Insilico Characterization, Phylogenetic Analysis and Homology based 3D Structural Modeling of HSPA4 Heat Shock Protein Family a (Hsp70) Member 4 from Homo Sapiens (Human).

J.Venkateshwara Rao

Abstract: HSP70 are the specialized Heat Shock Proteins that show their novelty presence inside the deepest part of Cellular Component Networks of Chaperons and involve with various heat Shock repairing up mechanisms as main Catalytic Components. They actively participate in proper alignment and resistance of all types pressures by their proper dealing with substrates and proper exposure of their hydrophobic peptide portions inside of their substrate molecules. The exchange of information cum process of information occurs between the cell components through HSP70’s low level of affinity ATP bind sites and High Affinity level of ADP binding sites. Therefore, ATP hydrolysis process plays a key role in fundamental mechanisms, in vitro and as well as in in vivo repairing mechanisms for the functioning of Chaperone like functionalities of HSP 70 proteins. In the current study analysis, Homology Modeling methods permits us for modeling the 3D structure of the protein by taking backbone of experimentally determined 3D-Structures of homologous proteins extracted in various structural formats like PDBs. We worked with HSPA4 heat shock protein family A (HSP 70) part 4 from Homo sapiens ( Human) and conducted 3-D structure prediction using an Automated Swiss Model Generation by employing the crystal structures of Template molecule. PDB Blast and Template search for required protein were conducted with help of all the parameters with respect to Impact and Psi Impact angles and that leads to for most noteworthy sequence identity, structural alignments and functionalities. Finally, the homology modeling were performed by using SWISS modeler and modeled proteins were fine tuned by using the Ramachandran Plot, by using PROCHECK and 3D-Check Software Programs.

Keywords : Homology Modelling, HSPA4, HSP 70 Protein, Phylogenetic Analysis, Procheck, Swiss Modeller.

I. INTRODUCTION

Heat Shock Proteins ( Or Stress repair) proteins (HSPs) are a special group of family members of cytoplasmic proteins that works crucially for to procure security from external stress bursts and sever metabolic dysfunction corrections in living organisms [ 1-3]. These proteins are classified into three major families: Hsp90 (84-92 kDa), Hsp 70 (66-74 kDa) and mono molecular weights Hsps (18-48 kDa) [4]. N-terminal ATPase of size 44 kDa and a C-terminal substrate of substantial space of size 24kDa which arranges to form a β structure-sandwich and a sub domain size of 14kDa along with C-terminal related α-helical sub domain. The main characteristics of Hsp 70 family include two diversified and independent set of qualities. A constitutive “housekeeping” heat shock cognate (hsc) 70 domain, and a physiological stress inducible hsp 70 domain. The Hsc70 protein domain plays a primary role in aspects of chaperon in wide extent of cellular activities such as bringing of proteins together, collapsing of proteins or degeneration of proteins, transportation of them across the ionic channels, translocation to remote places and denaturation [4-7].

Hsp70 proteins are crucially function to give proper anchorage for cellular homeostasis that help in maintenance of balance in various chemical reactions, in that way ensuring the protection from harmful UV or γ-irradiation or Chemical mutations or Heat Shock phenomenon. HSP70 is very crucially a potential activator of the intrinsic heat shock resistant framework of networks [10-12]. HSP 70 is one of the most basic and fundamental protein found in all living organisms from very smallest archae bacteria to plants and higher level mammals including highly evolved human beings. Both Hsc and Hsp 70 proteins have very conserved region of N-terminal ATPase adjoining by substrate binding space (SBD) which contains hydrophobic domains with a cap like structure and reasonable variable C-terminal space, which plays a key role in Hsp70 activities that are required for cell maintenance and development.

During the ATP-bound stage, the substrate would bind quickly and ATP hydrolysis takes place and cellular events will be processed. Return to the native or original stage establishes previous condition by encouraging the substrate to leave from the site [14, 15]. Hsp70’s are involved in housekeeping activities of the cell, they serve for the purpose of rearrangements, rebuilding and passing the signal mechanisms for repair and quality control mechanisms and repairing wrongly placed conformers. Most of these activities of repairs are based upon the key properties of HSP70 and their associated hydrophobic peptide components of proteins in ATP derived regions of controls. Almost every regulatory protein of higher organisms like Eukaryotes is regulated by means of their transitory relations with respect to Hsp70.

