J Mol Model (2013) 19:3993–4002
DOI 10.1007/s00894-013-1919-4

ORIGINAL PAPER

High temperature unfolding of a truncated hemoglobin by molecular dynamics simulation

Ravi Datta Sharma · Rajnee Kanwal · Andrew M. Lynn · Prerna Singh · Syed Tazeen Pasha · Tasneem Fatma · Safdar Jawaid

Received: 21 January 2013 / Accepted: 9 June 2013 / Published online: 10 July 2013

Abstract Heme containing proteins are associated with peroxidase activity. The proteins like hemoglobin, myoglobins, cytochrome *c* and micro-peroxidase other than peroxidases have been shown to exhibit weak peroxidase-like activity. This weak peroxidase–like activity in hemoglobin-like molecules is due to heme moiety. We conducted molecular dynamics (MD) studies to decipher the unfolding path of Ba-Glb (a truncated hemoglobin from *Bacillus anthracis*) and the role of heme moiety to its unfolding path. The similar unfolding path is also observed in vitro by UV/VIS spectroscopy. The data confirmed that the unfolding of Ba-Glb follows a three state process with a meta-stable (intermediate) state between the native and unfolded conformations. The present study is supported by several unfolding parameters like root-mean-square-deviation (RMSD), dictionary of protein secondary structure (DSSP), and free energy landscape. Understanding the structure of hemoglobin like proteins in unicellular dreaded pathogens like *B. anthracis* will pave way for newer drug discovery targets and in the disease management of anthrax.

Keyword *Bacillus anthracis* · High temperature unfolding · Molecular dynamics · Truncated hemoglobin

Ravi Datta Sharma and Rajnee Kanwal, contributed equally.

Electronic supplementary material The online version of this article (doi:10.1007/s00894-013-1919-4) contains supplementary material, which is available to authorized users.

R. D. Sharma · P. Singh
Department of Microbiology, C.C.S. University, Meerut, UP, India

R. Kanwal · T. Fatma (✉) · S. Jawaid
Department of Biosciences, Jamia Millia Islamia, New Delhi, India
e-mail: fatma_cbl@yahoo.com

R. Kanwal
Department of Urology, School of Medicine, Case Western Reserve University, Cleveland, OH, USA

A. M. Lynn
School of Information Technology, CCBB, Jawaharlal Nehru University, New Delhi, India

S. T. Pasha
Division of Biochemistry and Biotechnology, National Institute of Communicable Diseases, Ministry of Health and Family Welfare, Govt. of India, 22, Shamnath Marg, Delhi, India

S. Jawaid
Blanchette Rockefeller Neurosciences Institute, 9601 Medical Center Drive, Academic and Research Building, Rockville, MA, USA

R. D. Sharma (✉)
CEIT and TECNUN, University of Navarra, San Sebastián, Spain
e-mail: ravidattasharma@gmail.com

Introduction

Protein folding is a long standing problem in biology in which linear amino acid sequence fold into three-dimensional structure and play major role in cell signaling and metabolism. Protein folding follows specific pathway(s), that may involve in formation of compact intermediate or molten globule state [1, 2], i.e., native like structure and liquid-like interior usually larger than the native structure. If protein unfolds within molten globule like structure, additional validation would be necessary to ensure the protein unfolding pathway. The intermediate is the only transition state which shows preferred conformational state of the unfolded protein under refolding conditions. Therefore, confirmation of the transition state ensemble (TSE) is fundamental to determine the path in protein unfolding [3]. Also, the polarity and hydrophobic core play important roles in protein folding where non-polar amino acids have tendency to repel surrounding aqueous solvent

while polar amino-acids are exposed to the solvent. Hence, the exposure of hydrophobic core to the solvent is a vital feature in the measurement of protein unfolding [3]. An alternative approach is to determine the local energy minima projected through the global free energy landscape using principal component analysis (PCA) [4]. PCA, a covariance-matrix-based mathematical technique where a multidimensional complex set of variables are reduced in a lower dimension which is able to identify the diffusive properties (in our case, the meta-stable state) in dynamics of protein folding. Authors Gia et al. [4] successfully demonstrated that principal components (PCs) mainly identify the transition from unfolded state to native state and were able to plot a free energy landscape. They confirmed that the most prominent motion contributing to overall protein can be determined using only a very small number mode of fluctuations which contribute significantly to overall motion of the protein [4]. They also illustrated the significant contribution of PC1 and PC2 in the main feature of free energy landscape.

It is well-established that hemoglobin (Hb) is an oxygen carrier and the heme pocket structure of hemoglobin is optimized to bind and release oxygen reversibly. Despite the large differences in the tertiary structures, all hemoglobins from unicellular organisms discovered till date contain a histidine at the F8 position that coordinates with heme iron and a tyrosine at the B10 position in the distal pocket [5]. The distal E7 position, which is occupied by a histidine in mammalian Hbs, is occupied by a variety of different polar or nonpolar residues in these hemoglobins [6]. These positions (F8, E7 and B10) are explained in detail elsewhere [5]. Hemoglobins are also known to have peroxidase like activity. The discovery of the dioxygenase activity of flavohemoglobin (Hmp) from *Escherichia coli* [7], the nitric oxide-activated deoxygenase activity of *Ascaris lumbricoides* Hb [8], and the peroxidase activity of dehaloperoxidase (DHP) from *Amphitrite ornata* [9], clearly demonstrated that globins share a wide spectrum of catalytic activities with peroxidases. The results from the work on Flavohemoglobin (Hmp) suggest that it has a peroxidase-like active site structure that may promote O_2/NO chemistry instead of oxygen delivery [7].

Heme is the prosthetic group of numerous proteins involved in a wide variety of biological processes including oxygen carriers, redox enzymes and regulatory proteins. Heme-containing enzymes are abundant in many microorganisms and include cytochrome *c* oxidase, catalases, different types of peroxidases, and catalase-peroxidases [10, 11]. Heme-containing peroxidases are ubiquitous in living organisms and are one of the most extensively studied classes of enzyme [12, 13]. The importance of these peroxidases is their potential to use hydrogen peroxide (H_2O_2) as the electron acceptor to catalyze a number of oxidative reactions and often provide protection under conditions of oxidative stress [14–16]. The weak peroxidase like activity in heme containing protein cytochrome *c* has been well exploited as a probe for structural study by Diederix et al. [17] to evaluate the role of heme moiety in the unfolding pattern and its relation with the peroxidase activity.

