PLOS ONE

RESEARCH ARTICLE

Structural and thermodynamic analysis of factors governing the stability and thermal folding/unfolding of SazCA

Shashi Kumar, Parag A. Deshpande*

Quantum and Molecular Engineering Laboratory, Department of Chemical Engineering, Indian Institute of Technology Kharagpur, Kharagpur, India

* parag@che.iitkgp.ac.in

OPEN ACCESS

Citation: Kumar S, Deshpande PA (2021) Structural and thermodynamic analysis of factors governing the stability and thermal folding/unfolding of SazCA. PLoS ONE 16(4): e0249866. https://doi.org/10.1371/journal.pone.0249866

Editor: Jie Zheng, University of Akron, UNITED STATES

Received: February 4, 2021

Accepted: March 19, 2021

Published: April 15, 2021

Data Availability Statement: All relevant data are within the paper and its Supporting information files.

Funding: This work was supported by the Department of Biotechnology, Ministry of Science and Technology, Government of India (BT/PR7054/BID/7422/2012) awarded to PAD. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Molecular basis of protein stability at different temperatures is a fundamental problem in protein science that is substantially far from being accurately and quantitatively solved as it requires an explicit knowledge of the temperature dependence of folding free energy of amino acid residues. In the present study, we attempted to gain insights into the thermodynamic stability of SazCA and its implications on protein folding/unfolding. We report molecular dynamics simulations of water solvated SazCA in a temperature range of 293-393 K to study the relationship between the thermostability and flexibility. Our structural analysis shows that the protein maintains the highest structural stability at 353 K and the protein conformations are highly flexible at temperatures above 353 K. Larger exposure of hydrophobic surface residues to the solvent medium for conformations beyond 353 K were identified from H-bond analysis. Higher number of secondary structure contents exhibited by SazCA at 353 K corroborated the conformations at 353 K to exhibit the highest thermal stability. The analysis of thermodynamics of protein stability revealed that the conformations that denature at higher melting temperatures tend to have greater maximum thermal stability. Our analysis shows that 353 K conformations have the highest melting temperature, which was found to be close to the experimental optimum temperature. The enhanced protein stability at 353 K due the least value of heat capacity at unfolding suggested an increase in folding. Comparative Gibbs free energy analysis and funnel shaped energy landscape confirmed a transition in folding/unfolding pathway of SazCA at 353 K.

1 Introduction

Carbonic anhydrases (CAs) constitute a family of Zn-containing metalloenzymes which catalyze the reversible hydration of CO_2 to bicarbonates and protons [1–4], as shown in Eq (1).

$$CO_2 + H_2O \rightleftharpoons HCO_3^- + H^+ \quad (1)$$

The enzymes of this family are known to exhibit one of the fastest reaction rates known so far in nature [1]. CAs are ubiquitously found in bacteria, algae, plants as well as in eukaryotic

animals. These are enconded by six evolutionarily distinct gene families *viz*., α, β, γ, δ, ζ, and η [5, 6]. These different classes of CAs have different structural folds and share low sequence homology [7, 8]. α-CAs are the most studied CAs, in which Zn^{2+} ions are present as centres in their active sites, coordinated to three histidines residues [9–11]. In humans, CAs are vital for physiological processes such as regulation of blood pH, regulation of pressure of retinal fluids, and nourishment during bone growth [12]. Recently, CAs have been reported as potential biocatalysts for industrial applications in carbon capture, storage and sequestration (CCUS) [13]. These enzymes are also of interest as biocatalytic components in biomedical devices such as biosensors and artificial lungs [14]. Unfortunately, the use of CAs in these applications is severely limited by their poor thermal stability [15]. More robust and stable CAs are expected to benefit these applications. Thus, the employment of enzymes extracted from microorganisms living at high temperature range of 323-363 K (*i.e. thermophiles*) can be effectively and efficiently used to overcome these limitations [13]. Recently, several investigators have focused on the development of thermostable CAs to achieve a synergism of their thermostability and high enzymatic activity [16]. In the past, the major focus of the scientific community was dedicated to human CA II (hCAII) isoform to understand its structural and functional features [17]. In this study, we have focussed our attention on another α-CA of same CA family, namely SazCA, because of its potential in high temperature biotechnological CCUS applications. We describe the rationale behind the selection of SazCA in the text to follow.

SazCA is a thermostable CA extracted from *Sulfurihydrogenibium azorense*, a thermophilic bacterium found in hot springs of Azores. It is a dimeric protein with Zn^{2+} active center, as shown in Fig 1. It is 2.3 times more efficient than hCAII for catalyzing the reversible hydration of CO_2 [18, 19]. The kinetic parameters ($k_{cat} = 4.40\times10^6$ s^{-1}; $k_{cat}/K_M = 3.5\times10^8$ M^{-1} s^{-1} at 293 K and pH 7.5) for the reaction catalyzed by SazCA show this enzyme to be the most active one among all CAs to date. The thermostability and activity studies have shown this enzyme to be active in a temperature range of 273 to 373 K, with an optimal working temperature for the catalytic activity of about 353 K [19]. SaCA has evolved as a promising bio-catalyst for CCUS because of its exceptional thermostability and owing to the fact that it is the most active α-CA known till date for CO_2 hydration [20]. However, limited information is available on its biophysical behaviour. The folding/unfolding pathways of SazCA are not well explored. Comprehensive data need to be produced on the structure and dynamics of folded and unfolded intermediates of this protein so that a thorough understanding of folding/unfolding pathways of SazCA can be made to provide insights into this class of enzymes which can set a basis for mutations in hCAII or other class of enzymes to improve their stability and/or activity. It is also likely to improve our understanding of protein aggregation and misfolding in general. With this in consideration, we report a molecular dynamics (MD) study of thermal denaturation of SazCA in the present study. We have investigated the effect of temperature on the folding/unfolding pathways of SazCA. We report MD simulations in a range of 293-393 K to unravel the relationship of conformational stability and flexibility in this protein.

Several investigators have studied folding/unfolding characterstics of hCAII under moderate denaturation conditions in the past [21, 22]. Also, studies on unfolding pathways of bovine CA and CAIX in association with conformational stability have been reported [23]. Such details on thermostable SazCA are missing and we provide them by quantifying the secondary structure descriptors, H-bonding descriptors, root mean square deviation (RMSD), root mean square fluctuations (RMSD), radius of gyration (R_G), percentage of secondary structure contents, solvent accessible surface area (SASA) and pair correlation function (RDF) in this study. Thermodynamic analysis can aid the understanding of reasons behind protein stability and thermal denaturation [24–26]. Hence, thermodynamic analysis of protein stability and pathways of folding/unfolding was done to assess the thermal stability of SazCA. This

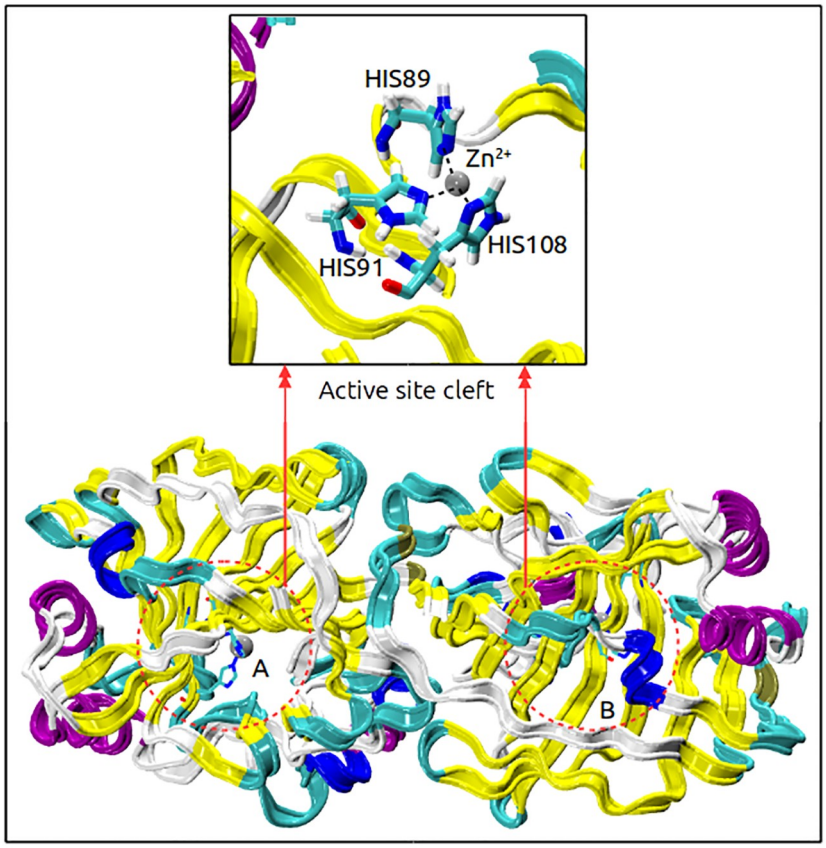

Fig 1. Dimeric structure of thermostable SazCA. Structural representation of dimeric SazCA analysed in this study. The dimer is shown with two chains A, B in Ribbon representation. Zn^{2+} ions are shown as grey spheres and histidine residues involved with direct interactions (active site cleft) with Zn^{2+} are shown in Licorice representation. α-helices are shown with violet color, 3_{10}-helices with blue color, β-sheets with yellow color, loops (turns and coils) are shown in white and cyan respectively.

https://doi.org/10.1371/journal.pone.0249866.g001

comprehensive study on the thermal denaturation of SazCA is expected to provide a molecular basis for folded and unfolded ensembles of SazCA, leading to the understanding of structure-function relationships in this protein.

