International Journal of Sciences

Research Article Volume 2 - October 2013 (10)

Homology Modeling and Molecular Dynamic Study of Candida Albicans Heat Shock Protein 21, a Potential Target Protein for Antifungal Drug Development

Abdul Wadood[1], Nasir Ahmed[1], Muhammad Riaz[1], Sulaiman Shams[1], Javed Anwer[1], Ayaz Ahmed[2]

[1]Computational Medicinal Chemistry Laboratory, Department of Biochemistry, Abdul Wali Khan University Mardan, Pakistan
[2]Depatment of Biotechnology, Abdul Wali Khan University Mardan, Pakistan

Abstract

***Background*:** Species of *Candida* are ranked as the third most frequently isolated pathogens from blood. Although, *Candida albicans* (*C. albicans*) are one of the main etiologic agent for candidiasis, the small heat shock protein 21 (sHsp21) in *C. albicans* showed its pivotal role for environmental stress adaptation and fungal virulence. In *C. albicans* Hsp21 is a necessary factor for thermal and oxidative stress tolerance.

***Results*:** In the present work a homology model of Hsp21 from *C. albicans* was developed and evaluated using validated methods. Ramachandran plot for the model demonstrated that 98.02 percent of residues are in most favorable region, indicating that the model is reliable. The computed energy value, instability index and root mean square deviation (RMSD) fluctuation of back bone alpha carbon of the model, confirming the stability of the model. Molecular dynamics simulations in explicit solvent environments were carried out for the entire protein by using Molecular Operating Environment (MOE) with AMBER99 force field with the aim to characterize the dynamics of the protein. The results showed an open motion of the protein in solvent.

***Conclusions*:** Until the determination of three-dimensional structure of Hsp21 experimentally, the predicted model will serve as a supportive reference for exploring the interactions between Hsp21 and its antagonists. This research might help to understand the mechanism of action of Hsp21 protein and might facilitate the design of new and potent chemo-types to combat infection caused by *Candida albicans*.

Keywords: *Candida albicans*, Heat shock protein, Homology modeling, MD simulation.

Introduction:

Candida albicans is a fungal pathogen of human that continues to be of considerable medical significance. *C. albicans* has the ability to infect skin as well as mucous membranes, and it also involves in life-threatening systemic infections, especially in the immune-compromised host [Calderone et al, 2002; Odd et al, 1988]. The rise in the immune-compromised population has brought about an increase in the frequency of candidiasis caused by numerous species of *Candida* [Jabra-Rizk et al, 2005; Walsh et al, 1999]. In the past decates the infection rates of *Candida* species has increased by over a factor of 20 and *Candida* species are currently considered as the third most commonly isolated pathogens from human blood. The mortality rate caused by *Candida* related infections is about 35–50% and only in the US the treatment costs exceed US$10 billion annually [Colombo et al, 2003; Mukherjee et al, 2005]. *C. albicans* is dimorphic specie have the ability to change their morphology from the yeast form to the hyphal form. During infection the presence of either of these forms demonstrates that the metabolic state of these species and ability of regulating their morphology, play a pivotal role development of disorder [Staib et al, 1999; Oh et al, 2001; Saville et al, 2003]. Filamentous hyphae are developed at the time of initial stages of tissue invasion that shows 'thigmotropism' or contact guidance. Cultured epithelial cells show a phenomenon, 'cellular internalization', where phagocytosis has been shown to be induced by yeast. This phenomenon, together with invasion of hyphae and adhesion mediated by, may alter the development of invasive *Candidia* infection [Hornby et al, 2001]. Heat shock proteins (Hsps) are found virtually in all living things, together with humans and fungi. They fulfill a

Abdul Wadood (Correspondence)
awadood@awkum.edu.pk
+92 937 542195

plethora of cellular functions, including folding, unfolding or refolding of other proteins (clients), client proteins translocation across membranes, clients activation, and prevention of uncontrolled aggregation of proteins [Whitesell et al, 2005]. Heat shock proteins are constitutively found within cells, however during stress their expression enhances dramatically; indeed, the concentrations of heat shock protein can reach over 20% of total protein in the cell [Lindquist et al 1992]. Thermal stress application to the host *Drosophila melanogaster* model led to the finding of the heat shock response [Ritossa et al, 1996]. Later on different studies showed that expression of heat shock is strongly induced in response to heat and other stresses. Heat shock proteins are classified into five different classes (Hsp100, Hsp90, Hsp70, Hsp60 and sHsps) depending on the mass of molecule [Richter et al, 2010].

Small heat shock proteins (sHsps) has historically received little attention as compare to heat sock protein having higher molecular mass e.g., Hsp90. Hsp12 and Hsp21 are the only small heat shock proteins that are recently investigated in *C. albicans*. Although Hsp12 is strongly up-regulated in response to different types of environmental stresses, both on transcriptional and protein levels, Hsp12 was shown to be dispensable for resistance of stress, morphogenesis and virulence in a model of *Drosophila* infection [Buchner et al, 1996]. Hsp21 is also strongly induced upon a wide range of environmental stresses [Lorenz et al, 2004; Thewes et al, 2007; Ramsdale et al, 2008; Nicholls et al, 2009; Enjalbert et al, 2009]. It is has been shown that Hsp21 is essential for thermal and oxidative stress tolerance in *C. albicans* [Fradin et al, 2005]. Furthermore, Hsp21 was shown to play an important role in normal filament formation, to regulate intracellular levels of the stress-protective molecule trehalose, and in mitogen-activated Protein (MAP) kinase Cek1 activation. The mutant species Hsp21 D/D had impaired capacity of *in vitro* damaging endothelial and epithelial cells, had increased sensitivity to neutrophils, and was strongly attenuated in virulence in two different *in vivo* models infection: an embryonated hen egg infection model and a mouse infection model of hematogenously disseminated candidiasis [Mayer et al, 2012].

