Published in final edited form as:

IEEE Trans Terahertz Sci Technol. 2017 March ; 7(2): 131–143. doi:10.1109/TTHZ.2016.2637380.

Terahertz Frequency Spectroscopy to Determine Cold Shock Protein Stability upon Solvation and Evaporation - A Molecular Dynamics Study

Saravana Prakash Thirumuruganandham¹, Edgar A. Gómez², Shanmugamurthy Lakshmanan^{3,4}, and Michael R. Hamblin^{5,6}

¹Department of Physics, Yachay Tech University, San Miguel de Urququí, Ecuador

²Programa de Física, Universidad del Quindío, Armenia, Colombia

³Department of Dermatology, Harvard Medical School, Boston, MA 02114, USA.

⁴Wellman Center for Photomedicine, Massachusetts General Hospital, Boston, MA 02114, USA

⁵Department of Dermatology, Harvard Medical School, Boston, MA 02114, USA.

⁶Wellman Center for Photomedicine, Massachusetts General Hospital, Boston, MA 02114, USA

Abstract

Infrared (IR) and Terahertz (THz) spectroscopy simulations were carried out using CHARMM35b2 to determine protein stability. The stabilities of three bacterial cold shock proteins (Csps) originating from mesophiles, thermophiles and hyper- thermophiles respectively were investigated in this study. The three different Csps were investigated by Normal-Mode analysis and Molecular Dynamics simulation of THz spectra using the Hessian matrix for solvated systems, interpreted in the harmonic approximation at optimum near-melting temperatures of each homologue, by incorporating differences in the hydrous and anhydrous states of the Csps. The results show slight variations in the large scale protein motion. However, the IR spectra of Csps observed at the low frequency saddle surface region, clearly distinguishes the thermophilic and mesophilic proteins based on their stability. Further studies on protein stability employing low-frequency collective modes have the potential to reveal functionally important conformational changes that are biologically significant.

Keywords

molecular dynamics; vibrational modes; desolvation spectra; cold shock proteins; thermologues

1. Introduction

Microbial cells (and in particular bacteria) have evolved to survive, grow and reproduce in vastly different environments, even at elevated temperatures where the basic reactions of most living cells have to be adjusted to cope with altered reaction kinetics and changed

designed enzyme analogues, thus identifying the molecular origin of high temperature adaptations. The former study suggest that lysine residues would prefer to stabilize their native state entropically, and amino acid substitutions would result in a noticeable stabilization of proteins.

The key motivation for studying the internal motions of protein dynamics at molecular level comes from two fundamental aspects of protein action: (i) Low-frequency large scale motions of proteins that are related to functional properties and (ii) Protein function, such as protein dynamics, structural transitions, stability rates and unfolding pathways [27]. Conformational changes can be achieved by activating the different modes to cross low-energy barriers by ligand binding or environmental changes like temperature, pH and the use of organic solvents [28, 29]. In terms of protein stability, enzymes that have evolved from thermophilic (40°C optimal growth 80°C) to hyperthermophilic (optimal growth $< 80^{\circ}\text{C}$) organisms have attracted considerable mainly because they are paradigms for delineating protein stability. Moreover, no standard theory has been formulated so far to evaluate protein thermostability because this involves the interplay of several factors including secondary structures, cavities, polar surface fractions, hydrogen bonding networks, improved packing and ion pair formation. For instance it was found that the electrostatic contributions to the thermal stability of proteins in terms of salt bridges and network of ion pairs have a predominant role in ferredoxin and DNA-binding protein Sac7d [30, 31]. Alessandro et. al. [32] performed 7ns molecular dynamic (MD) simulations on hyperthermophilic and mesophilic iron-sulfur rubredoxin. Their MD simulations show that thermostability in hyperthermophilic rubredoxin increases remarkably compared to their mesophilic counterpart. It is noted that increased hydrophobic core packing density and decreased cavities led to enhanced thermostability in aldehyde ferredoxin oxidoreductase [33]. In a recent MD study Juha et. al. [34] showed that using a combination of twelve xylanases (five mesophilic and seven thermophilic enzymes) mutagenesis takes place on potential targets in the protein, enhancing thermostability. Another study by Mombelli et. al. confirms single amino acid of phenylalanine-31 mutation enhanced the thermostability in the ribonuclease P2 [35]. A collection of 64 mesophilic and 29 thermophilic proteins derived from 25 source organisms were studied by Szilgyi et. al. [1]. The strongest correlation between thermostability and structural parameter is observed in the case of electrostatic interactions and ion-network formations.

Unlike mesophilic Csp, hyperthermophilic (*Thermotoga maritima*) Csps (Tmcsp) have optimum free energy difference ($G = -0.3\text{kcal/mol}$ at $T_{\text{opt}} = 353\text{K}$) that is smaller than the non-hyper thermophilic Csps of *E. coli* ($G = -2.2\text{kcal/mol}$ at $T_{\text{opt}} = 310\text{K}$) and *Bacillus subtilis* ($G = -1.5\text{kcal/mol}$ at $T_{\text{opt}} = 310\text{K}$). In contrast, at room temperature hyper thermophilic TmCsps have $G = 4.7\text{kcal/mol}$, approximately 1.5 times higher than its Csp homologue [36]. In contrast with hyper-thermophilic TmCsps, both mesophilic *B. subtilis* BsCsps and thermophilic *Bacillus Caldolyticus* (Bc-Csps) undergo reversible two-state unfolding. Furthermore, hyper-thermophile Csps have 20 times slower unfolding rate compared to its other counterparts [37]. These unique differences between Csps homologues provide further motivation for employing structural dynamics (large scale globular motion) at THz regime that has potential to yield IR spectral signatures. In this paper the lowest-frequency vibrational motion of three functionally different proteins were investigated with

special emphasis given to their thermal responses in terms of IR absorption spectra harmonics. Another motivating factor for the current study comes directly from recent work of structurally similar yet functionally different GPCR membrane proteins such as, bacteriorhodopsin (bR), rhodopsin (Rh), and cone opsins [20, 36, 37, 38]. Applications of MD simulations and other computational techniques for analyzing the dynamics of biomolecules with focus on areas most closely connected to IR spectroscopy [39, 40, 41] includes the calculation of the spectral density [42] the energy flow in biomolecules [32], the analysis of biomolecules-water interactions, water-water interactions, spectral density of hydrated and dehydrated biomolecule, and investigations on the biomolecular geometry [43, 44, 45, 46]. The thermophilic and mesophilic proteins were derived from the X-ray results obtained by Heinemann et. al. hyper thermophile, thermophile and mesophile served as the starting structure for setting up the simulations, shown CSP homologues with secondary structure sequence: Tm-csp, Bc-csp, Bs-csp (Red: $\beta 1$, Orange: $\beta 2$, Blue: $\beta 3$, Green: $\beta 4$, Cyan: $\beta 5$, Violet: loop (Fig. 1). This paper is largely focused on vacuum simulations of solvated systems at six different optimum growth temperatures of organisms with different Csp. The dynamics of the structure as a function of time and how the Csp transitions from a solvated to a vacuum environment affect the spectral features are reported here. Thermodynamic and structural parameters, such as radius of gyration (Rgyr), root mean square deviation (rmsd), solvent accessible surface (SASA) excluding surface area (SESA), and H-bonds for formation of protein-water network are analyzed.

