Biochimica et Biophysica Acta 1840 (2014) 2281–2288

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Molecular basis of thermal stability in truncated (2/2) hemoglobins

Juan P. Bustamante [a], Alessandra Bonamore [b], Alejandro D. Nadra [c], Natascia Sciamanna [b], Alberto Boffi [b], Darío A. Estrin [a,*], Leonardo Boechi [d,**]

[a] Departamento de Química Inorgánica, Analítica y Química Física, INQUIMAE-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
[b] Instituto Pasteur, Fondazione Cenci Bolognetti, Department of Biochemical Sciences, University of Rome "La Sapienza", Italy
[c] Departamento de Química Biológica, IQUIBICEN-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
[d] Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 21 October 2013
Received in revised form 14 March 2014
Accepted 25 March 2014
Available online 2 April 2014

Keywords:
Truncated hemoglobin
Thermostability
Thermobifida fusca
Mycobacterium tuberculosis
Molecular dynamics
Folding

ABSTRACT

Background: Understanding the molecular mechanism through which proteins are functional at extreme high and low temperatures is one of the key issues in structural biology. To investigate this phenomenon, we have focused on two instructive truncated hemoglobins from *Thermobifida fusca* (*Tf*-trHbO) and *Mycobacterium tuberculosis* (*Mt*-trHbO); although the two proteins are structurally nearly identical, only the former is stable at high temperatures.
Methods: We used molecular dynamics simulations at different temperatures as well as thermal melting profile measurements of both *wild type* proteins and two mutants designed to interchange the amino acid residue, either Pro or Gly, at E3 position.
Results: The results show that the presence of a Pro at the E3 position is able to increase (by 8°) or decrease (by 4°) the melting temperature of *Mt*-trHbO and *Tf*-trHbO, respectively. We observed that the ProE3 alters the structure of the CD loop, making it more flexible.
Conclusions: This gain in flexibility allows the protein to concentrate its fluctuations in this single loop and avoid unfolding. The alternate conformations of the CD loop also favor the formation of more salt-bridge interactions, together augmenting the protein's thermostability.
General significance: These results indicate a clear structural and dynamical role of a key residue for thermal stability in truncated hemoglobins.

1. Introduction

In nature there are many examples of organisms adapted to high temperatures. They can be classified as thermophiles with optimal growth temperatures between 333 K and 353 K, and hyperthermophiles with optimal growth temperatures between 353 K and 383 K, as opposed to the non-adapted mesophiles with optimal growth temperatures between 298 K and 323 K. Generally, enzymes belonging to these adapted organisms are also thermostable enzymes with T_m, the temperature at which 50% of the proteins are folded, close to the organism's optimal growth temperature [1]. The most studied thermostable protein is rubredoxin, from *Pyrococcus furiosus*, which presents an optimal growth temperature of 373 K [2].

Homologous enzymes from adapted and non-adapted organisms share the same catalytic mechanisms, high sequence identity and have a highly similar overall structure [3–5]. Thus, the differential adaptation to a high temperature is not easily explained at the molecular level. Indeed, an understanding of the physicochemical features that determine thermal stability still remains as one of the key issues in protein biophysics. Moreover, there is increasing interest in the use of modified proteins/enzymes with enhanced thermal stability for solving different tasks in biotechnology [6–8].

It has been proposed that thermostability is strictly related to the flexibility of the protein. Assuming that both thermostable and non-thermostable homologous enzymes perform the same function at different temperatures, and that the protein function is related to a certain degree of flexibility, the thermostable protein may have less flexibility at the same temperature. Increased rigidity of the folded state at room temperature would thus be related to an increased thermal stability. The lower flexibility would also explain why thermophiles and hyperthermophiles are inactive at low temperatures. Although many theoretical and experimental approaches do lend support to the presence of a more rigid environment in thermostable proteins at room temperature [9–13], new results coming from amide exchange experiments oppose this hypothesis [14], thus opening the discussion about the relationship between flexibility and thermal stability.

According to the folding landscape theory, the melting temperature (T_m) is related to the gap between the effective or free energy (averaged

* Corresponding author. Tel.: +54 11 4576 3378, Int. 123.
** Corresponding author. Tel.: +54 11 4576 3375.
E-mail addresses: dario@qi.fcen.uba.ar (D.A. Estrin), lboechi@ic.fcen.uba.ar (L. Boechi).

http://dx.doi.org/10.1016/j.bbagen.2014.03.018
0304-4165/

over the solvent degrees of freedom) of the folded state ensemble vs the unfolded state ensemble [15]. Therefore, a large number of hydrophilic as well as hydrophobic (related to the water entropic gain) contacts in the folded state, as opposed to those observed in the unfolded state, would favor larger T_m values. Along these lines, the number of hydrogen bonds (H-Bonds), as well as the number of salt-bridges were found to correlate with thermal stability [16].

Within the same context of the folding landscape, proteins have to lose configurational entropy in order to reach the folded state, since their configurational space is enormous in the unfolded state, but very narrow in the folded state. Any process that helps decrease the configurational entropy of the unfolded system also decreases the free energy gap between both the folded and unfolded ensembles, thus increasing T_m [17]. Considering that each backbone residue is able to explore different configurational space depending on the residue, it was proposed that the presence of more rigid backbone residues, such as prolines, are responsible for an increased thermostability through reduction of the configurational entropy [17]. Upon performing a more robust analysis of available protein sequences, however, no significant trends were found for proline residues [18]. In this context, although many hypotheses were proposed, none of them are conclusive or absolute to explain the thermostability.

From yet another point of view, we can consider the issue of thermostability as the effect of temperature on protein motion, and thus on protein function. In the present article we focused only on the first part of the problem; that is, how the temperature affects conformational sampling, by specifically considering the folding profiles and conformational spaces of the thermostable and non-thermostable proteins.

In this work we employ a combination of in silico Molecular Dynamics (MD) simulations and experimental techniques to investigate protein thermostability. In silico studies have shown to be useful to shed light on the molecular basis of protein thermostability in a variety of systems [19,20]. Specifically, we propose to study two members of the subfamily of two-over-two globins, namely the truncated hemoglobins (trHb): the trHb belonging to the thermophilic microorganism *Thermobifida fusca* (*Tf*-trHbO), with an optimal growth temperature of 333 K [21], and the trHb from the mesophilic microorganism *Mycobacterium tuberculosis* (*Mt*-trHbO) [22], with an optimal growth temperature of 298 K. The proteins have 58% sequence identity and a highly similar tertiary structure [21]. By performing MD simulations and thermal unfolding measurements of these trHbs, we were able to address a new mechanism through which these proteins acquire thermostability, helping to understand the differences between thermostable and non-thermostable proteins. We studied the *wild type* (wt) and E3-swapped mutants, in which the residues at position E3 (Pro of *Tf*-trHbO and Gly in *Mt*-trHbO) were interchanged in order to test the role of this single residue in the protein thermostability. We propose a mechanism that accounts for the thermostable protein being able to avoid unfolding at high temperatures. The *Tf*-trHbO has a flexible loop through which it can concentrate most of its fluctuations, thus avoiding the unfolding process. We found that the flexibility of the loop can be controlled by the presence of a proline residue. The lack of a flexible loop in the non-thermostable protein spreads the fluctuations over the entire protein, especially the termini, thus starting the protein unfolding process at high temperatures.