The above mentioned lines of evidences suggest that Hsp21 plays important roles in causing various *Candida albicans* infections in humans, therefore; targeting fungal-specific Hsps might represent an additional / alternative strategy to combat candidiais. It has been proposed that the Hsp21 of *C. albicans* is a promising candidate for such an approach for a number of reasons; however, unavailability of the 3D structure of this protein is a hindrance in elucidating its mechanisms of action and developing potential drugs for candidiasis.

The availability of web based techniques and different servers provide a tremendous opportunity to characterize *Candida albicans* Hsp21 physiochemical properties in addition to its secondary and three dimensional structural properties. The objective of the current work was primarily to report 3D structural analysis and characterization of Hsp21 that might helpful in understanding its mechanism of action and designing new drugs to combat infections caused by *Candida albicans*.

Material and methods

Protein sequence retrieval

The primary sequence of *Candida albicans* Hsp21 was obtained in FASTA format from *Candida* genome database (www.candidagenome.org). The 3D structure of this Hsp21 was not available in protein data bank (http;//www.rcsb.org/pdb), the aim was to develop 3D model of Hsp21.

Secondary structure prediction

The prediction of secondary structures was carried out through on line servers in which the sequence of Hsp21 was submitted and analysis of the secondary structures elements was done. The task was accomplished with online tools (DPM) [Deleage et al, 1987], discrimination of protein secondary structure class (DSC) [King et al, 1996], PHD [Rost et al, 1993], PREDATOR [Frishman et al, 1996], SIMPA96 [Levin et al, 1986], SOPM [Geourjon et al, 1994], self-optimized prediction method with alignment, HNN and GOR4 [Heringa et al, 2000].

Prediction of intrinsic disorder

Protein disordered regions prediction is a requisite phenomenon to explain the function of protein and also explore the protein folding pathway [Patel et al, 2001]. The online tools DisEMBLE [Linding et al, 2003], Globplot [Linding et al, 2003], regional order neural network (RONN) [Yang et al, 2005] were used to find out the regions of disordered and higher flexibility.

Homology modeling of heat shock protein 21

MOE 2011-12 comparative homology modeling technique was used for the generation of model to know about the structural characteristics of the Hsp21. To construct the model of Hsp21 selection of template, sequence alignment, model generation and finally refinement and evaluation of the model were carried out.

Template selection

Similarity search was carried out employing position specific iteration-basic local alignment search tool (PSI-BLAST) 24 against PDB database with default parameters like E-value threshold 10, word size 3 and Blosum 62 Matrix. Total three iterations of PSI-BLAST were considered as the BLAST search results

converged after three iterations. As a result of that high resolution X-ray crystallography structure of the (The Protease Domain of an Atp- Independent Heat Shock Protease Htra) was selected as template protein showing 43% identity with the target protein. It was used to construct model on MOE 2011-12.

Target template alignment

The MUSCLE SERVER (http://www.ebi.ac.uk/Tools/msa/muscle/) was used for target template alignment in which the CLUSTAL W program was selected to get alignment [Edgar et al, 2004]. The result was obtained in the form of Blosum scoring matrix which was selected with a gap penalty of 10 and a gap extension of 0.05.

Model generation

The comparative modeling program of MOE 2011-12 software was used to construct the structure of Hsp21 which constructed a refined 3D homology model of a protein sequence based on given sequence alignment and selected template. Adjustments were made to various parameters including intermediate in the refining of model was set to medium, final model to medium, using scoring function Generalized Born/Volume Integral (GB/VI). The force field Amber99 with Solvation R-Field was selected and the MOE align application was selected to align the sequence before starting the homology modeling. A total of 10 models were generated, and a stable model was selected for molecular dynamics simulation.

Molecular dynamics simulation

The dynamics simulation was performed by MOE 20011-12 software and it was proceeded with the selection of force field AMBER99 by concerning calculation toward solvation energy with Bornimplisit solvation. The energy of the system was minimized to RMSD Gradient 0.1 and the partial charge option in MOE 2011-12 was used to optimize the total charge of protein. Further all the parameters were adjusted to default values that were ensemble at NVT and algorithm at NPA for creating ensemble trajectory. The acceleration, velocity and position were saved after each 0.5 picoseconds (ps). The addition of water molecules was done by Water Soak option with soak mode BOX and layer width 5 prior to MD simulation and it was followed by energy minimization. The system was heated from 0K to 300K in 20 picoseconds (heat time) followed by production time of 1300 picoseconds. The system was cooled back to 0K in 20 picoseconds.

Evaluation and validation of the refined model

The structure of Hsp21 was built with the MOE 2011-12 Software and its evaluation and validation was carried out on RAMPAGE SERVER and ERRATE PROGRAM [Colovos et al, 1993] to get the steriochemical quality of the model.