2. Method

2.1. Molecules

B. subtilis has a family of three Csps (CspB, CspC and CspD) with > 70% homology but the major constituent is BsCspB (hereafter referred to as BsCsp). Although Csp from mesophilic and thermophilic organisms differ widely in their stabilities, they show close structural similarities. The Csp from hyperthermophilic *T. maritima*, TmCsp (1G6P) shows 76% homology to the Csp from thermophilic *B. caldolyticus*, BcCsp (1C9O) and 61% homology to mesophilic *B. subtilis* BsCsp (2ES2) whereas the thermal stability (T_{melt} : 82° C, T_{melt} denotes the melting temperature) of TmCsp exceeds that of BcCsp and BsCsp (T_{melt} : 77° C and 54° C respectively). The structure of hyperthermophilic TmCsp was derived from the NMR structure data [40], and the structures of the thermophilic (BcCsp) and mesophilic (BsCsp) proteins were derived from the X-ray results [41, 42]. RCSB (USA) Protein Data Bank (PDB) entries 1G6P (TmCsp), 1C9O (BcCsp) and 2ES2 (BsCspB) served as the starting structures for setting up the simulations as shown in Fig. 2. All the Csps have five beta sections: two C-terminal sections followed by β – III and then by the β -hairpin formed by β – II and β – I. The program CHARMM35b2 with the CHARMM22 parameter set [43] was employed for all the calculations.

2.2. Thermophilic Csp vs non thermophilic Csp

2.3. Normal-Mode Analysis of hydrated vs non hydrated

The crystal structures of these proteins were inappropriate for direct harmonic analysis due to poor nonbonding contacts and un-optimized internal geometries. For normal-mode calculations it was necessary to first relax the molecule via optimization. The mass-weighted

Hessian was therefore diagonalized to obtain the harmonic normal-mode frequencies. The resulting eigenvalues from this diagonalization were weighted until the eigenvalues attained a stable minimum. A series of energy minimization steps were carried out to satisfy this requirement. These steps were carried out using cycles of the steepest-descent and the adopted basis Newton-Raphson algorithms [44, 45]. Harmonic constraints were initially applied on the protein backbones, and were progressively reduced at each cycle ($10 \text{ kcal/mol}^{-1} \text{ \AA}^{-2}$, $5 \text{ kcal/mol}^{-1} \text{ \AA}^{-2}$ and $2 \text{ kcal/mol}^{-1} \text{ \AA}^{-2}$); finally the constraints were removed and free minimization steps were carried out for an additional 20000 iterations using adopted-basis Newton Raphson algorithm until the average gradient of the potential energy reached $10 - 6 \text{ kcal/mol}^{-1} \text{ \AA}^{-2}$ (i.e. 0.00001). The non-bonded interactions were truncated at $13 \text{ \AA}$ and the Van-der-Waals interactions were switched off between $10 \text{ \AA}$ and $12 \text{ \AA}$. Normal modes were obtained using VIBRAN option a commonly used approach. Similarly, for solvation of protein water, all the proteins were soaked in TIP3 water. From the base of the protein surface a sequences of thickness, $2 \text{ \AA}$, $2.5 \text{ \AA}$, $3 \text{ \AA}$, $3.5 \text{ \AA}$, $4 \text{ \AA}$ and $4.5 \text{ \AA}$ of hydrated TIP3P was added to the homologue. Further minimization and diagonalization schemes were employed similar to vacuum calculations reported [46, 47, 48].

2.4. FIR absorbance

Terahertz FIR absorbance was calculated for all three proteins based on the absorbance of the fundamental vibrational transitions represented by $V' = V'' + 1$ of each normal mode, Q_i is proportional to the square of the transition dipole moment derivative, $\delta\mu_i/\delta Q_i$. Anharmonicities are not included in this approach. The absorbance is numerically evaluated according to the following [49, 50, 51, 52] equation:

$$I_i = |\langle u' | \frac{\delta\mu_i}{\delta Q_i} | u'' \rangle|^2 \cong \left| \sum_{j=1}^{3N} \frac{e_j^0 (q_{ij} - q_{ij}^0)}{\Delta Q_i} \right|^2, \quad (1)$$

where the sum is over all $3N$ cartesian coordinate displacements, $q_{i,j} - q_{i,j}^0$ is obtained from the eigenvectors of the Hessian, by dividing it by the square root of the atomic masses. Here

Q_i represents the displacement of atoms in the mode Q_i arbitrarily set to one in the approximation of Eq. (1), since CHARMM assigns a static partial charge, e_j^0 , according to the type of atom regardless of the bond length. The electrical anharmonicity is by definition zero, making the aforementioned approximation necessary. The Eq. (1) yields the integrated intensity and therefore, for comparison with experiment, the Lorentzian function was used to describe each line shape according to the following equation

$$I(v) = I_i \frac{\Gamma^2}{\pi(v - v_i)^2 + \Gamma^2}. \quad (2)$$

Here I_i is inserted from Eq. (1) and is the full width at half maximum (FWHM). In this study these lines were uniformly broadened to a width of $\Gamma = 0.2 \text{ cm}^{-1}$ without including anharmonicities.

2.5. Molecular dynamics simulation: Protein in water

Csps were immersed in water (3000 TIP3) molecules in an octahedral box of side length $55\text{\AA}$ with periodic boundary conditions excluding molecules less than $2.2\text{\AA}$ from a protein. The water molecules in the system are in a perturbed environment surrounded by a protein molecule. Initially, a short energy minimization step (to remove Vander-Waals interactions) using harmonically reduced force constraint was applied to the protein. Subsequently, a steepest descent analysis using Newton-Raphson steps was performed. All simulation systems were equilibrated with the following protocol: In the first 100 ps , the system was heated to the working temperature of 300 K using velocity scaling. For the next $50\text{--}250\text{ ps}$ a proper equilibration was carried out in which the pressure was kept constant by adapting the extended-pressure algorithm function followed by a production phase step at 300 ps . Simulations were carried out in the isothermal-isobaric (NPT) ensemble at 300 K in which atomic coordinate resolution was $1\text{\AA}$, and temperature (Nose-Hoover thermostat) was fixed at $T = 300\text{ K}$, with a pressure $p = 1\text{ atm}$ (extended-pressure algorithm implemented in a leap-frog Verlet integrator). In addition, long-range electrostatic interactions were modeled with the particle-mesh Ewald criteria. The width of the Gaussian distribution 'K' used in this method was set to 0.34. The grid size of the particle-mesh Ewald method was $64 \times 64 \times 64\text{\AA}^3$ with a Bsp line order of $6\text{\AA}$ and nonbonding cut off of $12\text{\AA}$. During the simulations SHAKE algorithm function was deliberately inactivated to prevent all bond lengths free of their equilibrium values. The integration time step was set at 0.5 fs with every two steps (1 fs) recorded in a trajectory file for further evaluation.

