

Molecular Dynamics and Conformational Flexibility in Heat Shock Protein 60.2 of *Mycobacterium tuberculosis*

Pavan Gollapalli*, Manjunatha Hanumanthappa

Department of Biotechnology and Bioinformatics, Kuvempu University,
Shankaraghatta, Shivamogga, Karnataka, India.

Article Info

Article history:

Received Apr 25th, 2015

Revised May 20th, 2015

Accepted May 24th, 2015

Keyword:

Heat Shock Proteins
Molecular Dynamic Simulations
Principal Component Analysis
Clustering
Protein folding
Normal mode analysis

ABSTRACT

HSP 60.2 plays important role in pathogenesis of *Mycobacterium tuberculosis* causing tuberculosis. This chaperonin comprises of three domains namely- apical, intermediate and equatorial which assists in proper protein folding thus preventing aggregation of unfolded polypeptides. To evaluate the structural changes during protein folding, conformations of HSP 60.2 were monitored during 10 ns time scale. Molecular dynamic simulations are used to study the large amount of molecular and biomolecular conformations with the use of high end computational assistance. The Principal component analysis and clustering techniques are used for revealing major conformational changes that occur in the MD simulation. Normal mode analysis was also performed to study the conformation and direction of motion of a protein under study for a large time scale simulation. These studies suggest that functional behavior of protein that depends on the structure. Chaperonin 60.2 is not only plays a role as protein folding machinery, but also an immunologically important biomolecule. Hence it is provided and drawn a clear path between role of chaperon in protein folding and their role in the infection showing the immunological importance of Heat Shock Protein 60.2.

Copyright © 2015 *International Journal for Computational Biology*,
<http://www.ijcb.in>, All rights reserved.

Corresponding Author:

Gollapalli Pavan
Department of Biotechnology and
Bioinformatics, Jnana Sayhadri,
Kuvempu University,
Shivamogga, Karnataka, India.
Email : gollapalli.pavan@gmail.com



How to Cite:

Pavan Gollapalli *et. al.* Molecular dynamics and Conformational flexibility in Heat Shock Protein 60.2 of *Mycobacterium tuberculosis*. IJCB. 2015; Volume 4 (Issue 2): Page 31-45.

1. INTRODUCTION

The diseases caused by *Mycobacterium* are important sources of morbidity and mortality in the world today [1]. *Mycobacterium tuberculosis* is the causative agent of the tuberculosis that kills more than 2 billion people i.e., one third of world's population. Chaperonin 60 (Cpn60), also commonly referred to as heat shock protein 60 (Hsp60), is one of the major molecular chaperones that are present ubiquitously in all forms of life. These Molecular chaperones are known to assist the folding, assembly, and transport of several cellular proteins [2]. Invasion of a host is an apparent form of a stress, and induction of Cpn60s has also been observed in pathogenic organisms. The over expressed pathogen derived Cpn60s act as major antigens that result in strong immune responses from the host [3]. Its immunodominant epitopes have recently been identified. It has also been explored for possible inclusion in are combinant vaccine [4] Interestingly, the mycobacterial chaperonins have previously been shown to be secreted into the extracellular environment, although their role as molecular

chaperones is limited to the cytosol [5]. The existence of chaperonins in the extracellular environment thus suggests a possible alternate functional role [6].

However, recent literature suggests that the heat shock proteins or chaperonin (HSP 10 and HSP 60.1, HSP 60.2) play key role in pathogenesis of *Mycobacterium tuberculosis*. These chaperonins assist in the correct folding of most protein under both normal and stress conditions. Before their role as protein folding molecules was discovered, the chaperonin of certain pathogenic bacteria viz., *M. tuberculosis* were identified as major immunogens [7].

Totally structure comprises 447 amino acid residues in chain A and 445 residues in chain B. The overall architecture comprises three distinct domains-equatorial, intermediate and apical. The HSP 60.2 structure reveals large hydrophobic regions are exposed on the surface, with an elegant conformational change involving reorientation of the N-terminal helix [Residues 60 to 74]. In both A and B chain, the asymmetric unit shields hydrophobic residues of the equatorial domain. In addition, this conformational change appears to promote the packing of the α -helices spanning residues 87 to 107 and 60 to 97. Another hydrophobic patch occurs in the apical domain and has been shown to bind substrate proteins. These two hydrophobic patches in each of HSP 60.2 monomer have role in binding unfolded polypeptides [8].

To better understand the molecular mechanism of protein folding assisted by HSP 60.2 of *Mycobacterium tuberculosis*, we have studied the flexibility of backbone conformations by the methods of molecular dynamic simulations (MD). In the analysis of protein dynamics, an important goal is the description of slow large-amplitude motions in large proteins. These motions typically describe rearrangements of domains which are essential for the function of the protein. In the present work, the principal component analysis and clustering techniques are used for MD simulation trajectory data to investigate the major conformational changes in the HSP 60.2 during protein folding. In order to examine the inter- and intra-domain motion of *Mycobacterium tuberculosis* HSP 60.2, normal mode analysis was performed [9].

2. RESEARCH METHOD

2.1. Molecular Dynamic Simulation

The crystal structure of *Mycobacterium tuberculosis* HSP 60.2(PDB ID: 1sjp) was used to model unliganded molecular dynamic studies. The protein was solved in a cubic box of SPC pre-equilibrated water molecules with a minimum distance of 5Å between the solvent and each face of the box. The final system contained 79809 atoms within a box. The solute and solvent was separately coupled to temperature reservoir of 298.15 K with a coupling time of 0.4 ps. All the minimization and molecular dynamic simulation step were performed using GROMCS force field 43a1 and GROMACS program suite 4.5.5 at a stable volume and temperature of 300 K [10, 11].

The equation of motion was integrated by using a Leapfrog algorithm with a time step of 2fs. Covalent bond length between hydrogen and heavy atoms were constrained using SHAKE [12] with a relative geometric tolerance of 0.0001. The equilibration protocol was considered for 700 frames by steepest-descent minimization and not used in the main analysis for the reason described in results. Only the subsequent 10,000 frames (10ns) were used in the main analysis [13].