2 Results and discussion

2.1 Structural analysis

Hydrogen bonding governs the protein secondary structure, and is responsible for its structural stability. There are two types of H-bonds in a protein *viz*., backbone-backbone (B-B) H-bonds and backbone-water (B-W) H-bonds. In this study, we determined both type of H-bonds to probe protein rigidity, thermostablity and unfolding of SazCA, as detailed in the text to follow.

2.1.1 Backbone-backbone H-bonds. B-B H-bonds are responsible for preserving protein rigidity which in turn helps in improving its thermostability. With an increase in temperature, proteins become unstable due to a decrease in the strength and number of H-bonds. The lesser the number of H-bonds, the higher the flexibility of the protein [27]. The average numbers of B-B H-bonds in well-equilibrated ensembles of the thermophilic SazCA obtained in this study were 381, 368, 372, 366, 363 and 346 at 293, 313, 333, 353, 373 and 393 K, respectively, which

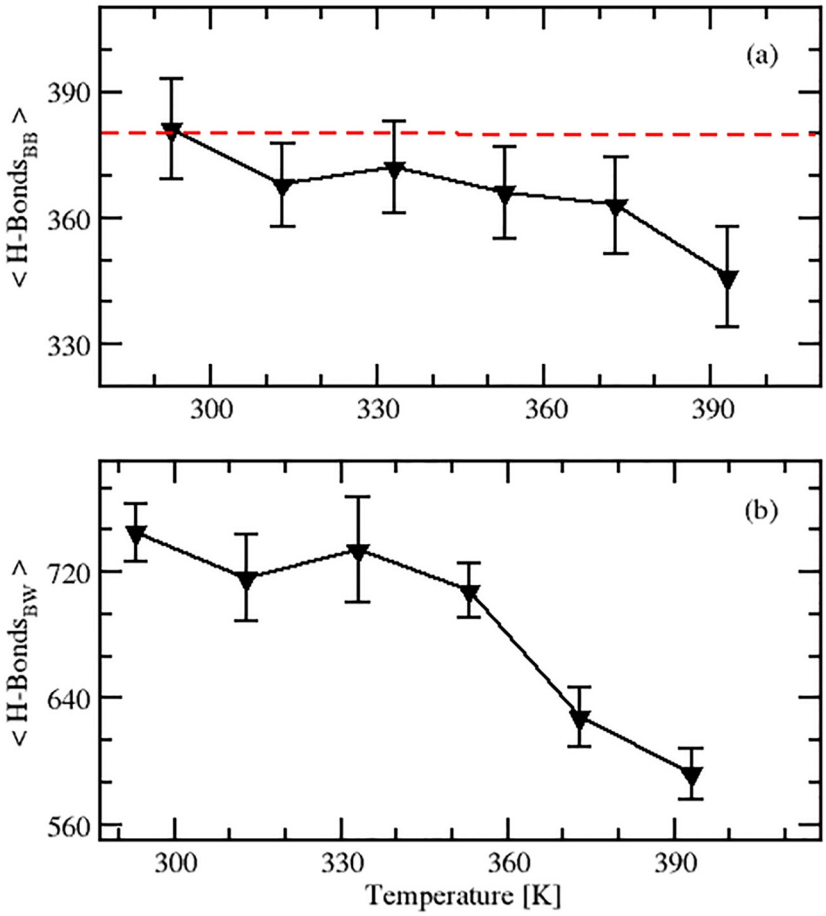

Fig 2. Average number of backbone-backbone and backbone-water H-bonds in SazCA. Analysis of effect of temperature on the number of backbone-backbone and backbone-water H-bonds in SazCA. The variation of average number of backbone-backbone H-bonds with temperature is shown in (a) and the variation of number of backbone-water H-bonds with temperature is shown in (b). The dashed red line in plot (a) represents the number of H-bonds present in the protein crystal structure.

https://doi.org/10.1371/journal.pone.0249866.g002

indicated a non-monotonous change in the number of B-B H-bonds in SazCA with an increase in temperature. The evolution of the average number of B-B H-bonds as a function of temperature is shown in Fig 2(a). The protein crystal structure had 380 H-bonds, as has been indicated by a red dashed line in Fig 2(a). The number of H-bonds initially decreased with an increase in temperature from 293 to 313 K. This was primarily due to the loss of native conformations in the protein. However, after 313 K, the number of H-bonds increased upto a temperature of 333 K. This increase in H-bonds can be attributed to the high dynamism in the protein which led to the formation of some new bonds at elevated temperatures. Such states were not accessible at low temperatures. A rise in temperature increased the thermal energy of the molecules. Due to the increased thermal energy, the side chains of the protein came closer resulting in interactions leading to the formation of H-bonds. This trend of non-monotonous change in the number of H-bonds with a change in temperature has also been reported by several researchers previously [27–30] for other proteins. Between 333 and 353 K, there was a little change in the average number of H-bonds in the protein. But a sharp decrease in the number of H-bonds was observed between 353 to 393 K. This was due to the fact that the number of B-B H-bonds which were present initially which maintained the conformational stability as

well as the thermostability decreased subsequently with an increase in temperature. From Fig 2(a), it was evident that the increase in temperature beyond 353 K induced unfolding in the protein by causing a significant loss in B-I H-bonds which destabilised the native conformations of the protein. The protein maintained its native conformations upto 353 K (optimal working temperature of SazCA). A similar observation of significant loss of B-B H-bonds and a larger exposure of hydrophobic surface residues to the solvent beyond 353 K has been reported to be responsible for unfolding of human isoforms of CA also [23].

2.1.2 Backbone-water H-bonds. B-W H-bonds are responsible for supporting protein-solvent interactions [27]. The interaction of the protein core with the solvent improves when the hydrophobic region of the protein starts unwrapping from its native (wild) folded state to the unfolded state. Larger exposure of hydrophobic surface residues enhances the affinity of protein core to the solvent molecules. The average numbers of B-W H-bonds in SazCA were 745, 716, 734, 708, 628, 592 at 293, 313, 333, 353, 373 and 393 K, respectively. This was similar to the trend for B-B H-bonds. The B-W H-bonds for SazCA first decreased from 293 to 313 K and thereafter increased upto 333 K. The reason behind this behaviour has been explained in the text above. With an increase in temperature from 333 to 353 K, number of H-bonds descended from 734 to 708. With a rise in temperature beyond 353 K, there was a sharp decline in the number of H-bonds as shown in the Fig 2(b). At high temperatures, when the protein unwrapped from its native folded state, the number of B-W H-bonds started decreasing rapidly, as has also been observed by others [27]. The number of B-W H-bonds in the thermophilic SazCA changed only slightly in a temperature range of 293 to 353 K as can be seen in the Fig 2(b) while in the temperature range of 353 to 393 K, the number of native conformations altered significantly due to a remarkable decreased in the number of H-bonds. This indicated that protein was stable upto 353 K and after that unfolding initiated.

2.1.3 Secondary structure assignments. Change in the secondary structure of a protein is an important indicator of the folding/unfolding process. To explore folding/unfolding of SazCA, we calculated the percentage amount of each secondary structure in SazCA as a function of temperature. The comparative secondary structure analysis of SazCA at different temperatures is shown in Table 1, where B, E, T, G, I, H, C represent the percentage of isolated β-sheets, extended β-sheets, turns, 3_{10}-helices, π-helices, α-helices and coils, respectively. It can be observed that between 293 and 353 K, there was a significant increase in the percentage amounts of α-helices (including 3_{10}-helices) and β-sheets (including isolated β-sheets). The corresponding decrease in the amount of turns and coils can also be observed. At 293 K, the percentage amount of helices (α-helices and 3_{10}-helices) was 6.5%. With a rise in temperature, an increase in the percentage amount of helices was observed, and these were 9.9% at 353 K. Similarly, the percentage amount of β-sheets (extended β-sheets, isolated β-sheets) was 35.5%

Table 1. Percentage contribution of secondary structure contents in SazCA as a function of temperature obtained from MD simulations, where B, E, T, G, I, H, C represent the amounts of isolated β-sheets, extended conformations (β-sheets), turns, 3_{10}-helices, π-helices, α-helices and coils, respectively. The secondary structure assignment data of crystal structure (experimental) of SazCA have also been shown for comparison.

Temperature (K)	B (%)	E (%)	T (%)	G (%)	I (%)	H (%)	C(%)
293	0.7	34.8	30.1	1.3	0	5.2	27.9
313	0.9	35.3	29.4	2.3	0	5.4	26.5
333	1.0	35.7	28.7	2.6	0	5.5	26.5
353	**1.3**	**35.9**	**27.2**	**3.9**	**0**	**6.0**	**25.6**
373	0.9	34.6	29.4	2.8	0	4.4	27.9
393	0.7	33.9	30.3	1.3	0	4.0	29.8
Experimental	**1.3**	**35.9**	**28.3**	**2.9**	**0**	**5.4**	**26.1**

https://doi.org/10.1371/journal.pone.0249866.t001

at 293 K and 37.2% at 353 K. On comparing the helices and sheets, it was observed that there was a small (approximately 2%) increase in the α-helices and β-pleated sheets from 293 to 353 K. The increase in the percentage amount of helices and sheets suggested higher structural rigidity and a decrease in the flexibility of the protein. A possible reason behind an increase in percentage of helices and sheets was that with an increase in temperature, protein became more flexible gaining a higher access to H-bonds. The secondary structures of the protein, however, lost their integrity with a further increase in temperature. With an increase in temperature, due to the high rate of dynamism in the protein, formation of new interactions took place leading to the formation of stronger bonds [29, 30] upto moderate temperatures. The highest percentage amount of helices and sheets were found at 353 K. The percentage contributions from α-helices and β-pleated sheets had the highest impact on the structure of the protein at 353 K resulting in retaining the maximum amount of secondary structure assignments. This contribution was not found for rest of the conformations at 373 and 393 K, which led to significant alterations in the secondary structures at these temperatures. This decrease in the percentage amount of helices and sheets can be observed in the Table 1. An increase in the formation of secondary structure assignments induced folding and subsequently unfolding with a decrease in helical and sheets contents. Experimental secondary structure assignments data obtained from the crystal structure of SazCA have also been presented in Table 1 for comparison. 353 K conformation was found to be in good agreement with available experimental data. The highest percentage of secondary structure contents found at 353 K configuration suggested that the structure was in its stable folded state and unfolding started after this temperature.