The fact that hemoglobin possesses peroxidative properties was first noticed more than 30 years ago when Tappel showed in vitro studies that methemoglobin and metmyoglobin can react with preexisting lipid hydroperoxides to yield lipid-associated free radicals, which resulted in the propagation of free radical-mediated reactions that produced extensive lipid peroxidation [1, 18, 19]. Human erythrocytes have also been shown to be susceptible to hemoglobin-catalyzed, H_2O_2-dependent lipid peroxidation [20–25]. On reaction with H_2O_2 hemoglobin forms a protein derivative that is capable of oxidizing a wide range of reducing substrates such as phenol and aromatic amines. The specificity of hemoglobin in catalysis of such reactions is, of course, far less than that of peroxidases [26].

We have performed molecular dynamics (MD) simulations with supportive experimental data to show that the weak peroxidase activity (due to heme moiety) in Ba-Glb can change the path of unfolding. So far many MD simulations have been carried out in different proteins [27–31] which gave the insights of two-state or multi-state unfolding. Only the transition state is accessible by classical biophysical experiments [32]. The intermediates (transition and meta-stable states) in protein folding differ minutely at atomic and molecular level which is currently captured by NMR for equilibrium structures in solution, kinetics on engineered mutants and MD simulations at atomic level [32]. MD simulation allows us to explore protein unfolding at microscopic time scale which cannot be measured by using classical biophysical techniques such as fluorescence spectroscopy and UV/VIS spectroscopy [33].

MD follows the classical Newton's method to generate trajectories, a large amount of binary data that contain the changes in structure over time. These trajectories can be further analyzed. Temperature and computational power play important roles in simulations. Here, we followed the principle introduced by Daggett & Levitt [34] that the unfolding of protein is reversible of folding process. These types of protein unfolding simulation are generally carried out at high temperature (498 K or above) resulting in loss of secondary structure in few nanoseconds. Later, Day et al. [35] confirmed that the high temperatures such as 500 K accelerate protein unfolding without changing its path. Although the unfolding does occur at low temperature it needs high computational power to capture the process which makes temperature an important constraint in MD simulation [35]. Therefore, the unfolding of Ba-Glb has been carried out at three temperatures −300 K (as a function of native protein), 400 K a middle temperature and 500 K (preferred for unfolding simulation). We carefully selected the parameters preferred for high temperature unfolding simulation (see Materials and methods section).

We focused to analyze the changes of unfolding pattern in truncated hemoglobin with heme moiety which is responsible for weak peroxidase activity. We investigated the unfolding path of Ba-Glb using root-mean-square-deviation (RMSD) -both overall and secondary structure components, the denaturation of secondary structure with dictionary of secondary structure of protein (DSSP) [36] and the global energy minima landscapes to find the meta-stable and transition states. We captured structural changes during MD simulation at different time steps and correlated with solvent accessible surface area (SASA) of heme moiety and protein. These analyses confirm that the unfolding of the Ba-Glb is indeed a three-state process (native, intermediate and unfolded) in agreement with the data obtained with melting curve analysis at UV/VIS spectroscopy.

Material and methods

In order to achieve structural analysis, the protein homology model structure was predicted using Modeller 8v2 [37] with 1UX8 from Protein Data Bank (PDB) as template and submitted earlier in PDB as ID# 2IIS. Sequence alignment of Ba-Glb with 1ux8 has been shown in Supplementary Fig. 2. The stereo-chemical quality was checked on PROCHECK [38] and post validation at Research Collaboratory for Structural Bioinformatics (RCSB) it was deposited under PDB ID: 2IIS. The overall G-factor obtained was 0.10 indicating the high quality of modeled structure. The Ramachandran plot contained 96.1 % of the residues in the most favored region, 2.9 % in additional allowed region and 1 % residue in generously allowed region. The structural analysis was accomplished using Visual Molecular Dynamics (VMD) molecular graphics program [39]. Table 1 shows the amino acid residues of Ba-Glb corresponding to each helix of the protein.

In molecular dynamics (MD) simulation, the protein was placed in a box of volume 244.00 (nm^3) and 7389 water molecules were added using explicit simple point charge (SPC) water model [40] which is suitable for high temperature unfolding simulation. G43a1 force field was applied to the system. Initially the system had a non-zero negative nine charge which was neutralized by randomly replacing solvent molecules with Na^+, resulting in 7380 water and nine Na^+ surrounding the protein. The resulting box consisted of a density 1.003 g/cc with 120 protein residues. The protein was kept fixed in the center of water box using periodic boundary conditions. An energy minimization simulation was applied to equilibrate the energies among the constituent atoms, which was achieved in 336 time steps. Position restrained MD simulations were performed after energy minimization which allowed water molecules to soak into protein, while partial freezing was applied to atom positions of protein.

Table 1 Structure designations of Ba-Glb helices

S. No.	Residues		Structure	Name
	Start	End		
1.	2	6	Helix	H1
2.	9	25	Helix	H2
3.	37	52	Helix	H3
4.	57	62	Helix	H4
5.	67	72	Helix	H5
6.	78	95	Helix	H6
7.	99	116	Helix	H7

MD simulations were carried out in order to achieve the stability of Ba-Glb at three temperatures, i.e., 300, 400 and 500 K. All the simulations were set on 10 ns using 2 fs time steps. It should be noted that before conducting unfolding simulations at 500 K and at middle temperature 400 K the systems were equilibrated for 1 ns at 300 K. In order to achieve robustness in the experiment and statistics, all simulations (500 K and 300 K) were repeated three times with different velocities and calculated the experimental errors using RMSD that were in acceptable range 0–.5 of standard deviation (See Supplementary Fig. 3). In contrast, since Ba-Glb was found to remain stable till 10 ns at 400 K, we set a long run twice till 50 ns to achieve unfolding. However, the protein remained stable even until 50 ns at 400 K simulation (see supplementary data). All simulations were subjected to isobaric-isothermal simulation at 1 atm with a Berendsen thermostat-barostat [41]. We performed the simulation using particle mesh Ewald (PME) electrostatics, which is known to be the best method for computing long range electrostatics. It provides more reliable energy estimation especially for the system possessing counterions like Na^+, Cl^-, Ca^{+2}, and so on [42]. Lincs method [43] was used to constrain all bonds. Gromacs program [44] was used for all simulations running in a fedora core operating system on eight precision machines. Trajectories were analyzed using inbuilt modules of Gromacs.