2.6. Desolvation dynamics of protein in Vacuo

In order to understand the different degrees of local stability at different regions of the protein during the transition from solution to gas phase, MD simulations were performed followed by IR spectra of proteins, during the dehydration process. These simulations were performed by starting from solvated system; a set of $4.5\text{\AA}$ thickness of water shell of proteins were created from the 250 ps of NPT simulation, i.e., droplets consisting of 529, 497 and 552 water molecules were created. The system was then subjected to minimization; a heating phase temperature of 300 K in steps of 5 K over 50 ps with time steps of 0.5 fs was established. During the equilibration phase, the velocities of atoms were scaled to bring the temperature back to 300 K whenever it deviated by more than 10 K from the reference. This procedure was continued until a time of 250 ps in vacuum was achieved. Finally, the production run was started, in which the system configuration was recorded every 1 fs over a total time of 1 ns . These values were found to increase the accuracy of the simulation without significant impact on the computational step. A simulation time step of 0.5 fs was allowed in order to simulate the high frequency bond vibrations between hydrogen and heavy atoms in the protein. During the simulations the bonds between the hydrogen atoms and the heavy atoms were maintained in the non-rigid mode using the SHAKE algorithm. The rate of sampling corresponded to a maximum frequency of 3600 cm^{-1} (normally the range is between $2500\text{--}3600\text{ cm}^{-1}$). Only a few bond-stretching vibrations occur above this frequency and therefore it was necessary to use a very short time step. These conditions in turn lead to a very time consuming simulation. The last 500000 points (500 ps) of the simulation were used for the analysis. Similarly, all homologues were subjected to initial temperature of 600 K , 400 K , 350 K , 340 K and 310 K to run the production dynamics of

500 ps. Using MD simulations, the density of states (DOS) was calculated applying IR spectrum $S(V)$ of the vibrational motion. The IR spectrum $S(V)$ is determined from the Fourier transform $\tilde{C}_{\mu\mu}(v)$ of the dipole auto-correlation function $\tilde{C}_{\mu\mu}(v)(t)$ using the following equation:

$$S(v) = \frac{2}{K_B T} \sum_{j=1}^{3N} M_j \tilde{C}_{\mu\mu}(v). \quad (3)$$

Here, M_j is the mass of atom j . The dipole auto-correlation function is calculated from MD using:

$$\tilde{C}_{\mu\mu}(v)(t) = \frac{\langle \mu_i(t + \tau) \mu_i(\tau) \rangle}{\langle \mu_i(\tau) \mu_i(\tau) \rangle}. \quad (4)$$

Where $\mu_i(\tau)$ is the dipole vector of the i -th atom in a protein molecule at time t and the notations within the angular brackets denote averaging over all atoms in the system. The Fourier transform of the dipole auto-correlation function provides a direct measure of the derived IR spectra as a function of the frequency. Sampling of the dipoles was carried out every 1 fs to obtain the high-frequency modes and less often to obtain the low-frequency modes (5 – 10 fs).

3. Results and Discussion

3.1. Normal-mode Analysis and Spectral Intensity Variation upon Hydration

Normal mode analysis is based on a harmonic approximation of the potential energy surface at global minimum. Fig. 2 illustrates the terahertz FIR absorption spectra of each Csp homologues with the corresponding crystallographic structure in vacuo with respect to two different surface water thicknesses (2 Å and 2.5 Å). The effect of solvent molecules on protein dynamics is an important issue that has been previously addressed by experimental and computational studies [3, 25, 28, 53]. Fig. 2 (a) shows that mesophilic BsCsp has increased absorption compared to their other two counterparts. Compared to BsCsp, the other two counterparts show fewer additional peaks at the low frequency region around 0 – 100 cm^{-1} . Furthermore, Fig. 2 (b-c) clearly displays the differences in absorption of protein in vacuo and in solvated system at 2 Å and 2.5 Å water thickness. An increase in amplitude was seen for TmCsp and BcCsp in the solvated system.

Subsequently, upon employing a solvation layer of 2 Å (thickness) the minimal absorption peaks of BsCsp indicate that the damping of surface water may influence the collective motion of the mesophilic Csp owing to the lack of ion pairs and weak hydrogen-bond network in the protein-surface water interactions in mesophilic Csp. By contrast the low absorption peak profile was not observed in the case of BcCsp and TmCsp. Upon addition of a simple water shell of 2 Å (thickness) there was no significant change in the low frequency collective motions in their in vacuo (native) state. Furthermore, upon progressive addition of

water (in $2\text{\AA}$ thickness steps when all homologues tend to form water IR modes), it was observed that the process of suppression of protein modes was faster in the mesophiles (BsCsp) but slower in thermo and hyper thermophilic homologues (BcCsp and TmCsp). Normal mode analysis of a $2\text{\AA}$ surface water shell mimics the cage effect of water, which in turn affects the protein conformational state (Fig. 2 (a)). This behavior confirms that cage effect of water might strongly influence the stability and hence the collective motion of the mesophilic BsCsp. In Fig. 2 (c) the increased IR absorption of TmCsp at low frequency region may be attributed to the ten lysine rich side chains associated with the molecule. For both TmCsp and BcCsp strong hydrogen bonding was observed. The ion-pair network protects the protein from globular conformational transitions for solvation less than $2.5\text{\AA}$, indicating that due to solvation, the large scale motion increases in thermophilic Csp as shown in Fig. 2 (a c). In Fig. 2 (c), it is noticed that as a function of increased solvation, thermophilic TmCsp and BcCsp have gradually increased its amplitude distribution ($0 - 100\text{ cm}^{-1}$). In contrast to thermophiles, BsCsp increased the overall fluctuations found in vacuo; the degree to which these low frequency modes have been increased varies among the Csp. Overall, the trend behaves as if the TmCsp and BcCsp can undergo concerted motions more often than the BsCsp. However, DOS started increasing upon addition of water to Csp, which was measured as a function of hydration in the THz frequency range as is shown in Fig. 3. No strong differences were seen for all the three homologues when the water concentrations were varied. In general, at levels above 85 cm^{-1} water molecules undergo a critical transition of low frequency modes of protein, corresponding to the point at which surface water sites are covered and bulk water begins to appear. This trend could be seen clearly at $30 - 75\text{ cm}^{-1}$ as a broadened peak for 272, 382 and 459 water molecules (see Fig. 3). The other striking difference between Csp homologues can be described through direct MD based calculations of free energy and configurational entropy from the solvation effects at high temperatures. Using MD based quasi-harmonic approximations, several authors have explained the important role of charge distribution for the thermostability of various thermophilic and non-thermophilic homologues [53, 54, 55, 56, 57]. Here, the configurational entropy S and specific heat capacity C_v from the mass weighted Hessian matrix were measured. The heat capacity of hydrous proteins, due to the presence of intrinsic covalent interactions, carry an additional component contributed by the solvent. This quantities can be computed directly by considering the following equations:

$$S = K_B \sum_{i=1}^{3n-6} \frac{\alpha}{e^\alpha - 1} - \ln(1 - e^\alpha), \quad (5)$$

$$c_v = K_B \sum_{i=1}^{3n-6} \frac{\alpha^2 e^\alpha}{(e^\alpha - 1)^2}, \quad (6)$$

where $\alpha = (\hbar \nu_i)/K_B T$ with K_B and $\hbar$, the Boltzmann and Planck constants, respectively.

