Computational Analysis of HSP-60 Protein with Structural Insights into Chaperonin Containing TCP-1 Subunit 5 in Bos Taurus

Abstract

Heat shock protein 60 kDa (HSP60) is a crucial chaperone that guides appropriate folding of denatured proteins under heat stress conditions. HSP60 provides assistance in correct folding of a multitude of denatured proteins. The Group II eukaryotic chaperonin TCP-1 ring complex is (TRiC or CCT) important cytosolic chaperonin which plays important role in folding of essential subset of cytosolic proteins and is believed to be highly conserved among different eukaryotic species. In this study, an extensive in silico analysis has been performed ranging from sequence comparison among species to homology modeling of Bos taurus Cct5 protein and determination of Hsp60 interacting partners. The comparative analysis of the protein revealed certain significant variations in Bos taurus. The experimental protein structure for Cct5 in Bos taurus is unknown till date, therefore the putative protein structure was determined by utilizing several scrutiny tools such as PROCHECK, VERIFY3D and RAMPAGE to ensure model accuracy. The mode of interaction between HSF1 and Cct5 protein was evaluated by molecular docking studies.

Keywords: Heat shock protein; Cct5; Bos Taurus; Homology modeling; Docking

Abbreviations: HSP: Heat Shock Protein; CCT: Chaperonin Containing TCP-1 Ring Complex; TRIC: TCP-1 Ring Complex; Chaperonin Containing TCP-1 Subunit 5

Introduction

Proteins being structural and functional blocks of all significant biological processes are synthesized as polypeptides comprising of amino acids. The molecular activities and functions are believed to be carried out via the protein machinery and other accessory components in vivo. The catalytic activity of an enzyme typically relies on its active site and the binding affinity to a compound. These sites permit the interactions among different proteins, their subunits or general structures which are devised for specific purposes. In order to retain the activity of active sites in a protein, correct folding of that particular protein is a vital process. In rare cases, the correct folding of these proteins is accomplished spontaneously, but frequently it requires the protein to interact with other proteins while it is being synthesized or while they are denatured under stress conditions. So these proteins which guide the appropriate folding of other proteins are called chaperones. The chaperone’s machinery basically locates those features in proteins that are not exposed in matured confirmed structures, typically detecting hydrophobic stretches/regions that usually aggregate in the center of mature protein structures but are exposed during unfolding or under denatured conditions [1].

After translation, the newly synthesized proteins leave ribosomes as linear amino acid chains and then chaperone recognize region those which have ability to aggregate the chain. By binding to these interacting regions, chaperones prevent the inappropriate interactions within different chains of proteins and monitor for the correct folding confirmation. The chaperonin folds the substrate in a closed central cavity, thereby providing the specific microenvironment favoring the folding reaction towards the phase where potential of folding is maximized and entropy gets minimized for the unfolded substrate. In eukaryotes the TCP-1/tcp-1 ring complex (CCT/TRiC) is one of the most important cytosolic subset of proteins required for living cells for maintaining their viability. This complex interacts with approximately ten percent of the freshly synthesized proteins and is involved in the folding of both skeletal and non-skeletal proteins [2]. It possesses a unique ability to fold some specific proteins such as actin, which are not folded by any other chaperonin. With the advent of proteomics studies, it was evident that CCT substrates include a number of proteins containing Tryptophan-Aspartic acid repeats involved in the formation of β-propeller domain. The size of CCT/TRiC substrates varies drastically ranging from small to very large and even some with molecular weight higher than 100 kDa [3]. These large proteins are folded by a mechanism of domain-wise folding where a substrate can be folded outside the cavity and thus, preventing encapsulation of the whole substrate.

Structurally, chaperonins are classified mainly into two
groups: Group I and Group II chaperonins. The subunits of both groups share a similar basic structural pattern and consist of three domains: a) an apical domain which is related to the binding of substrate, b) equatorial domain, and function related to ATP binding, c) hinge domain which enables the connection between rests of the two domains [4]. Endosymbiotic organelles and prokaryotic cells were found to contain in Group I chaperonin and Group II chaperonin in Eukaryotes and Archaea [5].

