JEB
Journal of Environmental Biology

Journal Website : www.jeb.co.in
E-mail : editor@jeb.co.in
ISSN: 0254-8704 (Print)
ISSN: 2394-0379 (Online)
CODEN: JEBIDP

Interpretation and analysis of functional domain and family prediction of Hsp90 protein-an insilco approach

Raid-Al-Akeel*

Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, Riyadh-11433, Kingdom of Saudi Arabia

*Corresponding Author E-mail: raalakeel@ksu.edu.sa

Publication Info

Paper received:
11 June 2015

Revised received:
05 August 2015

Re-revised received:
14 September 2015

Accepted:
28 September 2015

Abstract

Heat shock proteins (Hsp) are cellular proteins that are present under normal conditions but their expression level increases when cell is under stress or shock. In the present study, theoretical structure of HSP protein was built using ESYPred3D server by submitting the target sequence. Based on the template structure, it was clearly observed that the theoretical structure generated was structurally similar to the template structure which was highly sufficient for the development of specific ligand for HSP Protein 90 (p value 0.870). The present study would aid in studying protein-protein interactions in future and could be used further for drug target studies.

Key words

Hsp90 Protein, Protein-Protein interaction, Secondary structure

Introduction

Heat shock proteins , also known as 'Stress Proteins' are among the most abundant intracellular proteins present in a variety of microbial agents and mammals. These proteins have been classified into different families like Hsp110, Hsp90, Hsp70, Hsp60, Hsp40, Hsp20-30 and Hsp10 according to molecular mass (Rajaiah and Moudgil, 2009). Usually HSPs are present in low concentration and show distinctive noticeable differences in their expression in most organisms under different range of temperatures (Place and Hoffman, 2001), but their level increases under stress conditions such as exposure to high temperature, poison, oxyradicals, heavy metals, pathological conditions like ischemia and reperfusion inflammation and tissue damage, infection, mutant proteins associated with genetic diseases and cold temperature. Heat shock proteins prevents assembly under stress conditions through hydrophobic binding to proteins. To a large extent, molecular chaperones are not part of final structure of proteins folded. The release of the pillars of molecular chaperone at the end of the process. Many molecules associated with the performance of the form of ATP-dependent on the substrate interaction, such as Hsp100, Hsp90, HSP70, and HSP60 (Mayer *et al.,* 2000). Several reports have linked heat shock response with fungal pathogen of humans with *Candida albicans* (Alistair *et al.,* 2010) and in the leading human fungal pathogen, *Candida albicans*, only two interactions have been identified (Stephanie *et al.,* 2012).

Hsp90 is a vital and conserved molecular chaperone in eukaryotes that assist with folding diverse proteins, specially regulators of cellular signaling. By activating signaling in response to environmental cues, Hsp90 has a reflective impact on myriad aspects of biology. In fungi, Hsp90 influences development, drug resistance and evolution, and in *Saccharomyces cerevisiae*, Hsp90 interacts with ~10% proteins. Hsp90 protein preserves the tertiary structure of the proteasome which degrades ubiquitinated proteins involved in polyubiquitation pathway to degrade eukaryotic proteins that are no longer needed or misfolded or otherwise damaged in the cell. Investigations carried out with heat sensitive Hsp90 mutants and 26S proteasome suggests that Hsp90 is responsible for most of the ATPase activity in proteasome (Imai *et al.,* 2003). Edward *et al.* (2003) recognized 47-kDa antigen as sub fragment of *C. Albicans* Hsp90 and found it in the serum of patient suffering with candidiasis. This fragment was first identified as one of the main targets of immunity in patients with systemic candidiasis. In light of the above the present study was carried out to predict the protein structure, family, domain, location, localisation cellular

role and enzyme class of Hsp90 protein.