2. Materials and methods

2.1. Computer simulations

Crystal structures of both wt trHbs, *Tf*-trHbO (PDB ID: 2BMM [21]) and *Mt*-trHbO (PDB ID: 1NGK [22]) were used as starting points for the MD simulations. The *Tf*-trHbO crystal structure does not have the 34 N-terminal and 25 C-terminal amino acid residues resolved. These tail segments are predicted to be unstructured and have been shown to be easily cleaved in solution [21]. Despite the lack of the N- and C-terminal segments, structural and functional properties computed and measured by simulations and experimental methods in several previous works [23–26] are in concordance. In the *Mt*-trHbO case, only 6 N-terminal residues are not resolved in the crystal structure. Mutant proteins, i.e. ProE3Gly *Tf*-trHbO and GlyE3Pro *Mt*-trHbO, were built in silico by changing the corresponding amino acid in the original structure and allowing the system to equilibrate in order to avoid any possible clash. Fixed protonation states were assumed to correspond to those at physiological pH, all solvent-exposed His were protonated at the N-δ delta atom, as well as HisF8, which is coordinated to the heme iron. The systems were immersed in a pre-equilibrated octahedral box with ~4910 TIP3P water molecules, where the minimum distance between the protein and the extreme of the box was 10 Å. Residue parameters correspond to AMBER ff99SB force field [27] except for the heme which correspond to those developed [28] and widely used in several heme–protein studies [29–35]. All simulations were performed using periodic boundary conditions with a 9-Å cutoff and particle mesh Ewald (PME) summation method for treating the electrostatic interactions. The bond lengths involving hydrogen atoms were kept at their equilibrium distance using the SHAKE algorithm, while temperature and pressure were kept constant with the Langevin thermostat and barostat, respectively, as implemented in the AMBER11 package [36]. A minimization of the crystal structures was performed to optimize any possible structural clashes using the Sander module of AMBER11 package carrying out 1000 cycles of Steepest Descent algorithm. Equilibration protocol was as follows: the systems were heated slowly from 0 to 300 K for 20 ps at constant pressure (1 atm), with harmonic restraints of 80 kcal per mol Å^2 for all C_α atoms and then a pressure equilibration of the entire system simulated for 2 ns at 300 K with the same restrained atoms. In the case of the systems simulated at high temperature (360 K), after the equilibration at 300 K, the systems were heated from 300 K to 360 K for 20 ps and equilibrated for 2 ns. These equilibrated structures were the starting points for the production MD simulations. We performed 100 ns of MD of deoxy forms of wt and in silico mutants of both proteins using the AMBER11-parm99sb force field [27], each at both 300 K and 360 K. All production simulations were performed with the pmemd module of the AMBER11 package [36].

2.2. Generic engineering procedures

ProE3Gly *Tf*-trHbO and GlyE3Pro *Mt*-trHbO mutants were obtained by PCR on plasmid pET28b-trtfHb and pEt28b-MthbO respectively as a DNA template. Site-directed mutagenesis was conducted with the Quick change mutagenesis kit (Stratagene) according to the manufacturer's instructions, using complementary oligonucleotide pairs introducing the amino acid substitutions. The mutants were transformed into *Escherichia coli* XL1 Blue competent cells, selected on kanamycin plates, and screened by DNA sequencing. Plasmid DNA bearing the gene with the desired mutation was then transformed into *E. coli* BL21 (DE3) cells (Novagen) for expression.

2.3. Protein expression and purification

E. coli cells (BL21 DE3) expressing the wt trHbs and their mutants were grown in Luria–Bertani medium containing 30 mg/L kanamycin at 37 °C. When $OD_{600} = 0.6$ protein expression was induced overnight with 1 mM IPTG at 37 °C for wt *Tf*-trHbO and ProE3Gly *Tf*-trHbO and at 25 °C for wt *Mt*-trHbO and GlyE3Pro *Mt*-trHbO. Cells were recovered by centrifugation at 13,000 rpm, re-suspended in a minimum volume of lysis buffer (20 mM phosphate buffer pH 7 supplemented with complete Protease Inhibitor Cocktail (Roche)) and sonicated until the supernatant was reddish and clear. After centrifugation at 12,000 rpm for 40 min at 4 °C, the pellet was re-suspended in 20 mM phosphate buffer pH 7 containing 6 M urea under stirring for 1 h and centrifugated again in the same condition. The supernatant was loaded on a DEAE-cellulose column (Whatman International Ltd., Maidstone, UK) equilibrated with

20 mM phosphate buffer pH 7 (buffer A) and extensively washed with the same buffer. Protein was eluted with a linear gradient of NaCl in buffer A, dialyzed versus 100 mM phosphate buffer pH 7 and titered with hemin while monitored by an UV–Vis spectrophotometer. Purified reconstituted proteins were dialyzed versus buffer A and loaded again on DEAE-cellulose column in order to eliminate the excess of hemin. The protein purity was confirmed by a single band at about 15,000 Da on a SDS-PAGE.

2.4. Thermal unfolding analysis

Thermal unfolding experiments in the far-UV region were carried out by measuring circular dichroism spectra on a JascoJ-715 spectropolarimeter as reported in [37]. Melting profiles were measured on the cyano-met derivatives of all proteins in a 0.1 cm quartz cell at protein concentrations of 10–20 μM heme by following the circular dichroism signal at 225 nm. Denaturation free energies were determined in temperature-melting experiments at different guanidinium chloride (GdmCl) (Sigma-Aldrich, St. Louis, MO) concentrations, assuming a two-state model, in the framework of the Gibbs–Helmholtz equation:

$$\Delta G_{unf} = \Delta H_{T_m} - \frac{T\Delta H_{T_m}}{T_m} + \Delta C_p^{\circ}[T - T_m - T\ ln(T/T_m)].$$

Accordingly, the fitting of the melting profiles as a function of T yielded the unfolding enthalpy, ΔH_{T_m}, the melting temperature, T_m, and the heat capacity change associated with the unfolding process, ΔC_p°. Two sloped baselines were also fitted to take into account temperature-dependent effects uncorrelated with the cooperative folding process [37].

GdmCl denaturation experiments were carried out in 50 mM phosphate buffer containing 2 mM KCN at pH 7.0 and 25 °C. Experiments were carried out by monitoring the far-UV circular dichroism spectra. The circular dichroism signal at 225 nm as a function of GdmCl concentration was followed with a slit aperture of 4 nm. In the experiments (T = 25 °C) samples were equilibrated for 4 h in GdmCl before measurement. Full reversibility of the denaturation process was observed under these conditions. Solvent denaturation curves were fitted to the simple two state model according to the formalism described by Clarke et al. [38]:

$$\Delta G_{unf}^{\mathrm{GdmCl}} = \Delta G_{unf}^{water} - m[\mathrm{GdmCl}]$$

where $\Delta G_{\mathrm{unf}}^{\mathrm{GdmCl}}$ is the solvent-dependent denaturation free energy, $\Delta G_{\mathrm{unf}}^{water}$ is the unfolding free energy in water and m is a fitting constant representing the differential binding of GdmCl between the folded and unfolded state. Least squares minimizations were carried out by the MATLAB 6.0 program (South Natick, MA). Chemical unfolding experiments were also carried out by following the relevant heme absorption spectra in the UV–Vis region, in 20 mM phosphate buffer containing 2 mM KCN at pH 7.0 and 25 °C. The spectra thus obtained were deconvoluted by using a standard Singular Value Decomposition procedure, as implemented by Matlab 6.0, in order to identify possible heme coordination intermediates that contribute to the observed overall signal.