Results and discussion

Retrieval of primary sequence

The primary sequence of the alpha Hsp21 of *Candida albicans* was obtained in FASTA format from the *Candida* genome database (www.candidagenome.org) and it was characterized *in silico* using Expasy-proparam tool (www.expasy.ch/tools/protparam.html). The protein has 189 amino acids and the estimated molecular weight of 21494.6 Dalton. The isoelectric point (*pI*) is 5.15. The extinction coefficient of protein is 18450 which indicate that at a specific wavelength how much light is absorbed by the protein [Gill et al, 1989]. The instability index (II) is computed to be 29.69 which classify that the protein is stable [Guruprasad et al, 1990]. The aliphatic index is 54.13s [Kyte et al, 1982]. The value of Grand average of hydropathy (GRAVY) is -0.857 showing that it is hydrophilic in nature [Atsushi et al, 1980].

Secondary structure prediction

Secondary structure of the Hsp21 was determined through various online servers in order to find out the position of amino acids that are present in helix, strand and coil chains. After comparison of results obtained through various secondary structure predictions severs it was found that random coils dominated among the secondary structure elements followed by alpha helix, extended strand and beta turns. The details of secondary structures are given in Table 1.

Intrinsic disorder identification in protein:

The online servers RONN, GLOBPROT, DISEMBLE were used to identify the intrinsic disorders in Hsp21.The results showed the regions (37-55,170-175,183-186, and (189-189) as common. The intrinsic disorders profile of Hsp21 obtained using different servers is illustrated in Table 2 and Figure 1.

Target-template alignment

An optimal sequence alignment is essential to the success of homology modeling. We performed the sequence alignment on clustalW program at MUSCLE SERVER using default parameters. The sequence alignment results of the Hsp21 revealed 43% identity with the template. The Protease Domain of an ATP independent Heat Shock Protease Htra and the results are given Figure 2.

Homology modeling

Homology modeling technique was utilized to construct the three dimensional structure of Hsp21. The MOE 2011-12 software was used for homology modeling which is comprised of the construction of a new model on MOE 2011-12 software on the basis of relative objective functional values, molecular dynamics simulation step and finally the validation

step.

The construction of model on MOE 2011-12 was carried through following procedure; initially identity of residues was conserved by copying initial partial geometry of target sequence from template through MOE program. Insertions and deletions tools were used to treat those residues having no assigned backbone coordinates [Needleman et al, 1970]. The loops were molded in random order. The contact energy function which is based on Boltzmann weighted averaging was chosen to analyze a list of possible candidates [Sippl et al, 1993; Sippl et al, 1995]. Finally for model generation adjustment of the parameters was carried out, like Model Scoring to Generalized Born/Volume Integral (GB/VI) methodology [Labute et al, 2008]. In MOE force field AMBER99 is recommended for protein homology applications [Summa et al, 2007] so, it was selected. Force field AMBER99 was used to minimize energy to 0.1 gradients for the structure formed, which is specifically used for the proteins and nucleic acids and not recommended for small molecules. At the end of the process a new model was formed which was saved in PDB format with a proper name for recognition (Figure 3A). The superimposed model was obtained by superimposition of native model on the template and root mean square deviation (RMSD) of 1.742 Angstrom indicated close homology (Figure 3B).

Molecular dynamic simulation

The molecular dynamics simulation tool of MOE 2011-12 was used to refine the model [Smelcerovic et al, 2013]. Briefly, initially the saturation of the constructed model with partial charges was done then the energy was minimized to 0.1 RMS with the selected force field AMBER 99 [Weiner et al, 1986]. The NVT ensemble in which N stands for number of atoms, V for volume and T for temperature, in MD simulation all these values were kept constant [Cornell et al, 1995]. in order to generate true simulation trajectories the most accurate and sensitive method NPT algorithm was used the Nose-Poincare-Anderson (NPA) method generate theoretically correct NVE, NVT, NPH and NPT ensembles (N, V, T, H, E, P shows Number of atoms, volume, temperature, enthalpy, energy and pressure respectively [Sturgeon et al, 2000].

The MD simulation parameters were adjusted, which involved the fixation of nominal temperature in Kelvin at 0K to start the process of simulation and 20K for heating while 20K for cooling of the system, then the simulation was run at 1300 picoseconds [Bond et al, 1999]. Finally an equilibrium state was established between 1200 to 1300 picoseconds which indicated the stability of protein at human body temperature and system was cooled down for 20 picoseconds to get the stable bond energies. The interpretation of MD simulation results can be done through many ways while in this study we selected the potential energy plot which was drawn between the protein conformations against time. The obtained conformation explained the three dimensional structure of protein which can be changed without fluctuating covalent bonds. The potential energy plot of protein conformation verses time is given in Figure 4. While the RMSD value 1.742 angstrom was obtained as result of superposition of initial and final structure after MD Simulation which showed that the protein possessed greater stability at 300K (Figure 5).

