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