2.1.1. Principal component analysis (PCA)

Principal component analysis is an unsupervised statistical technique for finding patterns in high-dimensional data. It is often used a tool in exploratory data analysis to reveal the internal data structure in a way that best explains its variance [14-16].

A typical MD trajectory consists of the information of time-evolution of the coordinates of the entire constituent forming the system being studied. This MD time steps are an order of 1 fs while the simulation time ranges from few to ten nanosecond. The resultant trajectory consist a huge amount of data which can be splitted. When performed on a set of experimental structures, the principal components describe concerted atomic

displacement and can highlight major conformational changes between the structure [17, 18]. Because these motions are often essential for protein function [19]. For an N-atom system, the input dataset for PCA can be constructed as trajectory matrix in which each column contain a Cartesian coordinate for a given atom of each output time step. Apart from long-time MD simulations of generate sufficient trajectory data, the diagonalization of the 3N covariance matrix poses the most computationally exhaustive step during PCA. MD-PCA was carried out with a version of GROMACS 4.5.5 using the same thermostat as the reference simulation but with a time step dependent. This results in a simulation without overall rotation and translation [20].

PC analysis, performed on Cartesian coordinates or dihedral analysis (dPCA) has proved to be a valuable tool for studying conformational change. Mathematically, the Principal components are obtained by a diagonalization of the data covariance matrix C:

$$C=V\Lambda V^T$$

This results in the diagonal matrix ‘ Λ ’ containing the eigenvalues as diagonal entries and the matrix ‘V’ containing the corresponding eigenvectors [21].

2.1.2. Covariance matrix

The first step in the PCA is the construction of the covariance matrix, which capture the degree of colinearity of atomic motion for each pair of the atoms. This covariance matrix is subsequently diagonalized yielding a matrix of eigenvectors and a diagonal matrix of eigenvalues. Each of the eigenvectors describes collective motions of particles; when the values of the vectors indicated how much the corresponding atom participate in the motion. The associated eigenvalues gives equal the sum of the fluctuation described by the collective motions per atom and thus is a measure for the total modeling associated with an eigenvectors [22].

2.1.3. Clustering

Clustering analysis is another unsupervised technique for finding patterns within data. Clustering algorithm group similar objects into subgroups (i.e., clusters) by minimizing intra-cluster and maximizing inter-cluster difference. Therefore, most clustering algorithms require a measure of similarity or distance, of object. Clustering algorithm can be divided into partitional and hierarchical clustering algorithms [23]. Partitional clustering divides the objects from non-overlapping clusters; hierarchical clustering allows nested clustering and results in a hierarchical tree. The average value within a cluster is called the centroid: for clustering coordinate data, the centroid represents the conformation that best describe the conformations with a cluster [24].

2.2. Normal Mode Analysis

In order to examine the nature of collective motion of atoms and conformational changes of *Mycobacterium tuberculosis* HSP 60.2 [8], normal mode analysis was carried by using ELNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement [25] to compute the 100 lowest frequency modes of *Mycobacterium tuberculosis* HSP 60.2 [26]. The normal mode theory is based on the harmonic approximation of the potential energy function, around a minimum energy conformation. This approximation allows the analytic solution of the equation of the motion by diagonalizing the Hessian matrix, which yield eigenvalues (normal modes) and are the squares of the associated frequencies [27].

The classical semi-empirical potential energy function used in the all-atom force field is replaced by a simple parameter potential (a Hookean potential):

$$E_p = \sum_{d_{ij}^o < R_c} C(d_{ij} - d_{ij}^o)^2$$

Where d_{ij} is the distance between two atoms i and j , d_{ij}^0 is the distance between these atoms in the three dimensional structure, 'c' is a constant assumed to be the same for all interacting pairs which refers to the spring constant of the potential and R_c is an arbitrary cutoff parameter, beyond which interactions are not taken into account [28].

2.2.1. Distance fluctuation

The map highlights residual pair 'i' and 'j' with a strongest variation in the distance between their C-alpha atoms in a given model 'k'. In these maps, rigid and flexible blocks of amino acid residues can be identified along with their relative movement, which are represented by blue and red colored respectively [29].

2.2.2. Overlaps

This measures the degree of similarity between the direction of a chosen conformation change and the direction values for the squares of the overlap, starting from the normal modes form a basis, the cumulative sum reaches a value of one when it is computed overall modes [30, 31].

2.2.3. Collectivity:

Degree of collectivity measures the fraction of residues that are significantly affected by a given mode. For maximal collective movements the degree of collectivity tends to be the value of one whereas for localized motions, where the normal mode movement only involves from atoms, the degree of collectivity approaches zero. The low frequency normal modes are expected to have collective characters, especially those related to functional conformational changes of protein [32, 33].

3. RESULTS AND ANALYSIS

3.1. MD simulations

The potential energy of the 10ns simulation was shown it's decreased and fluctuation to -555226 J at 700 ps and hence we considered the trajectory after this time for the analysis of simulation.

3.1.1. Root Mean Square Deviation (RMSD)

The root mean square (RMSD) from the crystal structure through 10 ns trajectory is shown in Figure 1. The RMSD for the α -carbon tom kept stable around 9000 to 10000 ps which observed a deep decrease from 1 ns trajectory. It is also observed a short stability in the RMSD at 2000ps. The overall average RMSD value was shown as 0.31 nm. The potential energy, total energy, pressure and temperature were maintained all along the 10ns simulation were as indicated in table 1.

Table 1: Energy averages and their RMSD values for 10ns MD simulation

ENERGY	AVERAGE	RMSD
Potential energy	-555226	545.011
Kinetic energy	100007	377.977
Total energy	-455219	425.434
Pressure	1.00066	1.12279
Density	1010.2	98.4869
Temperature	299.783	1.50635

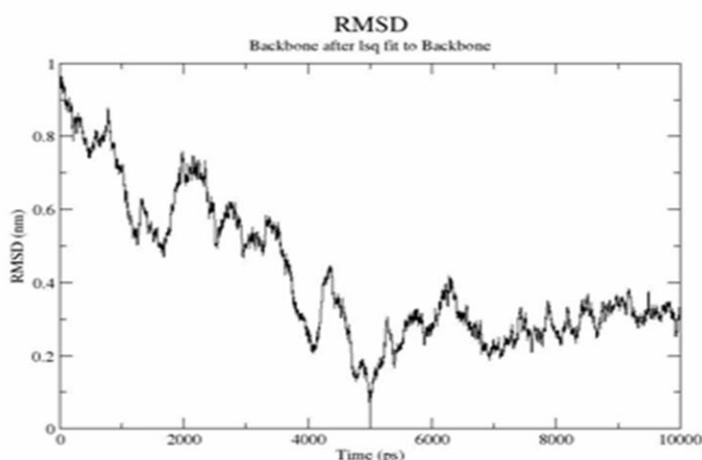


Figure 1: Root mean square deviation (RMSD) of the backbone atoms from the X-ray structure over simulation for 10 ns .