Evaluation and validation of refined model

In order to evaluate the overall quality and accuracy of the model, Ramachandran plot was obtained from rampage server which explained the sterochemical quality of the model .Stereo chemical evaluation of backbone Psi and Phi dihedral angles divulged that 69.5% were falling within the most favored regions, 23.5%of residues were found in additionally allowed regions , and 7.0% residues were falling within disallowed regions of Ramachandran plot respectively (Figure 6). Besides, Ramachandran plot analysis tool in MOE that several residues GLy-40, PRO-26, VAL-109 were placed out of energetically favored regions in the plot. Remaining residues are in the core regions of Ramachandran's plot, which indicated that the final structure is highly reliable for further studies, total, 98% of the residues are in the most favored and allowed region. The Errate value explained the statistics of non-bonded interaction between different atoms and a score of 50 is normally acceptable and for model of Hsp21 its value is 69.613% which indicated that the model is reliable and stable (Figure 7).

Conclusion

The homology modeling methodology was used to generate a model of the Hsp21 from *Candida albicans*. The modeling studies provided an insight into the folding pattern and the arrangement of amino acids. Until the complete structure of Hsp21 is determined through experimental assays, the predicted model will serve as a helpful reference for examining the interactions between Hsp21 and its antagonists. The predicted 3D structure of Hsp21 might help in understanding its mechanism of action and might open new door for the development of new and potent drug candidates to combat infection caused by *C. albicans*.

Conflict of interests:

The authors declare that there is no conflict of interests regarding publication of this article.

References

1. Atsushi I. Thermostability and aliphatic index of globular proteins. J Biochem 1980, 88:1895-1898.
2. Bond SD, Leimkuhler, BJ, Laird, BB. The Nose-Poincare method for constant temperature molecular dynamics. J

Comput Phys 1999, 151:114-134.
3. Buchner J. Supervising the fold: functional principles of molecular chaperones. *FASEB J*, 1996, 10:10-19.
4. Calderone RA, Clancy CJ. Candida and candidiasis. Vol. 3. 2002: ASM Press Washington, DC.
5. Colombo AL, Guimar TS. Epidemiology of hematogenous infections due to *Candida spp*. *Revista da Sociedade Brasileira de Medicina Tropical* 2003, 36:599-607.
6. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sc 1993, 2:1511-1519.
7. Cornell WD, Piotr C, Christopher IB, Ian RG, Kenneth MM, David MF, David CS, Thomas F, James WC, Peter AK. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 1995, 117:5179-5197.
8. Deleage G, Roux B. An algorithm for protein secondary structure prediction based on class prediction. Protein Eng 1987, 1:289-294.
9. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl Acids Res 2004, 32:1792-1797.
10. Enjalbert, B, Moran GP, Vaughan C, Yeomans T, Maccallum DM, Quinn J, Coleman DC, Brown AJ, Sullivan DJ. Genome-wide gene expression profiling and a forward genetic screen show that differential expression of the sodium ion transporter Ena21 contributes to the differential tolerance of *Candida albicans* and *Candida dubliniensis* to osmotic stress. *Mol Microb* 2009, 72:216-228.
11. Fradin C, De GP, MacCallum D, Schaller M, Klis F, Odds FC, Hube B. Granulocytes govern the transcriptional response, morphology and proliferation of *Candida albicans* in human blood. Mol Microb 2005, 56:397-415.
12. Francois LM, Duncan W, Ilse DJ, Pedro M, Silvia S, Iryna MB, Alistair JPB, Bernhard H. Small but crucial: the novel small heat shock protein Hsp21 mediates stress adaptation and virulence in *Candida albicans.* PloS One 2012, 7:38584-38594.
13. Frishman D, Argos P. Argos, Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. *Protein Eng* 1996, 9: 133-142.
14. Geourjon C, Deleage G. SOPM: a self-optimized method for protein secondary structure prediction. *Protein Eng* 1994, 7:157-164.
15. Gill SC, Hippel PHV. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 1989, 182:319-326.
16. Guruprasad K, Reddy BVB, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng* 1990 4:155-161.
17. Heringa J. Computational methods for protein secondary structure prediction using multiple sequence alignments. Curr Protein Peptide Sci 2000, 1:273-301.
18. Hornby JM, Ellen CJ, Amber DL, Joseph JT, Brandon J, Richard S, Patick D, Kenneth WN. Quorum sensing in the dimorphic fungus *Candida albicans* is mediated by farnesol. *Appl Environ Microb* 2001, 67:2982-2992.
19. Jabra-Rizk MA, Johnson JK, Forrest G, Mankes K, Meiller TF, Venezia RA. Prevalence of Candida dubliniensis fungemia at a large teaching hospital. *Clin Infect Dis* 2005, 41:1064-1067.
20. Ki-Bong O, Hiroshi M, Toshimichi N, Hideaki M. Purification and characterization of an autoregulatory substance capable of regulating the morphological transition in *Candida albicans*. *Proceedings of the National Academy of Sciences* 2001, 98:4664-4668.
21. King RD, Sternberg MJE. Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. *Protein Sc* 1996, 5: 2298-2310.
22. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 1982, 157:105-132.
23. Labute P. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. J Comput Chem 2008, 29:1693-1698.
24. Levin JM, Robson B, Garnier J. An algorithm for secondary structure determination in proteins based on sequence similarity. *FEBS Lett* 1986, 205:303-308.
25. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russel RB. Protein disorder prediction: implications for structural proteomics. *Structure* 2003. 11:1453-1459.
26. Linding R, Russel RB, Neduva V, Gibson TJ. GlobPlot: exploring protein sequences for globularity and disorder. Nucl Acids Res 2003, 31:3701-3708.
27. Lindquist S. Heat-shock proteins and stress tolerance in microorganisms. *Curr Opin Gen Dev* 1992, 2:748-755.
28. Lorenz MC, Bender JA, Fink GR. Transcriptional response of *Candida albicans* upon internalization by macrophages. Eukar Cell 2004, 3:1076-1087.
29. Mukherjee PK, Zhou G, Munyon R, Ghannoum MA. Candida biofilm: a well-designed protected environment. *Medical Mycology* 2005, 43:191-208.
30. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 1970, 48:443-453.
31. Nicholls S, Leach MD, Priest CL, Brown AJ. Role of the heat shock transcription factor, Hsf1, in a major fungal pathogen that is obligately associated with warm-blooded animals. *Mol Microb* 2009, 74:844-861.
32. Odds FC. *Candida* and candidosis: a review and bibliography. 1988: Bailliere Tindall.
33. Walsh TJ, Groll AH. Emerging fungal pathogens: evolving challenges to immunocompromised patients for the twenty • first century. *Transplant Infect Dis* 1999, 1:247-261.
34. Patel PH, Loeb LA. Getting a grip on how DNA polymerases function. *Nature Struct Biol* 2001. 8:656-658.
35. Richter K, Haslbeck M, Buchner J. The heat shock response: life on the verge of death. *Mol Cell* 40,:253-266.
36. Ritossa F. Discovery of the heat shock response. Cell Stress Chaper 1996, 1:97-98.
37. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* 1993, 232:584-599.
38. Staib P, Nichterlein MKT, Michel GKS, Hof H, Hacker J, Joachim M. Host-induced, stage-specific virulence gene activation in *Candida albicans* during infection. Mol Microb 1999, 32:533-546.
39. Saville SP, Anna LL, Carol M, Jose LL. Engineered control of cell morphology in vivo reveals distinct roles for yeast and filamentous forms of *Candida albicans* during infection. *Eukaryotic Cell* 2003, 2:1053-1060.
40. Sippl MJ. Knowledge-based potentials for proteins. *Curr Opin Struct Biol* 1995, 5: 229-235.
41. Sippl MJ. Recognition of errors in threeâ€ • dimensional structures of proteins. *Proteins Struct Funct Bioinform* 1993, 17:355-362.
42. Smelcerovic A, Ranqelov M, Smelcerovic Z, Veljkovic A, Cherneva E, Yancheva D, Nikolic GM, Petronijevic Z, Kocic G. Two 6-(propan-2-yl)-4-methyl-morpholine-2, 5-diones as new non-purine xanthine oxidase inhibitors and anti-inflammatory agents. *Food Chem Toxicol* 2013, 55:493-497.
43. Summa CM, Levitt M. Near-native structure refinement using in vacuo energy minimization. *Proceed Nat Acad of Sci* 2007, 104:3177-3182.
44. Thewes S, Kretschmar M, Park H, Schaller M, Filler SG, Hube B. In vivo and ex vivo comparative transcriptional profiling of invasive and non-invasive *Candida albicans* isolates identifies genes associated with tissue invasion. Mol Microb 2007, 63:1606-1628.
45. Weiner SJ, Peter AK, Dzung TN, David AC. An all atom force field for simulations of proteins and nucleic acids. *J Comput Chem* 1986, 7:230-252.
46. Whitesell L, Lindquist SL. HSP90 and the chaperoning of cancer. *Nature Rev Cancer* 2005 5:761-772.
47. Yang ZR, Thomson R, McNeil P, Esnouf RM. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* 2005, 21:3369-3376.

