Transworld Research Network
37/661 (2), Fort P.O.
Trivandrum-695 023
Kerala, India

Recent Advances in Pharmaceutical Sciences, 2011: 133-153 ISBN: 978-81-7895-528-5
Editor: Diego Muñoz-Torrero

6. Molecular simulations of globins: Exploring the relationship between structure, dynamics and function

Flavio Forti[1], Leonardo Boechi[2], Ana Novo de Oliveira[1], Damian Bikiel[2]
Pau Arroyo[2], Alejandro Nadra[3], Luciana Capece[2], Axel Bidon-Chanal[1]
Marcelo A. Martí[2,3], Darío Estrín[2] and F. Javier Luque[1]

[1]Departament de Fisicoquímica, Facultat de Farmàcia, and Institut de Biomedicina, Universitat de Barcelona, Avda. Diagonal 643, 08028 Barcelona, Spain; [2]Departamento de Química Inorgánica, Analítica, y Química Física, INQUIMAE-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
[3]Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales Universidad de Buenos Aires, Buenos Aires, Argentina

Abstract. The discovery in the last two decades of novel members of the globin superfamily has challenged the conventional view about the structure and function of globins. Thus, peculiar structural differences are expected to have direct influence on properties related to ligand migration, binding affinity and heme reactivity. Molecular simulations are a valuable tool to gain insigth into the molecular mechanisms that underlie those structural differences, and their relationship with the diversity of functional roles. In this work, the impact of molecular simulations in exploring

Correspondence/Reprint request: Dr. F. J. Luque, Departament de Fisicoquímica, Facultat de Farmàcia Universitat de Barcelona, Avda. Diagonal 643, 08028 Barcelona, Spain. E-mail: fjluque@ub.edu
Dr. D. Estrín Departamento de Química Inorgánica, Analítica y Química Física, INQUIMAE-CONICET Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
E-mail: dario@qi.fcen.uba.ar

the linkage between structure, dynamics and function is highlighted for three representative cases: the migration of ligands through the protein matrix of truncated hemoglobins, the modulation of binding affinity by heme distortion in protoglobin, and finally the functional implications due to the equilibrium between penta- and hexacoordination of the heme with distal histidine in neuroglobin.

Introduction

Globins are a family of heme-containing proteins found in all kingdoms of life. They belong to the hemeprotein superfamily, though they share some distinctive characteristics. In the Protein Data Bank around 2300 hemeproteins can be found, and they can be clustered in 34 different groups with very diverse structural and functional characteristics [1]. As members of this superfamily, globins are supposed to have evolved from a common ancestor and their characteristic tertiary structure is typically known as *globin fold*. This fold was identified in 1958 in myoglobin (Mb) [2], which was the first protein whose structure became solved by X-ray diffraction. In this sense, the *globin fold* was the first protein fold to be discovered. Though it originally consisted of a bundle of eight alpha helices, its generalized definition has been challenged in the last decades due to the discovery of new globins, which have a number of peculiar structural features. The *globin fold* is an all-alpha protein fold, since the only secondary structure found is the alpha-helix. Though primary sequences of globins can have as low as 16% sequence identity, the *globin fold* is highly conserved throughout the family.

Globins have evolved to play a variety of biological roles, such as transport and sensing of gases and catalysis of reactions between nitrogen and reactive oxygen species [3-5]. Some of them are present as monomers under physiological conditions, though others form multimeric species, as illustrated by the prototypical cases of mammalian Mb –monomeric– and hemoglobin (Hb) –tetrameric- [5-8]. As mentioned before, globins generally adopt a common *globin fold* characterized by a 3-over-3 helical sandwich (A/BC/E and F/G/H helices), which contains the hydrophobic pocket that accommodates the heme group. The heme iron is coordinated to the only fully conserved residue along this family: the proximal HisF8 [9], leaving the sixth coordination position in the distal side usually free for binding of the exogenous ligand. Typical exogenous ligands are NO, CO and O_2, being molecular oxygen the most abundant and the one with the lowest affinity for free heme. Therefore, O_2 affinity is a key parameter for gaining insight into the function of globins.

The affinity of a protein for a ligand is characterized by the equilibrium constant K, which in turn can be related with the ratio between the apparent kinetic rate constants for the association and dissociation processes, called k_{on} and k_{off} respectively (see Fig. 1 for the binding of O_2 to the ferrous form of a hemeprotein).