Moreover, ν_i denotes the i -th vibrational mode. Fig. 4 (a-b) shows that hyperthermophilic TmCsp has higher heat capacity and configurational entropy after increase in temperature,

compared to mesophilic BsCsp and the partially hydrated and non-hydrated structures as well. However, there is only a small difference observed between BcCsp and BsCsp, which are more compact for the fully minimized structure, based on time dependent simulated trajectories for the folded state, the direct calculation from the diagonalization approach shows that the behavior of native TmCsp > native BsCsp and native BsCsp, which clearly distinguish between the three entities on a molecular level. It can be deduced that TmCsp are able to absorb more heat compared to their mesophilic counterpart (BsCsp) for each temperature increment. Looking at proteins in another context, on the basis of these structures, molecular adaptations that increase thermostability could be identified both within and between the subunits. In particular, combining positive and negative mutations leads to increased thermostability. Scholtz et. al. [58] showed that introducing a single lysine residue to the surface of ribonuclease can increase the stability of the protein; it is possible that this addition disrupts the existing electrostatic repulsive interactions. Similarly, Ladenstein et. al. [59] confirmed that entropic contributions affect configurational states of freedom of similar energy. The authors also showed that the thermophilic glutamate dehydrogenase have higher stability due to the formation of multiple microstates from the contribution of a higher number of aromatic amino acid side chains such as histidine (His), phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp). They further showed that the electrostatic contribution of five arginine residues raise the degenerated energy level of the thermophilic proteins, whereas a single arginine of the mesophilic protein contributed to the decreased entropy in glutamate dehydrogenase. The findings in the current paper can be correlated with the quasi harmonic analysis of Ladenstein et. al. [60]. It can be seen that the higher entropy and heat capacity in TmCsp is an indication of a higher degeneracy of the accessible energy states. Uniformly formed ionic interactions between ten positively charged lysines and nine aromatic side chains are present in TmCsp, compared to BsCsp with only five lysines and eight aromatic side chains. These interactions mean a lower number of ion pairs is formed which decreases the entropic and specific heat capacities of BsCsp at elevated temperatures. From our findings it can be estimated that the local internal protein motions of ionic and aromatic amino acids may increase stability by increasing the entropy of the native state. Such motions can also raise the melting temperature of a protein by increasing the heat capacity of the native state and thus decreasing the heat capacity differences between the native and its intermediate structures. In order to obtain a quantitative measure of the engagement of the various protein domains in a vibrational mode, we calculated the normalized squared atomic displacements for each C_α atom by considering the following equation [61, 62]

$$D = \frac{d_i^2}{\sum_i d_i^2}, \quad (7)$$

where d_i is a component of the eigenvector corresponding to the i -th C_α atom. The normalized atomic displacements for modes numbered as 7, 8, 10, and 15 are represented in Fig. 5 as a function of the amino acid sequence for vacuo, primary ($2.5\text{\AA}$) and secondary ($3.5\text{\AA}$) solvation shells respectively. Atomic fluctuations can be separated into local oscillations in the range close to picosecond time scale superimposed on motions with a

more collective character. These collective motions are responsible for the variation in the magnitude of the fluctuation that is characteristic of different parts of a protein and thus likely to be of primary importance to its biological function. The highest peaks in the plots of Fig. 5 (a-i) correspond to the most displaced residues in the corresponding mode. Overall, TmCsp have a similar fluctuation profile both in vacuo and with a 2.5 Å water shell, whereas in BcCsp and BsCsp all the five beta sheets contribute to the low frequency modes. Although the mode distribution is identical in vacuo, the three different Csp differ in their displacement after addition of a proximal water thickness of 2.5 Å attributed by the strong sub-domain network and non-bonding interactions that stabilize the protein conformation to produce a dampening of amplitude that is less favorable in TmCsp. Moreover, mode 7 has a small delocalized motion indicating that primary water shells influence low frequency collective behavior of mesophiles and thermophiles. This behavior indicates conformational transitions are different in each homologue. Moreover, the inclusion of surface water dampens the amplitude of atomic displacements. However, no significant change was observed for TmCsp upon adding proximal water thickness of 2.5 Å possibly due to the strong subdomain network and non-bonding interactions that stabilize the protein conformation damping the amplitude in TmCsp. Moreover, mode 7 shows a small delocalized motion at β -II- β -III (Euler's equations for the motion of a symmetric top), and was dominant at the loop that connects the IV and V β sheets. Mode 8 is dominant at the loop between II and III- β sheets, whereas mode 9, 15 and 150 (not shown here) are conspicuously delocalized at β -III to β -I. On the other hand, with a surface water shell (2.5 Å) similar distributions were observed, except for a few smaller distortions in mode 7 which significantly damped the delocalized motion of β -II to β -III. Also, it could be estimated that as the mode number increases, the distribution of modes travels starting from the C terminus to N terminus for all the three homologues. In addition, 3.5 Å systems showed comparatively higher peaks of overall domain motions than the system with surface water (2.5 Å) and the in vacuo system. The absence of amplitude damping and enhanced amplitude confirms that low frequency motions are affected significantly in the bulk water case, therefore 2.5 Å can be considered as a critical region for the transition of surface to bulk water ($> 2.5 \text{ Å}$).

3.2. Dehydrated Spectral Intensity and Regional stability

The amino acid sequences of the three Csp in terms of globular motion are compared in Fig. 6. BcCsp differs from BsCsp at 12 positions (that said, include a deletion at the C terminus), whereas TmCsp differs from BsCsp at 24 positions (includes three deletions, one insertion and one addition of residue at the C terminus). 42 out of 67 residues are conserved in all three proteins. The ability to measure THz absorbance of biomolecules in solution has the potential to experimentally confirm MD calculations on the frequency scale of biologically relevant macromolecular motions. Studies on the effects of the solvent pH, and of denaturing agents are in progress. Studies related to vibrational density of states (VDOS) of lysozyme in its native and denatured states were performed using diagonalization schemes [63]. In the present study we present the THz signature at elevated temperature and measure the induced transition states and denatured state based on time-dependent trajectory analysis. The desolvation process is accompanied by a strong loss of water molecules due to the high temperature. Since the system was simulated over 1 ns, the first 1000 ps can be considered

as a primarily favorable evaporation process, after that, the water evaporation rate is steady and slower from 1000 – 2000 *ps* which is considered as a secondary evaporation process step. Birge et. al. [6] reported experimental observation showing that the absorbance of myoglobin excited too many modes at room temperature (294 *K*), whereas at low temperature (77 *K*) the modes were suppressed. Fig. 4 show the difference in the THz absorbance profile between the native conformation and the thermally induced denatured meso-, thermo- and hyperthermo-Csp profiles at 0–600 *cm*⁻¹, and the impact of low frequency globular modes on the individual contributions of thermostability in each Csp homologue. In BsCsp, the spectral regions are shifted in the 300 *K* hydrated system, compared to the systems at 400 *K* and 600 *K*, in addition to the simultaneous excitation of too many modes (results not shown here). In addition, different absorption bands were shifted by different amounts, which can be attributed to the temperature dependent behavior of the different coupled lattice vibrational modes. In contrast to BsCsp, for the case of BcCsp and TmCsp (Fig. 6 (a-d)), the low frequency regions are less affected due to the evaporation and high temperature (> 400 *K*). Table 1 shows that the heated system at 400 *K* reached an average of 331 *K* and 329 *K* during the standard evaporation phase with a loss of 25 and 17 water molecules. Similarly, the 600 *K* system reached an average of 381 *K* and 395 *K* with a loss of 250 water molecules, i.e., the systems have a tendency to shift from liquid to gas phase transitions. The results suggest that the hydration and dehydration mechanisms are quite different for TmCsp compared to BsCsp. To investigate the Csp stability, it was further explored if there were any significant differences in the collective mode region during the dehydration process. It has to be noted that, to date, a modeling for the temperature dependent DOS in Csps has not been reported. As shown in Fig. 7 (a-h) all homologues shift their low frequency modes at higher temperatures, i.e., the protein desorption process at 400 *K* and 600 *K* dissipates the collective modes, and gives an increased absorption for the second peak of the saddle surface. In Fig. 7, it can be seen that with a low temperature desorption process (300 – 400 *K*) the mesophilic absorption shifts are greater compared to thermophilic ones; this hypothesis is supported by the fact that desorption process impacts the nonbonding interactions of proteins, and gives extremely weak temperature dependent frequency shifts in BsCsp. The differences found in the temperature shifts of low frequency modes are moderately less in BcCsp and TmCsp (Fig. 7 (a-f)). At higher temperature, it is more likely that thermal motions lead to variation in both the frequency shift and the evaporation process of the system; as the evaporation process in the current study may not have allowed the system time to become completely free of water (see Table 1). It was observed that there are a few water molecules still frozen at the surface of the protein; some of these could possibly conserve the nonbonding interactions of the protein and some could possibly change the collective behavior of the protein due to the biological effect of the water network. Even if this is the case, it is also likely that the frozen solvent matrix possibly induces the variation of absorption phenomena. TmCsp (Fig. 7 (g-h)) showed high-temperature (600 *K*) spectra which does not change the shape of the saddle peak compared those at low temperature (300 *K*), whereas BsCsp and TmCsp both show a wider deviation and shift. Furthermore, due to the nature of high temperature resistivity, the spectral heterogeneity is comparatively smaller in the thermophilic Csp. On the other hand, the mesophilic BsCsp showed strong evidence of spectral heterogeneity. There also exists the possibility of anharmonic contribution to the atomic motions that are protein-dependent