Materials and Methods

Protein family prediction analysis using P fam : Proteins are usually composed of one or more functional regions, generally termed as domains. Different combination of domains give rise to various range of proteins originate with in nature. In order to determine the functional annotation and to investigate gene family, heat shock protein 90 (Hsp90) were used to search protein families using Pfam search. Pfam (27.0) database, which is a large group of protein families, each characterized by multiple sequence alignments and hidden Markov models (HMMs) was used. In Pfam, each family is represented by multiple sequence alignments and Hidden Markov models (HMMs) (Finn *et al.*, 2008).

The submitted Hsp90 sequences at Pfam server predicted the protein families, motifs and repeats at default pfam parameter. Identification of domains that occur within proteins can, consequently, provide an insight into their function. Pfam also produces higher-level alliances of related families, known as clans. A clan is an assortment of Pfam-A entries which are related by similarity of sequence, structure or profile-HMM (Punta *et al.,* 2012).

Prediction of protein location : Automated prediction of protein subcellular localization is an important tool for bacterial genome annotation and drug discovery. PSORT is one of the most widely used computational methods for protein localization and analysis. PSORT is a computer database for the likelihood expectation of protein localization sites in cells. For PSORT, the user is asked to submit one or more amino acid sequences in FASTA format. Results are returned on a new web page. In the present study, PSORT gathered information of HSP90 sequences and its source origin, *e.g.,* Gram-negative bacteria, as inputs. Then, it estimated the input sequence by applying stockpiled rules for various sequence features of identified protein sorting signals. Lastly, it reported the possibility for input protein to be localized at each candidate site with further information. Thus, computational prediction of subcellular location from amino acid sequence information would help in annotation and functional prediction of protein coding genes in complete genomes. Protein subcellular localization likelihood expectation involves computational prediction of where protein exist in a cell. The function of protein such as cellular role, enzyme class and gene ontology of query protein sequence can be expected using the tool Protfun. Cellular role provides information about the functional category to which the query protein belongs, while enzyme class gives information if protein is enzyme or non- enzyme.

Domain Analysis : Domain analysis was performed to find the functional regions of protein by using tool SMART which allows identification and annotation of genetically mobile domains and analysis of domain architectures.

Results and Discussion

Heat shock proteins (HSPs') are evolutionally conserved from microorganism to mammals and play important role in many biological processes including thermal tolerance (Park *et al.,* 2008).Heat shock protein 90 (Hsp90) is one of the common heat-related proteins present in bacteria and all branches of eukarya, but it is actually absent in archaea (Chen *et al.,* 2006). *Candida* species account for 88% of all hospital-acquired fungal infections (Pfaller and Diekema, 2010). *C. albicans* is an important fungal pathogen in humans with mortality rates upcoming 50%, and is the fourth common source of hospital-acquired infection (Pfaller and Diekema, 2007; Zaoutis *et al.,* 2005). Fungi not only offer the most prevailing eukaryotic genetic model systems but also act a key threat to human health, and in view of this Hsp90 holds great aptitude as a therapeutic target (Cowen, 2008). Invasive fungal contagions are primary cause of mortality among immuno compromised individuals, comprising those with cancer and HIV. Compromising Hsp90 function can transform antifungals from Ineffective to highly efficacious proteins otherwise lethal infections can caused by the most prevalent fungal pathogens of humans, like *Candida albicans* and *Aspergillus fumigates* (Cowen *et al.,* 2009). In *C. albicans*, Hsp90 not only controls drug resistance, but also morphogenesis and virulence (Shapiro *et al.,* 2009). ProtFun 2.2 server was used to create the estimates of protein function from sequence (Q71QT8). This technique queries a large number of other feature expectation servers to obtain evidence on various post-translational and localization features of protein, which are integrated into final predictions of cellular role which involves Functional category of Transport binding (OR=1.885) and belongs to enzyme class Ligase (OR= 1.285) (Table 1)**.** Molecular chaperone Hsp90 showed a crucial role in folding and maturation of regulatory proteins. ProtFun 2.2 server uniformly predicted the function of HSP90 to be transport binding, enzymes and also voltage gated ion channel etc. Key features of Hsp90's molecular mechanism and its adenosine-5'-triphosphate (ATP)-controlled active cycle remained indefinable**.** Earlier studies have verified that HSP90 may have a role in pathogenesis and virulence in C*. neoformans* (Nooney *et al.,* 2005). On the other hand, production of HSP 90 was noticed to vary among the cells grown in same growth circumstances. Therefore, production of HSP90 in cell is likely to be provoked by other factors, such as post-translation regulation and gene expression. In addition, incidence of Hsp90 in capsule materials helped in utilizing this protein as target for new antifungal therapy,