3. 1. 2. Root Mean Square Fluctuation (RMSF)

The RMSF captures, for each atom, the fluctuation about its average position. This gives insight into the flexibility of regions of the protein and corresponds to the crystallographic b-factors (temperature factors). Figure 2.A. shows the RMSF plot obtained was observed was stable only at few atoms around 500-900 and 1600-2000 atoms. The remaining atoms show a large fluctuation of maximum at 1340 atom. The energy component was inspected to ensure the stability of the trajectory.

For HSP 60.2 which has high flexibility of the protein backbone. The protein can change to conformation, often to assist protein folding. The apical domain that binds to HSP 60.2 shows high fluctuation of about 0.98 nm from original conformations. The apical domains form the opening to the solvent of the central channel. The segments of these domains that form the top surface of the molecule as well as those that face the upper regions of the channel are flexible and not very well ordered. The flexibility of these segments probably accounts for the promiscuous binding of a wide range of unfolded polypeptides [35]. It suggests the possibility of it to account for large intra-molecular motion while the incoming of proteins as well as release of proteins.

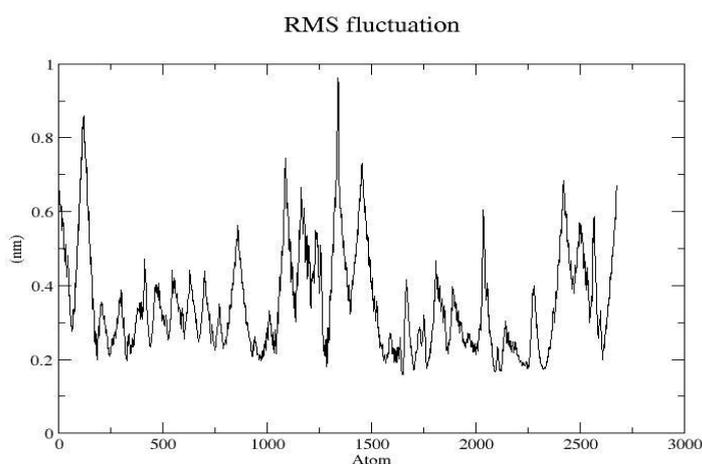


Figure 2.A: Root mean square fluctuation (RMSF) of the backbone atoms from X-ray structure over 10 ns simulation

3. 1. 3. Hydrogen bonds

During the simulation, several hydrogen bonds broke and formed. It is found that the number of hydrogen bonds range 700 during simulation. The hydrogen bond network is weak, because there is no major variation in the number of hydrogen bonds suggesting the molecule is flexible. Figure 2.B. shows the plot of hydrogen bonds in HSP 60.2 pairing within 0.35nm. The presence of hydrogen bond donors (336) and acceptors (672) is very essential for the protein ATPase activity during the course of protein folding. Due to a weak K⁺ stimulated ATPase activity, an essential component of the chaperonin-mediated protein folding reaction, the *Mycobacterium tuberculosis* HSP 60.2 shows weak ATPase activity [8]. This analysis also provides a measure for the formation of α -helix or β -turns or strands.

3. 1. 4. Radius of Gyration

Radius of gyration gives an indication of the shape of the molecule at each time. The radius of gyration compares to the experimentally obtainable hydrodynamic radius. This indicates that the first individual component corresponds to the lowest axis of molecule, which the last corresponds to smallest. In effect, the three axes give a global indication of the shape of the molecule. It is a rough measure for the compactness of a structure. All around the timescale of simulation, we found some changes in R_g . The large fluctuation in R_g is due to the conformation change in apical domain during folding of a protein. In the first 2000 ps of simulation we see R_g is stable which increases to 3.55 nm at 3000 ps time scale. In contrast, from 4000 ps the R_g decreases gradually indicating the release of folded protein (Figure 3).

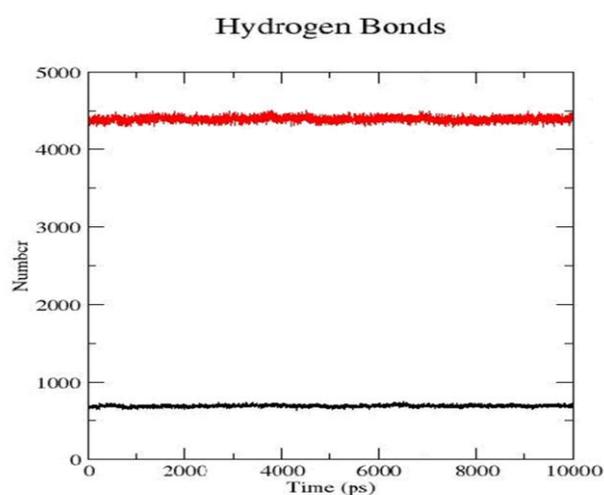


Figure 2.B: Radius of Hydrogen bonds for MD simulation for 10 ns. Black lines indicate hydrogen bonds and the red line shows pairs within 0.35 nm.