It was observed that CCT interacts with some very important proteins like actin, tubulin and von Hippel-Lindau tumor suppressor proteins etc. in subunit-specific and geometry dependent-manner [6]. By using electron microscopy, three dimensional reconstruction of alpha helix showed the interaction of both small and large domain of alpha actin with different subunits of CCT, whereas Cct5p subunit found to interact with the large domain. Furthermore, electron microscopic studies and immunoprecipitation experiments of actin CCT complexes has revealed that Cct5 has highest substrate affinity than other subunits [7]. On the other hand, tubulin comprises of total eight binding sites with CCT. Three binding sites are located at N-terminal and five binding sites at C-terminal, among the five binding sites at C-terminal two were found to interact with Cct5 [8].

Heat shock transcription factors play important role in regulation of expression of molecular chaperones that are expressed during heat stress. The heat shock factors are the family of some transcription factors which controls the regulation of gene expression of proteins involved in folding of damaged or improper folded proteins during stress conditions. Among various transcription factors, one of the most important transcription factors is heat shock factor 1 [9]. Under normal conditions heat shock proteins bind with HSF1 and act as its repressor and in addition, maintaining its nonactive and monomeric confirmation [10]. Under stress conditions, these heat shock proteins start to bind with damaged and denatured proteins thus releasing the HSF1, which undergoes homotrimerization (active form) and now HSF1 is able to translocate to the nucleus for DNA binding to up regulate the gene expression of heat shock proteins. These interactions between HSF1 and Cct5 have been extensively analyzed in the current study using molecular docking studies.

Cattle, being an indispensable livestock of India is reared in most of the areas bearing hot climatic conditions and in response to such extreme conditions, these animals express many heat shock proteins as a part of general adaptation phenomenon. Considering the fact, present study attempts to describe the 3D model of chaperonin containing TCP-1 subunit 5 of Bos taurus which provides a rational opportunity to obtain a reasonable 3D model in the absence of an experimentally determined crystal structure. The study also focuses on the physiochemical properties of HSP 60 protein and crucial interactions observed in Bos taurus.

Materials and Method

Categorization of Hsp60 Proteins

All the 20 available hits for Hsp60 proteins from HSPIR database [11] were compiled and tabulated on the basis of different selection criterions. These include the type of chaperones, absence/presence of domain, motif and signal peptide, their sub cellular location, their involvement in the biological processes and their molecular function.

Retrieval and Characterization of target sequence

The 541 amino acid long sequence of Cct5 with Accession number (NP.001029767) was obtained from the NCBI protein reference database. The physiochemical properties of the Cct5 protein were then predicted. The transmembrane helices of Cct5 were also obtained by TMHMM 2.0 [12], while the molecular weight, total number of negatively (Asp + Glu) and positively charged residues (Arg + Lys), grand average of hydropathicity (GRAVY), aliphatic index, isoelectric point, instability index and extinction coefficient of Cct5 were computed using Expasy’s ProtParam proteomics server [13].

Comparative sequence analysis for Cct5 protein

The homologous protein sequences of Cct5 from 10 species i.e. B.T (Bos taurus), R.N (Rattus norvegicus), M.M (Mus musculus), H.S (Homo sapiens), B.B (Bubalus bubalis), O.A (Ovis aries), C.H (Capra hircus), SS (Sus scrofa), E.C (Equus caballus), C.F (Camelus ferus) were retrieved from NCBI protein reference database. The comparative analysis for the protein was carried out using MAFFT multiple sequence alignment tools [14], which revealed important variations in Bos taurus Cct5 protein.

Model quality assessment

The information on beta strands and alpha helix constituting the secondary structure of Cct5 was retrieved via Phyre2 server; this server is based on Hidden Markov model [15]. The protein sequence of Cct5 deduced the secondary structure and domain composition. The proteins interact with one another and perform various cellular functions in tertiary form. The tertiary structure for Cct5 protein in Bos taurus is unknown till date so an attempt was made to generate a homology model for the same in the study. The template (1DB1 - 3lygE) was used for comparative model building of Cct5 in Bos taurus using Phyre2 server. The model after being constructed using Phyre2 server was subjected to refinement and energy minimization via ModRefiner program. Energy minimization is an important aspect, as the comparison, determination and the evaluation of structure can be done by the criteria of energy minimization, even if there is unavailability of experimental structure [16]. The accuracy and stereo chemical quality of predicted model was observed from RAMPAge, QMEAN value and verify 3D score. Additionally, the overall G-factor was estimated using PROCHECK program from SAVES server [17]. Finally, the protein was visualized in Swiss-PDB viewer [18].