From the comparison of all sampled configurations with secondary structure data of the experimental conformation of SazCA, it was found that 353 K configuration was associated with the maximum number of native conformations. Our computational findings are in a close agreement with H-bond analysis discussed above in this study as well as with the experimental findings. Therefore, it can be deduced that the structural conformation at 353 K supported a compact structure with the highest thermostability.

Ramachandran plot is an important means to check the quality of the computed 3-D structure of a protein. Ramachandran plot analysis was carried out for well equilibrated conformations of SazCA at different temperatures to analyze the protein structures with energetically allowed conformations. 2-D Ramachandran plots of various conformations of SazCA at different temperatures are shown in Fig 3. The blue regions indicate the most populated and favoured conformations, the green regions represent the allowed conformations, while the white regions correspond to disallowed conformations. Quadrant I shows a region where some conformations are allowed, corresponding to left-handed α-helices. Quadrant II shows the largest region in the Ramachandran plot where the most favoured conformations of proteins lie and it corresponds to the β-strands. The III quadrant corresponds to the right-handed α-helices and it is the second largest region in the Ramachandran plot. The IV quadrant has no specific outlined region and it contains the most disallowed conformations. From Fig 3, it is clear that there was a noticeable difference in the amount of residues present in the most favoured and allowed regions. On comparing, it was found that most of the residues corresponding to 353 K configuration were located in the most favoured region as shown in Fig 3(d). Also, less number of residues were present in the disallowed region at this temperature. A possible reason behind this is lesser steric hindrance among the residues for SazCA at 353 K. We quantified the trends shown in the Ramachandran plot of Fig 3 to get a clearer idea of the changes in the population of the conformations in favoured and disallowed regions as a function of temperature. The quantification has been given in Table 2. It can be seen that the percentage of favoured regions go through a maxima with the highest percentage of allowed

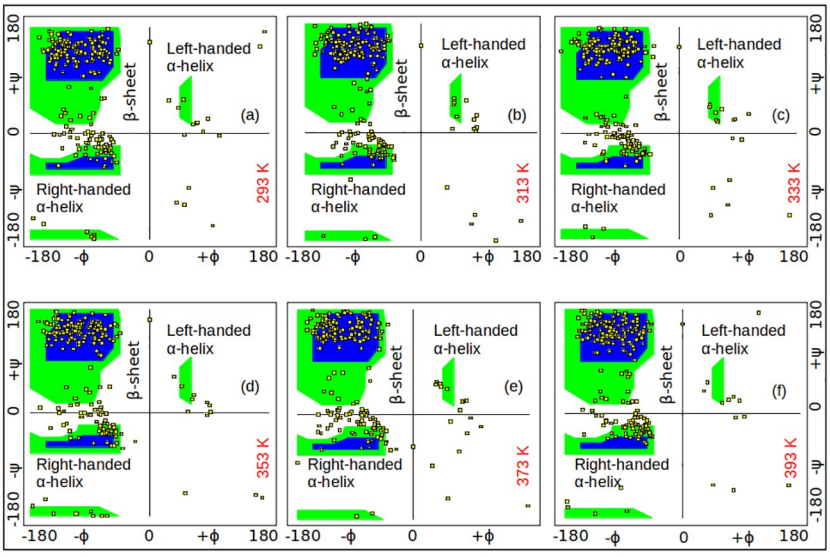

Fig 3. Ramachandran plot analysis. Ramachandran plot of simulated conformational ensembles of SazCA at different temperatures. The blue regions, green regions and white regions represent favoured, allowed and disallowed conformations, respectively. The square symbols represent the simulation data.

https://doi.org/10.1371/journal.pone.0249866.g003

regions observed at 353 K. A close correspondance was observed for the population of the conformations in disallowed regions with the minimum population of disallowed conformations at 353 K. This supported our observation that the protein exhibited the most stable geometry at 353 K. The findings of Ramachandran plots were also in a good agreement with the quantitative analysis of secondary structure assignments where maximum number of helices and sheets were observed for this conformation. Therefore, it can be concluded that the system geometry at 353 K was the most stable.

2.1.4 Analysis of the solvation shell. Proteins perform their functions in an aqueous environment [31]. In general, water around proteins can be divided into three categories: (1) the bulk water that surrounds the protein, (2) the bound water that forms H-bonds, and (3) the hydration water at the protein surface which has direct interactions with the protein [32–35]. Protein-bound water (*i.e.* bound water and hydration water) form strong interactions with proteins. Protein-bound water molecules are essential to the structure and functions of the protein and they also affect the local protein structure as they contribute to several properties of the protein such as protein stability, protein folding, drug docking and oligomeric formation [36]. Therefore, in order to describe the structure-dynamics of solvent layer around the protein surface, radial distribution function (RDF) of the protein backbone C_α atoms with

Table 2. Percentage occurence of residues in favoured and disallowed regions of SazCA, as obtained from the Ramachandran plot at various investigated temperatures.

Temperature (K)	Favoured region	Disallowed region
293	92.13	2.02
313	92.13	1.37
333	92.36	1.34
353	94.60	1.31
373	91.68	1.62
393	91.68	1.79

https://doi.org/10.1371/journal.pone.0249866.t002

respect to the water oxygen was determined. The dynamics of water molecules at the protein interface in terms of probability distribution and water mobility for the simulated frames were analyzed using the following equation [37]:

$$g(r) = \frac{V}{4\pi N_f N(N-1) r^2 \delta r} \sum_{n=frames} \sum_{i} \sum_{j \neq i} \delta(r - r_{ij}) \quad (2)$$

where $N(N\text{-}1)/V$ = number of atoms pairs per unit volume, $4\pi r^2 \delta r$ = volume sampled at distance r.

We calculated the RDFs for SazCA at the aforementioned temperatures *viz*., 293, 313, 333, 353, 373 and 393 K. The distribution of the solvent layers around the protein surface is indicated by the RDF shown in Fig 4. The calculated RDFs showed that there were two maxima representing two hydration shells around protein surface, at a radial distances of 3.85 Å and 4.95 Å, respectively. The first hydration shell consisted of solvent molecules interacting with the charged or polar species of the protein whereas the second shell indicated the presence of van der Waals interactions between the protein and water molecules [38]. A comparison of the RDFs at investigated temperatures showed that there were negligible differences in the first peak positions but significant differences in their amplitudes. The highest intensity peak in the first solvation shell can be observed for SazCA at 353 K compared to those at the rest of the temperatures. The reason behind this high intensity peak at 353 K was the high local structuring of water molecules around the protein surface. These local ordering of water molecules in the vicinity of protein surface was due to the steric hindrance and specific interactions between the protein and water molecules [38]. It also signified the high ordering of solvent molecules with protein backbones atoms through strong hydrophilic interactions. Also, on comparing the second solvation shell, a sharp and high intensity peak was observed at a temperature of 353 K but relatively at a larger distance compared to the peak in the first solvation shell. It suggested the existance of ordered water molecules near the protein surface. After the second

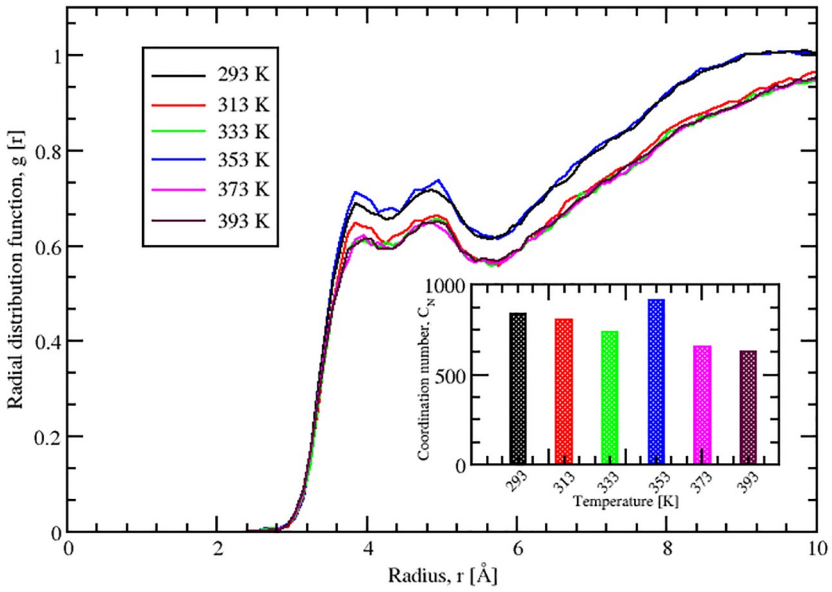

Fig 4. Radial distribution function analysis. Radial distribution function of the mass center of solvent molecules with respect to the protein surface at different temperatures. The variation of the coordination number of water molecules around the protein in the first solvation sphere with temperature is shown in the inset.

https://doi.org/10.1371/journal.pone.0249866.g004

solvation shell, a further increase in $g(r)$ extended the peaks close to 1 at $g(r) = 10$, indicating the hydration water to have reached to the bulk properties. The RDF of SazCA at 353 K was sharper indicating that protein bound water was more ordered and less mobile than the bulk water. Therefore, it can be concluded that at 353 K, SazCA had a better and stable representative structural ensemble elucidating the structural stability, folding and dynamical properties compared to structural ensembles at other temperatures. The variation of coordination number of water molecules around the protein in the first solvation sphere with temperature was quantified. The variation is shown in the inset of Fig 4. It can be seen that the highest coordination number was observed at 353 K showing the maximum ordering of the water molecules around the protein resulting in its stabilisation at 353 K.