3. Results and discussion

3.1. Protein dynamics upon temperature shift

In order to address the differences in thermal stability between *Tf*-trHbO and *Mt*-trHbO, 100 ns of unconstrained MD simulations for each protein at room (300 K) and high temperature (360 K) were performed. Analysis of the Root Mean Square Deviation and Fluctuations (RMSD and RMSF), as well as a close monitoring of MD simulations, indicate no difference, in terms of protein structural motions, between the thermostable and non-thermostable proteins at room temperature. The average ΔRMSF value between the two simulations was 0.08 ± 0.22 Å, in disagreement with the hypothesis that the thermostable species be less flexible than its thermostable homologue at room temperature. An analysis of the RMSD from the simulations confirms that *Tf*-trHbO conserves its overall fold, even at high temperatures (Table 1). Furthermore, these results reveal that wt *Tf*-trHbO concentrates a high degree of fluctuations in a flexible loop between C and E helices, called the CD loop (Figs. 1A, 2A). On the other hand, wt *Mt*-trHbO presents a very rigid CD loop and concentrates many fluctuations at both amino and carboxyl-terminal regions at high temperature (Figs. 2B, 1B, respectively). Despite the limited flexibility of the CD loop, the protein still manages to sample the two conformations found in the *Mt*-trHbO crystal structure. The increased flexibility at the terminal residues of the protein might be the onset of the unfolding process.

The CD loop distortion observed in *Tf*-trHbO at 360 K is correlated with an incipient breakage of the E α-helix, contiguous with the CD loop (Fig. 3A). As is well known, due to side chain constraints and sterics, prolines are unable to satisfy the Ramachandran angles requisite for α-helix formation. On one hand, then, the proline side chain is cramped into the α-helix backbone space and, on the other, the methylene group is in the space that would normally be occupied by an H-bonding amide proton, thus disrupting the H-bond network of the helix. As a consequence, ProE3 seems to be responsible for both destabilizing the E α-helix formation by breaking the H-bond between AspCD5 and AlaE4 residues (Fig. 3), and for increasing the flexibility of the CD loop (Fig. 2). *Mt*-trHbO has a glycine at E3 (GlyE3) that maintains the H-bond between AspCD5 and AlaE4, thus stabilizing the E α-helix.

It is important to note that both proteins keep a very stable structure similar to that found in the crystal structures when the simulations were performed at 300 K. Both CD loops and helices remain conserved.

Table 1
Average RMSD and standard deviation values (Å) for wt and mutant proteins, i.e. ProE3Gly *Tf*-trHbO and GlyE3Pro *Mt*-trHbO at different temperatures, 300 K and 360 K. RMSD values were computed using the complete protein, all but the terminal residues, only the terminal residues, or both the terminal and CD loop residues. Terminal residues included the first and last 12 residues.

		Tf-trHbO		*Mt*-trHbO	
		300 K	360 K	300 K	360 K
Wild type	Complete	1.10 ± 0.13	1.98 ± 0.33	1.08 ± 0.14	3.03 ± 0.62
	No terminal	1.00 ± 0.11	1.48 ± 0.24	0.87 ± 0.14	1.35 ± 0.23
	Only terminal	1.25 ± 0.19	2.18 ± 0.44	1.37 ± 0.24	5.21 ± 1.30
	Terminal and CD loop	1.31 ± 0.20	2.58 ± 0.47	1.35 ± 0.22	4.87 ± 1.17
Mutant	Complete	1.21 ± 0.24	1.91 ± 0.41	1.51 ± 0.32	2.03 ± 0.37
	No terminal	1.08 ± 0.23	1.35 ± 0.18	1.15 ± 0.14	1.46 ± 0.22
	Only terminal	1.37 ± 0.30	2.65 ± 0.90	1.63 ± 0.31	2.84 ± 0.76
	Terminal and CD loop	1.40 ± 0.31	2.46 ± 0.75	1.59 ± 0.27	2.69 ± 0.65

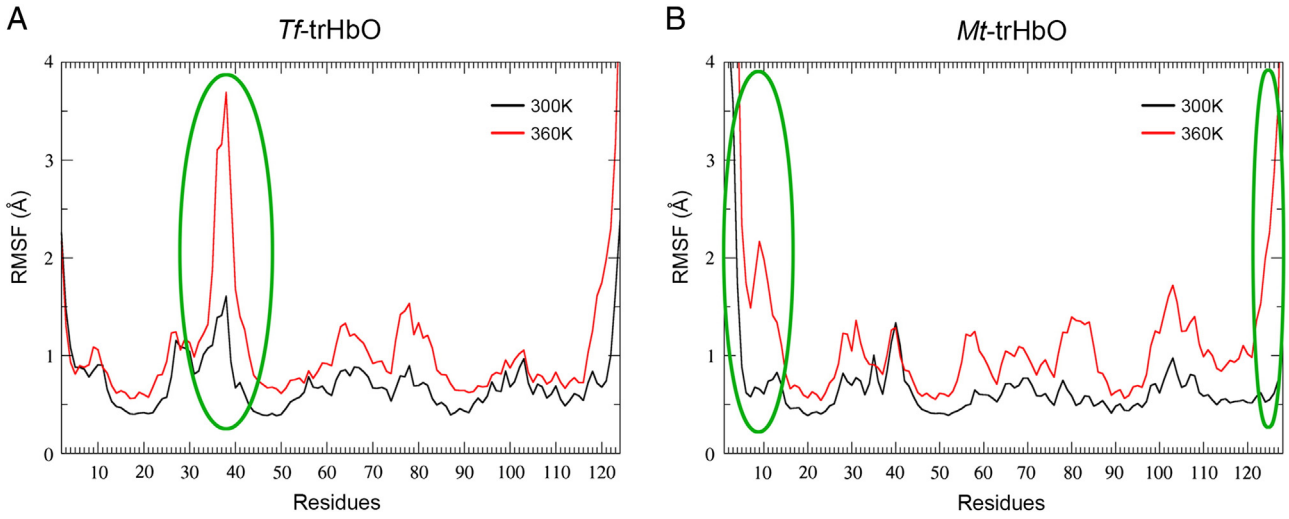

Fig. 1. RMSF analysis for wt *Tf*-trHbO (A) and *Mt*-trHbO (B) proteins at 300 K (black lines) and 360 K (red lines). Significant differences are circled in green. Regions between residues 35 to 42 in *Tf*-trHbO (A), and 37 to 44 in *Mt*-trHbO (B), correspond to the CD loop. Preferred magnification factor: full width.