Elucidation of interacting partners of Hsp 60 and docking studies

For determining the vital interactions demonstrated by Hsp60 in Bos taurus for performing various cellular activities, STRING database [19] was utilized and information on protein-protein interactions were deduced. In the present study, efforts are made
to investigate the possible mode of interaction between HSF1 and CCT5. For the same, HSF1 was docked within the CCT5 homology model using Clus Pro 2.0 docking server where CCT5 was used as a receptor and HSF1 was used as a ligand. Clus Pro 2.0 is an automated server based on an algorithm which can discriminate thousands of confirmations (2000 confirmations) of two proteins on the basis of different electrostatic and desolvation potential. The latter can be further discriminated via clustering and finally give the best fit structure between two proteins which was generally found closest to native structure from x-ray crystallography results [20].

Results and Discussion

Since, CCT5 protein form an integral component of CCT which are involved in folding of denatured proteins under stress conditions in Bos taurus therefore a comprehensive in silico structure elucidation & stereo chemical properties evaluation of Cct5 protein was performed. In addition, classification & protein-protein interaction of Hsp 60 protein was studied. Table 1 depicts all the reported Hsp 60 proteins along with their types, domains, cellular locations and functions. The majority of Hsp60 proteins belongs to Group II Chaperonin class and only six from Group I Chaperonin class and is implicated mainly in ATP binding.

Table 1: HSP 60 proteins along with their types, domains, cellular locations and functions.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Heat Shock Proteins</th>
<th>Type of Chaperone</th>
<th>Domain</th>
<th>Motif</th>
<th>Signal Peptide</th>
<th>Sub Cellular Location</th>
<th>Biological Process</th>
<th>Molecular Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HSP60_1329</td>
<td>Group I chaperonin</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Mitochondrion</td>
<td>Protein Folding</td>
<td>ATP binding</td>
</tr>
<tr>
<td>2</td>
<td>HSP60_1330</td>
<td>Group I chaperonin</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Mitochondrion</td>
<td>Cellular metabolic process</td>
<td>ATP binding</td>
</tr>
<tr>
<td>3</td>
<td>HSP60_1331</td>
<td>Group I chaperonin</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>Nucleus</td>
<td>Cellular metabolic process</td>
<td>ATP binding</td>
</tr>
<tr>
<td>4</td>
<td>HSP60_1332</td>
<td>Group I chaperonin</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>Cytoplasm</td>
<td>Cellular metabolic process</td>
<td>ATP binding</td>
</tr>
<tr>
<td>5</td>
<td>HSP60_1333, 1335, 1336 and 1337</td>
<td>Group II chaperonin</td>
<td>Chaperonins TCP</td>
<td>-</td>
<td>No</td>
<td>Cytoplasm</td>
<td>Protein Folding</td>
<td>ATP binding</td>
</tr>
<tr>
<td>9</td>
<td>HSP60_1338</td>
<td>Group II chaperonin</td>
<td>Chaperonins TCP</td>
<td>-</td>
<td>No</td>
<td>Melanosomes</td>
<td>Protein Folding</td>
<td>ATP binding</td>
</tr>
<tr>
<td>10</td>
<td>HSP60_1339</td>
<td>Group II chaperonin</td>
<td>Chaperonins TCP</td>
<td>-</td>
<td>No</td>
<td>Cytoplasm</td>
<td>Protein Folding</td>
<td>ATP binding</td>
</tr>
<tr>
<td>11</td>
<td>HSP60_1340</td>
<td>Group II chaperonin</td>
<td>Chaperonins TCP</td>
<td>-</td>
<td>No</td>
<td>Cytoplasm</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>HSP60_1341, 1342, 1343, 1345</td>
<td>Group II chaperonin</td>
<td>Chaperonins TCP</td>
<td>-</td>
<td>No</td>
<td>Cytoplasm</td>
<td>Protein Folding</td>
<td>ATP binding</td>
</tr>
<tr>
<td>16</td>
<td>HSP60_1346</td>
<td>Group II chaperonin</td>
<td>Chaperonins TCP</td>
<td>-</td>
<td>No</td>
<td>Microtubule</td>
<td>Protein Folding</td>
<td>ATP binding</td>
</tr>
<tr>
<td>17</td>
<td>HSP60_1347</td>
<td>Group II chaperonin</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>Cytoplasm</td>
<td>Protein Folding</td>
<td>ATP binding</td>
</tr>
<tr>
<td>18</td>
<td>HSP60_1348</td>
<td>Group I chaperonin</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>Cytoplasm</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>HSP60_1349</td>
<td>Group II chaperonin</td>
<td>DEP domain, Phosphatidylinositol phosphate kinase, (PIPK) domain</td>
<td>Five finger FY VE / FY VE</td>
<td>No</td>
<td>Nucleus</td>
<td>Cellular metabolic process</td>
<td>Phosphatase kinase activity</td>
</tr>
<tr>
<td>20</td>
<td>HSP60_1350</td>
<td>Group I chaperonin</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>Mitochondrial matrix</td>
<td>Protein Folding</td>
<td>ATP binding</td>
</tr>
</tbody>
</table>