2.1.5 Further structural details. In order to develop further insights into the conformational stability and flexibility in SazCA, we did a set of analyses such as RMSD, RMSF, R_G and SASA of protein to figure out the deviations of the protein from the crystal structure. Simulations at elevated temperatures showed that SazCA was stable in a temperature range of 293 to 353 K, indicated by the low RMSD observed between the simulated conformations and initial crystal structure. On the other hand, SazCA displayed a sharp increase in RMSD from 353 to 393 K, suggesting the onset of an increase in fluctuations and denaturation of this enzymes at high temperatures [27]. The average values of RMSD obtained at different configurations as a function of temperature are shown in the S1 Table in S1 File. RMSF analysis also revealed fluctuations of lower magnitude in a temperature range of 293 to 353 K. It suggested lesser structural flexibility at this temperature range. Beyond this temperature, there was a significant decrease in the structural rigidity and an increase in the flexibility. The associated residues responsible for higher RMSF at elevated temperatures were VAL98, ASN99, GLY100, LYS101, GLU145 and HIS207. [39] These highly flexible amino acid residues and their variation in length with simulation temperature are shown in S2 Table in S1 File. These residues were identified for the enhanced flexibility and unfolding of conformational ensembles of SazCA at higher temperatures. Further analysis of R_G indicated a gradual increase in the magnitude of R_G form 293 to 353 K, suggesting partial expansion in the protein structure. A large increase in the value of R_G at 373 and 393 K, suggested larger expansion in the protein structure. The breaking of the native contacts [23] at higher temperature can be a possible reason behind the expansion of the protein. This can be correlated with the analysis of secondary structure assignments described above in this study. The average R_G of SazCA at different temperatures are shown in S1 Table in S1 File. The complete analysis of MD simulations involving the structural parameters (RMSD, RMSF and R_G) with an increase in temperature can be find out in our earlier investigations on this enzyme [39].

To further develop an understanding of the structural stability, we performed SASA analysis at different investigated temperatures. SASA is considered to be a crucial parameter in the determination of stability of proteins [40]. We examined the SASA of all the sampled conformations and the variation of SASA of each sampled conformations as a function of simulation temperature is shown in S3 Fig in S1 File. From the analysis of S3 Fig in S1 File, it is clear that there were significant changes in the magnitudes of SASA obtained at different temperatures. The average SASA obtained for 293, 313, 333, 353, 373 and 393 K configurations were $1184 \times 10^2 \pm 6.45$ Å^2, $1147 \times 10^2 \pm 3.26$ Å^2, $1129 \times 10^2 \pm 3.11$ Å^2, $1104 \times 10^2 \pm 2.72$ Å^2, $1154 \times 10^2 \pm 3.28$ Å^2 and $1189 \times 10^2 \pm 5.14$ Å^2, respectively. From the comparison of average SASA, we found that least SASA was associated with 353 K configuration. Our reported results of SASA were consistent with the analysis of secondary structure assignments, hydrogen bonding pattern and radius of gyration. Lesser SASA obtained for 353 K configuration inferred less structural disruption during the simulations runs. It can also be deduced that the amino acid residues were intact inside the protein core and less surface area was exposed to solvent

medium. A possible reason was higher folding [41] for this conformation as we have discussed in the analysis of unfolding pathways. From SASA analysis of 353 K configuration, it again appeared to be the best conformation for the structural stability due to the high folding and less initial drift from the native structure.

2.2 Thermodynamic analysis of unfolding pathways

Various analyses detailed till now all point out to the fact that the protein exhibited the highest stability at 353 K and hence, thermal denaturation of the protein is expected afterwards with an increase in temperature. We carried out the thermodynamic analysis of the unfolding process, described in the text to follow. We firstly analysed the pathway followed by the protein during its unfolding. Structural snapshots of all the sampled conformations of SazCA throughout the MD trajectories at 0, 20, 40, 60, 80 and 100 ns have been presented in Fig 5. There were no significant changes in α-helix contents at 293 K but a significant increase in the β-sheets was observed between 60 to 100 ns. An increase in the amount of α-helices and β-sheets were observed at 313 and 333 K. Also the increase in size of α-helices and β-sheets can be observed between 20 to 100 ns. At 353 K, there was a significant increase in the secondary structure contents compared to its initial conformation, especially in 3_{10}-helices. Also, an increase in the sizes of the β-pleated sheets and α-helices were seen between 20 to 100 ns in 353 K configuration. A subsequent decrease in the amount of turns and coils can also be observed. The maximum amount of secondary structures found at 353 K conformation suggested lesser flexibility and higher thermal stability of the protein structure at this temperature. The other reason associated with the higher thermostability and lesser flexibility of 353 K configuration was strong salt bridge networks and interactions between the residues, as described elsewhere [39]. These strong salt bridges networks were responsible for increased amount of helices and sheets in SazCA at 353 K. After 353 K, there were structural distortions at 373 and 393 K. Significant losses in α-helices and β-sheets were observed at both the temperatures. The protein was intact at 353 K but unfolding prevailed after 353 K, which resulted into lesser contents of helices and sheets and more number of turns and coils at 373 and 393 K. This phenomena of unfolding can be seen in Fig 5. High flexibility and compromised thermostability at elevated temperatures (373 and 393 K) were due to five unstable salt bridges (ASP113-LYS81, ASP115-LYS81, ASP115-LYS114, GLU144-LYS143 and GLU144-LYS206) [39]. Our findings of unfolding pathways were in agreement with the quantitative analysis of secondary structures analysis reported above in this study. Hence, it can be inferred that the structural conformations at 353 K were the most thermostable conformations with their higher resemblance to the native state and the association of higher contents of secondary structure assignments.

There exist close relationships among the protein stability, structure, function, and the associated energetics. Proteins undergo conformational changes during their functional performance, resulting in changes in properties which are governed by the energetics. The contributions of stability interactions to the protein structure, which are defined by the structure, are known to be temperature dependent, where some interactions may be more dominant than others at high temperatures and less dominant than others at low temperatures or *vice versa* [42–45]. The energetics not only determines the stability of the structure but also identifies the domains of protein which undergo large conformational changes with increasing temperatures when compared to the native state of the protein. Therefore, in order to get a better understanding of the thermal denaturation of SazCA, we carried out a macroscopic description of the protein stability in terms of thermodynamic parameters, as described in detail in the sections to follow.

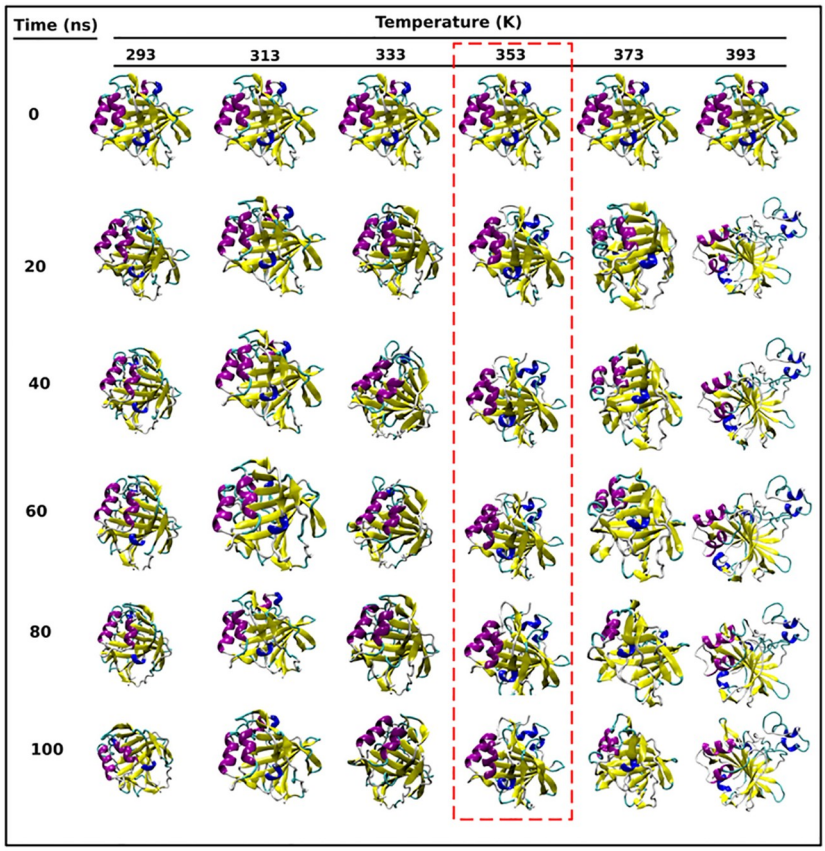

Fig 5. Unfolding pathways of SazCA. The structural snapshots taken at 0, 20, 40, 60, 80 and 100 ns of different configurations of SazCA at 293, 313, 333, 353, 373 and 393 K, describing the unfolding pathways with respect to its native state. α-helices are shown with violet color, 3_{10}-helices with blue color, β-sheets with yellow color, loops (turns and coils), are shown in white and cyan color. The native-like structure found at 353 K is represented with red color dashed rectangle.

