle[7].

[+] Corresponding author. Tel.: +0060103701097.
E-mail address: rohinik23@gmail.com.

2. Materials and Methods

2.1. Sequence retrieval

The Sequence retrieval for the protein desmocollin-2 isoform 2a from Uniprot database (http://www.uniprot.org/). The protein contains 901 amino acids and molecular weight of 99,962 Daltons.

2.2. Domain

The domain analysis of query desmocollin-2 was done using Pfam (http://pfam.sanger.ac.uk/). The option sequence search was used to find the domain for the query protein.

2.3. Secondary structure prediction

The 2D structure was predicted using GOR IV tool. The algorithm used in GOR IV was information theory.

2.4. Tertiary structure prediction and validation

The 3D structure model was generated by I-TASSER software by Zhang. I-TASSER represents threading or fold recognition method with template as fold library. The structure validation was done in SAVES server by PROCHECK.

3. Results and Discussion

3.1. Domain

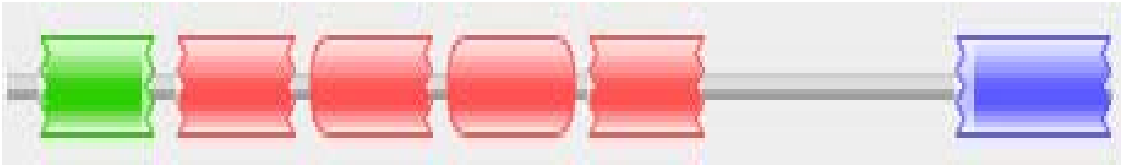


Fig. 1: Result of Pfam domain.

The Pfam results of desmocollin-2 showed the query protein contains six domains. The N terminal of the protein consists of domain belongs to cadherin pro-domain like class. Then four consecutive cadherin domains were found in the middle of the protein to perform function of cell adhesion. The terminal C domain consists of cadherin cytoplasmic region.

3.2. Secondary structure prediction

The secondary structure prediction wasperformed by using tool GORIV. The results showed more random coil of 54%, extended strand of 27% and alpha helix17% which in turn acts as transmembrane helix.

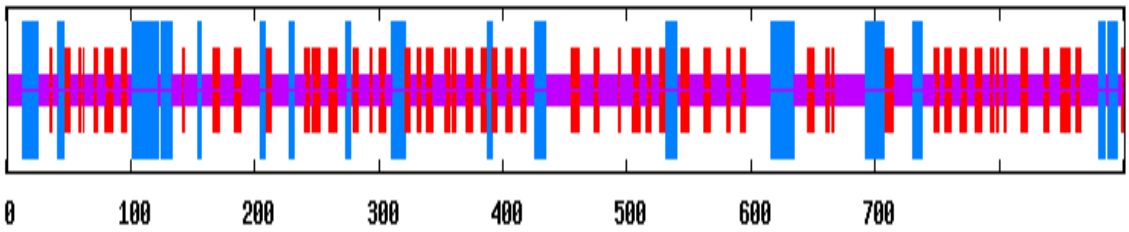


Fig. 2: Result of GOR IV secondary structure prediction

3.3. Tertiary Structure Prediction and Validation

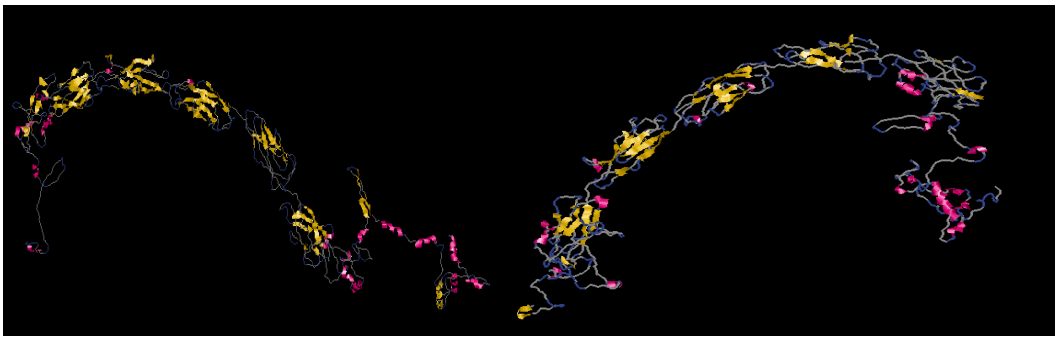
The protein 3D structure of human desmocollin-2 was predicted by using I-TASSER of threading method. I-TASSER (Iterative-Threading / ASSembly / Refinement) found to be a better server for protein structure prediction assessed by CASP9 experiments. The structural models from I-TASSER have good quality and resolution. The server generates five models and the best one can be selected based on C-score .C-score, a measure to observe the quality of resulting models showed the correlation quality of the model prediction results. From the C- score,model 1 was selected as the best predicted structure model of human desmocolloin-2. The overall model predictions were evaluated by using RMSD and TM-score. The