which become evident at high temperatures. Fig. 7 (g-h) show that the anharmonicities of motion induce the transitions of different conformational substrates of the protein and functional behavior from low to high temperatures. The influence of the initial structure upon desolvated conditions on the MD simulations in proteins has not been extensively studied; however some previous research [63, 64], has indicated there are differences in protein structure when MALDI and laser spray experiments and simulations were used. As the temperature increases, unfolding exposes more hydrophobic groups, causing an increase in hydrophobic interactions in both the primary and secondary stages of evaporation. A common practice to enhance the thermal resistance is to increase the rigidity of the molecular structure; several studies have shown that thermophilic proteins are more flexible than mesophilic ones [65]. In the current study (Table 1) the average RMSD of N and C terminal residues increase with temperature. Moreover, with respect to Cloud Point Temperature (CPT) and $T = 300\text{ K}$ the maximum deviation is observed at 400 K and 600 K for all three homologues. Since all homologues pass through uniform and marked increments of N and C terminal RMSD, these do not contribute to any difference in the thermal resistivity. Despite large RMSD differences observed in the molecular structure at $T = 600\text{ K}$, none of the proteins undergo any significant unfolding processes, as evidenced by the radius of gyration (Fig. 6 (a)) The radius of gyration (Rgyr) is a measure of the compactness of a protein in the liquid-vacuum transition, indicating the volume, orientation and shape of the protein. The increased Rgyr correlate to the expansion of the system during the desorption process; for temperatures higher than 400 K , the systems maintain their crystallographic/native secondary structures. Sudden fluctuations of Rgyr at $400\text{ K} < T < 600\text{ K}$ (Fig. 6 (a)) indicates that all the Csp homologues pass through two structural stages during the desorption process. In particular, during the primary and secondary evaporation process, the mesophilic-Csp have a larger deviation of Rgyr when observed at 600 K compared to TmCsp and BcCsp. This trend shows that the low thermal resistivity unfolding process is faster in mesophilic Csp compared to TmCsp and BcCsp, indicating less thermostability in BsCsp. Interestingly, it is noticed that the Rgyr transition or pseudo-melting of the protein structure starts at 410 K , 450 K and 500 K for BsCsp, BcCsp and TmCsp, respectively (Fig. 6 (a)). Additional measures to quantify protein conformation were carried out using solvent accessible surface area (SASA) shown in Fig. 6 (b), the solvent excluding surface area (SESA) in shown in Fig. 6 (c), and the number of hydrogen bonds as a function of temperature is shown in Fig. 6 (d). Since the balance between the forces dominating protein and protein-solvent interactions is crucial to establish the secondary and tertiary structures of proteins this concept is widely used in protein structural analysis. At $T = 300\text{ K}$ the SASA of TmCsp, BcCsp and BsCsp are maintained at $48\text{ \AA}^2$, $46\text{ \AA}^2$ and $44\text{ \AA}^2$, respectively. Later on, these values moderately increases upon temperature dependent desorption process ($300\text{ K} < T < 350\text{ K}$). For temperatures between 340 K and 350 K the SASA continues to increase in BsCsp, but did not increase for the other two homologues. At the same time, SESA (the area of the corresponding surface which is not accessible by solvent), start to shrink due the proteinwater collisions during the evaporation process. Interestingly, during the primary evaporation process, between $300\text{ K} < T < 350\text{ K}$, TmCsp do not show any extreme increment of SASA as expected, but by contrast there is a shrinkage of SESA (red line). On the other hand there is a sharp increase in SASA and decrease of SESA for both BcCsp and BsCsp. It can be predicted that the desorption

mechanism between $300\text{ K} < T < 350\text{ K}$, does not influence the conformational changes in all three homologues; and the hydration shell of the protein is still more or less intact. However, in the case of the evaporation process between $350\text{ K} < T < 600\text{ K}$, due to the temperature rise in both protein and solvent (Table 1) and the conformational fluctuations of protein Fig. 6 (b-c), SESA is larger and SASA is smaller. These values are considerably smaller than protein in bulk, whose behavior assumes that protein is either completely or partially free of water. Moreover, once the protein is completely exposed to vacuum (gas state) SESE and SASA no longer exist; this condition is important to determine the protein structure. Similar for Rgyr, Fig. 6 (a) compared to the primary and secondary evaporation process, mesophilic BsCsp show a large deviation of SASA and SESA for the evaporation process $400\text{ K} < T < 400\text{ K}$, which clearly indicates that mesophilic Csp undergoes the denaturation process much faster than thermophilic and hyperthermophilic Csps. In addition, Fig. 6 (d) illustrates how formation of a protein-protein hydrogen bond network profoundly increases as the number of water molecules decrease. As the temperature (Table 1) is gradually increased, intra-protein hydrogen bonds are formed. It appears that the loss of protein-solvent bonds is compensated by the creation of new protein-protein bonds. At $300\text{ K} < T < 350\text{ K}$, TmCsps maintain the hydrogen bond network during the primary and secondary evaporation process, while this trend is not found in either BsCsp or Csp. In each protein at temperatures 600 K and above, an abrupt increase in intramolecular hydrogen bonding of the order of 25% takes place, confirming that the formation of intra-protein hydrogen bonds stabilizes the protein in its solvent-free vacuum phase and provides a different conformational state for each homologue.