These HSP70 proteins are having of some key molecular components (Steroid Hormone Components), Kinases like (Raf, elf2 α-kinase, CyclinB1/Cdk1) and protein production components (HSF, c-Myc, pRb).
Insilico Characterization, Phylogenetic Analysis and Homology based 3D Structural Modeling of HSPA4 Heat Shock Protein Family a (Hsp70) Member 4 from Homo Sapiens (Human).

These are common components in many of the developed Chaperone-Substrate based complexes. The formation of reactive compound complexes depends on the interactions of various components discussed as like above. This complex plays a high role in finding solutions in neurodegenerative, auto immune, viral infection and aging processes.

There are two main locations in the protein which are very important in carrying out several functional activities of the protein 1) Peptide Bonding Domain (PBD) and 2) Amino terminal ATPase Binding Domain (ABD). For carrying out functional activities like hydrophobic protein repairing mechanisms PBD plays a major role and to hold the proteins properly using a lid like structure ABD is crucial. The top portion of HSP70 protein remains open in condition when it tied with protein which it wants to repair but closes as soon as it is bound with ADP. In a dangerous disease condition like melanoma, this kind of protein is produced in large quantities [14]. But more expression we can in can in case of Renal Cell Cancers [15]. Most of the threat signals to immune system deliberated by extracellular portions of HSP70 against responses to disease conditions. They provide signals for empowerment of MHC class I and II and upgrades activation T Cells. The interactions of HLA-DR with mammalian cytoplasmic Hsp70 molecules gives us a clue of how the HLA-DR exchanges the antigenic bond peptides with ternary complexes in the channel of HLA-DR atomic molecules by HSP70 protein residues. Rohrer et al., emphasized on the interactions between HSP70 and HLA-DR, exterior groove of HLA-DR is allowed to interact with ATPase space of Heat Shock Proteins 70 and gives a chance of improved performance of antigenic displaying cell promoting for T cells proliferation [16]. The resources for taking homologues are based on X-ray structures or NMR. The precise structures give us accurate information regarding positional residues. In homology modeling, the tool has to find out a suitable format of structure with recognizable similar sequence identity. In case, if the suitability or matching falls below 35% the yield likely shows like imperfect nature of match or structure. We have to construct or rebuild a perfect model by considering all the constitutive parameters of the corresponding to group of homologues proteins with their functional groups [18].

For many modern biological applications now a day’s generation of 3Dcomputational models is an essential part, even to develop a better understanding of how these molecules interact with other molecules during 3D model development of repairs is a key option. The development of these 3D models depends upon the existing natural homology based structures or X-ray generated database models

Software Used: BLAST, PDB BLAST, RASOMOL, PDDBIWER, SWISS MODELLER, PROCHECK and 3D CHECEK.

II. RESULT AND DISCUSSION

Extraction of Protein Sequence

As being our part of study we extracted Hsp70 protein from Heat Shock protein family A part 4 from Homo sapiens (Human) with NCBI accession code: NCBI NM_002154.4. It is having 4774 nucleotide bases in length and 1148 peptide bases or amino acids in length (Table 1). The extracted sequence can be accessed from the following the URL link addressed at website page https://www.ncbi.nlm.nih.gov/nuccore/NM_002154.4

| Table 1: Target protein sequences are searched against the template sequences for Model Building |

Fig 1: Multiple Sequence Alignments of HSP 70 Family members
Table 2: BLAST HIT:

<table>
<thead>
<tr>
<th>Accession</th>
<th>Protein names</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01314</td>
<td>Heat shock 70kDa protein 4 (from M. tuberculosis)</td>
</tr>
<tr>
<td>D85148</td>
<td>Heat shock protein 60kDa (from B. subtilis)</td>
</tr>
<tr>
<td>P01315</td>
<td>Heat shock 70kDa protein 6 (from C. albus)</td>
</tr>
<tr>
<td>P01308</td>
<td>Heat shock protein 60kDa (from E. coli)</td>
</tr>
</tbody>
</table>

Table 3: Template Selection by PDB-BLAST.