3.2. In silico mutant protein dynamics

Taking into account that the residue at the E3 position seems to be responsible for the CD loop flexibility at high temperatures, and thus involved in thermal stability, we performed MD simulations of the two swapped mutants: a *Tf*-trHbO mutant in which ProE3 was replaced by Gly (ProE3Gly *Tf*-trHbO), and *Mt*-trHbO mutant in which GlyE3 was replaced by Pro (GlyE3Pro *Mt*-trHbO). The presence of GlyE3 in *Tf*-trHbO restricts in a very significant way the CD loop flexibility and increases its N-terminal flexibility (Fig. 4A), also disposing a reasonable local structural accommodation for the E α-helix. On the other hand, the presence of ProE3 in *Mt*-trHbO decreases its C-terminal flexibility degree, and increases slightly its CD loop motion (Fig. 4B). In the GlyE3Pro *Mt*-trHbO case, no clear breakage of E α-helix was detected, perhaps due to the simulation timescale and/or force field limitations. Although it is important to note that all-atom classical MD simulations (typically on the ns timescale) may not capture some larger conformational changes, the structural motions observed in our simulations were enough to address at least one significant dynamical difference between thermostable and non-thermostable studied proteins that allow us to reach conclusions that are in agreement with the experimental measurements mentioned in Section 3.4.

These results revealed by in silico mutant simulations support a key role of the E3 residue in the dynamics of the CD loop and suggest an interesting connection to the global fold's thermal stability.

3.3. Impacts on polar interactions

Sequence analysis highlights an overall enrichment in charged residues for the thermostable over the mesophilic homologue (54 vs 41). Despite the observed differences in the total number of charged residues, we note that the number of polar interactions (H-bonds and salt-bridges) observed through MD simulations is similar at room temperature (Table 2). Upon increasing the temperature to 360 K, both proteins undergo a loss of polar interactions, mainly the protein-solvent ones (Table 2). The observed trend pertaining to salt-bridge interactions is, however, quite different: wt *Tf*-trHbO shows an increase of ~6 salt-bridge interactions at 360 K due to the gain in CD loop flexibility, which facilitates salt-bridge interactions between the CD loop and B α-helix (Fig. 6). The GlyE3Pro *Mt*-trHbO follows the same trend and confirms that a single residue mutation promotes increased CD loop flexibility and consequently new salt-bridge interactions (Table 2). In summary, both proteins offer comparable amounts of polar interactions in the folded state, and H-bond interactions seem to be similarly affected by the increase in temperature, regardless of their thermostability.

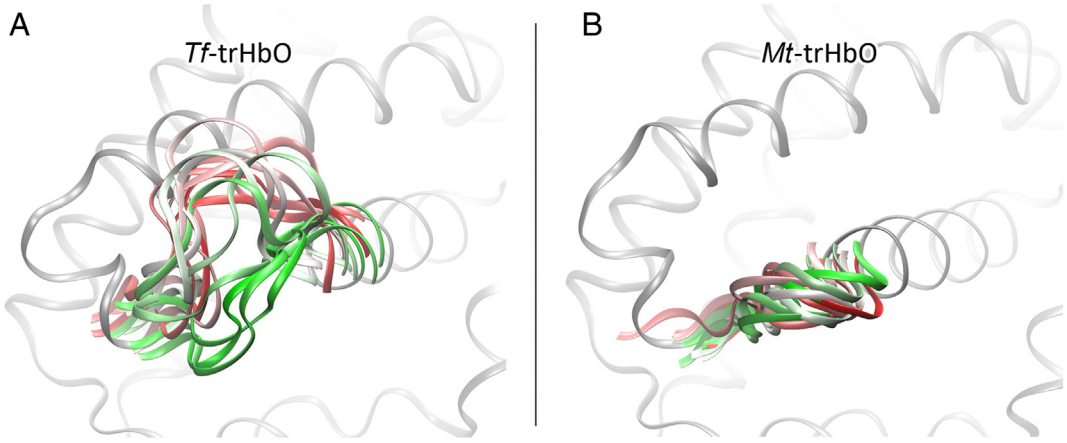

Fig. 2. CD loop conformations during MD simulations for wt *Tf*-trHbO (A) and *Mt*-trHbO (B) at 360 K. Snapshots are colored according to timestep, on a green–white–red color scale. Preferred magnification factor: full width.

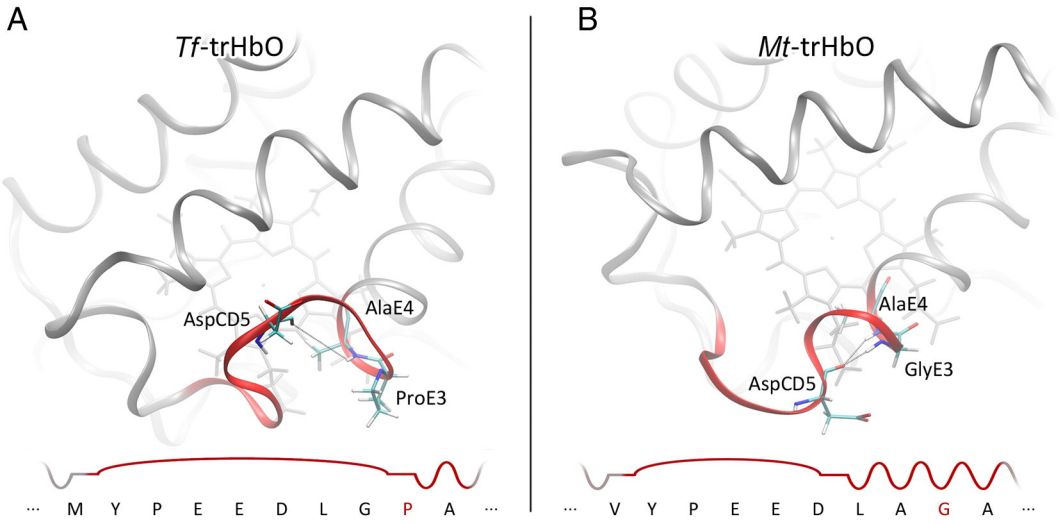

Fig. 3. The breakage and formation of E α-helix complex (red) after MD simulations for *Tf*-trHbO (A) and *Mt*-trHbO (B), respectively, at 360 K. The H-bond interaction between AspCD5 and AlaE4 in *Mt*-trHbO resulted in a reduction of the CD loop length and an extension of the E α-helix. Inset: amino acid sequence and secondary structure information for the CD loop and beginning of E α-helix. Pro and Gly residues at E3 position are highlighted in red. Preferred magnification factor: full width.

On the other hand, a clear difference was detected in the number of salt-bridge interactions which tended to increase in the presence of ProE3 and at higher temperature.