https://doi.org/10.1371/journal.pone.0249866.g005

2.2.1 Prediction of folding free energy change. Thermodynamic stability of a protein at a given temperature is governed by the folding free energy change ΔG(T) at that temperature and the thermal stability via the melting (denaturation) temperature T_m. Knowledge of ΔG is important because all the other thermodynamic parameters that characterize the protein folding transition can be extracted from it. The protein folding transitions (stability curves) can be described by the Gibbs-Helmholtz equation [46, 47]:

$$\Delta G(T) = \Delta H_m(1 - \frac{T}{T_m}) - \Delta C_p(T_m - T + T\ln[\frac{T}{T_m}]) \quad (3)$$

where T_m is the melting temperature, which is the midpoint of the thermal denaturation, ΔH_m is the enthalpy of unfolding measured at T_m, and ΔC_P is the heat capacity change measured at unfolding. We determined the free energy change for all the conformational ensembles of SazCA in a temperature range of 293-393 K. The individual stability curves of different conformational ensembles of SazCA illustrating the temperature dependence of change in free energy, ΔG, are shown in Fig 6. By convention, ΔG(T), ΔH_m and ΔC_P have been defined as the difference between folded (native) and the unfolded (denatured) state and hence, the nature of resulting stability curve is concave up. As the temperature deviates from the maximal stability

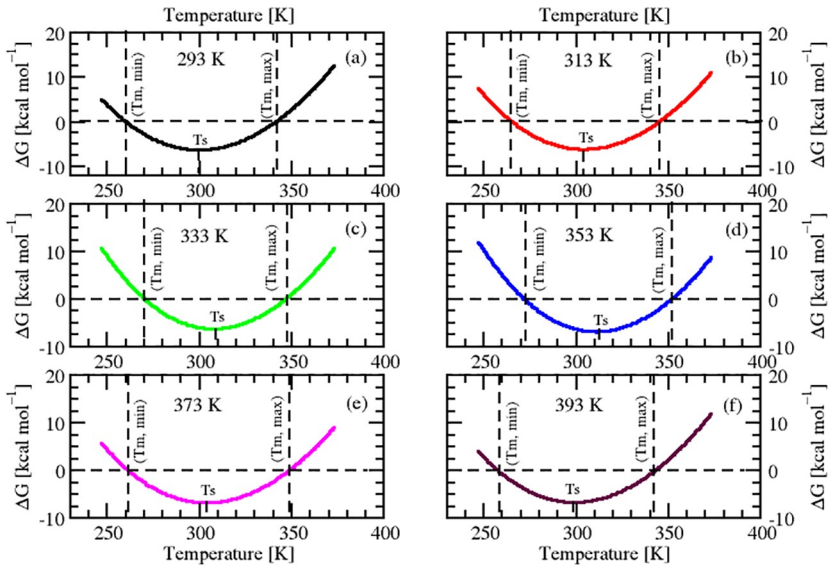

Fig 6. Stability curves of SazCA. Stability curves of different conformational ensembles of SazCA illustrating the temperature dependence of the folding free energy ΔG. Temperatures of cold unfolding $T_{m,min}$, hot unfolding $T_{m,max}$ and temperature of maximum stability T_s have been indicated in the plots.

https://doi.org/10.1371/journal.pone.0249866.g006

temperature, the protein can unfold either due hot denaturation or cold denaturation. The temperature corresponding to the maximum stability has been marked by T_S in Fig 6. The temperatures at the two transitions *i.e.*, hot and cold unfolding have been labelled as $T_{m,max}$ and $T_{m,min}$ respectively. From Fig 6, it can be observed that there was temperature T_S at which the protein conformation was in most stable form. As the temperature increased or decreased, ΔG also changed to favour denaturation and protein conformations became more unstable.

From the comparison of the individual stability curves at different temperatures in Fig 6, it was found that the temperature of maximum stability, T_S, initially underwent a shift towards the high temperature, and then retracted back to lower temperatures with an increase in temperature. For conformational ensembles from temperatures 293 to 353 K, there was a right shift in maximum stability towards high temperature region and after 353 K upto 393 K, it showed a decrease. The maximum shift towards high temperature and maximum stability temperature was found for conformational ensemble at 353 K. This change in maximum stability can be easily visualized in Fig 7 where the stability curves of all the sampled configurations have been superimposed. Also, it was found that there was a global decrease in free energy change as the temperature increased from 293 to 353 K and the least value was observed at 353 K.

The funnel shaped energy landscape of folding/unfolding pathways of SazCA is shown in the Fig 8. It can be observed from the Fig 8 how SazCA approached into its most stable folded (native-like) structure by minimizing the free energy. Unfolding prevailed as temperature rose beyond 353 K. In this canonical depiction of the energy landscape, the depth of the funnel indicated the stability of the protein and width of the funnel represented the conformational entropy of the protein. Hence, these finding suggested 353 K configuration to be a most stable conformation and having the highest resemblance to the native folded state when compared to other tested conformations. Moreover, our findings were in a close agreement with the experimental observations which also showed 353 K to be the temperature with the most stable configuration. The stability curve for experimental conformation has been shown with regular

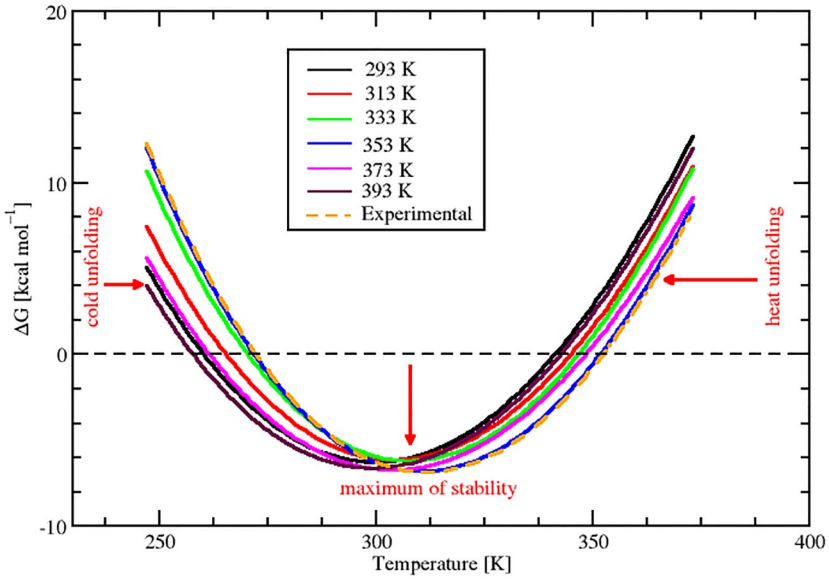

Fig 7. Superimposed stability curves. Comparative stability diagram of predicted Gibbs free energy for different conformational ensembles of SazCA. The experimental Gibbs free energy has also been plotted with dashed line for comparison.

https://doi.org/10.1371/journal.pone.0249866.g007

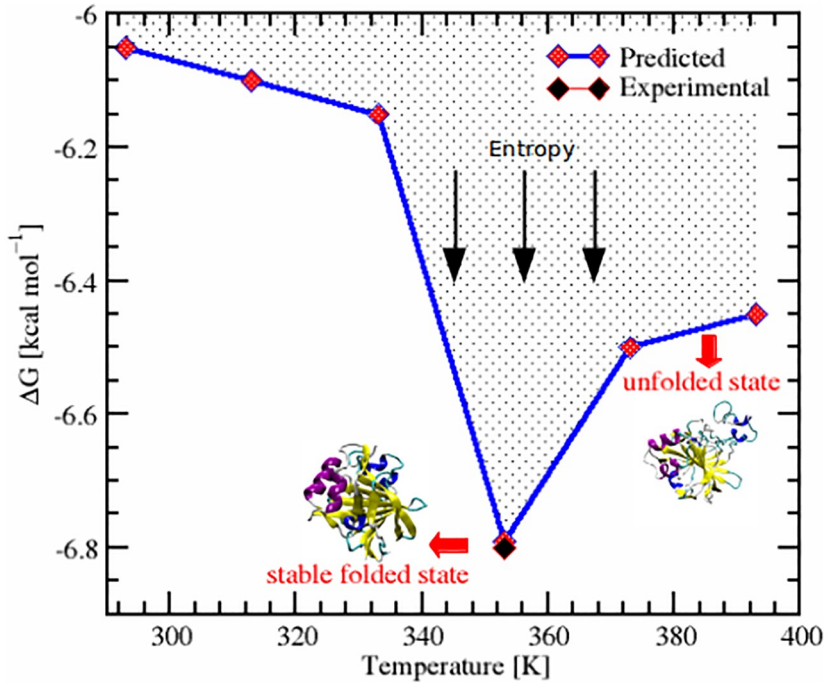

Fig 8. Energy landscape of SazCA. Funnel shaped energy landscape of SazCA representing the relationship of folding/unfolding to the free energy as a function of temperature. The native-like structure found at 353 K is shown in secondary structure representation as the stable folded structure.