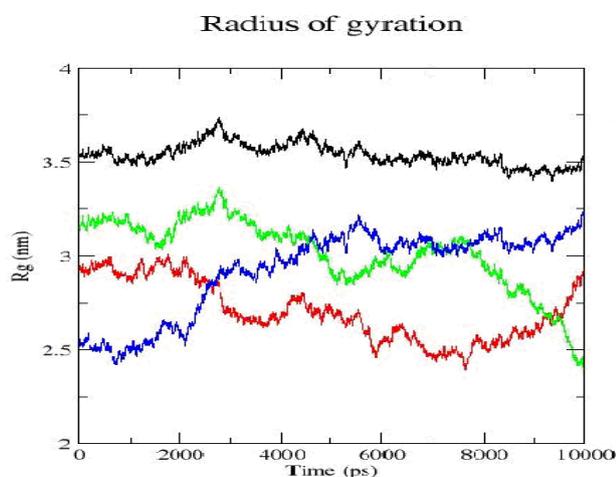


Figure 3: Radius of gyration (R_g) for MD simulation for 10 ns. . In the obtained plot the red, green and grey lines indicate the change in shape of molecule along x, y and z axis. The blue line indicates the overall change i.e., global change in shape of molecule.

3. 1. 5. Minimum distance matrix

To get information about contacts in the protein one can plot the distance between two atoms or the minimum distance between two groups or atoms. If one plots the distances between all residues of the protein, one obtains a symmetrical matrix. Plotting these matrices for different time frames (see Fig.4.), one can analyze changes in the structure. In our analysis a total of 19 matrices for 892 residues with 8187 atoms for every 1000ps were plotted to study the changes in the structure that occurs. All along the simulation time there is a large ratio of increase in the number of contacts between the residues 400-580 with nearly 1250 ratio and a gradual decrease in the number of contacts from the residues ranging from 650-892 with ratio nearly 250; when compared to the experimental x-ray structure where there is very low ratio of nearly 270 of *M. tuberculosis* HSP 60.2. The Figure 4 shows along the timescale of simulation, there is a very low minimum distance between the residues of the protein groups at nearly 7800 ps with 0.099999.

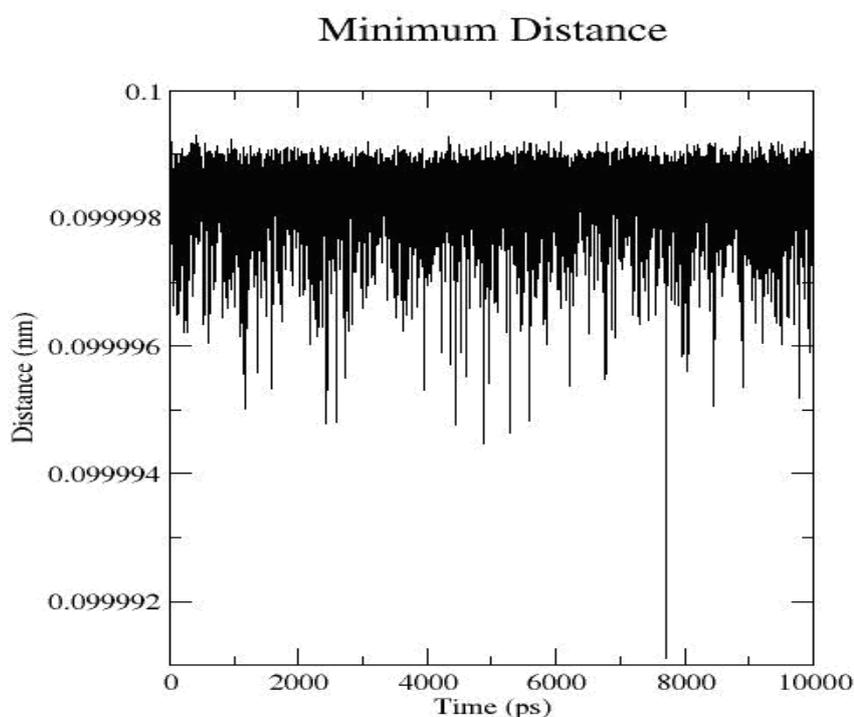


Figure 4: Minimum distance of contacts between residues based on the time scale of MD simulation for 10 ns. Black lines indicate minimum distance between protein-protein

3.2. Principle component Analysis

Principal component analysis of MD simulation of protein have indicated the collective degree of freedom dominate protein conformational fluctuation [36]. The large scale collective motions are essential for the essential dynamics of the protein [37]. Distance bonds are defined on the basis of inter atomic interactions within the starting conformation of the protein. PCA involves diagonalization of the covariance matrix of positional fluctuations. Resulting eigenvectors describe modes of the collective fluctuation of which the corresponding eigenvalues is a measure of the mean square fluctuation along the mode [38]. The subsequent equilibration time was required for the model from HSP 60.2 crystal structure, its immersion into water. Equilibration was established by monitoring the system's RMSD and radius of gyration [39] (Figure 5). It is useful to split the dynamics into different modes of motion and analyze these modes individually. In principle,

in a system of N atoms, $3N-6$ of such modes of internal fluctuations exist (6 degrees of freedom are required to describe the external rotation and translation of a system. Matrix showing coordinate covariance's between $c \alpha$ atoms. Red means that two atoms move together, so it is reasonable that on diagonal there is a red line. Blue means that they move in opposite directions. The intensity of colors indicates the amplitude of the fluctuations. From the covariance matrix it is possible to see that group of atoms move in a correlated or anti-correlated manner.

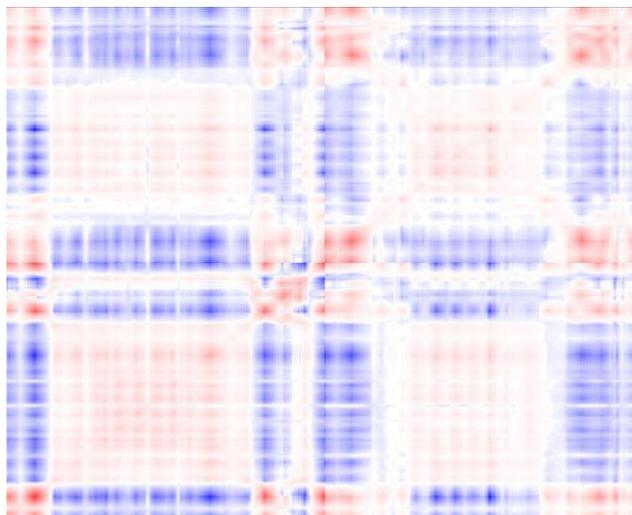


Figure 5: covariance matrix of PCA for MD simulation for 10 ns.