Citation: Singh R, Kumar V, Rajesh C, Gurao A, Kulshrestha A et al. (2017) Computational Analysis of HSP-60 Protein with Structural Insights into Chaperonin Containing TCP-1 Subunit 5 in Bos Taurus. MOJ Proteomics Bioinform 6(1): 00183. DOI: 10.15406/mojpb.2017.06.00183
Comparative analysis and characterization of target sequence

The multiple sequence alignment among the 5 species has been presented in Figure 1 where most of the sequence was found to be conserved with minor variations in context to B.T (Bos taurus), R.N (Rattus norvegicus), M.M (Mus musculus), H.S (Homo sapiens), B.B (Bubalus bubalis), O.A (Ovis aries), C.H (Capra hircus), S.S (Sus scrofa), E.C (Equus caballus), C.F (Camelus ferus). The observed variations in CCT5 have been highlighted in red color in the alignment. The primary structure of CCT5 as estimated from Expasy’s ProtParam server showed that the subunit has a total of 541 amino acids with molecular weight of 59615.0. Total number of negatively charged residues (Asp+Glu) in the sequence is 81, whereas 69 positively charged residues (Arg+Lys) were found. The grand average of hydropathicity (GRAVY) is -0.175 which is calculated by summing up of all the hydropathy values of amino acids and then dividing them by number of residues in the whole protein. Thus, the lower value of GRAVY indicates the better interaction of CCT5 protein with water i.e., more hydrophilic nature [21]. Interactions in solution and also help in calculating the concentration of a particular protein. TMHMM server reported that no region of CCT5 located in transmembrane (Figure 2).

Determination of CCT5 tertiary structure

The protein sequence of CCT5 was used to deduce the secondary structure and disorder prediction using Phyre2 program. The results showed the alpha helix region dominating other secondary structure (50%) followed by beta strand (12%) and disorder region is found to be very less (10%) in the protein. The high percentage of alpha helix region in protein indicates the stability and conservation of protein structure.

The three dimensional (3D) model for CCT5 protein (Figure 3) was generated by Phyre2 server and the model refinement and energy minimization was done by ModRefiner program that utilizes an algorithm for atomic level to develop and refine protein structures, critically based on a two-step, atomic-level energy minimization [22]. Ramachandran plot of model as illustrated in Figure 4 was analyzed by RAMPAGE server and overall G factor from PROCHECK. If the three dimensional profile score S for the amino acid sequence is high only then, the 3D protein model is considered correct [16] and Verify 3D analysis of best model revealed that 80.96% of the residues had an average 3D-1D score <0.2, indicating that the model is compatible with its sequence (Figure 5). Overall G-factor was found to be -0.41 with 94.4% residues were present in favored region of Ramachandran plot. The numbers of residues in allowed region were 4.6%. None of the residues was present in the disallowed region of plot as, per the acceptance criteria that the quality of model is reliable and good if the 90% of residues fall under allowed region. Qualitative Model Energy Analysis termed as QMEAN analysis [23,24] is a composite scoring function describing the major geometrical aspects of protein structures. Our study predicts this model has comparable qualities to experimental structures, its value being 0.562, as for a good model QMEAN value lies between 0-1 (Figure 6).
Protein-protein interactions and docking analysis

A single protein may not perform all the tasks individually therefore, these proteins interact with one another and form complexes that may initiate or regulate numerous biological processes. The HSP 60 protein interacts with the members of heat shock proteins and often other proteins to guide the folding process. These interactions are important to identify for extensive understanding of intricate biological mechanisms and pathways. The HSP 60 protein is found to be interacting with many proteins such as HSPA9, HSP90AA1, HSP90AB1, HSPCB, TLR4, HSPA1A, HSPA8, HSPE1 and HSPA4 respectively (Figure 7). These proteins are mainly molecular chaperones and are involved in cell-cycle regulation and signal transduction (Table 2). The Figure 7a represents the evidence view from STRING database where the edges depict the type of available evidence for the interaction i.e. experimental, text-mining, co-expression, co-occurrence, gene fusion, and neighborhood or through database. Whereas, Figure 7b shows the confidence view where the thickness of the edge represents the confidence in interaction.