https://doi.org/10.1371/journal.pone.0249866.g008

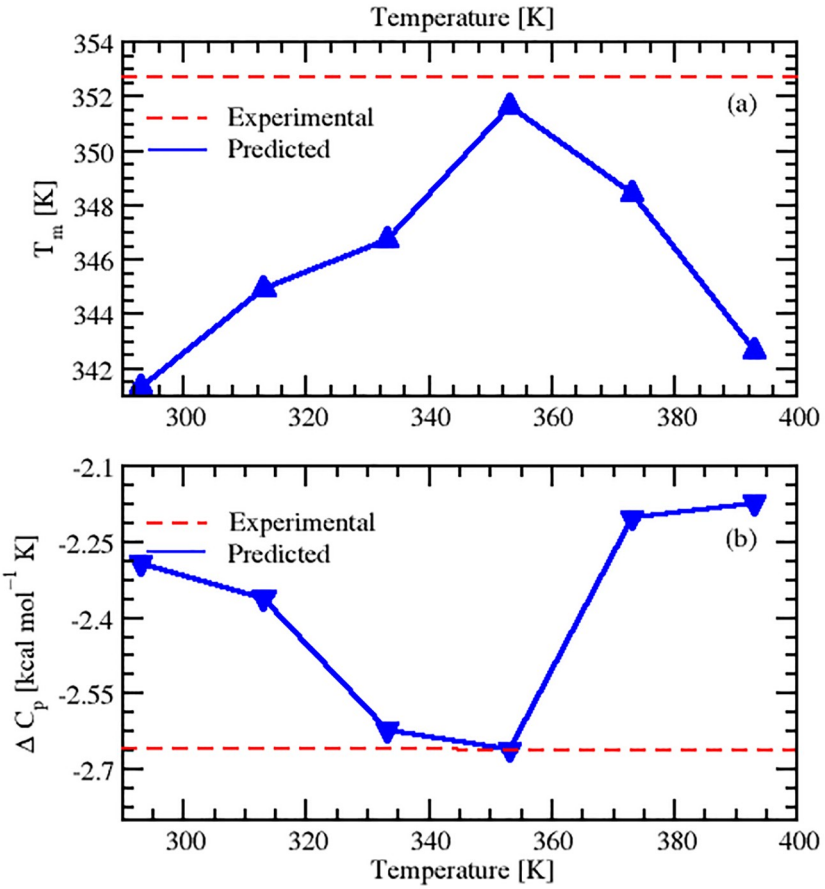

Fig 9. Variation of melting temperatures and heat capacities as a function of temperature. Predicted melting temperatures T_m (a) and heat capacities (b) at unfolding for different conformations of SazCA. The experimental T_m and ΔC_P of the crystal structure have also been plotted for comparison with dashed line.

https://doi.org/10.1371/journal.pone.0249866.g009

dashed line in Fig 7. Our findings corroborate with the results reported by Pucci *et al*., that the protein conformation which belongs to maximum stability temperature and lowest value of ΔG has the most stable form and exhibits the highest thermostability (thermoresistance) [48]. Hence, 353 K conformations were the conformations with the highest stability.

2.2.2 Prediction of melting temperature T_m. Apart from the thermodynamic stability in terms of ΔG, the thermal stability in terms of melting temperature, T_m, can be calculated from the stability diagram as an increase in T_m is associated with free energy of maximum stability. A global decrease in the ΔG implies an increase in T_m. From Figs 6 and 7, it can be observed that there was a significant increase in T_m of sampled configurations from 293 K to 353 K, but after this temperature, there was a reduction in T_m. It can also be observed that the configuration sampled at 353 K was associated with the highest T_m (351.65 K) and this temperature (353 K) has been reported to be the experimental optimum temperature. The experimental configuration has a value of T_m (352.75 K ≃ 353 K) which is close to the our finding of 353 K. Fig 9a shows the variation of T_m of all sampled conformational ensembles at different temperatures. The corresponding data for thermal stability for different sampled conformations at various temperatures can be found in Table 3. The protein conformations denaturing at higher values of T_m tend to adopt greater maximum thermal stability [49]. Therefore, our findings suggested 353 K configuration to be the best descriptor for thermal

Table 3. Predicted and experimental values of thermodynamic and thermal parameters of ensembles of SazCA at different temperatures. The thermal and thermodynamic parameters of crystal structure (experimental) of SazCA are also shown for comparison.

Temperature (K)	ΔH_m (kcal mol^{-1})	ΔC_P (kcal mol^{-1} K^{-1})	ΔG (kcal mol^{-1})	T_m (K)
293	-101.1	-2.29	-6.05	341.35
313	-101.8	-2.36	-6.10	344.95
333	-108.3	-2.62	-6.15	346.75
353	**-115.4**	**-2.66**	**-6.79**	**351.65**
373	-103.2	-2.20	-6.50	348.45
393	-102.4	-2.17	-6.45	342.65
Experimental	**-116.2**	**-2.66**	**-6.80**	**352.75**

https://doi.org/10.1371/journal.pone.0249866.t003

stability because of its association with highest denaturation temperature. It also implied that structural ensemble at 353 K was in a stable folded state and unfolding started after this temperature. A decrease in magnitude of T_m after 353 K configuration can be observed in Fig 9a and corresponding data in Table 3.

2.2.3 Prediction of heat capacity ΔC_P at unfolding. To gain more understanding into the thermodynamic stability and to make the analysis more quantitative, we predicted another thermodynamic parameter, heat capacity (ΔC_P) at unfolding. We predicted the ΔC_P at all investigated temperatures and its variation at different temperatures is shown in Fig 9b. The corresponding data of these ΔC_P values are shown in Table 3. On comparing ΔC_P, we found a similar trend as we found in the analysis of ΔG. From temperatures 293 to 353 K, there was a decrease in the magnitude of ΔC_P values followed by subsequent increase from 353 to 393 K. The maximum negative ΔC_P was found for 353 K conformations. The experimental ΔC_P has also been plotted and is shown as the regular dashed line in Fig 9b. It was found that our predicted value of ΔC_P and the experimental were in agreement. From Figs 7 and 9(b), it can be observed that lesser negative value of ΔC_P at 353 K configuration broadened the stability curve and resulted in an up-shift towards high temperatures. From this, it can be deduced that there was a close relationship between ΔC_P and T_m of the protein. The lesser the negative value of ΔC_P, the higher the magnitude of T_m. The protein associated with the less negative value of ΔC_P had higher denaturing temperature (T_m) and exhibited enhanced thermoresistance [48, 50]. The association of highest melting temperature and least ΔC_P value at 353 K resulted these conformations to be the most thermoresistant and thermostable. The enhanced protein stability at 353 K suggested an increase in folding [51]. The decrease in melting temperature after 353 K altered the native conformations and suggested a gain in conformational disorder (entropy) due the increase in the disorder of the protein-solvent interactions at higher temperatures. Hence, it can be concluded that at 353 K, SazCA has a transition in the folding/unfolding pathways. The stable folded conformation with highest thermostability was found for 353 K and beyond this temperature, denaturation prevailed.

3 Method

The crystal structure of SazCA isolated from *Sulfurihydrogenibium azorense* (PDB ID: 4X5S) [20] was used as the initial guess structure for setting up MD simulations. The protein reported in this PDB had a dimeric arrangement and Zn^{2+} ions were present in each of the monomeric units. The resolution of crystal structure of this PDB of thermostable SazCA was 1.95 Å. The amino acid numbering from the published crystal structure of SazCA was used in this study. Hydrogens were added to the hydrogen-less crystal structure conforming to the following protocol.

SazCA has been reported to be catalytically active for CO_2 hydration at 293 K and 7.5 pH [19]. Conforming to these conditions, the protonation states of amino acid residues were decided. This effect was characterized by the pK_a of the amino acid side chain. PSFGEN package [52] was used to add hydrogens to the crystal structure. The protonated protein was then immersed in a periodic cubic water box of dimensions 120 × 120 × 120 Å conforming to the constraint that there was a distance of at least 15 Å between the edge of the water box and the nearest protein fragment. This was done to minimize the interactions of periodic images of the protein. Further, Na^+ and Cl^- counterions were added to the system so as to render the system a zero net charge.

To study the structural and dynamical behaviour of SazCA, MD simulations of the solvated protein were carried out at six different temperatures in a range of 293-393 K, each for a period of 100 ns, with AMBER99SB* [53] force-field using NAMD [54] package. AMBER99SB* was chosen for the protein description following a detailed study by us reported elsewhere [39]. The water model TIP3P [55] was used to solvate the protein as TIP3P water model has been reported to support the sampling of metastable β-structures, turns and helical structures [56] which are important in determining the conformational stability of a protein. The steepest descent method [57] of energy minimization was applied for 1000 cycles to minimize the system energy. This was done to eliminate all initial bad stresses and contacts. Long range electrostatic interactions were calculated using Particle Mesh Ewald method and a cut off distance of 12 Å was employed to handle real-space interactions [58]. Once these preliminary steps were completed, the system was allowed to equilibrate using 1 fs time step of integration and 1 bar pressure in isothermal-isobaric (NPT) ensemble. Barendsen's coupling method was applied to control pressure fluctuations in the system [59]. The SHAKE algorithm was employed to constraint the bonds between the heavy atoms and hydrogens [60]. Isothermal conditions were maintained throughout the MD simulations by Nosé-Hoover's [61] method where Langevin dynamics [62] was used to control the barostat fluctuations. The results reported in this study are based on well equilibrated final 20 ns simulation trajectories for each system.

The tools provided in the VMD [52] were utilised to analyse different MD trajectories and to render the protein structures. The measurements of the RMSD, RMSF, R_G and their convergence were computed and the same have been provided in the S1 File. Hydrogen bonding is a key parameter for maintaining protein secondary structure. To calculate H-bonds, we used a geometrical criterion according to which the cutoff distance between the donor and acceptor was less than 3.4 Å and angle was less than 30˚ for an H-bond to exist [63]. In this study, we considered both intrapeptide as well as interpeptide H-bonds. Analysis of MD trajectories also included the measurement of secondary structure assignments and SASA. STRIDE algorithm was used to measure these parameters. The folding free energies were computed to explore the folding/unfolding pathways of SazCA. SCooP [64, 65] web server was used to calculate the thermodynamic parameters for the stability curves of the protein.