<table>
<thead>
<tr>
<th>Template</th>
<th>Seq identity</th>
<th>Dali state</th>
<th>Oggi found by</th>
<th>Method</th>
<th>Resolution</th>
<th>Seq similarity</th>
<th>Coverage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3c7n.1A</td>
<td>39.68</td>
<td>monomer</td>
<td>HHBlits</td>
<td>X-ray</td>
<td>2.12A</td>
<td>0.39</td>
<td>0.77</td>
<td>heat shock protein 70kDa</td>
</tr>
<tr>
<td>3c7n.1A</td>
<td>39.63</td>
<td>monomer</td>
<td>HHBlits</td>
<td>X-ray</td>
<td>2.35A</td>
<td>0.38</td>
<td>0.77</td>
<td>heat shock protein 70kDa</td>
</tr>
<tr>
<td>2px1.1A</td>
<td>39.67</td>
<td>dimer</td>
<td>HHBlits</td>
<td>X-ray</td>
<td>2.41A</td>
<td>0.58</td>
<td>0.78</td>
<td>heat shock protein 70kDa</td>
</tr>
<tr>
<td>5a6d.1A</td>
<td>33.50</td>
<td>monomer</td>
<td>HHBlits</td>
<td>X-ray</td>
<td>2.98A</td>
<td>0.57</td>
<td>0.71</td>
<td>heat shock protein 70kDa</td>
</tr>
<tr>
<td>5a6d.4A</td>
<td>33.50</td>
<td>monomer</td>
<td>HHBlits</td>
<td>X-ray</td>
<td>2.98A</td>
<td>0.57</td>
<td>0.71</td>
<td>heat shock protein 70kDa</td>
</tr>
</tbody>
</table>

Retrieval Number: A9476109119/2019©BEIESP
DOI: 10.35940/ijeat.A9476.109119

Fig 2: Phylogeny Tree of HSP 70 Family members

Template search and selection
To the concerned target proteins, formats were searched and looked for most suitable HHBlits against the Swiss model layout Library (SMTL) [18-22]. A total of 98 formats were selected. For every suitable layout, the structural identity and quality parameters of template-target were analyzed for the selection process. Among all the formats chosen 3c7n.1 was selected which is a most outstanding resemblance showing with target sequencing with grouping personality score of 39.88 and with other similar structural formats of template molecules are 3d2e.1 chain A and 2qx1.1 chain with the grouping scores of 39.63 and 39.91 respectively, but their grouping scores are much lesser than 3c7n.1. Their scope is found to be less than 0.79. So, from above data 3c7n.1 was chosen as template molecule for structure building of a new model.

Fig 3: Target vs Template Alignment.

Protein Modelling of Target Molecule:
The procedure followed here is based on the building of target molecule on target-template arrangement using the Software Insight Builder II. The arrangement of atoms is based on the highly narrowed selection based on similar alignments. The template molecule 3c7n.1 is found to be the most extreme closely related template format compared to any one and the alignments that were tested between targets and template structural layouts. All the functionalities from template molecules are retrieved by proper alignments preserved in template molecule form. For this purpose the inbuilt library functionalities were taken for help and in addition, modification, edition, deletion and realignments and rebuilding of side chains were also made. Drive field forces were employed to properly orient the geometry of the model. The resulted model was tested and visualized by using Deep Viewer, Rasmol and PDBviewers. The global and mean value scores of all the residues were evaluated by using QMEAN4 software using it’s scoring matrices [22].

Fig 4: Template Characteristics and Q Mean Values.

Modelling of Ligand Molecules:
It was observed that may key functional groups of ligands such as BEF, ADP, SO4 are within the allowed regions or boundaries of the template structure and these
are not differing from the structure or functional nature of target molecule.