Consequently the time frame of 379.684 ns was considered as production data in order to fully remove the modal's equilibration period. The resulting 8028×8028 covariance matrix was subjected to the PCA; yielding eigenvectors describe the overall translation and rotation of the system. The eigenvalues from the PCA (Table 2), ranked by magnitude, decrease rapidly: the 14th largest eigenvalue is less than one tenth (0.1) of the largest and 233rd largest is less than one-hundredth (0.01) of the largest.

Table 2: Eigenvalues of the covariance matrix

Eigenvalues Index	Nm/S2/N
1	207.512
2	84.0253
3	25.1492
4	13.5509
5	8.50485
6	5.10276
7	3.70947
8	2.85243
9	2.46254
10	2.24156
11	1.69513
12	1.5786
13	1.26809
14	1.15748

3. 3. Cluster Analysis

In addition to the production simulation trajectory data, two artificial sets of data were constructed from independent trajectories of the HSP 60.2 to create trajectories containing nearly 2900 configurations at 1ps interval. We found, that the results of the clustering analysis were unambiguous, in terms of the optimal number of clusters of conformation. The first represent 19 different sized clusters created for 100 ps (Table. 3). The

RMSD ranges from 0.0434 to 1.14 nm and average clustering RMS value is 0.4913 with energy of the matrix 1949.29 nm.

Table 3: Number of clusters with size (number of structures in each cluster)

Cluster	Structures	Transition
1	2950	24
2	1401	50
3	1396	60
4	895	10
5	882	48
6	749	6
7	688	26
8	253	12
9	194	50
10	131	29
11	126	5
12	118	18
13	71	14
14	51	10
15	46	14
16	35	6
17	31	8
18	5	4
19	1	2

The total number of the transition in the structure is 223 by representing a maximum 18 between two specific clusters (Figure 6.A). The set is more different to cluster as it has both very small clusters with small variance and relatively large clusters with large variance. The average distance of each conformation in the cluster to its centroid spans a large range. The plot (Figure 6.B) shows that each cluster in the dissimilar from its neighbors and also the large clusters are may best represented. The diagonal elements (White) on left hand top represent self-comparison or zero RMSD and the black (right hand side of diagonal) shows the largest pairwise RMSD. The clustering matrix used was the best-fit RMSD of the residues defining the proper folding of proteins by HSP 60.2, specifically the apical and equatorial domains.

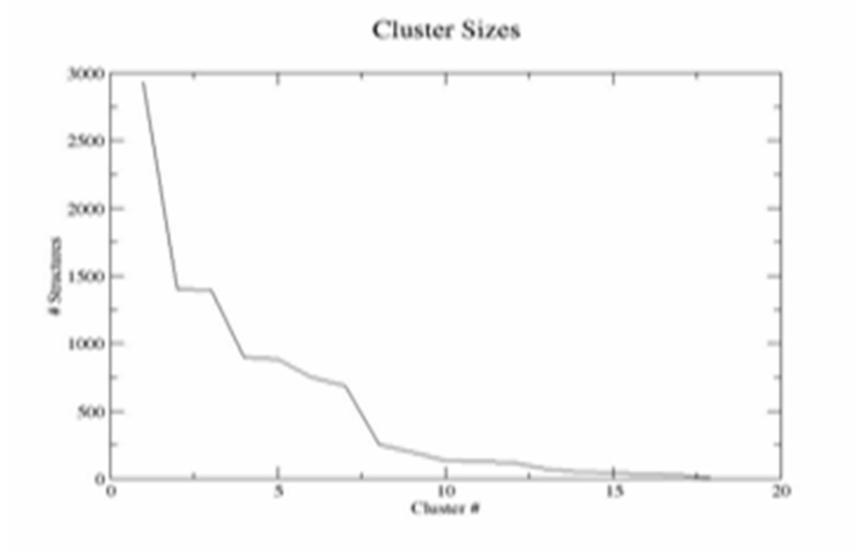


Figure 6. A: Show Size of clusters for MD simulation for 10 ns.

Clustering can identify different conformational states sampled during the simulation by grouping molecules structures into subsets based on the conformational similarity [40]. This requires a definition of similarity by a distance matrix and a definition of the space where the clustering should occur. In this study, we used only Euclidean distance expressed by the root mean square deviation (RMSD) between conformational and focus on the different options for the subspace that will be used for clustering. Here we considered only backbone carbon atoms since over focus on studying large scale motion of the model protein.

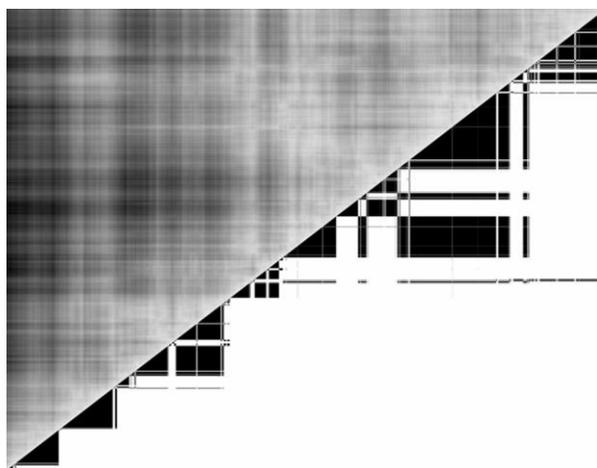


Figure 6. B: 2D RMSD plot for all frames from HSP 60.2 simulation with 19 differentially sized clusters. Note that only 18 clusters are readily visible since other one cluster is very small and negligible.

3. 4. Normal mode analysis

The 3.20 Å⁰ X-ray crystallographic coordinates were used to perform dynamic fitting using normal mode analysis. This method uses Hookean potential and linear combination of low frequency normal modes in the interactive manner to deform the structure optionally to conform to the low-resolution structure. Normal modes are used for the flexible fitting between they represent the large conformational change observed in biological system [32, 41]. One advantage of this analysis over the independent fitting of disconnected domains [42] is that it incorporates the structural constraints of the connected hinge region, thereby restraining domain separation to energetically reasonable limits. Only the lowest frequency normal modes that 7-fold symmetry is selected for functional analysis (Table 4).