Table 1 : Results of ProtFun2.2predictions

ProtFun2.2predictions	Prob	Odds
Functional category transport binding	0.773	1.885
Enzyme/nonenzyme Enzyme	0.368	1.285
Gene Ontology categoryVoltage -gated_ion_channel	0.279	12.68

such as human recombinant antibody against HSP90 and this supported the idea of using Hsp90 for serodiagnosis of cryptococcsosis (Cowen, 2005).

Target P 1.1 server anticipated the results of the query sequence using non-plant networks - predicting the location secretory pathway (SP), mitochondrion (mTP) and other location. Highest p-value score revealed in the any other location p-value (0.870) comparatively among all location. Reliability class (RC) was found to be with in the value of 2 which showed strongest expectation of sub cellular location (Table 2). All of the important domain matches acquired from the search sequence were found by means of Pfam-A under default settings. Pfam-Hsp90 chaperones were exclusive in their ability to control a specific subset of cellular signalling proteins that have been implicated in disease courses, including steroid hormone receptors, intracellular protein kinases and growth factor receptors and functions as unfolded protein binding and ATP binding protein. The ATPase-binding region of Hsp90 is currently under intense study, because it is the principal binding site of drug targeting this protein (Chiosis *et al.,* 2006).

Pfam quest results of Hsp90 query sequence results revealed two domains those considered by the site to be significant. First domain showed characterizes to be HATpase. Pfam found this domain to begin at position 22 and end at position

Table 2 : Results of Protein Location (Target p) Analysis

Name	Len	mTP	SP	other	Loc	RC
tr_Q71QT8	691	0.056	0.143	0.870	——	2

142, with an e-value of 3.7e-12. Second domain characterized Hsp90 domain. Pfam found this Hspp90 domain to start at position 1 and end at position 515, with an e-value of 9.3e-251. The results of protein localisation were anticipated using Psort 6.4 version, which showed p-value for two localisation sites; nucleus p-value (0.6) and mitochondrial matrix p-value (0.1) of the query protein sequence (Table 3). The overall performance and certainity values for nucleus, endoplasmic reticulumn, mitochondrial matrix and lysosomes were predicted to be 0.600, 0.000,0.100 and 0.100 respectively using PSORT. PSORT-B represents a powerful tool for prediction of protein subcellular localization for Gram-negative bacteria. High precision allows for confident predictions, and prevents propagation of erroneous predictions.

Various softwares are available for prediction of domains which have been developed using different approaches such as SVM, HMM, Neural Network etc. Thus, for same input they give different result and also differ in accuracy. This variation in result and accuracy leads to dilemma of choosing software for prediction of domains. SMART, simple modular architecture

Table 3 : Results of Psort Analysis

Organelles	Certainty value
Nucleus	0.600
Endoplasmicreticulum (membrane)	0.000
Mitochondrial matrix space	0.100
Lysosome (Lumen)	0.100