4 Conclusions

In the current study, MD simulations were carried out for understanding the molecular basis of thermal stability and folding/unfolding of SazCA. Different stability-determining variables, such as RMSD, RMSF, R_G, SASA, H-bonds, RDF, Gibbs free energy, melting temperatures, secondary structure assignments, unfolding pathways, associated with structural stability and folding/unfolding of the protein were quantified. The effect of elevated temperatures on RMSD, RMSF and R_G confirmed that the enzyme maintained the structural stability upto 353 K and beyond this temperature, protein conformations were highly flexible. Due to higher resemblance of protein conformations at 353 K to the native state, SazCA exhibited the highest

rigidity and thus was the most stable at 353 K. The SASA analysis revealed 353 K conformations to show stable dynamics due the higher rate of folding and less initial drift of the protein structure from the wild structure. An analysis of H-bonds showed that SazCA maintained its native state upto 353 K due the lesser exposure to unfolding for this temperature range. The conformational ensembles of the protein were in the most folded state and unfolding prevailed after 353 K due to a remarkable decrease in the backbone-backbone as well as backbone-water H-bonds and larger exposure of hydrophobic surface residues to the solvent medium. RDF analysis showed 353 K conformations with higher ordering of solvent molecules with the protein backbones atoms through strong hydrophilic interactions to be the stable structural ensemble to depict folding and dynamical properties. Due to the occurrence of higher percentage amount of α-helices and β-sheets at 353 K, these conformations exhibited the highest thermostability in SazCA.

In addition to providing the insights into the thermal stability of SazCA, we also addressed the effect of temperature on the thermodynamic stability of SazCA. From the thermodynamic stability analysis, it was found that SazCA had the most stable folded conformation at 353 K. The mechanism of maximum stabilisation exhibited by SazCA at 353 K can be summarised into three major outcomes: first, its association with highest T_s and maximum negative ΔG, second, the highest T_m which was 351.65 K, close to the experimentally reported optimum working temperature (353 K) of SazCA, and third, the maximum negative ΔC_P of SazCA at 353 K, which upshifted the stability curve towards high temperature regime. After 353 K, SazCA had a transition in folding/unfolding pathways. Hence, it can be concluded that at 353 K, the enzyme SazCA has most stable folded ensemble with the highest thermostability.

Supporting information

S1 File.
(PDF)

Author Contributions

Conceptualization: Parag A. Deshpande.

Data curation: Shashi Kumar.

Formal analysis: Shashi Kumar, Parag A. Deshpande.

Funding acquisition: Parag A. Deshpande.

Investigation: Shashi Kumar.

Methodology: Shashi Kumar, Parag A. Deshpande.

Supervision: Parag A. Deshpande.

Writing – original draft: Shashi Kumar.

Writing – review & editing: Parag A. Deshpande.

References

1. Supuran CT. Structure and function of carbonic anhydrases. Biochem J. 2016; 473(14):2023–2032. https://doi.org/10.1042/BCJ20160115 PMID: 27407171
2. Kararli T, Silverman DN. Kinetics of the Hydration of CO2 Catalyzed by Carbonic Anhydrase III from Skeletal Muscle of the Cat a. Ann NY Acad Sci. 1984; 429(1):129–136. https://doi.org/10.1111/j.1749-6632.1984.tb12323.x

3. Wells JW, Kandel SI, Koenig SH. The pH dependence of solvent proton relaxation in carbonic anhydrase solutions: paramagnetic and diamagnetic effects. Biochemistry. 1979; 18(10):1989–1995. https://doi.org/10.1021/bi00577a022
4. Villafuerte FC, Swietach P, Patiar S, Harris AL, Vaughan-Jones RD. Comparison of pH-dependence of Carbonic Anhydrase Activity in vitro and in Living Cells. Biophys J. 2009; 96(3):625a. https://doi.org/10.1016/j.bpj.2008.12.3305
5. Smith KS, Jakubzick C, Whittam TS, Ferry JG. Carbonic anhydrase is an ancient enzyme widespread in prokaryotes. PNAS. 1999; 96(26):15184–15189. https://doi.org/10.1073/pnas.96.26.15184 PMID: 10611359
6. Capasso C, Supuran CT. An overview of the alpha-, beta-and gamma-carbonic anhydrases from Bacteria: can bacterial carbonic anhydrases shed new light on evolution of bacteria? J Enzyme Inhib Med Chem. 2015; 30(2):325–332. https://doi.org/10.3109/14756366.2014.910202 PMID: 24766661
7. Liljas A, Laurberg M. A wheel invented three times. EMBO Rep. 2000; 1(1):16–17. https://doi.org/10.1093/embo-reports/kvd016 PMID: 11256616
8. Lindskog S. Structure and mechanism of carbonic anhydrase. Pharmacol Ther. 1997; 74(1):1–20. https://doi.org/10.1016/S0163-7258(96)00198-2 PMID: 9336012
9. Zheng YJ, Merz KM Jr. Mechanism of the human carbonic anhydrase II-catalyzed hydration of carbon dioxide. JACS. 1992; 114(26):10498–10507. https://doi.org/10.1021/ja00052a054
10. Satoh D, Hiraoka Y, Colman B, Matsuda Y. Physiological and Molecular Biological Characterization of Intracellular Carbonic Anhydrase from the Marine DiatomPhaeodactylum tricornutum. Plant Physiol. 2001; 126(4):1459–1470. https://doi.org/10.1104/pp.126.4.1459 PMID: 11500545
11. Nishimori I, Onishi S, Takeuchi H, Supuran CT. The α and β classes carbonic anhydrases from Helicobacter pylori as novel drug targets. Curr Pharm Des. 2008; 14(7):622–630. https://doi.org/10.2174/138161208783877875 PMID: 18336307
12. Di Fiore A, Alterio V, Monti SM, De Simone G, D'Ambrosio K. Thermostable carbonic anhydrases in biotechnological applications. Int J Mol Sci. 2015; 16(7):15456–15480. https://doi.org/10.3390/ijms160715456 PMID: 26184158
13. Alterio V, Monti SM, De Simone G. Thermal-stable carbonic anhydrases: A structural overview. In: Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications. Springer; 2014. p. 387–404.
14. Boone CD, Habibzadegan A, Gill S, McKenna R. Carbonic anhydrases and their biotechnological applications. Biomolecules. 2013; 3(3):553–562. https://doi.org/10.3390/biom3030553 PMID: 24970180
15. Fisher Z, Boone CD, Biswas SM, Venkatakrishnan B, Aggarwal M, Tu C, et al. Kinetic and structural characterization of thermostabilized mutants of human carbonic anhydrase II. Protein Eng Des Sel. 2012; 25(7):347–355. https://doi.org/10.1093/protein/gzs027 PMID: 22691706
16. Capasso C, De Luca V, Carginale V, Cannio R, Rossi M. Biochemical properties of a novel and highly thermostable bacterial α-carbonic anhydrase from Sulfurihydrogenibium yellowstonense YO3AOP1. J Enzyme Inhib Med Chem. 2012; 27(6):892–897. https://doi.org/10.3109/14756366.2012.703185 PMID: 22803664
17. Kean KM, Porter JJ, Mehl RA, Karplus PA. Structural insights into a thermostable variant of human carbonic anhydrase II. Protein Sci. 2018; 27(2):573–577. https://doi.org/10.1002/pro.3347 PMID: 29139171
18. Akdemir A, Vullo D, De Luca V, Scozzafava A, Carginale V, Rossi M, et al. The extremo-α-carbonic anhydrase (CA) from Sulfurihydrogenibium azorense, the fastest CA known, is highly activated by amino acids and amines. Bioorg Med Chem Lett. 2013; 23(4):1087–1090. https://doi.org/10.1016/j.bmcl.2012.12.009 PMID: 23294703
19. De Luca V, Vullo D, Scozzafava A, Carginale V, Rossi M, Supuran CT, et al. An α-carbonic anhydrase from the thermophilic bacterium Sulphurihydrogenibium azorense is the fastest enzyme known for the CO2 hydration reaction. Bioorg Med Chem Lett. 2013; 21(6):1465–1469. https://doi.org/10.1016/j.bmc.2012.09.047
20. De Simone G, Monti SM, Alterio V, Buonanno M, De Luca V, Rossi M, et al. Crystal structure of the most catalytically effective carbonic anhydrase enzyme known, SazCA from the thermophilic bacterium Sulfurihydrogenibium azorense. Bioorg Med Chem Lett. 2015; 25(9):2002–2006. https://doi.org/10.1016/j.bmcl.2015.02.068 PMID: 25817590
21. Jagannadham M, Balasubramanian D. The molten globular intermediate form in the folding pathway of human carbonic anhydrase B. FEBS Lett. 1985; 188(2):326–330. https://doi.org/10.1016/0014-5793(85)80396-3 PMID: 3928403
22. Krishnamurthy VM, Kaufman GK, Urbach AR, Gitlin I, Gudiksen KL, Weibel DB, et al. Carbonic anhydrase as a model for biophysical and physical-organic studies of proteins and protein- ligand binding. Chem Rev. 2008; 108(3):946–1051. https://doi.org/10.1021/cr050262p PMID: 18335973

23. Prakash A, Dixit G, Meena NK, Singh R, Vishwakarma P, Mishra S, et al. Elucidation of stable intermediates in urea-induced unfolding pathway of human carbonic anhydrase IX. J Biomol Struct Dyn. 2018; 36(9):2391–2406. https://doi.org/10.1080/07391102.2017.1355847 PMID: 28705076