3.4. Thermal melting measurement of wt and swapped mutants

For the purpose of determining and characterizing possible thermodynamic properties that could be modified by changing only the E3 residue, the temperature dependence of the circular dichroism signal at 225 nm was examined for wt and mutant *Tf*-trHbO and *Mt*-trHbO cyano-met holoproteins, respectively, as a function of GdmCl concentration. Reversibility of the unfolding process was also measured at the higher guanidine concentration used (see supporting information, Fig. S1B). The observed profiles, fitted to a single, two-state unfolding process under the assumption of temperature independent ΔC_p (Fig. S1), bring about a clear picture on the thermodynamic determinants of thermostability in both proteins and their E3-swapped mutants. As expected, analysis of the thermal melting profiles data indicate that *Tf*-trHbO is indeed significantly more thermostable than *Mt*-trHbO under all conditions examined, showing a T_m of 348 K compared to 330 K. It is also noteworthy that *Tf*-trHbO, *Mt*-trHbO, and their mutants are characterized by similar ΔC_p values (see Table 3) and comparable solvent stabilities, as indicated by isothermal GdmCl denaturation experiments (Fig. S2). The thermal behavior of *Tf*-trHbO with respect to the ProE3Gly mutant is intriguing in that a sizeable decrease in T_m and ΔC_p values is observed in the mutated protein (see Fig. 5). In particular, the T_m values of wt and mutated proteins, 348 K and 344 K respectively, are strongly divergent in buffer but merge by increasing GdmCl concentration (see Fig. 5). In other words, GdmCl appears to "quench" the effect of the mutation on the thermal melting profile, such that, at GdmCl concentrations higher than 2 M, the unfolding line shapes are almost superimposable in the native protein and in the mutant. In structural terms, the new salt-bridge interactions that stabilize the protein native state appear to be destabilized by the presence of a highly polar cosolvent, where GdmCl molecules have a complete access from solvent environment (Fig. 6). In this way, the same result as in ProE3Gly mutation was achieved, where salt-bridge interactions between CD loop and B α-helix are destabilized by other reason, the reduction of the CD loop length (Fig. 3B). In this context, the thermodynamic interpretation is consistent with the role of

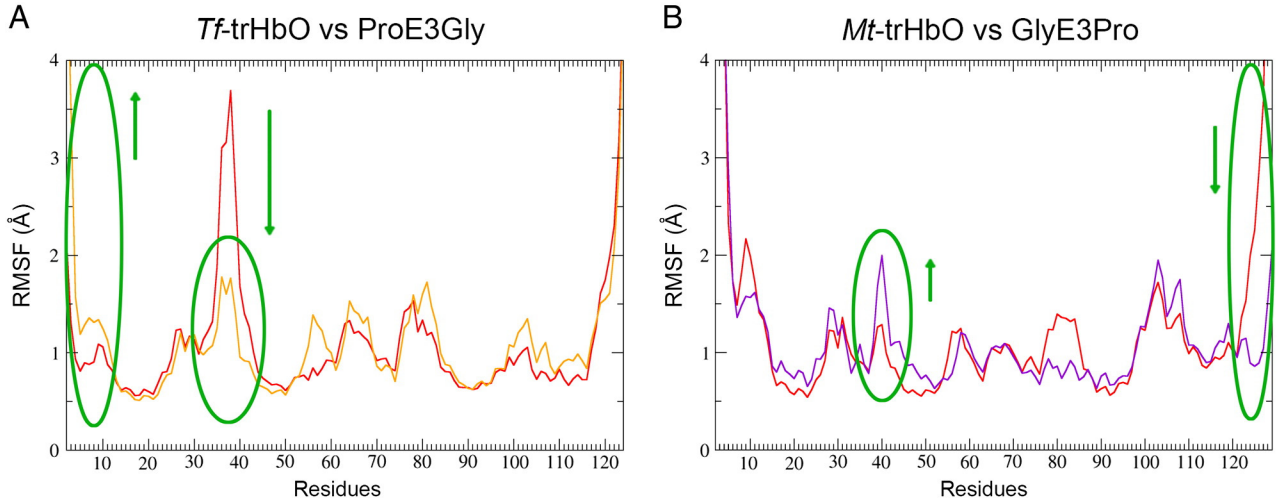

Fig. 4. RMSF analysis for wt (red) vs ProE3Gly (orange) *Tf*-trHbO (A) and wt (red) vs GlyE3Pro (violet) *Mt*-trHbO proteins (B), all at 360 K. Significant differences in both CD loops and extremes fluctuations are circled in green. Arrows indicate changes in fluctuations upon in silico mutation. Preferred magnification factor: full width.

Table 2
Number of polar interactions (salt-bridges and H-bonds) calculated for wt and mutant proteins at 300 K and 360 K during 100 ns of unconstrained MD simulations are shown. Salt-bridges and H-bonds are defined considering a cut-off distance less than 3.3 Å between the two charged or electronegative atoms, respectively.

		Tf-trHbO		*Mt*-trHbO	
		300 K	360 K	300 K	360 K
Wild type	Salt-bridges	17 ± 1	23 ± 2	17 ± 2	16 ± 2
	H-bonds-intra	53 ± 5	50 ± 5	57 ± 5	46 ± 4
	H-bonds-total	230 ± 12	189 ± 11	241 ± 13	201 ± 12
Mutant	Salt-bridges	17 ± 1	18 ± 2	17 ± 2	23 ± 1
	H-bonds-intra	54 ± 5	47 ± 5	55 ± 5	50 ± 5
	H-bonds-total	233 ± 14	190 ± 12	238 ± 14	201 ± 12

relevant polar interactions and their dependence on CD loop extension as mentioned before. The energetic balance, i.e. the $\Delta\Delta G_{unf}$ (wt *minus* mutant) of about 1 kcal/mol, is also consistent with the raised scenario.

Thermal and chemical unfolding data for *Mt*-trHbO and its GlyE3Pro mutant indicate a reverse, though not specular, picture with respect to *Tf*-trHbO and its ProE3Gly mutant. Actually, the substitution of GlyE3 with proline, in a topologically analogous position with respect to *Tf*-trHbO, brings about a clear increase in the T_m, from 330 K to 338 K, ΔH_{unf} and $\Delta\Delta G_{unf}$ (wt *minus* mutant) increase of 0.5 kcal/mol. At variance with *Mt*-trHbO (see Fig. 5B), however, the quenching effect of GdmCl is observed to a lesser extent with respect to *Tf*-trHbO. On this basis, the structural determinants for the increase in thermostability upon the GlyE3Pro mutation in *Mt*-trHbO is most likely explained in terms of an increase in CD loop flexibility that confers an additional structure with polar stabilizing interactions in the former protein (as is shown in Fig. 6 for wt *Tf*-trHbO). The topological positions within the CD loop and closest regions (ten residues in length, see Fig. 3, at the inset) thus appear to play a pivotal role in the unfolding process. This fact confirms that at higher temperatures, the presence of a ProE3 imposes a structural constraint by impairing the E α-helix formation and conferring higher flexibility to the CD loop, allowing new salt-bridge interactions with B α-helix that stabilize the protein native state.

Control experiments were also carried out by following GdmCl-dependent denaturation within the relevant UV–Vis absorption spectrum of the heme (Fig. S3). The measurements indicated that the overall process is certainly more complex than that described by a simple two-state model in that it entails cyanide release and most likely heme dissociation from the unfolded protein. Interestingly, from analysis of the deconvoluted absorption spectra (reported in Fig. S3 A and B for wt *Tf*-trHbO), a transient species bearing features of a five coordinated ferric heme is clearly identified in all proteins examined. The GdmCl concentration dependence of the observed UV–Vis spectra, reported in Fig. S3E, is qualitatively similar to that observed in CD experiments with almost identical apparent GdmCl concentration midpoints. However, the observed profiles, indicate a contribution from cyanide/heme dissociation that is apparent at high GdmCl concentrations. Cyanide release and accompanying heme coordination changes are not detectable in the CD experiments, due to the small contribution of heme absorption in the far UV region. The thermodynamic contribution of the cyanide dissociation process is thus comprised within the unfolding process and is assumed to be comparable in all proteins investigated in the present work.