TM score for human desmocollin-2 model was 0.46±0.15 and RMSD was 14.1±3.8Å. The results showed the quality model generation by I-TASSER for the human desmocollin-2 protein.

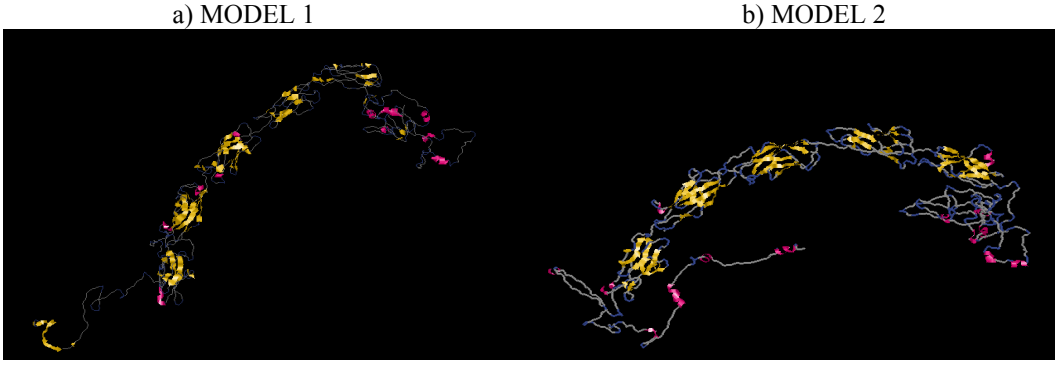
Table 1: C- score of I-TASSER models

STRUCTURE MODEL	C- Score
MODEL1	-2.17
MODEL2	-2.69
MODEL3	-2.73
MODEL4	-2.94
MODEL5	-3.03

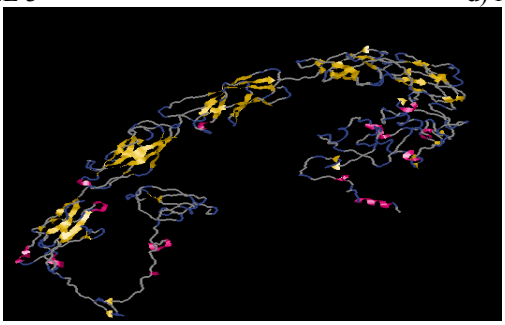
The best modeled structure of human desmocollin-2 was validated by PROCHECK in SAVES server (http://nihserver.mbi.ucla.edu/SAVES/). Ramachandran plot constructed by the PROCHECK explained the regions of amino acids in the plot and the overall quality of the model based on the phi-psi angles. The results showed human desmocollin-2 model was good based on the overall geometry. The amino acids arranged in most favored regions in plot were 80% and some of the amino acids in disallowed regions of 1.5%.



a) MODEL 1 b) MODEL 2



c) MODEL 3 d) MODEL 4



e) MODEL 5

Fig. 3: Result of structure models of human desmocollin-2 by I-TASSER.

4. Conclusion

Our study on structural modeling of human desmocollin-2 revealed the information about cadherin domain, and secondary structure arrangements which includes transmembrane helices. The tertiary structure model of human desmocollin-2 which contains five cadherin domain and ICS domain by I-TASSER

observed as the first report. The results confirmed the ICS domain in the cytoplasm of desmocollin-2 'C' terminal which specifically present in isoform 2a. The modeled structure may favor to carry out mutation analysis and drug design step towards the disease Arrhythmogenic right ventricular cardiomyopathy (ARVC).