We performed the following analysis -RMSD for overall protein and secondary structure of the protein to ensure path of unfolding, DSSP -a separate program to analyze the secondary structure denaturation and, PCA analysis to find the meta-stable and transition states during unfolding. In PCA analysis, we extracted eigenvalues through a covariance matrix using trajectory at 500 K. The free energy landscape was projected with first two dominant PCs to determine the local energy minima. Next, we elaborated the role of heme moiety with the help of SASA values of protein and prominent residues of heme moiety. We selected the following prominent residues of BA-Glb that participated to form the surface inside the heme moiety: Tyr51, Leu48 and Leu109 those hinder the passage between two cavities, Val16 and Phe19 (helix H2), Gln44 and Leu48 (helix H3), Trp84 and Met88 (helix H6), C-terminal

residues of helix H4 (Tyr51 and Leu52) and helix H7 residues Phe104, and Arg108.

In experimental analysis, the truncated hemoglobin protein Ba-Glb from *B. anthracis* obtained from its over-expression in *E.coli* and successive purification to homogeneity. The quality of the protein was confirmed on 15 % SDS-PAGE, by silver staining and HPLC as mentioned in the laboratory manual of Sambrook [45]. The thermal melting curve was obtained using 0.25 mg/ml Ba-Glb protein in Tris-Cl buffer pH 8.0 in a Varian Carry UV/VIS spectrophotometer by measuring the absorbance at 280 nm between 25 °C to 98 °C with 1 °C temperature increment steps.

Results

Structural analysis

Three-dimensional structure of BA-Glb is essentially similar to that of X-ray crystallographic structure of the Bs-trHb of *B. subtilis* [5]. The overall folds and relative position of conserved sequences and heme coordinating residues are juxtaposed in both the molecules. The structures are highly super-imposable and the proximal and distal heme pockets are essentially similar. This fact is also supported with the structural comparison of the two by superimposition yielding a low RMSD 0.18 A°. The three dimensional structural model of BA-Glb (Fig. 1a, red color) is superimposed with Bs-trHb (yellow color). In the proximal heme pocket most characteristic bonding is between His71 to the heme at a short bond length of 1.9 A° at F8 position. Tyr20, Gln44 and Trp84 also surround the proximal heme pocket and are known to help in the binding and stabilization of ligand to heme moiety [5]. These residues surrounding the proximal heme pocket, inter-residue bonds have been shown in Fig. 1b.

Unfolding pattern of Ba-Glb with UV/Vis spectrophotometry

Denaturation curves of truncated hemoglobin protein Ba-Glb were measured three times. It is assumed that the transition between N (native) and D (denatured) states follows a two-transition-state mechanism in the protein. In the UV/Vis spectrocopy thermal unfolding data of protein (Fig. 2a), the process of unfolding started at a slower rate from 20 to 55 °C. As the temperature rose above 55 °C the protein unfolds quickly at a faster rate till 73 °C following an intermediate meta-stable state. The reasons behind the fast unfolding and an intermediate state could be: 1) the unfolding process accelerates as heme is exposed to the solvent and 2) the heme gets packed around with some amino acid residues. When the heme moiety is accessible to solvent, it increases the instability of the protein that leads to unfolding of the protein at faster rate. However, the intermediate state arises due to the unavailability of heme to the solvent. It has been observed that when temperature increases near to 80 °C proteins again starts unfolding at high rate similar to unfolding pattern between 55–73 °C which confirmed that the heme was exposed to the solvent again and protein unfolded by 92 °C. We further correlate these assumptions with MD simulation data.

Unfolding pattern of Ba-Glb with MD simulations

RMSDs were calculated over different trajectories (Fig. 2b). The protein remained stable (native like) throughout the simulation conducted at 300 K temperature with RMSD value around 2.0 Å (Fig. 2b, RMSD in red color). In unfolding simulation at 500 K temperature (Fig. 2b, RMSD in blue color), most of the secondary structures were denatured within 1.5 ns of simulation as confirmed by increase in RMSD up to 9 Å. After slight stabilization at 8 Å at around 3.5 ns, the RMSD continued increasing to 13 Å till 5.5 ns. However,

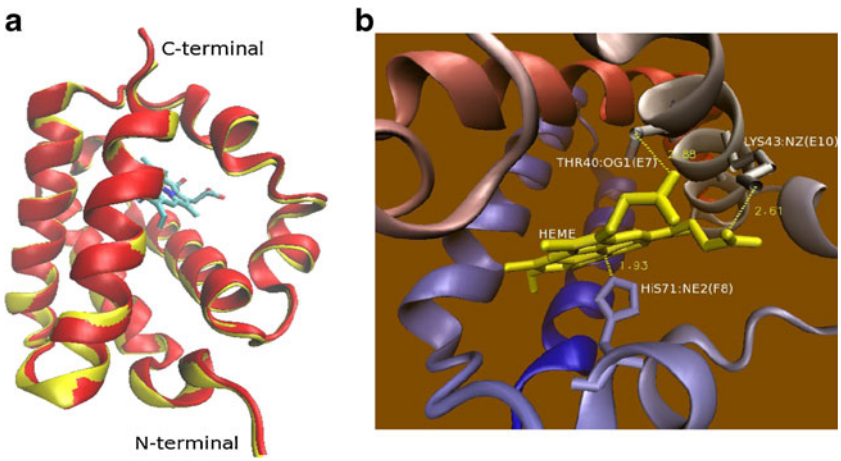

Fig. 1 **a** Super-imposed structure of Ba-Glb and Bs-trHb. **b** Structure of the proximal heme pocket