4. Conclusions

4.1. Flexibility at the temperature shift

The aim of this study was to shed light on the molecular mechanism of thermal stability in truncated hemoglobins. Our theoretical and experimental results demonstrated that besides some variations at sequence level, there are important dynamical and structural differences between thermostable and non-thermostable trHbs under high temperature perturbation. MD simulations allowed us to inquire whether thermostability correlates with the rigidity/flexibility of protein structure. Comparing wt trHbs belonging to thermophilic and mesophilic microorganisms, no significant differences on residue flexibilities were observed at room temperature. Similar results were observed previously by performing amide exchange experiments [14], although other theoretical and experimental evidence had shown differential degrees of flexibility for thermostable and non-thermostable proteins [9–13]. Here, on the other hand, a clear increase in CD loop flexibility was observed in the wt form of *Tf*-trHbO as compared to wt *Mt*-trHbO.

4.2. Polar interactions at different temperatures

No substantial differences were observed in the number of total polar interactions (H-bonds and salt-bridges) between thermostable and non-thermostable trHbs at room temperature. When the temperature is increased from 300 K to 360 K both proteins lose several H-bonds, mainly residue-water interactions. This means that both proteins are affected in a similar way by the temperature perturbation shift considering H-bond interactions. This fact may be related to a previous observation about the lack of first solvation shells at high temperatures [39,40].

It was mentioned that T_m is related to the difference between the energy of the folded state vs the unfolded but collapsed states. Although no data about the polar interactions in the unfolded (but collapsed) state were given, it is known that wt *Tf*-trHbO increases its number of salt_bridges upon temperature increase, which is possible due to structural accommodations where surface charged residues are now able to form salt-bridge interactions, such as shown in Fig. 6. This phenomenon involves a decrease in magnitude of the solvation free energy of a charged residue at high temperatures, which is in energetic balance with increased Coulombic and electrostatic interactions of side-chains

Table 3
Thermodynamic parameters obtained from the curve fitting of guanidine denaturation experiments shown in Fig. S2 (m and ΔG_{unf}) and from linear fitting of Fig. 1 (ΔC_p). Curve fitting for GdmCl experiments were carried out according to the formalism used by Clarke et al. [38]. Experimental conditions: T = 25 °C, 50 mM phosphate buffer at pH 7.0 containing 2 mM KCN.

		m (kcal · mol^{-1}) $_M^{-1}$	ΔG_{unf} kcal · mol^{-1}	ΔC_p kcal · mol^{-1} · K^{-1}	T_m K
Tf-trHbO	wt	2.7 ± 0.07	9.1 ± 0.02	1.56 ± 0.07	348
	ProE3Gly	2.3 ± 0.05	8.2 ± 0.02	1.22 ± 0.09	344
Mt-trHbO	wt	2.1 ± 0.05	7.4 ± 0.02	1.21 ± 0.07	330
	GlyE3Pro	2.4 ± 0.05	7.9 ± 0.01	1.28 ± 0.06	338

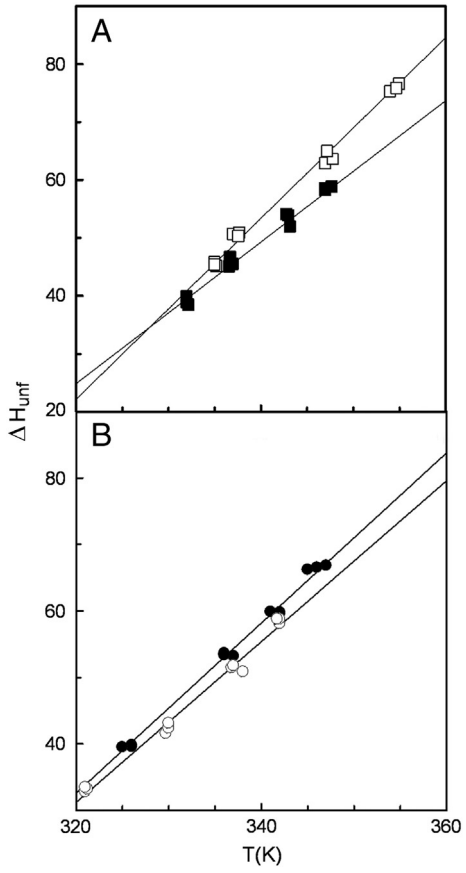

Fig. 5. Temperature dependence of ΔH_{T_m} for wt and ProE3Gly *Tf*-trHbO (A) and wt and GlyE3Pro *Mt*-trHbO (B) trHbs. The thermodynamic parameters ΔH_{T_m} and T_m, obtained from the fitting procedure of the melting profiles of Fig. 1S are shown as a function of GdmCl concentration (0.5, 1, 1.5 and 2 M, respectively). The slopes yielded values of ΔC_p (kcal · mol^{-1} · K^{-1}) of 1.56 ± 0.07 for wt *Tf*-trHbO (□), 1.22 ± 0.07 for ProE3Gly *Tf*-trHbO (■) and 1.21 ± 0.07 for wt *Mt*-trHbO (○), 1.28 ± 0.06 for GlyE3Pro *Mt*-trHbO (●), respectively. Preferred magnification factor: single column.

of surface charged residues, enhancing the global fold stability against temperature perturbation. However, the number of salt-bridge interactions does not increase in wt *Mt*-trHbO, due to a structural inability to form those salt-bridge interactions because of a short CD loop. Since in the GlyE3Pro *Mt*-trHbO protein the environment of the CD loop is almost the same as in wt *Tf*-trHbO, with a longer loop, the same increase in salt bridges was observed, gaining interactions that account for the folded state. Interestingly, a single residue, a proline, not only gives more flexibility to the CD loop, but also allows formation of new interactions that stabilize the protein native state.

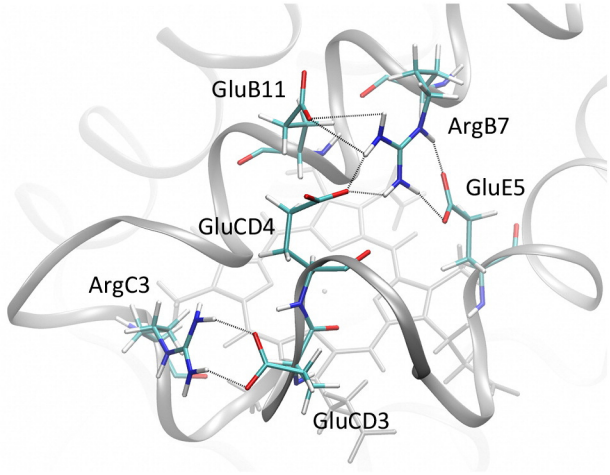

Fig. 6. The salt-bridge interactions between CD loop and B α-helix at the protein surface on *Tf*-trHbO is shown. It can be assumed that GdmCl molecules are able to disrupt this network interactions due to fact that they have complete access from solvent environment. Preferred magnification factor: single column.