Table 1: Secondary structure prediction calculated by different tools.

Secondary Structure Servers	DPM	HNN	SOPM	PHD	Predator	SIMPA96	GOR4
Alpha helix	28.57.%	17.99%	16.40%	4.23%	3.17%	11.05%	19.58%
3_{10} helix	0.00%	0.00%	0.00%	00.00%	0.00%	0.00%	0.00%
Pi helix	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Beta bridge	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Extended strand	19.58%	15.34%	22.22%	28.57%	17.99%	18.95%	14.81%
Beta turn	16.40%	0.00%	11.64%	0.00%	0.00%	0.00%	0.00%
Bend region	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Random coil	35.45%	66.67%	49.74%	67.20%	78.84%	69.47%	65.61%
Ambiguous states	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Other states	0.00%	0.00%	0.00%	0.00%	0.00%	0.53%	0.00%

Table 2: Intrinsic disorders predicted by different servers.

Server	Disordered	Disordered by REM465	Disordered by Loop/Coil definition	Disordered by Hot-Loop definition	Disordered by Russell/Linding definition
GLOBPLOT					1-54,75-80, 132-138.
DISEMBL		37-55	1-92,100-115, 127-189	1-13,23-42, 124-149,167-189.	
RONN	1-1,25-57,102-145, 170-175,183-186, 189-189.				