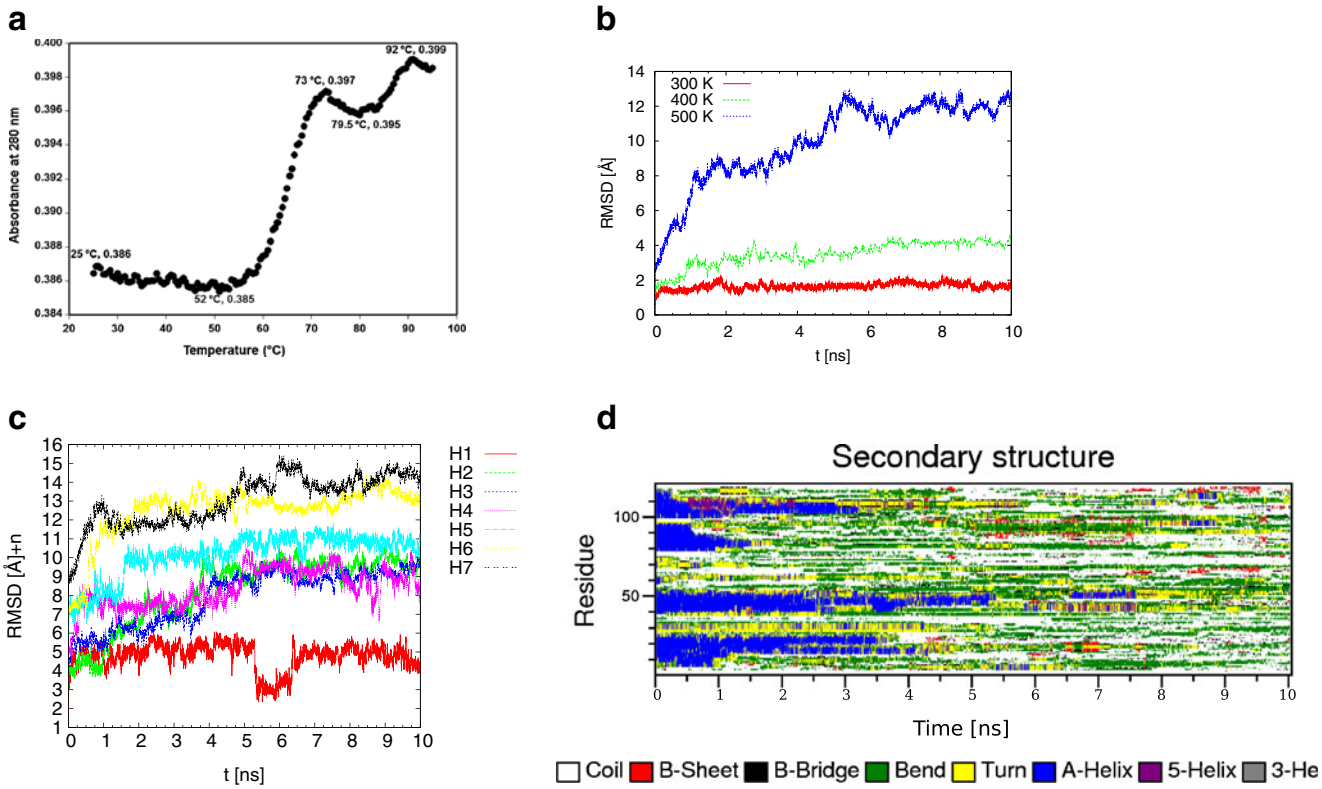

Fig. 2 **a** BA-Glb thermal melting curve (Tm) at 280 nm. **b** RMSD at 300 K (red color), 400 K (green color) and 500 K (blue color) simulations. **c** RMSD of individual helices at 500 K. **d** Denaturation of secondary structures at higher temperature (500 K) at different time steps

there was a drastic shift toward the lower RMSD of 10 Å observed near 6.7 ns from where it again rises up to its previous high value near 7.5 ns. The sudden compaction in RMSD leads to the reformation of helix H3 (Fig. 2d). On further simulation, no drastic changes were observed in RMSD. A meta-stable (intermediate) state in protein unfolding was observed between 1.5 and 3.5 ns, which is probably the preferred state for refolding where the RMSD value stabilized near 9 Å. However, it would be early to conclude this phase as meta-stable and thus it was further confirmed using DSSP and free energy landscapes results.

Furthermore, we included de-convolution method which was earlier implemented by authors [44] to gain insight in the unfolding path. We calculated the RMSD separately for all the helices and detected the path of unfolding (Fig. 2c). It should be noted that the RMSD is calculated on the basis of residues involved in the formation of a particular secondary structure. Therefore, in graph (Fig. 2c) the colored lines (RMSD) may represent the already lost secondary structure during the simulation. The RMSD of secondary protein structure was unable to provide the information about the time when these structures were lost in simulation. To evaluate the decay of secondary protein structure, Fig. 2d drawn using DSSP. The DSSP graph shows the unfolding pattern at 500 K temperature at different time steps and also confirms the consistency of protein secondary structures during meta-stable state (from 1.5 ns to 3.5 ns) observed in RMSD data. Table 1 contains the protein residues corresponding to particular secondary structures in Ba-Glb and hence, these residues can be correlated to the loss of secondary structure during simulation time course. H1 and H5 helices were the first secondary structures to disappear at the beginning of simulation within 0.3 ns, while helix H2 completely lost its secondary structure by 4.5 ns. As represented in the secondary structure of Ba-Glb (Table 1), H1 helix is a short helix of five residues present at a minimal distance of two residues from helix H2. H1 is one of the first secondary elements that unfolded at the beginning of high temperature simulation. Being on the terminal of protein, H1 should have shown highest flexibility and thus highest deviation but in contrast it showed the lowest deviation among all helices of Ba-Glb. This can be attributed to the fact that H1 is very near to H2 which makes core of the protein and unfolds at the end of simulation compared to other helices and thus it does not allow a free movement of H1. As a consequence, in the unfolding simulation of Ba-Glb, H1 helix exhibits lowest deviation at 500 K. The sixth helix (H6) was also found to be denatured within 1.5 ns and helix H4 was uncoiled at 1.1 ns. We did not observe complete decay of any secondary structure during the 1.5 ns to 3.5 ns and hence supported our

observation of a meta-stable state. At the end of the simulation (after 9.5 ns) only turns and bends could survive as observed in graph plotted with DSSP (Fig. 2d).