Table 4: Low Frequency and Collectivity Values of Normal modes obtained

R₂	Frequency	collectivity
Mode 7	1.00	0.5920
Mode 8	1.05	0.6282
Mode 9	1.32	0.5013
Mode 10	1.73	0.5576
Mode 11	2.07	0.5899
Mode 12	2.60	0.5646
Mode 13	2.72	0.6186
Mode 14	2.78	0.7123
Mode 15	2.78	0.5331
Mode 16	3.30	0.4599

The RMSD difference of the C α backbone between the two modes is zero. Interestingly, the calculated normal mode analysis of the HSP 60.2 subunit into the reconstruction positions shows that the unresolved equatorial residues of the N-terminal and C-terminal end of the beginning of the new inter-ring density connections (Figure 7). The NMA derived structures showed, that both the *trans* and *cis* apical domain rotations counterclockwise with respect to the 7-fold axis. The large mobile loop region (297-355) in both the *cis* and *trans* apical domains appear to rotate in the same direction [43].



Figure 7: Ribbon view of two chains in the HSP 60.2 extracted from a cluster (Red-chain A and Blue-chain B)

Although there is not always a clear one-to-one correspondence, there is a quantitatively good agreement between movements described by the first five modes [44] (i.e., mode 7, 8, 9, 10 & 11). The covariance of the motion between the α -carbons of HSP 60.2 of both inter- and intra-domain motions is shown in the form of covariance matrix (Figure 8). The eigenvalues of the first five modes denote a small variation in the movement (Figure 9).

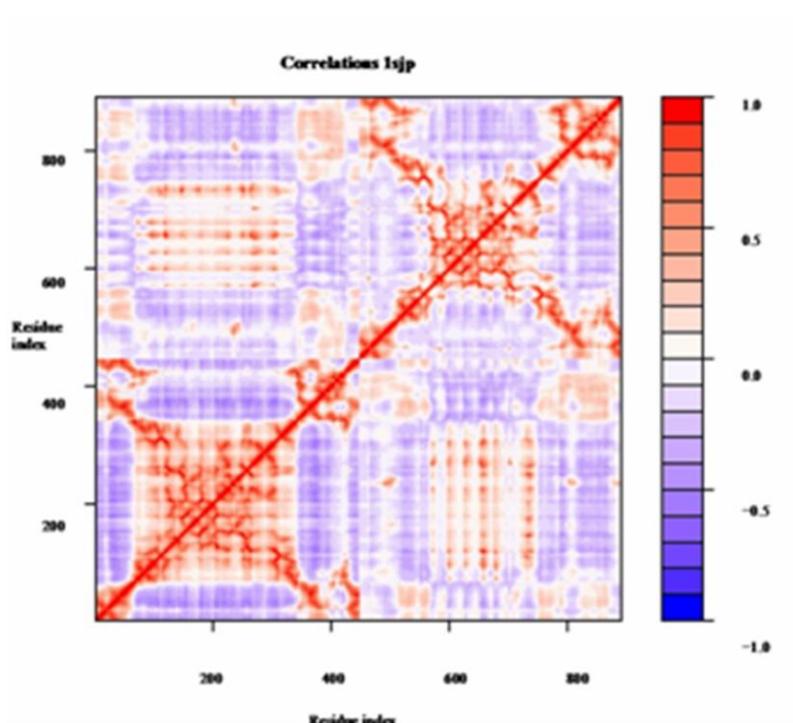


Figure 8: Correlation matrix of Normal mode analysis

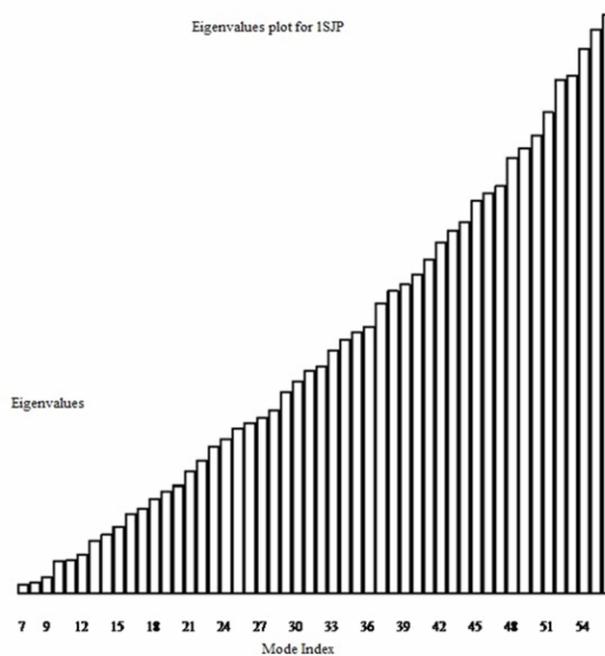


Figure 9: Histogram representing eigenvalues of each normal mode.

This shows both the apical and equatorial domain move in a direction towards intermediate domain. The blue and red colored blocks in the matrix indicate these residues for which the distance changes significantly in movement of 7th mode. Figure 10 shows the map for distance fluctuations between residue measure and the relative movements in the 7th mode. The rigid and flexible blocks of amino acid residues were identified along with their relative movement.