research tool is a web based tool that allows domain identification and annotation. The tool compares every sequence with its databases of domain sequences and multiple alignments, as well as, identifies compositionally biased regions such as signal peptide, transmembrane and coiled coil segments. SMART tools are used to predict the domains of query protein sequence**.** Two domains have been found; HATPase c which start at position 23 and end at position 178 with E-value (7.01e-9) and Pfam-Hsp90 which start at position 180 and end at position 690 with E-value (5.9e-248). From the databased used, a sequence-based on method that categorizes and integrates relevant features can be used to assign proteins of unidentified function to functional classes, and enzyme classifications for enzymes. Results showed that strategies for elucidation of protein function may benefit from a number of serviceable attributes that are more straight, related to linear sequence of amino acids, and hence easier to guess, than protein structure. These features include, features connected with protein sorting and post-translational modifications, but also much easier aspects such as isoelectric point, length and composition of polypeptide chain. HATPase c is found in several ATP-binding proteins for example: histidine kinase, DNA gyrase B, topoisomerases (Bellon *et al.,* 2004), heat shock protein (Immormino *et al*., 2004; Roe *et al.,* 2004), phytochrome-like ATPases and DNA mismatch repair proteins and functions as ATP-binding protein. Heat shock protein 90 (Hsp90) accounts for 1-2% of total proteins in normal cells and functions as a molecular chaperone that folds, assembles and stabilizes client proteins. Hsp90 is over expressed (3- to 6-fold increase) in stressed cells like cancer cells, and regulates over 200 client and co-chaperone proteins. Hsp90 client proteins are intricate in plethora of cellular signalling events comprising numerous growth and apoptotic pathways. Subsequently, pathway-specific inhibitors can cause problem in shutting down multiple pathways. Drug-resistant cancers at once is a promising approach when developing new therapeutics (Ardi *et al.,* 2011)**.**

Based on the template structure it is evident that theoretical structure created is structurally similar to the template structure, which is vastly sufficient for the advancement of specific ligand for HSP protein 90.The present study would encourage in studying protein-protein interactions in future and could be further used for drug target studies.

Acknowledgement

Author is thankful to the Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King

Saud University, Riyadh, Saudi Arabia for supporting the study.

References

Alistair, J.P., B. Michelle, D. Leach and S. Nicholls: The relevance of heat shock regulation in fungal pathogens of humans. *Virulence*, **1**, 330–332 (2010).

Ardi, V.C., L.D. Alexander, V.A. Johnson and S.R. McAlpine: Macrocycles that inhibit the binding between heat shock protein 90 and TPR-containing proteins, ACS *Chem. Biol.*, **16**, 1357-66 (2011).

Bellon, S., J.D. Parsons, Y. Wei, K. Hayakawa, L.L. Swenson, P.S. Charifson, J.A. Lippke, R. Aldape and C.H. Gross: Crystal structures of *Escherichia coli* Topoisomerase IV ParE Subunit (24 and 43 Kilodaltons): a single residue dictates differences in novobiocin potency against Topoisomerase IV and DNA Gyrase". *Antimicrob. Agents Chemother.*, **48**, 1856–64 (2004).

Cowen, L.E.: The evolution of fungal drug resistance: modulating the trajectory from genotype to phenotype. *Nat. Rev. Microbiol.*, **6**, 187–198 (2008).

Cowen, L.E., S.D. Singh, J.R. Kohler, C. Collins, A.K. Zaas, A.Wiley, Schell, H. Aziz, E. Mylonakis, J. R. Perfect, L. Whitesell and S. Lindquist: Harnessing Hsp90 function as a powerful, broadly effective therapeutic strategy for fungal infectious disease. *Proc Natl. Acad. Sci. U SA.,* **106**, 2818–2823 (2009).

Chiosis, G., E.C. Lopes and D. Solit: Heat shock protein-90 inhibitors: A chronicle from geldanamycin to today's agents. *Curr. Opin. Investig. Drugs*, **7**, 534–41 (2006).

Cowen, L.E. and S. Lindquist: Hsp90 potentiates the rapid evolution of new traits: drug resistance in diverse fungi. *Science*, **309**,2185-2189 (2005).

Chen, B., D. Zhong and A. Monteiro: Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. *BMC Genomics.*, **7**, 156 (2006).

Edward, T. Burt., D. Rhona, H. Deana, T. Yuri, E. Michael and L. Bryan: Isolation and Partial Characterization of Hsp90 from Candida albican. *Annals of Clinical and Laboratory Science*, **33**, 1 (2003).