Above said de-convolution of structure is useful especially in cases where the protein unfolds within a molten globule like structure. In Ba-Glb, the radius of gyration does not show clear shift throughly despite an initial remodeling (Fig. 3a, blue line at 500 K) within 2 ns, i.e., a compactness of protein toward native structure due to a possible meta-stable intermediate. The decrement in number of hydrogen bonds (Fig. 3b) is supportive to presence of a meta-stable state. Different structural changes in unfolding simulation were captured from trajectories using VMD software at 1 ns time interval in unfolding simulation at 500 K (Fig. 4). The unfolding process was observed in three direct states -native or folded (N), intermediate (I), and unfolded or denatured (D) states following a transition state ensemble (TSE). We correlated these transition and meta-stable states to observed structural changes during MD simulation. First transition was observed within 1.5 ns of acquiring large structural changes in simulation with a clear increment in the backbone RMSD to approximately 9 Å (Fig. 2b, in blue color) and the hydrophobic core became partially exposed to solvent (Fig. 3c) at 2 ns at 500 K simulation (See Fig. 4 at 2 ns). The only meta-stable state was found near 1.5 ns up to 3.5 ns and no significant structural changes were observed in DSSP graph (Fig. 2d). Second transition was observed when heme moiety was exposed to solvent during 3.5–5.6 ns simulation resulting in a rapid unfolding of Ba-Glb. Here, the exposure of heme moiety can be correlated with the data of SASA (Fig. 3c and d). During the observed meta-stable state between 1.5–3.5 ns (Fig. 3c), the SASA values declined to its lowest in the simulation but soon started increasing with the beginning of the transition state post 3.5 ns and reached its highest value at around 4.6 ns. This fact can be correlated to changes in SASA of prominent residues in heme moiety that are selected by structural analysis. We observed that the SASA values of prominent residues support the data obtained with overall protein SASA during meta-stable state. The last major structural changes were observed when the protein became fully denatured by 6 ns simulation (See Fig. 4 at 6 ns).

While analyzing the results of free energy landscape, two dominant eigenvectors (Fig. 5a) that are involved in concerted motion, were selected and projected through the Gibbs free energy. Three population of local free energy minima were observed reflecting native (N), meta-stable (I) and unfolded protein (D) structures (Fig. 5b, Supplementary Fig. 4 for data of replicates). Also, two transition states can be recognized between native to meta-stable (TS1) and meta-stable to unfolded protein (TS2) states. The first population represents the lowest free energy state that reflects the native structure of the protein where most of the interactions are intact. In the second population, Ba-Glb has lost some of interaction but the core of protein remained stable, since heme moiety was not disrupted. The total loss of the overall confirmation of the protein caused the unfolded state, i.e., the third population of minimum free energy. Interestingly, PCA analysis of simulation at 400 K showed a single population with a subset (denoted with Ns in Supplementary Fig. 1e) at the beginning of the simulation. In this analysis, the first vector PC1 captured the movement of few residues at N-terminal. The movement is

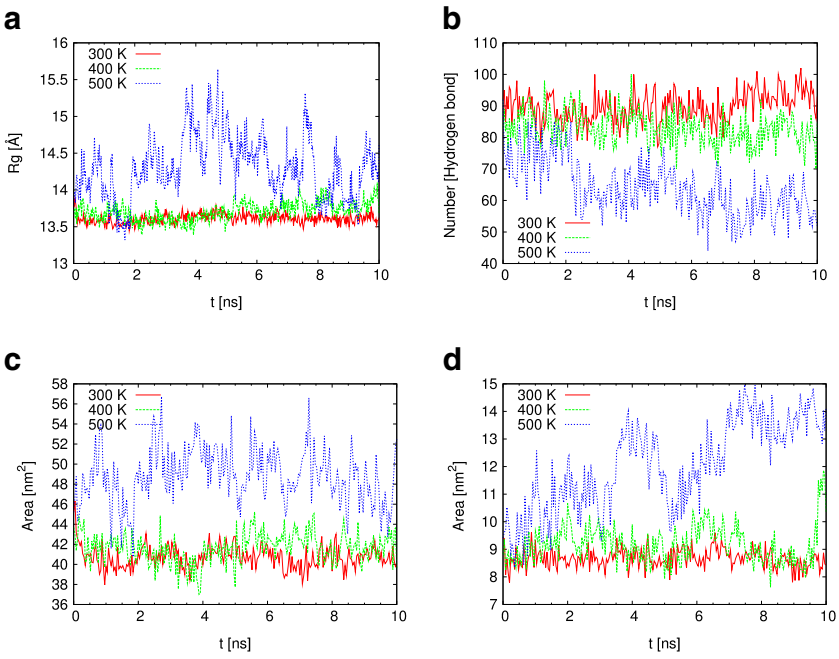

Fig. 3 Figure illustrates: **a** The radius of gyration of protein as a function of time. **b** Decrement in the number of hydrogen bonds during simulation **c** solvent accessible surface area (SASA) for whole protein and **d** SASA of prominent residues present in inner lining of heme moeity selected using structural analysis at three different temperatures (300 K- red color, 400 K- green color and 500 K- blue color) at 10 ns.