4.3. Influence of E3 residue composition in thermal stability

It was proposed that prolines increase T_m by reducing the configurational entropy of the proteins for having more rigid backbones [17]. This interesting hypothesis considers only the primary structure of the protein, and does not consider the secondary and tertiary structure changes that these residues cause. By performing MD simulations of wt and swapped mutants of trHbs, we found that the presence of a specific proline, ProE3, allows the CD loop to gain flexibility in the native state. Melting temperature measurements of wt and the swapped mutants GlyE3Pro *Mt*-trHbO and ProE3Gly *Tf*-trHbO confirm that a specific residue is able to increases (by 8°) or decreases (by 4°) T_m in the corresponding protein. Therefore, we propose that the increase of the measured melting temperature is related to a tertiary distortion produced only by the ProE3. Specifically, ProE3 unlocks the CD loop which can now fluctuate more actively and interact with B α-helix complex residues, thus stabilizing the folded state.

In concordance with our results, all previously studied non-thermostable globin proteins belonging from mesophilic microorganisms like *M. tuberculosis* (trHbN), *Bacillus subtilis* (trHbO) and *Physeter catodon* (sperm whale Mb), show the same trend of *Mt*-trHbO, their loops (and particularly the CD loop) seem almost rigid [41–43], without displayingthe two conformations characterized here observed in *Tf*-trHbO at high temperatures. On the other hand, with a long and highly flexible CD loop [41], the hexacoordinated human neuroglobin shows a T_m of 100 °C [44], where the CD loop appears to be a critical region in its thermal stability, a fact probably related to the endogenous hexacoordination [44]. In this way, this study elucidated the importance of CD loop flexibility in truncated hemoglobins and other globin folds thermal stability, and our structural/dynamical approach can be applied as a thermostabilization strategy at least for this subfamily of proteins.

Acknowledgments

This work was supported by CONICET (grant PIP 11220080101207) and UBA, Argentina–Italia grant Progetti di grande Rilevanza, Ministero degli Affari Esteri (grant number 68876), CoMIUR FIRB RBFR08F41U_002 and Progetto Ateneo 2012 C26A1234TJ (A. Bonamore), and by Institute Pasteur, Fondazione Cenci Bolognetti project "Novel redox sensing pathways in parasitic microorganisms" (A. Boffi). J.P.B. holds a CONICET PhD fellowship. L.B. is a Pew Latin American Fellow. D.A.E. and A.D.N. are members of CONICET. J.P.B. thanks Marcelo A. Martí for revealing and very helpful discussions. We thank Mehrnoosh Arrar for close reading and editing of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagen.2014.03.018.

References

[1] P.M. Hicks, M.W.W. Adams, R.M. Kelly, Enzymes, extremely thermostable, in: M.C. Flickinger, S.W. Drew (Eds.), Encycl. Bioprocess Technol. Ferment. Biocatal. Biosep, John Wiley Sons, Inc., New York, 1999, pp. 987–1004.
[2] M.W. Day, B.T. Hsu, L. Joshua-Tor, J.-B. Park, Z.H. Zhou, M.W.W. Adams, D.C. Rees, X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaebacterium Pyrococcus furiosus, Protein Sci. 1 (11) (1992) 1494–1507.