Finn R.D., J. Tate, J. Mistry, P.C. Coggill, S.J. Sammut, H.R. Hotz, G. Ceric, K. Forslund, S.R. Eddy, E.L. Sonnhammer and A. Bateman: The Pfam protein families database. *Nucleic Acids Res.,* **36**, 281-8 (2008)

Imai, J., M. Maruya, H. Yashiroda, I. Yahara and K. Tanaka: The molecular chaperone Hsp90 plays a role in the assembly and maintenance of the 26S proteasome. *EMBO J.*, **22**, 3557–67 (2003).

Immormino, R.M., D.E. Dollins, P.L. Shaffer, K.L. Soldano, M.A. Walker, Gewirth: Ligand-induced conformational shift in the N-terminal domain of GRP94, an Hsp90 chaperone. *J. Biol. Chem.*, **279**, 46162–71 (2004).

Mayer, M.P., S. Rüdiger and B Bukau: Molecular basis for interactions of the DnaK chaperone with substrates. *Biol Chem.,* **381**, 877-885 (2000).

Punta, M., C. Penny, R. Coggill, Y. Eberhardt, M. Jaina, T. John, B. Chris, P. Ningze, F. Kristoffer, C. Goran, C. Jody, H. Andreas, H. Liisa, L. Erik, L. Sonnhammer, S.R. Eddy, B.Alex and R.D. Finn: The Pfam protein families database. *Nucleic Acids Res.*, **40**, 290–301 (2012)

Nooney, l., R.C. Matthews and J. Burnie: Evaluation of Mycograb, amphotericin B, caspofungin and fluconazole in combination against *Cryptococcus neoformans* by checkerboard and time-kill methodologies. *Diagn. Microbiol. Infect Dis.,* **51**,19-29 (2005).

Park, E.M., Y.O. Kim , B.H. Nam, H.J. Kong, W.J. Kim, S.J. Lee, Y.J. Jee, I.S. Kong and T.J. Choi: Cloning and expression analysis of a small HSP26 gene of Pacific abalone (*Haliotis discus* hannai). *J. Environ. Biol.,* **29**,577-80 (2008).

Place, S.P. and G.E. Hofmann: Temperature interactions of the molecular chaperone Hsc70 from the eurythermal marine goby Gillichthys mirabilis. *J. Experim. Biol.*, **204**, 2675–2682 (2001).

Pfaller, M.A. and D.J. Diekema: Epidemiology of invasive mycoses in North America. *Crit. Rev. Microbiol.*, **36**, 1–53 (2010).

Pfaller, M.A. and D.J. Diekema: Epidemiology of invasive candidiasis: A persistent public health problem. *Clin. Microbiol. Rev.*, **20**, 133–163 (2007) .

Rajaiah, R. and K.D. Moudgil: "Heat-shock proteins can promote as well as regulate autoimmunity," *Autoimmunity Reviews*, **8**, 388–393 (2009).

Roe, S.M., M.M. Ali, P. Meyer, C.K. Vaughan, B. Panaretou, P.W. Piper, C. Prodromou and L.H. Pearl: "The Mechanism of Hsp90 regulation by the protein kinase-specific cochaperone p50 (cdc37). *Cell.*, **116**, 87–98 (2004).

Shapiro, R.S., P. Uppuluri, A.K. Zaas, C. Collins, H. Senn, J.R. Perfect, J. Heitman and L.E. Cowen: Hsp90 orchestrates temperature-dependent *Candida albicans* morphogenesis via Ras1-PKA signalling. *Curr. Biol.*, **19**, 621–629 (2009).

Stephanie, D., M. Magali, S. Rebecca, Gary. Shapiro, D. Bader Leah and E. Cowen: Mapping the Hsp90 genetic interaction network in *Candida albicans* reveals environmental contingency and rewired circuitry. *PLoS Genetics*, **8** , 3 (2012).

Zaoutis, T.E., J. Argon, J. Chu, J.A. Berlin and T.J. Walsh: The epidemiology and attributable outcomes of candidemia in adults and children hospitalized in the United States: A propensity analysis. *Clin. Inf. Dis.*, **41**, 1232–1239 (2005).