Table 2: HSP 60 interacting proteins during cell-cycle regulation and signal transduction.

<table>
<thead>
<tr>
<th>Interacting Protein</th>
<th>Protein Size (Aa Chain Length)</th>
<th>Protein Details</th>
<th>Combined Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSPE1</td>
<td>102</td>
<td>10 kDa heat shock protein, mitochondrial; Eukaryotic CPN10 homolog which is essential for mitochondrial protein biogenesis, together with CPN60. Binds to CPN60 in the presence of Mg-ATP and suppresses the ATPase activity of the latter</td>
<td>0.999</td>
</tr>
<tr>
<td>HSPA9</td>
<td>679</td>
<td>Stress-70 protein, mitochondrial; Implicated in the control of cell proliferation and cellular aging</td>
<td>0.988</td>
</tr>
<tr>
<td>HSP90AA1</td>
<td>733</td>
<td>Heat shock protein HSP 90-alpha; Molecular chaperone that promotes the maturation, structural maintenance and proper regulation of specific target proteins involved for instance in cell cycle control and signal transduction. Undergoes a functional cycle that is linked to its ATPase activity. This cycle probably induces conformational changes in the client proteins, thereby causing their activation. Interacts dynamically with various co-chaperones that modulate its substrate recognition, ATPase cycle and chaperone function</td>
<td>0.946</td>
</tr>
<tr>
<td>HSP90AB1</td>
<td>724</td>
<td>Heat shock protein HSP 90-beta</td>
<td>0.941</td>
</tr>
</tbody>
</table>

Citation: Singh R, Kumar V, Rajesh C, Gurao A, Kulshrestha A et al. (2017) Computational Analysis of HSP-60 Protein with Structural Insights into Chaperonin Containing TCP-1 Subunit 5 in Bos Taurus. MOJ Proteomics Bioinform 6(1): 00183. DOI: 10.15406/mojpb.2017.06.00183
Heat shock protein HSP 90-beta; Molecular chaperone that promotes maturation, structural maintenance and proper regulation of specific target proteins involved for instance in cell cycle control and signal transduction. Undergoes a functional cycle that is linked to its ATPase activity. This cycle probably induces conformational changes in the client proteins, thereby causing their activation. Interacts dynamically with various co-chaperones that modulate its substrate recognition, ATP cycle and chaperone function.

Toll-like receptor 4 precursor; Cooperates with LY96 and CD14 to mediate the innate immune response to bacterial lipopolysaccharide (LPS). Acts via MYD88, TIRAP and TRAF6, leading to NF-kappa-B activation, cytokine secretion and the inflammatory response.

Heat shock 70 kDa protein 1A; In cooperation with other chaperones, Hsp70s stabilize pre-existent proteins against aggregation and mediate the folding of newly translated polypeptides in the cytosol as well as within organelles. These chaperones participate in all these processes through their ability to recognize non-native conformations of other proteins. They bind extended peptide segments with a net hydrophobic character exposed by polypeptides during translation and membrane translocation, or following stress-induced damage.

Heat shock cognate 71 kDa protein; Acts as a repressor of transcriptional activation. Inhibits the transcriptional coactivator activity of CITED1 on Smad-mediated transcription. Chaperone. Component of the PRP19-CDC5L complex that forms an integral part of the spliceosome and is required for activating pre-mRNA splicing. May have a scaffolding role in the spliceosome assembly as it contacts all other components of the core complex.

Heat shock 70 kDa protein 4

Uncharacterized protein; Probably plays a role in facilitating the assembly of multimeric protein complexes inside the ER.
Acknowledgements

Authors are thankful to Indian Council of Agriculture Research-National Bureau of Animal Genetic Resources, Karnal, India for providing all the necessary facilities for carrying out the work.

References