4. Conclusions

In this study, vibrational spectra in hydrated and dehydrated phases for three geometrically similar yet functionally different beta-sandwich proteins namely, mesophilic, thermophilic and hyperthermophilic Csp were analyzed. The low frequency THz spectra and DOS of the three Csps were calculated as a function of hydrous and anhydrous states using both normal-mode analysis and MD simulations at a temperatures in the range $300\text{ K} - 600\text{ K}$. Taking into account the vibrational anharmonicities which affect molecular motion at these temperatures and hydration levels, especially those involving averages over many modes, the normalized atomic fluctuations calculated from the normal modes and default IR spectral calculations allowed the thermostability of each Csp homologue to be clearly distinguished and characterized. Furthermore, the frequency distribution of the modes that contribute significantly to the fluctuations ($< 100\text{ cm}^{-1}$) is in consensus with the estimates from MD. This correlation provides evidence for the utility of a normal-mode treatment for the analysis of the internal motions of Csp ($0 - 100\text{ cm}^{-1}$). It can be concluded that IR spectra of Csp during the evaporation process, in particular, low frequency saddle surface region (double peak; $0 - 500\text{ cm}^{-1}$) clearly indicates differences between the thermophilic and mesophilic proteins. In addition the absorptions associated with both amide-I and amide-II bands are sensitive to the secondary structural content of a protein. It is possible to extend our simulations with improved polarizable force fields which has potential to open pathways to the comparison study between the meso and thermophilic homologues at different melting temperatures. When the protein moves from a hydrate state to the gas phase, the temperature

drop as a function of radius of gyration confirms that mesophilic Csp undergoes faster denaturation than its other two counterparts. In our study it was observed that during the evaporation process of the system at 400 K and 600 K, protein-water clusters suddenly lost thermal contact with the water, while the average primary evaporation temperature still remained higher than the secondary evaporation temperature, due to the loss of the majority of the hydration shell. Taking this observation into account, it can be concluded that the temperature drop is relatively faster in mesophilic BsCsp, i.e. the temperature drops from 400 K to 355 K for primary and 400 K to 324 K for secondary evaporation processes which is relatively higher than observed with the other two counterparts.

Acknowledgements

SPT acknowledge Saguaro supercomputing facility at Arizona State University, Tempe, USA. S. Lakshmanan acknowledges World Institute for Scientific Exploration for funding. EAG acknowledges the financial support from Vicerrectoría de Investigaciones at Universidad del Quindío (Colombia) through research grant No. 752.

6. REFERENCES

- [1]. András S, Péter Z, Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey, *Structure*, 8 (2000).
- [2]. Kaledin AL, Kaledin M, Bowman JM, All-Atom Calculation of the Normal Modes of Bacteriorhodopsin Using a Sliding Block Iterative Diagonalization Method, *Journal of Chemical Theory and Computation*, 2 (2005) 166–174.
- [3]. Leitner DM, Havenith M, Gruebele M, Biomolecule large-amplitude motion and solvation dynamics: modelling and probes from THz to X-rays, *International Reviews in Physical Chemistry*, 25 (2006) 553–582.
- [4]. Fanconi B, Low-frequency vibrational spectra of some homopolypeptides in the solid state, *Biopolymers*, 12 (1973) 2759–2776. [PubMed: 4782552]
- [5]. Shotts WJ, Sievers AJ, Far infrared spectra between 4 and 200 cm^{-1} of poly-L-alanine, *Chemical Physics Letters*, 21 (1973) 586–588.
- [6]. Markelz A, Whitmire S, Hillebrecht J, Birge R, THz time domain spectroscopy of biomolecular conformational modes, *Phys Med Biol*, 47 (2002) 3797–3805. [PubMed: 12452570]
- [7]. Markelz AG, Roiberg A, Heilweil EJ, Pulsed terahertz spectroscopy of DNA, bovine serum albumin and collagen between 0.1 and 2.0 THz, *Chem Phys Lett*, 320 (2000) 42–48.
- [8]. Zhang C, Durbin SM, Hydration-Induced Far-Infrared Absorption Increase in Myoglobin, *The Journal of Physical Chemistry B*, 110 (2006) 23607–23613. [PubMed: 17107217]
- [9]. Hinsen K, Kneller GR, Solvent effects in the slow dynamics of proteins, *Proteins: Structure, Function, and Bioinformatics*, 70 (2008) 1235–1242.
- [10]. Hinsen K, Kneller GR, A simplified force field for describing vibrational protein dynamics over the whole frequency range, *The Journal of Chemical Physics*, 111 (1999) 10766–10769.
- [11]. Giraud G, Karolin J, Wynne K, Low-Frequency Modes of Peptides and Globular Proteins in Solution Observed by Ultrafast OHD-RIKES Spectroscopy, *Biophysical Journal*, 85 (2003) 1903–1913. [PubMed: 12944303]
- [12]. Balog E, Smith JC, Perahia D, Conformational heterogeneity and low-frequency vibrational modes of proteins, *Physical Chemistry Chemical Physics*, 8 (2006) 5543–5548. [PubMed: 17136269]
- [13]. Korter TM, Balu R, Campbell MB, Beard MC, Gregurick SK, Heilweil EJ, Terahertz spectroscopy of solid serine and cysteine, *Chemical Physics Letters*, 418 (2006) 65–70.
- [14]. Balu R, Zhang H, Zukowski E, Chen JY, Markelz AG, Gregurick SK, Terahertz Spectroscopy of Bacteriorhodopsin and Rhodopsin: Similarities and Differences, *Biophysical Journal*, 94 (2008) 3217–3226. [PubMed: 18199669]