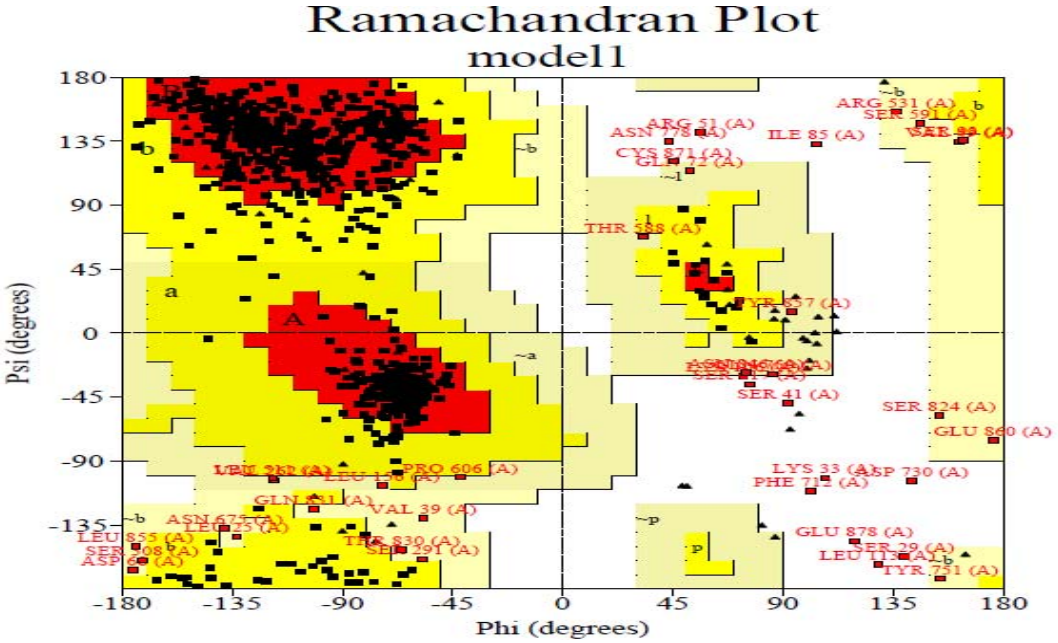


Fig. 4: Ramachandran plot for model 1.

5. References

[1] Ahmad F. The molecular genetics of arrhythmogenic right ventricular dysplasiacardiomyopathy.Clin Invest Med. 2003, 26: 167–178.

[2] Awad MM, Calkins H and Judge DP. Mechanisms of disease: molecular genetics of arrhythmogenic right ventricular dysplasia/cardiomyopathy. Nat Clin Pract Cardiovasc Med. 2008, 5: 258–267.

[3] Dusek RL, Godsel LM and Green KJ. Discriminating roles of desmosomal cadherins: beyond desmosomal adhesion. J Dermatol Sci. 2007, 45: 7–21.

[4] Lorimer JE, Hall LS, Clarke JP, Collins JE, Fleming TP and Garrod DR. Cloning, sequence analysis and expression pattern of mouse desmocollin 2 (DSC2), a cadherin-like adhesion molecule. Mol Membr Biol. 1994, 11: 229–236.

[5] Green KJ and Simpson CL. Desmosomes: new perspectives on a classic. J Invest Dermatol. 2007, 127: 2499 – 2515.

[6] Giorgia Beffagna, et al. Missense mutations in Desmocollin-2 N-terminus, associated with arrhythmogenic right ventricular cardiomyopathy, affect intracellular localization of desmocollin-2 in vitro. BMC Medical Genetics. 2007, 8: 65.

[7] M. Punta, P.C. Coggill, R.Y. Eberhardt, J. Mistry, J. Tate, C. Boursnell, N. Pang, K. Forslund, G. Ceric, J. Clements, A. Heger, L. Holm, E.L.L. Sonnhammer, S.R. Eddy, A. Bateman and R.D. Finn. The Pfam protein families database. Nucleic acids research. 2012, 40: 290-301.

[8] Garnier J, Gibrat J-F, Robson B .GOR secondary structure prediction method version IV Methods in Enzymology. R.F. Doolittle Ed. 1996, 266, 540-553.

[9] Ambrish Roy, Alper Kucukural and Yang Zhang. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols. 2010, 5, 725-738.

[10] Zhang, Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics.2008, 9: 40.

[11] Laskowski R A, MacArthur M W, Moss D S and Thornton J M. PROCHECK - a program to check the stereochemical quality of protein structures. J. App. Cryst.1993, 26: 283-29.

[12] Ramachandran, G.N., Ramakrishnan, C. and Sasisekharan. Stereochemistry of polypeptide chain configurations. J. Mol. Biol. 1963, 7: 95-99.