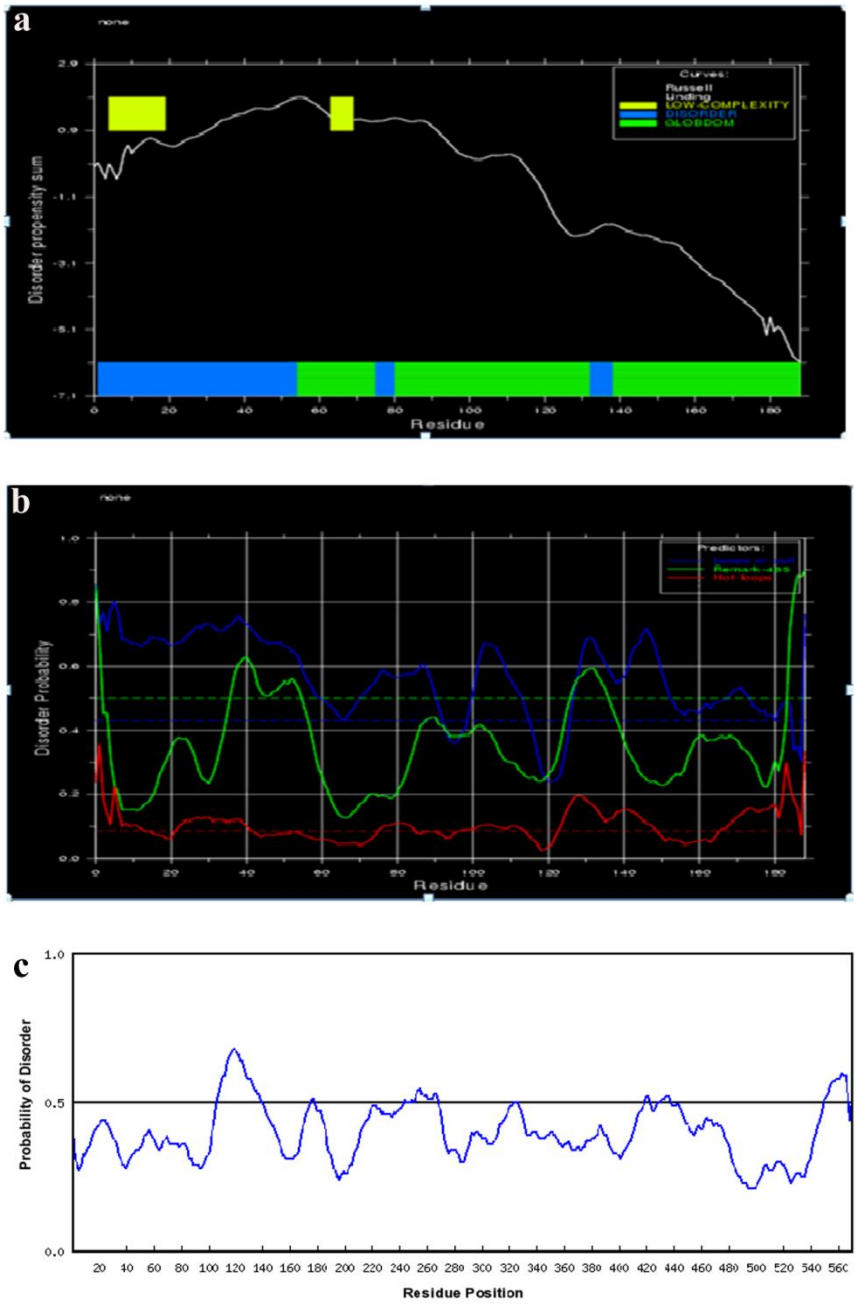

Figure 1: Protein intrinsic disorders predicted by different server **(A)** DisEMBL **(B)** Globplot **(C)** RONN.

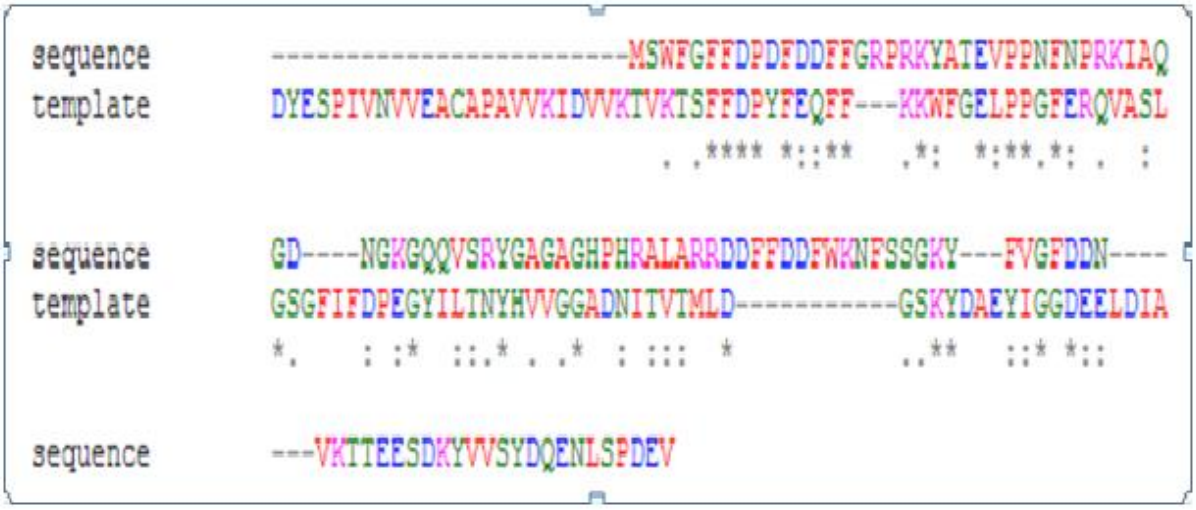

Figure 2: Sequence alignment between heat shock protein 21 and template 1L1J-A (The Protease Domain Of An ATP- independent Heat Shock Protease Htra.