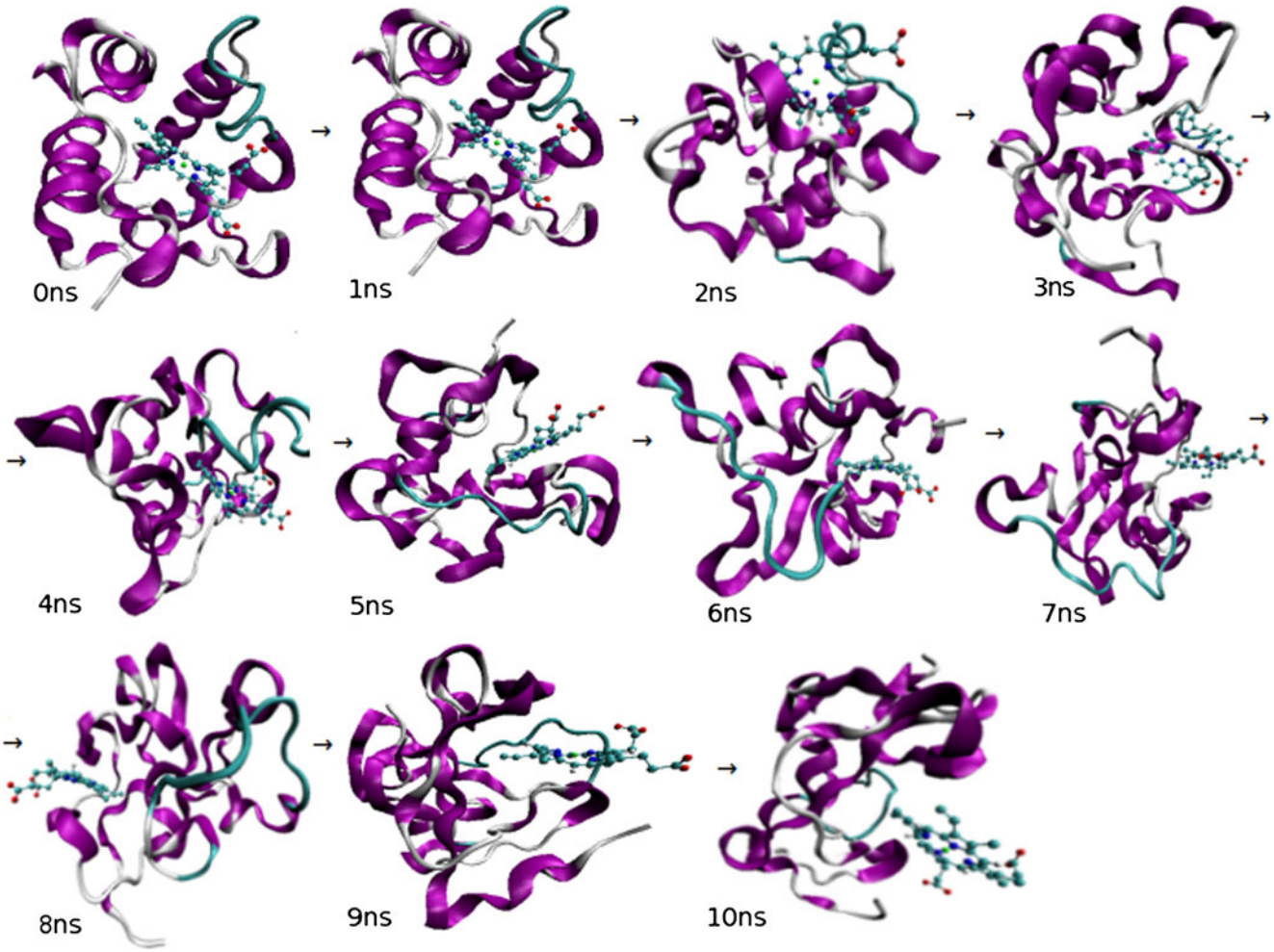

Fig. 4 Structural changes in protein at selected time points (1–10 ns)

toward N-terminal to C-terminal and facilitated by a change in confirmation of helices H6 and H7 which in turn arise due to denaturation of the 4th and 5th helices (H4 & H5). The denaturation of helices 4th and 5th are captured by PC2 (Ns). Indeed, these movements of residues occurred in early stage at 500 K simulation.

Discussion

Protein folding can be studied by characterizing the structure and stability of protein under equilibrium states and at high temperatures. In vitro and in silico methods have been used in this study to unfold the path of protein denaturation of Ba-Glb, a truncated hemoglobin from *Bacillus anthracis*. The thermal melting curve characterization of Ba-Glb was done by using UV/VIS spectroscopy. The spectroscopic study, which may distinguish the structural intermediates formed during the protein unfolding is evident in the changing absorbance pattern while transforming from native folded conformation to unfolded structure as shown (Fig. 2a). MD simulation has advantage to capture minute structural changes which are otherwise difficult to detect using classical optical method. The similar case of finding intermediates using peroxidase activity as unfolding probe in cytochrome *c* was reported by Diedrix et al. [17]. The changes in peroxiases activity of cytochrome *c* were correlated with the exposure of heme moiety during the course of unfolding of protein. The lowest free energy changes which are difficult to capture using available optical methods were detected by using peroxidase activity in cytochrome *c* [45]. We used MD simulation as probe to catch the minute changes with respect to transitions in protein states occurring during thermal unfolding of heme containing bacterial truncated hemoglobin Ba-Glb. The structure of Ba-Glb was predicted by using homology modeling (PDB ID-2IIS). To understand the unfolding process of Ba-Glb the MD simulations were set at three temperatures −300 K at which most of proteins are stable; 400 K, a middle temperature and, 500 K, a relatively higher temperature which is preferred for most protein