- [15]. Thirumuruganandham S, Urbassek H, Low-frequency vibrational modes and infrared absorbance of red, blue and green opsin, *J Mol Model*, 15 (2009) 1545–1545.
- [16]. Zhang H, Zukowski E, Balu R, Gregurick SK, A dynamics study of the A-chain of ricin by terahertz vibrational calculation and normal modes analysis, *J Mol Graph Model*, 27 (2009) 655–663. [PubMed: 19095477]
- [17]. Knab JR, Chen J-Y, Yunfen H, Markelz AG, Terahertz Measurements of Protein Relaxational Dynamics, *Proceedings of the IEEE*, 95 (2007) 1605–1610.
- [18]. Mathias G, Marx D, Structures and spectral signatures of protonated water networks in bacteriorhodopsin, *Proceedings of the National Academy of Sciences*, 104 (2007) 6980–6985.
- [19]. Rajamani R, Gao J, Combined QM/MM study of the opsin shift in bacteriorhodopsin, *Journal of Computational Chemistry*, 23 (2002) 96–105. [PubMed: 11913393]
- [20]. Ganim Z, Tokmako A, Spectral Signatures of Heterogeneous Protein Ensembles Revealed by MD Simulations of 2DIR Spectra, *Biophysical Journal*, 91 (2006) 2636–2646. [PubMed: 16844758]
- [21]. Pleiss J, Jhnig F, Collective vibrations of an alpha-helix. A molecular dynamics study, *Biophysical Journal*, 59 (1991) 795–804. [PubMed: 2065186]
- [22]. Wang Q, Wong CF, Rabitz H, Simulating Energy Flow in Biomolecules: Application to Tuna Cytochrome c, *Biophysical Journal*, 75 (1998) 60–69. [PubMed: 9649368]
- [23]. Siber A, , Vibrations of closed-shell Lennard-Jones icosahedral and cuboctahedral clusters and their effect on the cluster ground state energy, *Phys. Rev. B* 70 (2004) 075407.
- [24]. Pevsner A, Diem M, Infrared Spectroscopic Studies of Major Cellular Components. Part I: The Effect of Hydration on the Spectra of Proteins, *Appl. Spectrosc*, 55 (2001) 788–793.
- [25]. Wintrode PL, Zhang D, Vaidehi N, Arnold FH, Goddard WA, Protein dynamics in a family of laboratory evolved thermophilic enzymes, *J Mol Biol*, 327 (2003) 745–757. [PubMed: 12634066]
- [26]. Lazaridis T, Lee I, Karplus M, Dynamics and unfolding pathways of a hyperthermophilic and a mesophilic rubredoxin, *Protein Science*, 6 (1997) 2589–2605. [PubMed: 9416608]
- [27]. Dauber-Osguthorpe P, Osguthorpe DJ, Stern PS, Moult J, Low frequency motion in proteins, *J. Comput. Phys*, 151 (1999) 169–189.
- [28]. Xu J, Plaxco KW, Allen SJ, Collective Dynamics of Lysozyme in Water: Terahertz Absorption Spectroscopy and Comparison with Theory, *The Journal of Physical Chemistry B*, 110 (2006) 24255–24259. [PubMed: 17125398]
- [29]. Ming D, Wall ME, Quantifying allosteric effects in proteins, *Proteins: Structure, Function, and Bioinformatics*, 59 (2005) 697–707.
- [30]. Perutz MF, Raidt H, Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2, *Nature*, 255 (1975) 256–259. [PubMed: 1143325]
- [31]. Bakker P.I.d., Hunenberger PH, McCammon JA, Molecular dynamics simulations of the hyperthermophilic protein sac7d from *Sulfolobus acidocaldarius*: contribution of salt bridges to thermostability, *J Mol Biol*, 285 (1999) 1811–1830. [PubMed: 9917414]
- [32]. Grottesi A, Ceruso M-A, Colosimo A, Nola AD, Molecular dynamics study of a hyperthermophilic and a mesophilic rubredoxin, *Proteins: Structure, Function, and Bioinformatics*, 46 (2002) 287–294.
- [33]. Chan MK, Mukund S, Kletzin A, Adams MW, Rees DC, Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase, *Science*, 267 (1995) 1463–1469. [PubMed: 7878465]
- [34]. Purmonen M, Valjakka J, Takkinen K, Laitinen T, Rouvinen J, Molecular dynamics studies on the thermostability of family 11 xylanases, *Protein Engineering Design and Selection*, 20 (2007) 551–559.
- [35]. Mombelli E, Afshar M, Fusi P, Mariani M, Tortora P, Connelly JP, Lange R, The Role of Phenylalanine 31 in Maintaining the Conformational Stability of Ribonuclease P2 from *Sulfolobus solfataricus* under Extreme Conditions of Temperature and Pressure, *Biochemistry*, 36 (1997) 8733–8742. [PubMed: 9220960]
- [36]. Motono C, Gromiha MM, Kumar S, Thermodynamic and kinetic determinants of *Thermotoga maritima* cold shock protein stability: A structural and dynamic analysis, *Proteins: Structure, Function, and Bioinformatics*, 71 (2008) 655–669.

- [37]. Huang X, Zhou H-X, Similarity and difference in the unfolding of thermophilic and mesophilic cold shock proteins studied by molecular dynamics simulations, *Biophysical journal*, 91 (2006) 2451–2463. [PubMed: 16844745]
- [38]. Farantos SC, Non-Linear Vibrational Normal Modes of Biomolecules, *Computational Science and its Applications*, 2007. ICCSA 2007. International Conference on, 2007, pp. 444–450.
- [39]. Tamm LK, Tatulian SA, Infrared spectroscopy of proteins and peptides in lipid bilayers, *Quarterly reviews of biophysics*, 30 (1997) 365–429. [PubMed: 9634652]
- [40]. Kremer W, Schuler B, Harrieder S, Geyer M, Gronwald W, Welker C, Jaenicke R, Kalbitzer HR, Solution NMR structure of the cold-shock protein from the hyperthermophilic bacterium *Thermotoga maritima*, *European Journal of Biochemistry*, 268 (2001) 2527–2539. [PubMed: 11322871]
- [41]. Mueller U, Perl D, Schmid FX, Heinemann U, Thermal stability and atomic-resolution crystal structure of the *Bacillus caldolyticus* cold shock protein, *J Mol Biol*, 297 (2000) 975–988. [PubMed: 10736231]
- [42]. Max KEA, Zeeb M, Bienert R, Balbach J, Heinemann U, T-rich DNA single strands bind to a preformed site on the bacterial cold shock protein Bs-CspB, *J Mol Biol*, 360 (2006) 702–714. [PubMed: 16780871]
- [43]. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M, CHARMM: The biomolecular simulation program, *Journal of Computational Chemistry*, 30 (2009) 1545–1614. [PubMed: 19444816]
- [44]. Brooks BK, Martin, Harmonic Dynamics of Proteins: Normal Modes and Fluctuations in Bovine Pancreatic Trypsin Inhibitor, *Proc Natl Acad Sci U S A*, 1983, pp. 6571–6575. [PubMed: 6579545]
- [45]. Baudry J, Collective vibrations in the crystalline alanine dipeptide at very low temperature, *J. Mol. Struct. Dyn*, 7 (1997) 4954.
- [46]. Mouawad L, Perahia D, Diagonalization in a mixed basis: A method to compute low-frequency normal modes for large macromolecules, *Biopolymers*, 33 (1993) 599–611.
- [47]. Kitao A, Hirata F, Go N, The effects of solvent on the conformation and the collective motions of protein: Normal mode analysis and molecular dynamics simulations of melting in water and in vacuum, *Chem. Phys*, 158 (1991) 2–3.
- [48]. Tama F, Sanejouand Y-H, Conformational change of proteins arising from normal mode calculations, *Protein Engineering*, 14 (2001) 1–6. [PubMed: 11287673]
- [49]. Prabal K, Maiti A, Pascal TA, Vaidehi N, Goddard WA, The stability of Seeman JX DNA topoisomers of paranemic crossover (PX) molecules as a function of crossover number, *Nucleic Acids Res*, 32 (2004) 60476056.
- [50]. Sagnella DE, Straub JE, A Study of Vibrational Relaxation of B-State Carbon Monoxide in the Heme Pocket of Photolyzed Carboxymyoglobin, *Biophysical Journal*, 77 (1999) 70–84. [PubMed: 10388741]
- [51]. Nadler W, Brnger AT, Schulten K, Karplus M, Molecular and stochastic dynamics of proteins, *Proc Natl Acad Sci U S A*, 84 (1987) 7933–7937. [PubMed: 3479772]
- [52]. Hayward S, Kitao A, N. G, Harmonicity and anharmonicity in protein dynamics: A normal mode analysis and principal component analysis, *Proteins: Structure, Function, and Bioinformatics*, 23 (1995) 177–186.
- [53]. Polyansky AA, Kuzmanic A, Hlevnjak M, Zagrovic B, On the Contribution of Linear Correlations to Quasi-harmonic Conformational Entropy in Proteins, *Journal of Chemical Theory and Computation*, 8 (2012) 3820–3829. [PubMed: 26593023]
- [54]. Xie A, van der Meer AFG, Austin RH, Excited-State Lifetimes of Far-Infrared Collective Modes in Proteins, *Physical Review Letters*, 88 (2001) 018102. [PubMed: 11800992]
- [55]. Strawn R, Melicherik M, Green M, Stockner T, mail JC, Ettrich R, Symmetric Allosteric Mechanism of Hexameric *Escherichia coli* Arginine Repressor Exploits Competition between L-Arginine Ligands and Resident Arginine Residues, *PLoS Comput Biol*, 6 (2010) 6.