[3] G. Auerbach, R. Ostendorp, L. Prade, I. Korndörfer, T. Dams, R. Huber, R. Jaenicke, Lactate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: the crystal structure at 2.1 Å resolution reveals strategies for intrinsic protein stabilization, Structure 6 (6) (1998) 769–781.
[4] C. Vieille, J. Hess, R. Kelly, J. Zeikus, xylA cloning and sequencing and biochemical characterization of xylose isomerase from *Thermotoga neapolitana*, Appl. Environ. Microbiol. 61 (5) (1995) 1867–1875.
[5] L. Chen, M.F. Roberts, Characterization of a tetrameric inositol monophosphatase from the hyperthermophilic bacterium *Thermotoga maritima*, Appl. Environ. Microbiol. 65 (10) (1999) 4559–4567.
[6] R. Jaenicke, Protein stability and molecular adaptation to extreme conditions, Eur. J. Biochem. 202 (3) (1991) 715–728.
[7] M.W.W. Adams, R.M. Kelly, Enzymes from microorganisms in extreme environments, Chem. Eng. News 73 (51) (1995) 32–42.
[8] R. Jaenicke, G. Böhm, The stability of proteins in extreme environments, Curr. Opin. Struct. Biol. 8 (6) (1998) 738–748.
[9] T. Lazaridis, I. Lee, M. Karplus, Dynamics and unfolding pathways of a hyperthermophilic and a mesophilic rubredoxin, Protein Sci. 6 (12) (1997) 2589–2605.
[10] G. Manco, E. Giosuè, S. D'Auria, P. Herman, G. Carrea, M. Rossi, Cloning, overexpression, and properties of a new thermophilic and thermostable esterase with sequence similarity to hormone-sensitive lipase subfamily from the archaeon *Archaeoglobus fulgidus*, Arch. Biochem. Biophys. 373 (1) (2000) 182–192.
[11] A. Gershenson, J.A. Schauerte, L. Giver, F.H. Arnold, Tryptophan phosphorescence study of enzyme flexibility and unfolding in laboratory-evolved thermostable esterases, Biochemistry 39 (16) (2000) 4658–4665.
[12] P. Závodszky, J. Kardos, Svingor, G.A. Petsko, Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins, Proc. Natl. Acad. Sci. U. S. A. 95 (13) (1998) 7406–7411.
[13] T. Mamonova, A. Glyakina, O. Galzitskaya, M. Kurnikova, Stability and rigidity/flexibility—two sides of the same coin? Biochim. Biophys. Acta 1834 (5) (2013) 854–866.
[14] G. Hernández, F.E. Jenney Jr., M.W.W. Adams, D.M. LeMaster, Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature, Proc. Natl. Acad. Sci. U. S. A. 97 (7) (2000) 3166–3170.
[15] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: a synthesis, Proteins Struct. Funct. Genet. 21 (3) (1995) 167–195.
[16] G. Vogt, S. Woell, P. Argos, Protein thermal stability, hydrogen bonds, and ion pairs, J. Mol. Biol. 269 (4) (1997) 631–643.
[17] B.W. Matthews, H. Nicholson, W.J. Becktel, Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding, Proc. Natl. Acad. Sci. U. S. A. 84 (19) (1987) 6663–6667.
[18] G. Bohm, R. Jaenicke, Relevance of sequence statistics for the properties of extremophilic proteins, Int. J. Pept. Protein Res. 43 (1) (1994) 97–106.
[19] F. Sterpone, S. Melchionna, Thermophilic proteins: insight and perspective from in silico experiments, Chem. Soc. Rev. 41 (2012) 1665–1676.
[20] L. Bleicher, E.T. Prates, T.C.F. Gomes, R.L. Silveira, A.S. Nascimento, A.L. Rojas, A. Golubev, L. Martínez, M.S. Skaf, I. Polikarpov, Molecular basis of the thermostability and thermophilicity of laminarinases: X-ray structure of the hyperthermostable laminarinase from rhodothermus marinus and molecular dynamics simulations, J. Phys. Chem. B 115 (24) (2011) 7940–7949.
[21] A. Bonamore, A. Ilari, L. Giangiacomo, A. Bellelli, V. Morea, A. Boffi, A novel thermostable hemoglobin from the Actinobacterium *Thermobifida fusca*, FEBS J. 272 (16) (Aug. 2005) 4189–4201.
[22] M. Milani, P.-Y. Savard, H. Ouellet, P. Ascenzi, M. Guertin, M. Bolognesi, A TyrCD1/TrpG8 hydrogen bond network and a TyrB10TyrCD1 covalent link shape the heme distal site of *Mycobacterium tuberculosis* hemoglobin O, Proc. Natl. Acad. Sci. U. S. A. 100 (10) (May 2003) 5766–5771.
[23] F.P. Nicoletti, E. Droghetti, L. Boechi, A. Bonamore, N. Sciamanna, D. a Estrin, A. Feis, A. Boffi, G. Smulevich, Fluoride as a probe for H-bonding interactions in the active site of heme proteins: the case of *Thermobifida fusca* hemoglobin, J. Am. Chem. Soc. 133 (51) (Dec. 2011) 20970–20980.
[24] F.P. Nicoletti, E. Droghetti, B.D. Howes, J.P. Bustamante, A. Bonamore, N. Sciamanna, D.A. Estrin, A. Feis, A. Boffi, G. Smulevich, H-bonding networks of the distal residues and water molecules in the active site of *Thermobifida fusca* hemoglobin, Biochim. Biophys. Acta-Proteins Proteomics 1834 (2013) 1901–1909.
[25] E. Droghetti, F.P. Nicoletti, A. Bonamore, L. Boechi, P. Arroyo Mañez, D. a Estrin, A. Boffi, G. Smulevich, A. Feis, Heme pocket structural properties of a bacterial truncated hemoglobin from *Thermobifida fusca*, Biochemistry 49 (49) (Dec. 2010) 10394–10402.
[26] A. Marcelli, S. Abbruzzetti, J.P. Bustamante, A. Feis, A. Bonamore, A. Boffi, C. Gellini, P.R. Salvi, D.A. Estrin, S. Bruno, C. Viappiani, P. Foggi, Following ligand migration pathways from picoseconds to milliseconds in type II truncated hemoglobin from *Thermobifida fusca*, PLoS One 7 (7) (Jan. 2012) e39884.
[27] J. Wang, P. Cieplak, P.A. Kollman, How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J. Comput. Chem. 21 (12) (2000) 1049–1074.
[28] M.A. Marti, L. Capece, A. Bidon-Chanal, A. Crespo, V. Guallar, F.J. Luque, D.A. Estrin, Nitric oxide reactivity with globins as investigated through computer simulation, Methods Enzymol. 437 (2008) 477–498.
[29] M.A. Marti, A. Crespo, L. Capece, L. Boechi, D.E. Bikiel, D.A. Scherlis, D.A. Estrin, Dioxygen affinity in heme proteins investigated by computer simulation, J. Inorg. Biochem. 100 (4) (2006) 761–770.
[30] D.E. Bikiel, L. Boechi, L. Capece, A. Crespo, P.M. De Biase, S. Di Lella, M.C. González Lebrero, M.A. Martí, A.D. Nadra, L.L. Perissinotti, D.A. Scherlis, D.A. Estrin, Modeling heme proteins using atomistic simulations, Phys. Chem. Chem. Phys. 8 (48) (2006) 5611–5628.
[31] F. Forti, L. Boechi, D. Bikiel, M.A. Martí, M. Nardini, M. Bolognesi, C. Viappiani, D.A. Estrin, F.J. Luque, Ligand migration in *Methanosarcina acetivorans* protoglobin: effects of ligand binding and dimeric assembly, J. Phys. Chem. B 115 (46) (2011) 13771–13780.
[32] L. Capece, A. Lewis-ballester, M.A. Marti, D.A. Estrin, S. Yeh, Molecular basis for the substrate stereoselectivity in tryptophan dioxygenase, Biochemistry 50 (2011) 10910–10918.
[33] P. Arroyo Mañez, C. Lu, L. Boechi, M.A. Martí, M. Shepherd, J.L. Wilson, R.K. Poole, F.J. Luque, S.-R. Yeh, D.A. Estrin, Role of the distal hydrogen-bonding network in regulating oxygen affinity in the truncated hemoglobin III from *Campylobacter jejuni*, Biochemistry 50 (19) (2011) 3946–3956.
[34] D. Giordano, L. Boechi, U. Samuni, A. Vergara, M.A. Martı, A. Estrin, J.M. Friedman, L. Mazzarella, G. Prisco, L. Grassi, The hemoglobins of the sub-antarctic fish *Cottoperca gobio*, a phyletically basal species — oxygen-binding equilibria, kinetics and molecular dynamics, FEBS J. 276 (2009) 2266–2277.
[35] L.L. Perissinotti, M.A. Marti, F. Doctorovich, F.J. Luque, D.A. Estrin, A microscopic study of the deoxyhemoglobin-catalyzed generation of nitric oxide from nitrite anion, Biochemistry 47 (37) (2008) 9793–9802.
[36] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham III, S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, Comput. Phys. Commun. 91 (1–3) (1995) 1–41.
[37] A. Matouschelt, J.M. Matthews, C.M. Johnson, A.R. Fersht, Extrapolation to water of kinetic and equilibrium data for the unfolding of barnase in urea solutions, Protein Eng. 7 (9) (1994) 1089–1095.
[38] J. Clarke, A.R. Fersht, Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation, Biochemistry 32 (16) (1993) 4322–4329.
[39] P.I.W. De Bakker, P.H. Hünenberger, J.A. McCammon, Molecular dynamics simulations of the hyperthermophilic protein Sac7d from *Sulfolobus acidocaldarius*: contribution of salt bridges to thermostability, J. Mol. Biol. 285 (4) (1999) 1811–1830.
[40] G.I. Makhatadze, V.V. Loladze, D.N. Ermolenko, X. Chen, S.T. Thomas, Contribution of surface salt bridges to protein stability: guidelines for protein engineering, J. Mol. Biol. 327 (5) (2003) 1135–1148.
[41] L. Capece, M.A. Marti, A. Bidon-Chanal, A. Nadra, F.J. Luque, D.A. Estrin, High pressure reveals structural determinants for globin hexacoordination: neuroglobin and myoglobin cases, Proteins 75 (4) (Jun. 2009) 885–894.
[42] L. Giangiacomo, A. Ilari, A. Boffi, V. Morea, E. Chiancone, The truncated oxygen-avid hemoglobin from *Bacillus subtilis*: X-ray structure and ligand binding properties, J. Biol. Chem. 280 (10) (2005) 9192–9202.
[43] A. Crespo, M.A. Martí, S.G. Kalko, A. Morreale, M. Orozco, J.L. Gelpi, F.J. Luque, D.A. Estrin, Theoretical study of the truncated hemoglobin HbN: exploring the molecular basis of the NO detoxification mechanism, J. Am. Chem. Soc. 127 (12) (Mar. 2005) 4433–4444.
[44] D. Hamdane, L. Kiger, S. Dewilde, J. Uzan, T. Burmester, T. Hankeln, L. Moens, M.C. Marden, Hyperthermal stability of neuroglobin and cytoglobin, FEBS J. 272 (8) (2005) 2076–2084.