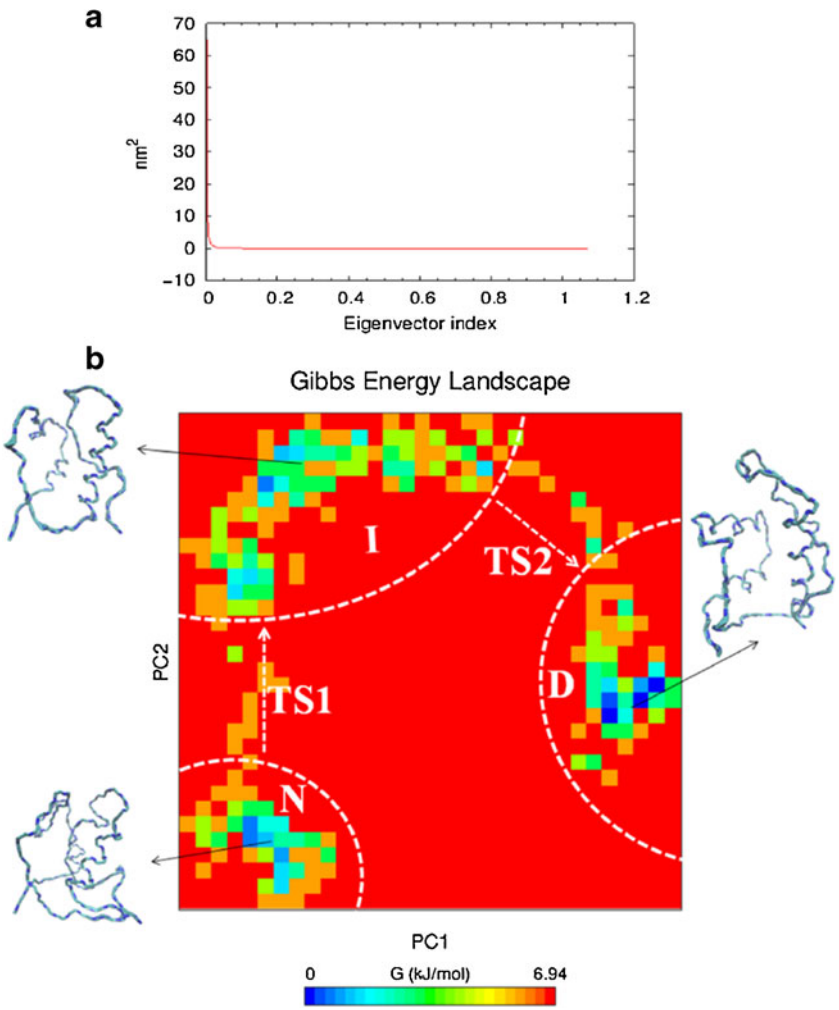

Fig. 5 **a** Eigen values of the covariance matrix representing the two dominant principal components (PCs). **b** Free energy landscape projected through the eigenvectors (nm) showing three local energy minima N-native structure, I-meta-stable and D-unfolded protein; and two transition states TS1 (N–>I) and TS2 (I–>D). Most clustered confirmations are represented by arrow for each population

unfolding processes [48, 49]. MD simulation study of Ba-Glb suggests that unfolding pattern with simulation can be directly correlated with the thermal melting curve data of protein. The RMSD of whole Ba-Glb protein (Fig. 2b) was calculated using hydrophobic residues. The hydrophobic residues are almost exclusively contained in the inner core of the protein. As the protein folds, the hydrophilic residues tend to stay at the water interface, while the hydrophobic residues are pushed together at inner part of the protein enabling it to achieve proper folding and increased stability. We also used de-convolution, the individual RMSDs of secondary structures shown in Fig. 2c to ensure the path of protein unfolding. De-convolution can be more useful if the protein unfolds in molten globule. Although RMSD is a useful parameter in unfolding simulation and shows that the obtained structure varies in coordinates with initial structure but does not clearly show the time step when the secondary structures were lost during simulation. A DSSP graph is able to impose the denaturation of secondary structure during the simulation. We found PCA analysis is an effective way to find meta-stable state where free energy landscapes are projected through dominant eigenvectors. In our case two dominant eigenvectors were used and we found three populations of local energy minima clearly emphasizing the presence of a meta-stable state (Fig. 5b).

In peroxidase-like activity of hemoglobins, heme ligand plays a crucial role as described elsewhere [6–13]. The free energy differences in these molecules between various states vastly depend on the strength of the axial heme ligand [50]. MD trajectories revealed presence of large fluctuations in heme moiety and residues forming heme access channel. These movements enlarge the opening and facilitate the access of substrates to enzyme active site which corresponds to an increase in unfolding rate. However, the closing of heme access channel causes an intermediate state (Fig. 4) of protein unfolding. Thus, the unfolding path of proteins may infer following a meta-stable state. Although, single domain proteins usually unfold in two states (native and unfolded states) as discussed by Daggett et al. [26] and in our previous report [46], we found that the unfolding of Ba-

Glb with MD simulation followed a three-state process (native, intermediate and unfolded) which supports our finding in the Ba-Glb thermal melting curve measured using UV–VIS spectroscopy. These findings can be correlated with the mechanism of peroxidase activity of hemoglobins which seems to be similar to that of peroxidases involving the formation of two intermediary compounds, compound I and compound II [50]. Hydrogen peroxidase reacts with the protein forming first intermediate compound, compound I by withdrawing two electrons from the protein heme group. This in turn withdraws two electrons from the substrate completing the peroxidase reaction. Both the intermediates are heme with four positive charges carrying iron in ferryl form [50]. The increase in the strength of the sixth heme ligand is directly proportional to the increase in peroxidase activity upon unfolding of protein [47]. The inhibition in the peroxidase activity is also caused by these protein-based ligands under fully unfolded condition as observed in cytochrome *c* [47]. Thus, heme is critical to hemoglobin and peroxidase activity and the current study shows that for the proteins with heme ligands, MD simulation is useful in detecting different states of protein unfolding via *in-silico* methods and can also be compared with the activity probes such as peroxidase activity of cytochrome *c* [47] to determine the unfolding path. In the current study, we correlated the unfolding path to weak peroxidase activity in single domain protein with heme (Ba-Glb) using MD simulation.