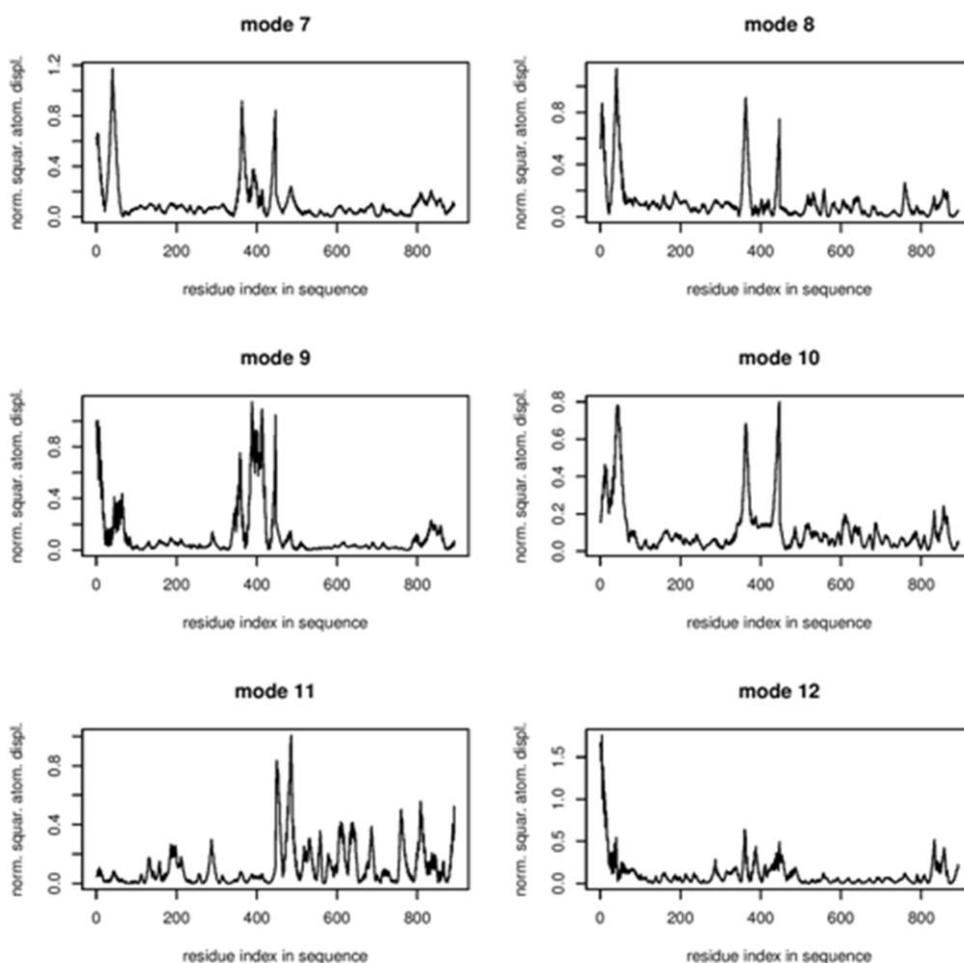


Figure 10: Normal modes 7-12 showing residue indexes based on normal square atomic displacement.

4. CONCLUSION

A molecular dynamic simulation are extensively used in order to understand the protein function in relation with structure and is particularly well-suited for studying the local minima in the free energy landscape and the transitions between these minima which characterize how biomolecules perform their requisite functions. Our MD simulation reveals that the HSP 60.2 of *Mycobacterium tuberculosis* flexibility and observe a conformational changes that occur during protein folding i.e., the protein subunit binding favors the conformer that has a strong affinity towards protein subunit since ATP-dependent conformational changes in GroEL which leads for proper folding of unfolded protein. Our analysis suggests the global movement may be significantly flexible, thus making it different by showing that all domains move in a strict fashion. Principal component analysis and Clustering methodologies applied to the results of MD simulations focus on partitioning structural ensembles into groups of structures which share similar conformational features. It is hoped that when applied to simulations of HSP 60.2, the clustering results in partitions which correspond to the descriptive-meta-stable and transition-states of the system as discussed in results. The PC analysis on molecular dynamics trajectory data revealed that major motions and conformational changes during simulation are dominant by the N-terminal helix, which is important property for protein folding. Normal mode analysis is used to predict the direction of intra-domain movement of proteins. The 7th lowest frequency mode suggests the possibility of *Mycobacterium tuberculosis* HSP60.2 large conformational changes that are regulated by binding and hydrolysis of ATP that are occurring due to flexible movement of the apical domain towards the equatorial domain. The large en bloc movement of apical

domain away from the equatorial domain and a contemporary inward movement of the intermediate domain suggesting the rigid body movement across two hinge regions (one at interface of apical and intermediate domain and the other between intermediate and equatorial domain). The distribution of negative charge on the surface of HSP60.2 and presence of hydrophobic regions on their surface might play key role in binding unfolded polypeptides and thus preventing their miss-folding.

ACKNOWLEDGEMENTS

We gratefully acknowledge for the award of INSPIRE fellowship from the Department of Science and Technology (DST), New Delhi, India for the financial support and encouragement.