The protein structure is evaluated by using many parameters by employing software programs like Structure approval computer program Suite (PSVS). This suite highly valuable in evaluating the structures that are coming from NMR, x-ray Crystallography and homology based models. PSVS measures the values coming from many parameters that are useful to check in programs like RPF, PROCHECK, 3D Modelling, Insight II, MolProbilty, PDB approval programs and various other structural accepted notions [23]. PSVS gives overall inputs about standard limitation investigations or explorations, checking goodness of fit for various approved structures, Z-score values and integration of all the databases regarding information based on structure based quality scores.

Display of 3D- Model Building: The resulted structure of origin of protein was thoroughly visualized in 3D parameters using RasMol atomic structure visualization. The RasMol program has the capacity of visualizing the data concerned side chain, particles and bonds between the residues and every minute information can be visualized and examined [24].

Analysis of the Model:
HSP70 protein was constructed based on target template alignments and arrangements since the structural arrangement of 3c7n.1 is more suitable and worthy consideration (Table 3) and based on that model.

The results from PSVS suite gives us properties related to stereo chemical nature of peptides. The atomic weight of target molecule is found to be 131896. The calculated RMS deviation of bond points is observed to be 1.4°. Several nearby contacts are found to be in the range of 2.2 Å, and bond lengths are found to be 0.011 Å. From the retrieval of database records mean and standard deviation from 251 structures of more than 500 build ups of the structural models ≤ 1.79 Å average determination values are, and R-free ≤ 0.27, R-factor ≤ 0.24. This positive trend demonstrates a better score for chosen buildup. 3A-547B shows the 3D protein structure as in Figure 3.

The Statistics of Ramachandran Plot:
The Ramachandran plot analysis revealed that the psi and phi spiral conformational points are within the target protein allowed regions of Hsp70 as shown in Figure 7. The red colour zone in Ramachandran Plot corresponds to core of the local regions and gives us a confirmation of combination of foremost positive nature of phi-psi values. Few of the residues are found in the permitted regions. The quality of buildup of the model is found in the central core portion of Ramachandran Plot as depicted in Figure 7.

![RMS distances from planarity](image-url)
In the present study, we have verified the 3D structures of selected Hsp70 proteins by using homology modeling methods, Tools and visualizing the same with the assistance of online available bioinformatics software tools and techniques. It was found that 3-D models of Hsp70 protein met the requirements and striking resemblance with target sequence alignments. Our homology modeling has compared the relatedness between the models and target molecules. In the current model template is homologous protein which can be distinguished with target molecule by 89.45% of identity and satisfying of all thermodynamic properties. The expectations of allowed regions of residues satisfies almost all the regions of protein surface. Ramachandran Plot graph examination reveals that from PROCHECK the maximum residues are found in the allowed regions, i.e., 80.20% of the buildup residues are found in the most favored regions and also structures from Richardson’s lab Molprobity display 86.2% in allowed region for given residues. The present study gives the long terms benefits of understanding the 3D atomic arrangements of protein molecule and develops the foremost accurate model for Hsp70 molecule of human and its functionalities.

REFERENCES
Insilico Characterization, Phylogenetic Analysis and Homology based 3D Structural Modeling of HSPA4 Heat Shock Protein Family a (Hsp70) Member 4 from Homo Sapiens (Human).

23. http://psvs-1_5-dev.nesg.org/results/HSP/OUTPUT.html

AUTHORS PROFILE

Dr. J. Venkateshwara Rao, M.Sc. B.Ed, M.Tech, Ph.D, Post Doc (USA), BOYSCAST Fellow, is an Asst.Professor in Dept of Zoology, Osmania University, joined in Osmania University in 2007. In 2009, he got Young Scientist and Investigator award from DBT under DBT-RGYI Scheme and in 2010 he got most prestigious DST BOYSCAST fellowship to pursue high end research in advanced areas. He went to Virginia Common Wealth University, Richmond and conducted research in Signal transduction cascades in smooth muscle cells of Gastrointestinal Muscles.