24. Mayor U, Johnson CM, Daggett V, Fersht AR. Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. PNAS. 2000; 97(25):13518–13522. https://doi.org/10.1073/pnas.250473497 PMID: 11087839

25. Peterson ME, Daniel RM, Danson MJ, Eisenthal R. The dependence of enzyme activity on temperature: determination and validation of parameters. Biochem J. 2007; 402(2):331–337. https://doi.org/10.1042/BJ20061143 PMID: 17092210

26. Daniel RM, Peterson ME, Danson MJ, Price NC, Kelly SM, Monk CR, et al. The molecular basis of the effect of temperature on enzyme activity. Biochem J. 2010; 425(2):353–360. https://doi.org/10.1042/BJ20091254

27. Bharatiy SK, Hazra M, Paul M, Mohapatra S, Samantaray D, Dubey RC, et al. In silico designing of an industrially sustainable carbonic anhydrase using molecular dynamics simulation. ACS Omega. 2016; 1 (6):1081–1103. https://doi.org/10.1021/acsomega.6b00041 PMID: 30023502

28. Scandurra R, Consalvi V, Chiaraluce R, Politi L, Engel PC. Protein thermostability in extremophiles. Biochimie. 1998; 80(11):933–941. https://doi.org/10.1016/S0300-9084(00)88890-2 PMID: 9893953

29. Kumar S, Tsai CJ, Nussinov R. Factors enhancing protein thermostability. Protein Eng. 2000; 13 (3):179–191. https://doi.org/10.1093/protein/13.3.179 PMID: 10775659

30. Szilagyi A, Zavodszky P. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. Structure. 2000; 8(5):493–504. https://doi.org/10.1016/S0969-2126(00)00133-7 PMID: 10801491

31. Prabhu N, Sharp K. Protein- solvent interactions. Chem Rev. 2006; 106(5):1616–1623. https://doi.org/10.1021/cr040437f PMID: 16683747

32. Purkiss A, Skoulakis S, Goodfellow JM. The protein–solvent interface: A big splash. Philos Trans R Soc London, Ser A. 2001; 359(1785):1515–1527. https://doi.org/10.1098/rsta.2001.0863

33. Roberts BC, Mancera RL. Ligand- protein docking with water molecules. J Chem Inf Model. 2008; 48 (2):397–408. https://doi.org/10.1021/ci700285e PMID: 18211049

34. Nakasako M. Large-scale networks of hydration water molecules around bovine β-trypsin revealed by cryogenic X-ray crystal structure analysis. J Mol Biol. 1999; 289(3):547–564. https://doi.org/10.1006/jmbi.1999.2795

35. Higo J, Nakasako M. Hydration structure of human lysozyme investigated by molecular dynamics simulation and cryogenic X-ray crystal structure analyses: On the correlation between crystal water sites, solvent density, and solvent dipole. J Comput Chem. 2002; 23(14):1323–1336. https://doi.org/10.1002/jcc.10100 PMID: 12214315

36. Raschke TM. Water structure and interactions with protein surfaces. Curr Opin Struct Biol. 2006; 16 (2):152–159. https://doi.org/10.1016/j.sbi.2006.03.002 PMID: 16546375

37. Lounnas V, Pettitt B, Phillips G Jr. A global model of the protein-solvent interface. Biophys J. 1994; 66 (3):601–614. https://doi.org/10.1016/s0006-3495(94)80835-5 PMID: 8011893

38. Chen X, Weber I, Harrison RW. Hydration water and bulk water in proteins have distinct properties in radial distributions calculated from 105 atomic resolution crystal structures. J Phys Chem B. 2008; 112 (38):12073–12080. https://doi.org/10.1021/jp802795a PMID: 18754631

39. Kumar S, Seth D, Deshpande PA. Molecular dynamics simulations identify the regions of compromised thermostability in SazCA. Proteins Struct Funct Bioinf. 2021; 89:375–388. https://doi.org/10.1002/prot.26022 PMID: 33146427

40. Ausaf Ali S, Hassan I, Islam A, Ahmad F, et al. A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states. Curr Protein Pept Sci. 2014; 15 (5):456–476. https://doi.org/10.2174/1389203715666140327114232

41. Matthes D, De Groot BL. Secondary structure propensities in peptide folding simulations: a systematic comparison of molecular mechanics interaction schemes. Biophys J. 2009; 97(2):599–608. https://doi.org/10.1016/j.bpj.2009.04.061 PMID: 19619475

42. Folch B, Rooman M, Dehouck Y. Thermostability of salt bridges versus hydrophobic interactions in proteins probed by statistical potentials. J Chem Inf Model. 2008; 48(1):119–127. https://doi.org/10.1021/ci700237g PMID: 18161956

43. Folch B, Dehouck Y, Rooman M. Thermo-and mesostabilizing protein interactions identified by temperature-dependent statistical potentials. Biophys J. 2010; 98(4):667–677. https://doi.org/10.1016/j.bpj.2009.10.050 PMID: 20159163

44. Gromiha MM. Important inter-residue contacts for enhancing the thermal stability of thermophilic proteins. Biophys Chem. 2001; 91(1):71–77. https://doi.org/10.1016/S0301-4622(01)00154-5 PMID: 11403885

45. Kannan N, Vishveshwara S. Aromatic clusters: a determinant of thermal stability of thermophilic proteins. Protein Eng. 2000; 13(11):753–761. https://doi.org/10.1093/protein/13.11.753 PMID: 11161106

46. Hawley SA. Reversible pressure-temperature denaturation of chymotrypsinoge. Biochemistry. 1971; 10(13):2436–2442. https://doi.org/10.1021/bi00789a002

47. Privalov PL, Gill SJ. Stability of protein structure and hydrophobic interaction. In: Advances in Protein Chemisry. vol. 39. Elsevier; 1988. p. 191–234.

48. Pucci F, Rooman M. Stability curve prediction of homologous proteins using temperature-dependent statistical potentials. PLoS Comput Biol. 2014; 10(7):e1003689. https://doi.org/10.1371/journal.pcbi.1003689 PMID: 25032839

49. Rees DC, Robertson AD. Some thermodynamic implications for the thermostability of proteins. Protein Sci. 2001; 10(6):1187–1194. https://doi.org/10.1110/ps.180101 PMID: 11369857

50. Kumar S, Nussinov R. How do thermophilic proteins deal with heat? Cell Mol Life Sci. 2001; 58 (9):1216–1233. https://doi.org/10.1007/PL00000935 PMID: 11577980

51. Luke KA, Higgins CL, Wittung-Stafshede P. Thermodynamic stability and folding of proteins from hyperthermophilic organisms. The FEBS J. 2007; 274(16):4023–4033. https://doi.org/10.1111/j.1742-4658.2007.05955.x PMID: 17683332

52. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graphics. 1996; 14(1):33–38. https://doi.org/10.1016/0263-7855(96)00018-5 PMID: 8744570

53. Best RB, Hummer G. Optimized molecular dynamics force fields applied to the helix- coil transition of polypeptides. J Phys Chem B. 2009; 113(26):9004–9015. https://doi.org/10.1021/jp901540t PMID: 19514729

54. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005; 26(16):1781–1802. https://doi.org/10.1002/jcc.20289 PMID: 16222654

55. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983; 79(2):926–935. https://doi.org/10.1063/1.445869

56. Smith MD, Rao JS, Segelken E, Cruz L. Force-field induced bias in the structure of Aβ21–30: A comparison of OPLS, AMBER, CHARMM, and GROMOS force fields. J Chem Inf Model. 2015; 55(12):2587–2595. https://doi.org/10.1021/acs.jcim.5b00308 PMID: 26629886

57. Adcock SA, McCammon JA. Molecular dynamics: survey of methods for simulating the activity of proteins. Chem Rev. 2006; 106(5):1589–1615. https://doi.org/10.1021/cr040426m PMID: 16683746

58. Cheatham TI, Miller J, Fox T, Darden T, Kollman P. Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. JACS. 1995; 117(14):4193–4194. https://doi.org/10.1021/ja00119a045

59. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. MMolecular dynamics with coupling to an external bath. J Chem Phys. 1984; 81(8):3684–3690. https://doi.org/10.1063/1.448118

60. Lee KJ. Molecular dynamics simulations of a hyperthermophilic and a mesophilic protein L30e. J Chem Inf Model. 2011; 52(1):7–15. https://doi.org/10.1021/ci200184y PMID: 22168407

61. Feller SE, Zhang Y, Pastor RW, Brooks BR. Constant pressure molecular dynamics simulation: the Langevin piston method. J Chem Phys. 1995; 103(11):4613–4621. https://doi.org/10.1063/1.470648

62. Paquet E, Viktor HL. Molecular dynamics, monte carlo simulations, and langevin dynamics: a computational review. Biomed Res Int. 2015; 2015. https://doi.org/10.1155/2015/183918 PMID: 25785262

63. Pokhrel R, Bhattarai N, Baral P, Gerstman BS, Park JH, Handfield M, et al. Molecular mechanisms of pore formation and membrane disruption by the antimicrobial lantibiotic peptide Mutacin 1140. PCCP. 2019; 21(23):12530–12539. https://doi.org/10.1039/C9CP01558B PMID: 31147666

64. Pucci F, Kwasigroch JM, Rooman M. SCooP: an accurate and fast predictor of protein stability curves as a function of temperature. Bioinformatics. 2017; 33(21):3415–3422. https://doi.org/10.1093/bioinformatics/btx417 PMID: 29036273

65. Pucci F, Rooman M. Towards an accurate prediction of the thermal stability of homologous proteins. J Biomol Struct Dyn. 2016; 34(5):1132–1142. https://doi.org/10.1080/07391102.2015.1073631 PMID: 26198299