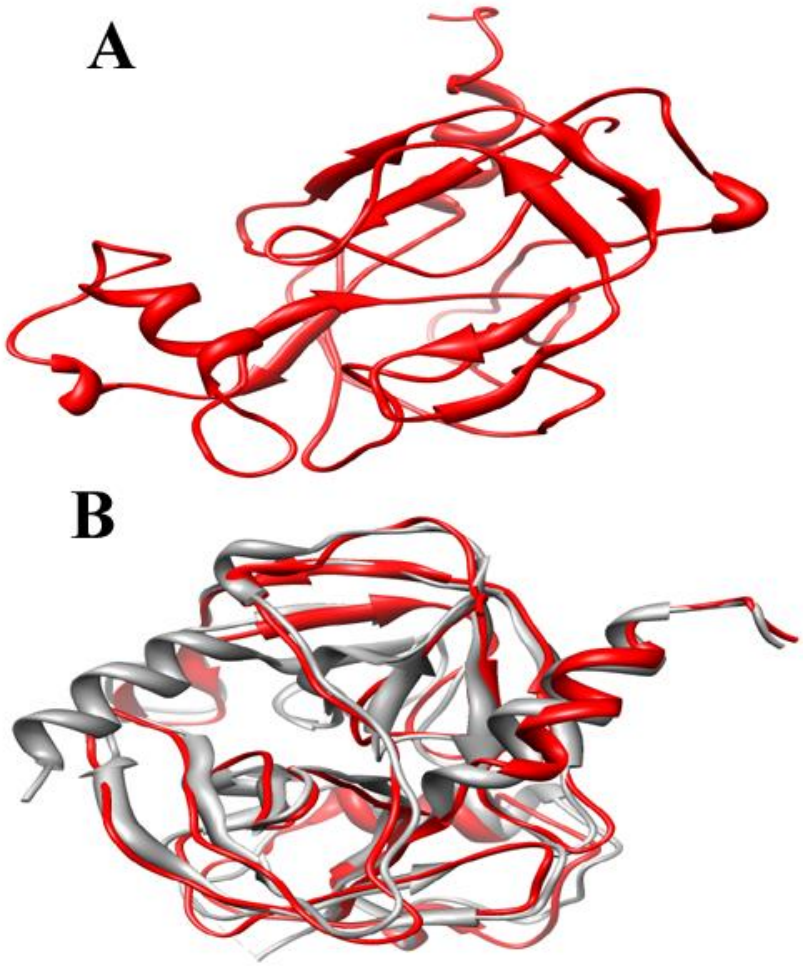

Figure 3: (A) Homology model for heat shock protein 21 (B) Superposition of target heat shock protein and template The Protease Domain of an Atp- Independent Heat Shock Protease Htra.

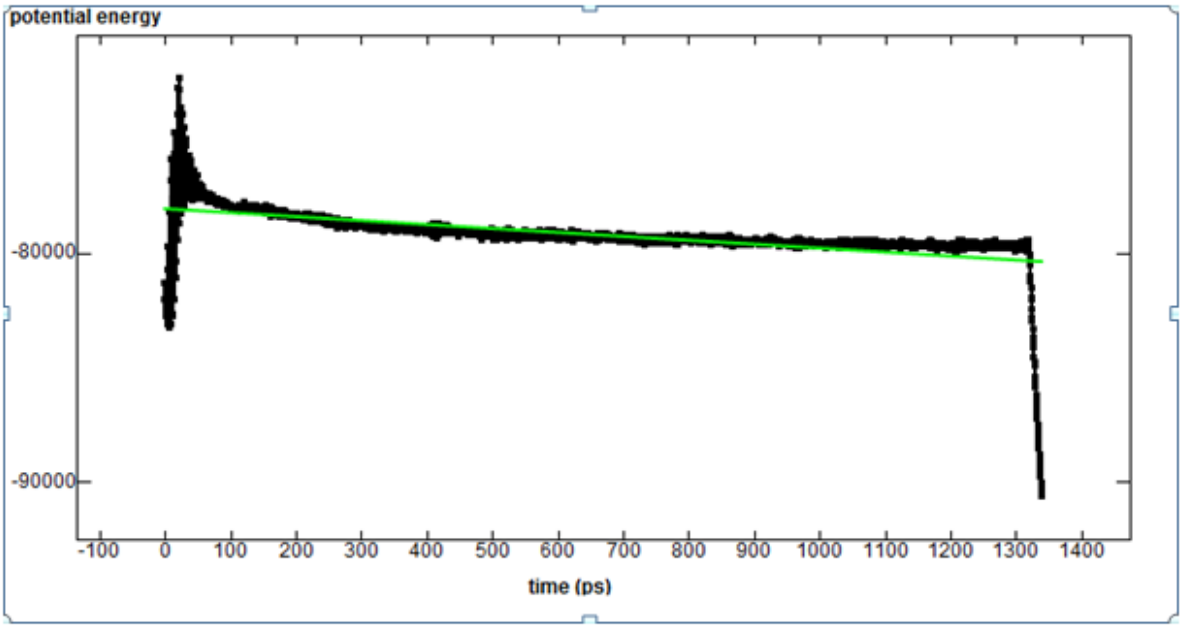

Figure 4: Potential energy plot during MD simulation at 300K.