- [56]. Batista PR, de Souza Costa MG, Pascutti PG, Bisch PM, de Souza W, High temperatures enhance cooperative motions between CBM and catalytic domains of a thermostable cellulase: mechanism insights from essential dynamics, *Physical Chemistry Chemical Physics*, 13 (2011) 13709–13720. [PubMed: 21713261]
- [57]. Vondrek J, Kuba T, Jenney FE, Adams MWW, Koek M, ern J, Sklen V, Hobza P, Dispersion Interactions Govern the Strong Thermal Stability of a Protein, *Chemistry A European Journal*, 13 (2007) 9022–9027.
- [58]. Razvi A, Scholtz JM, Lessons in stability from thermophilic proteins, *Protein Science*, 15 (2006) 1569–1578. [PubMed: 16815912]
- [59]. Ladenstein R, Heat capacity, configurational entropy, and the role of ionic inter- actions in the role of ionic interactions in protein thermostability, *Biotechnology and Biotechnological Equipment*, (2008).
- [60]. Reuter N, Hinsin K, Lacapre J-J, Transconformations of the SERCA1 Ca-ATPase: a normal mode study, *Biophysical journal*, 85 (2003) 2186–2197. [PubMed: 14507684]
- [61]. Gaillard T, Martin E, San Sebastian E, Cosso FP, Lopez X, Dejaegere A, Stote RH, Comparative normal mode analysis of LFA-1 integrin I-domains, *J Mol Biol*, 374 (2007) 231–249. [PubMed: 17919656]
- [62]. van Vlijmen HWT, Karplus M, Analysis of Calculated Normal Modes of a Set of Native and Partially Unfolded Proteins, *The Journal of Physical Chemistry B*, 103 (1999) 3009–3021.
- [63]. Patriksson A, Marklund E, van der Spoel D, Protein Structures under Electrospray Conditions, *Biochemistry*, 46 (2007) 933–945. [PubMed: 17240977]
- [64]. Marklund EG, Larsson DSD, van der Spoel D, Patriksson A, Coleman C, Structural stability of electrosprayed proteins: temperature and hydration effects, *Physical Chemistry Chemical Physics*, 11 (2009) 8069–8078. [PubMed: 19727514]
- [65]. Tang L, Liu H, A Comparative Molecular Dynamics Study of Thermophilic and Mesophilic Ribonuclease HI Enzymes, *Journal of Biomolecular Structure and Dynamics*, 24 (2007) 379–392. [PubMed: 17206853]

Figure 1.

DSSP sequence and tertiary structure of Csp homologues (top:BsCsp, middle:BcCsp and bottom:TmCsp) The thermophilic and mesophilic proteins were derived from the X-ray results obtained by Heinemann et. al. hyper thermophile (TmCsp), thermophile (BcCsp) and mesophile (BsCsp) served as the starting structure for setting up the simulations. Panel a) for BsCsp, panel b) BcCsp and panel c) TmCsp.

Figure 2.

Illustrates Terahertz FIR absorption spectra of each Csp homologues with the corresponding crystallographic structure in vacuo with respect to two different surface water thicknesses ($2.0\text{\AA}$ and $2.5\text{\AA}$). a) In vacuo, mesophilic BsCsp have increased absorption compared to their thermophilic TmCsp and BcCsp counterparts. Compared to BsCsp, both TmCsp and BcCsp show few additional peaks at low frequency region around $0-100\text{ cm}^{-1}$ and b) clearly display the differences in absorption of protein in vacuo and in solvated system at $2.0\text{\AA}$ and $2.5\text{\AA}$ thickness; an increase in amplitude is noticed for TmCsp and BcCsp in solvated system. c) The density of states started increasing upon addition of water to Csp, which was measured as a function of hydration in the Terahertz (THz) frequency range ($0-100\text{ cm}^{-1}$.)

Figure 3.
Density of states (DOS) for BcCsp at increased hydration shell.

Figure 4.

Illustrates confirmation that Csp homologues of hyperthermophilic TmCsps carry higher heat capacities and congruational entropy compared to mesophilic Csps of BsCsps for temperature enhancement as well as partially hydrated and anhydrated structures.

Figure 5.

The normalized atomic displacements for modes numbered as 7, 8, 10, and 15 are represented in as a function of the amino acid sequence for vacuo, primary ($2.5\text{\AA}$) and secondary ($3.5\text{\AA}$) solvation shells respectively. The highest peaks in the plots correspond to the most displaced residues in the corresponding mode.

Figure 6.

The amino acid sequences of the three cold shock protein in terms of globular motions are compared. BcCspB differs from BsCspB at 12 positions (including a deletion at the C terminus), whereas TmCspB differs from BsCspB at 24 positions (including three deletions, one insertion and one additional residue at the C terminus). 42 out of 67 residues are conserved in all three proteins. (a) Despite huge RMSD differences observed in the molecular structure at $T = 600$ K, all proteins do not undergo significant unfolding process, as evidenced by the radius of gyration. (b) Additional measures to quantify protein conformation were established by the concepts of solvent accessible surface area (SASA) and (c) solvent excluding surface area (SESA) since the balance between the forces dominating protein and protein solvent interactions is crucial for establishing the secondary and tertiary structures of proteins. (d) Number of hydrogen bonds as a function of temperature.

Figure 7.

Show the difference in THz absorbance profile between native conformation and thermally induced denatured meso, thermo and hyperthermo profile at 0 – 500 cm^{-1} and the impact of low frequency globular modes related to individual contributions of thermostability in each homologues (e-f). The differences found in the temperature shifts of low frequency modes are moderately less in BcCspS and TmCspS and (g-h) show that anharmonicity of motions induces the transitions of different conformational substrates of the protein and functional behavior from low to high temperatures.

Table 1.

Root Mean Square Deviation (RMSD) of ‘N’ and ‘C’ terminal regions. Number of $H_2O < 300\text{\AA}$ from surface of the protein, protein temperature, averaged over last 1000 ps for both at Primary and Secondary evaporation process.

Temp (K)	TmCsp	BcCsp	BsCsp	TmCsp	BcCsp	BsCsp
Primary evaporation			Secondary evaporation			
RMSD N- terminal						
CPT	0.56	0.65	0.41	0.56	0.65	0.41
300	0.50	0.47	0.50	0.46	0.55	0.41
310	0.57	0.25	0.68	0.59	0.46	0.66
340	0.51	0.28	0.34	0.30	0.95	0.27
350	0.72	0.73	0.54	0.52	0.93	0.60
400	0.78	0.81	0.76	0.87	0.82	0.93
600	0.97	0.95	0.96	0.97	0.97	0.97
RMSD C- terminal						
CPT	0.52	0.60	0.53	0.52	0.60	0.53
300	0.63	0.62	0.74	0.80	0.38	0.41
310	0.50	0.56	0.47	0.53	0.52	0.85
340	0.39	0.27	0.86	0.46	0.87	0.63
350	0.86	0.77	0.86	0.88	0.95	0.70
400	0.81	0.88	0.78	0.90	0.88	0.78
600	0.97	0.47	0.98	0.98	0.97	0.98
< No. H ₂ O > $T_{protein}^{(K)}$						
300	286/301	255/298	259/297	274/298	240/296	254/292
310	277/310	246/307	249/307	271/303	235/302	244/301
340	269/330	242/330	252/328	260/319	233/321	242/315
350	269/337	232/337	234/334	264/324	228/318	226/318
400	236/359	209/360	221/355	217/331	187/329	211/324
600	58/419	48/429	60/416	31/381	26/395	38/377