REFERENCES

- [1] Kaufmann SHE, Heat shock proteins and pathogenesis of bacterial infections. Springer Semin. Immunopathol. 1991; 13: 25-36.
- [2] Hartl FU, Molecular chaperones in cellular protein folding. Nature 1996; 391: 571-580.
- [3] U Zu gel U, Kaufmann SHE., Role of heat shock proteins in Protection from and pathogenes of infectious diseases. Clin Microbiol Rev, 1999; 12: 19-39.
- [4] Tascon RE, Colston MJ, Ragno S, Stavropoulos E, Gregory D, Lowrie DE., Vaccination against tuberculosis by DNA injection. Nat Med. 1996; 2: 888-892.
- [5] De Bruyn J, Soetaert K, Buysens P, Calonne DI, Coene JL, Gallet X, et al. Evidence for specific and non covalent binding of lipids to natural and recombinant *Mycobacterium bovis* BCG Hsp60 proteins, and to the Escherichia coli homologue GroEL. Microbiology 2000; 146: 1513-1524.
- [6] Fossati G, Isso G, Rizzi E, Gancia E, Modena D, Moras ML, et al. Mycobacterium tuberculosis chaperonin 10 is secreted in the macrophage phagosome: Is secretion due to dissociation and adoption of a partially helical structure at the membrane? J Bacteriol. 2001; 185: 4254-4267.
- [7] Mande SC, Santosh Kumar CM, Sharma A., Evolution of bacterial chaperonin 60 paralogues and moonlighting activity. In: B. Henderson (Ed.) Moonlighting cell stress proteins in microbial infection. Heat shock proteins book series, Springer Netherlands, 2003; Vol 7: 101-121.
- [8] Rohini Qamra, Shekar C Mande., Crystal structure of the 65-kilodalton Heat shock protein, Chaperonin 60.2 of Mycobacterium tuberculosis. J Bacteriol. 2004; 8105-8113.
- [9] Tama F, Sanejouand YH., Conformational changes in protéin arising from normal mode calculation. Protein Eng, 2001 Jan; 14(1): 1-6.
- [10]Schuler LD, Xavier Daura, Wilfred F Van Gunsteren., An improved GROMOS 96 force field for aliphatic hydrocarbons in the condensed phase, J Copul Chem, 2001; 22: 1205.
- [11] Oostenbrink C1, Villa A, Mark AE, van Gunsteren WF., A biomolecular force field based on the free enthalpy of hydration and salvation: the GROMOS force field parameter set 53A5 and 53A6. J Comput Chem. 2004; 25(13): 1656-76.
- [12] Ryckaert JP, Ciccotti G, Berendsen JH., Numerical integration of the Cartesian equations of motion of a system with constrains: molecular dynamics of n-alkanes, J Comput Phys 1977; 23: 327-41.
- [13] Tara S, Straatsma TP, McCammon JA., Mouse acetylcholinesterase unliganded and in complex with hyperzine A: a comparison of molecular dynamics simulations, Biopolymers, 1999; 50: 35-43.
- [14] Pearson K, On lines and planes of closest fit to system of points in space. Philos Mag.1901; 2: 559-572.
- [15] Hotelling H, Analysis of a complex of statistical variables into principal components, J Educ Psychol. 1933; 24: 417-441 and 498-520.
- [16] Jolliffe IT, Principal component analysis, New York, Springer 2002.
- [17] Van Aalten DM, Conn DA, de Groot BL, Berendsem HJ, Findlay JB, Amadei A. Protein dynamics derived from clustering on crystal structure. Biophys J. Dec 1997; 73 (6): 2891-96.
- [18] Anadei A, Linsen AB, Berendsem HJ., Essential dynamics of proteins, Proteins, 1993; 17: 412-425.
- [19] Hayward S, de Groot BL., Normal mode and essential dynamics, Methods Mol Biol, 2008; 443: 89-106.
- [20] Amadei A, Chillemi G, Ceruso MA, Grottesi A, Di Nola A., Molecular dynamics simulations with constrained roto-translation motions: theoretical basis and statistical mechanical consistency, J Chem Phys, 2000; 113: 9-23.
- [21] Strang G, Linear algebra and its application. San Diego: Harcourt Brace Jonanovich, 3rd ed., 1988.
- [22] Hugher G, On the mean accuracy of statistical patterns recognizers. IEEE Transaction on Information Theory, 1968; 14(1): 55-63.
- [23] Tan PN, Steinbach M, Kumar V. Introduction to data mining, Addison-Wesle, Chapter cluster analysis: Basic concepts and algorithms. 2006; 487-568.
- [24]Shao J, Tonner SW, Thopson N, Cheathan TE, Clustering molecular dynamic trajectories: Characterizing the performance of different clustering algorithms, J Chem Theor Comput, 2007; 3: 2312-2334.
- [25] Suhre S, Sanejouand YH., ELNemo: a normal mode web-server for protein movement analysis and the generation of templates for molecular replacement, Nucleic Acids Research, 2004; 32: 610-614.
- [26] Brook B, Karpleus M., Harmonic dynamics of protein: Normal mode and fluctuations in bovine pancreatic trypsin inhibitor, Proc Natl Acad Sci USA 1983; 80: 6571-75.

- [27] Tirion M, Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys Rev Letters*. 1996; 77: 1905-1908.
- [28] Nishikawa T, Go N., Normal mode of vibrations in bovine pancreatic trypsin inhibitor and its mechanical property, *Protein Struct Funct Genet*, 1987; 2: 308-329.
- [29] Delarue M, Sanejouand YH., Simplified normal mode analysis of the conformational transition in DNA-dependent polymerase: the elastic network model, *J Mol Biol*, 2005; 320: 1011-24.
- [30] Perahia D, Monawed L., Computation of low frequency model modes in macromolecules: improvements to the method of diagonalization in a mixed basis and application to hemoglobin, *Compt Chem*, 1995; 19: 241-246.
- [31] Tama F, Sanejouand YH., Conformational changes of protein arising from normal mode calculations. *Protein Eng*, 2001; 14: 1-6.
- [32] Tama F, Gadae FX, Marques O, Sanejouand YH., Building-block approach for determining low-frequency normal modes of macromolecules, *Protein Structure Funct Genet*, 2000; 41: 1-7.
- [33] Kleywegt GJ, Use of non-crystallographic symmetry in protein structure refinement. *Acta Cryst D*. 1996; 52: 842-857.
- [34] Ralph Z, Ashley MB, Sarah P, Christopher MJ, Fernando JC, Ralph G, et al. Chaperone activity and structure of monomeric polypeptide binding domains of GroEL. *Proc Natl Acad Sci USA*. Dec 24 1996; 93(26): 15024–29.
- [35] de Groot BL, Van Aalten DMF, Amadei A, Berendsen HJC., The Consistency of large concerted motion in proteins in Molecular Dynamics Simulation, *Biophysical J*, 1996; 21: 1707-13.
- [36] Amadei A, Linssen ABM, Berendsen HJC., Essential dynamics of protein, *Protein Struct Funct Genet*, 1993;7: 412-425.
- [37] Wolf A, Baumann S, Amdt HD, Kirschnner KN., Influence of theotropton binding on the ribosomal associated region characterized by molecular dynamic simulation, *Bioorg Med Chem*. 2012 Dec 15; 20 (24):7194-205.
- [38] Shenkin PS, McDonald DQ., Cluster analysis of molecular conformations, *J Comput Chem*, 1994;15: 899-915.
- [39] Tama F, Valle M, Frank J, Brooks CL., Dynamic reorganization of functionally active ribosome explored by normal mode analysis and cryo-electron microscopy, *Proc Natl Acad Sci USA*, 2003; 100(3): 9319-23.
- [40] Rosemn AM, Ranson NA, Gowen B, Fuller SD, Saibil HR., Structure of unliganded and ATP-bound states of the Escherichie Coli chaperon GroEL by cryoelectron microscopy, *J Struct Biol*, 2001; 135: 115-125.
- [41] Falke S, Tama F, Charles LB III, Edward PG, Mark TF., The 13 A0 structure of a chaperonin GroEL-protein substrate complex by cryo-electron microscopy, *J Jmb*, 2005; 02: 027.
- [42] Skjaerven L, Martinez A, Reuter N., Principal component analysis and Normal mode analysis of protein; a quantitative comparison using the GroEL subunit, *Protein*, 2011; 79(1): 252-43.