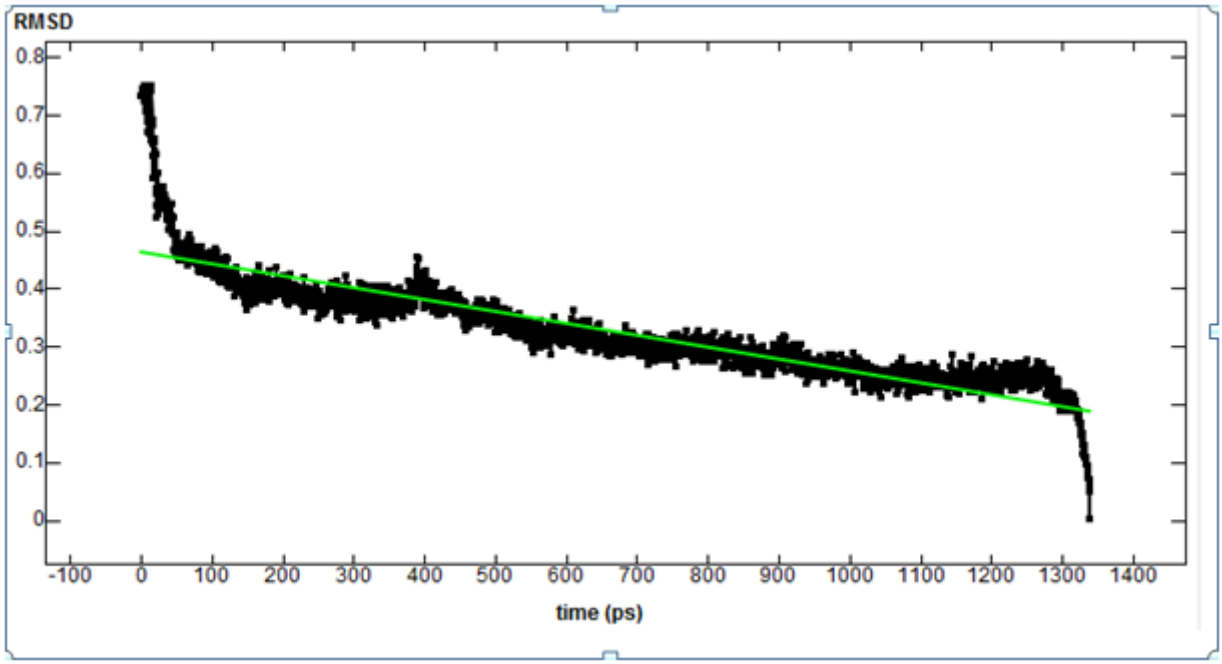

Figure 5: RMSD graph of heat shock protein 21.

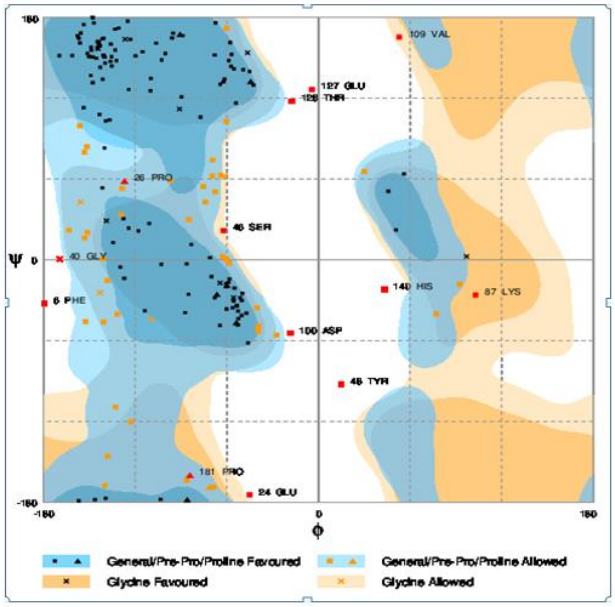

Figure 6: Ramachandran plot.

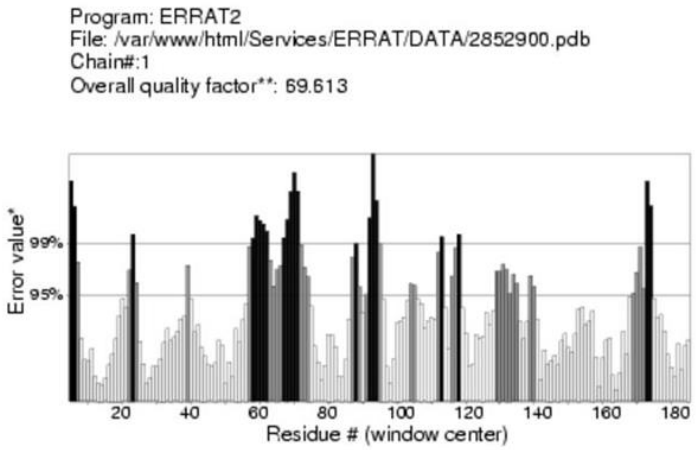

Figure 7: Overall quality factor of heat shock protein 21.