Conclusions

In the present study, we determined the path of unfolding of Ba-Glb, a truncated hemoglobin from *B. anthracis* using MD simulation and UV/VIS spectrophotometry. The 3D-staructure of Ba-Glb (PDB ID 2IIS) constructed using homology modeling was used for the molecular dynamics simulation. MD simulation study confirmed that the unfolding path in Ba-Glb is a three state process (folded, intermediate and unfolded) in the presence of heme moiety. The heme moiety in the truncated hemoglobin was pivotal in the unfolding path. In agreement with the experimental data of thermal melting curve using UV/VIS spectrophotometer, the path of unfolding of Ba-Glb with MD simulation confirms and validates present analysis. We believe that the detailed information of molecular dynamics would greatly facilitate the understanding of structural and functional aspects of this protein and related molecules. Clear understanding of the structural properties of truncated hemoglobin from a pathogenic bacterium like *B. anthracis* is crucial and will help in drug discovery and anthrax disease management by targeting the hemoglobin of the organism, specifically when it is within the host system.

Acknowledgments We sincerely acknowledge National Institute of Communicable Diseases, Delhi for providing facilities for the research. We also thank Prof. Rajiv Bhatt, Centre for Biotechnology, Jawaharlal Nehru University, Delhi for providing valuable suggestion and support for the work. Thanks are due to Dr. PK Sharma for facilities in Department of Microbiology, C.C.S., University, Meerut.

References

1. Richards FM (1991) Scientific American Jan: 34–41
2. Creighton TE (1990) J Biochem 270:1–16
3. Fersht AR (1995) Curr Opin Struct Biol 5(1):79–84
4. Maisuradze GG, Liwo A, Scheraga HA (2009) J Mol Biol 385(1):312–329
5. Giangiacomo L, Ilari A, Boffi A, Morea V, Chiancone E (2005) J Biol Chem 280:9192–9202
6. Mukai M, Savard PY, Ouellet H et al. (2002) Biochemistry 41:3897–3905
7. Mukai M, Mills CE, Poole RK, Yeh SR (2001) J Biol Chem 10:7272–7277
8. Minning DM, Gow AJ, Bonaventura J et al. (1993) Nature 401:497–502
9. LaCount MW, Zhang E, Chen YP et al. (2000) J Biol Chem 275:18712–18716
10. Woloszczuk W, Sprinson DB, Ruis H (1980) J Biol Chem 255:2624–2627
11. Zou M, Yesilkaya A, Ullrich Y (1999) Drug Metab Rev 31:343–349
12. Thomas EL, Bozeman PM, Learn DB (1991) In: Peroxidases in chemistry and biology. CRC, Boca Raton, pp 123–142
13. Magnussen RP (1991) Peroxidases in chemistry and biology. CRC, Boca Raton, pp 199–220
14. Degtyarenko KN, North AC, Findlay JB (1997) Protein Eng 10:183–186
15. Li H, Poulos TL (1994) Structure 2:461–464
16. Welinder KG (1991) Biochim Biophys Acta 1084:215–220
17. Diederix RE, Ubbink M, Canters GW (2002) Biochemistry 41:13067–13077
18. Tappel AL (1953) Arch Biochem Biophys 44:378–395
19. Tappel AL (1955) J Biol Chem 217:721–733
20. Benatti U, Morelli A, Damiani G et al. (1982) Biochem Biophys Res Commun 106:1183–1190
21. Chiu D, Lubin B, Shohet SB (1982) Pryor WA, (Ed.); Academic, New York 5: 115–160.
22. Clemens MR, Einsele H, Remmer H et al. (1985) Biochem Pharmacol 34:1339–1341
23. Koster JF, Slee RG (1983) Biochim Biophys Acta 752:233–239
24. Sauberman N, Fortier NL, Joshi W et al. (1983) Br J Haematol 54:15–28
25. Stocks J, Dormandy TL (1971) Br J Haematol 20:95–111
26. Daggett V, Li AJ, Itzhaki LS et al. (1996) J Mol Biol 257:430–440
27. Mayor U, Johnson CM, Daggett V et al. (2000) Proc Natl Acad Sci USA 97:13518–13522
28. Kazmirski SL, Wong KB, Freund SM et al. (2001) Proc Natl Acad Sci USA 98:4349–4354
29. Mayor U, Guydosh NR, Johnson CM et al. (2003) Nature 421:863–867
30. Petrovich M, Jonsson AL, Ferguson N (360) *J Mol Biol* 360: 865–881.
31. Interlandi G, Settanni G, Calflisch A (2006) Proteins 64:178–192
32. Fersht AR, Daggette V (2002) Cell 108:573–582
33. Allen MP (2004) NIC Series 23:1–28
34. Daggett V, Levitt M (1993) J Mol Biol 232:600–619
35. Day R, Bennion BJ, Ham S et al. (2002) J Mol Biol 322:189–203

36. Kabsch W, Sander C (1983) Biopolymers 22:2577–2637
37. Marti-Renom MA, Stuart A, Fiser A et al. (2000) Annu Rev Biophys Biomol Struct 29:291–325
38. Laskowski RA, McArthur MW, Moss DS et al. (1993) J Appl Cryst 26:283–291
39. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14:33–38
40. Berendsen HJC, Postma JPM, Gunsteren WFV et al. (1981) Intermolecular forces. Reidel, Dordrecht, pp 331–342
41. Berendsen HJC, Postma JPM, Dinola A et al. (1984) J Phys Chem 81:3684–3690
42. Essman U, Perela L, Berkowitz ML et al. (1995) J Chem Phys 103:8577–8592
43. Hess B, Bekker H, Berendsen HJC et al. (1997) J Comp Chem 18:1463–1472
44. Berendsen HJC, Spoel DVD, Drunen RV (1995) Comp Phys Comm 91:43–56
45. Sambrook J, Russell DW (2001) A laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
46. Sharma RD, Lynn AM, Sharma PK et al. (2009) Bioinformation 3–10:430–434
47. Xu Y, Mayne L, Englander SW (1998) Nat Struct Biol 5:774–778
48. Alonso DOV, Daggett V (1995) J Mol Biol 247:501–520
49. Settanni G, Fersht AR (2008) Biophys J 94:4444–4453
50. Everse J, Johnson MC, Marini MA (1994) Methods Enzymol 231:547–561