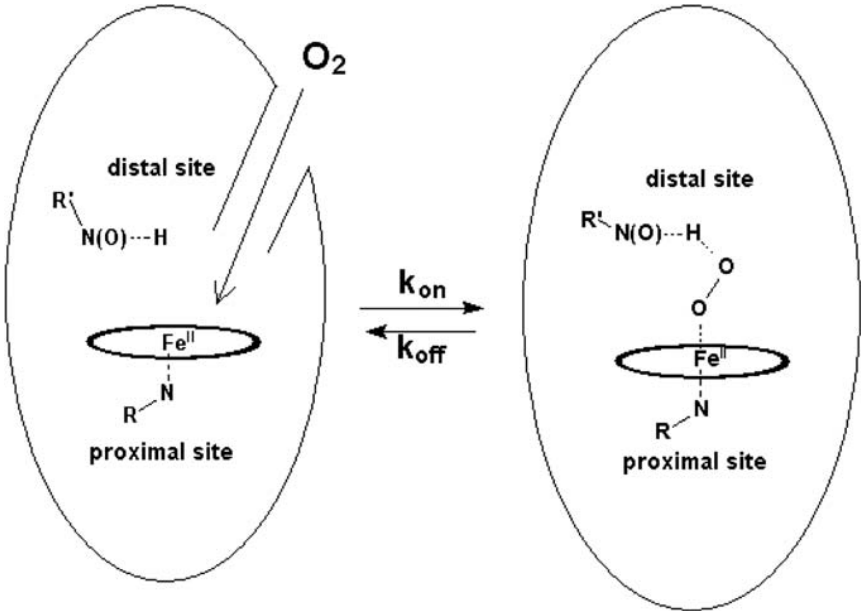


Figure 1. Schematic representation of the association and dissociation processes of small molecules to the heme moiety in globins.

In most globins, the association rate constant depends on two processes: i) ligand migration from the solvent to the heme active site, and ii) ligand coordination to the heme iron [3]. The ligand migration process is regulated by the presence of internal pockets or even tunnels [10-12], and the presence of specific residues that are capable to act as "gates" [7]. In some cases, a distal site residue (mainly His, Tyr) is able to bind to the Fe atom in the so-called *internal hexacoordination*, which affects the entire process [13-16]. On the other hand, non-coordinated water molecules located in the distal site can also modulate the k_{on}, adding another variable to the complexity of the association process [17-19]. Moreover, the coordination step depends on the spin state of the ligand and the relative in-plane position of the iron. For these reasons, in most cases the value of k_{on} for O_2 is higher than for CO, whereas the k_{on} value for NO is even higher [7]. The association rates span a wide range of values (spanning up to five orders of magnitude), starting at about 10^4 $M^{-1}s^{-1}$ in those systems with very low accessibility to the iron, and rising to 10^9 $M^{-1}s^{-1}$ when the association rate is mainly controlled by the diffusion from the solvent to reach the protein, as observed for isolated porphyrins [7,20].

The dissociation constant (k_{off}) involves two processes: i) thermal breaking of the protein-ligand interactions, and ii) ligand escape from the active site into the solvent. Dissociation rate constants are generally regulated mainly by the protein-ligand thermal breaking step, and span a range of roughly seven orders of magnitude (from 10^{-3} s^{-1} to 10^4 s^{-1} [3, 7, 21]). Both

diffusion and breaking of protein-ligand interaction are processes that vary from ligand to ligand, thus leading to a wide range of ligand affinities. Oxygen binds exclusively to ferrous (Fe_{II}) heme, and its dissociation rate constant is strongly influenced by interactions between the coordinated O_2 and the protein matrix [22]. For CO and NO, the dissociation from ferrous heme is mostly dominated by breaking of the Fe-ligand bond, and similarly low values ($\approx 1 \times 10^{-2}$ s^{-1}) are observed for many different proteins [3]. In the O_2 case, the energy required for breaking the protein-ligand bond is regulated by several factors, which include [23,24]:

- *distal effect*, which accounts for the interaction of the ligand with hydrogen-bond donor residues present in the distal cavity,
- *proximal effect*, which takes into account the influence of the local structure of the axial histidine, and
- *heme distortion*, which alters the strength of the Fe-ligand bond.

In the last decades new members of the globin family have been discovered, greatly expanding the globin world. Thus, globins are widely distributed and exhibit an intricate and complex phylogenetic network, which has been proposed to be divided in three main lineages [25]. Novel globins show distinct structural and functional features when compared to the emblematic mammalian Mb and Hb (Fig. 2). Apart from the canonical 3/3 Mb fold, which encompasses many different globins, another lineage shows a characteristic fold denoted as 2/2 Hb, which has been found in the three domains of life [25]. The 2/2 Hb fold, also referred to as *truncated* Hb, is around 20-30 residues shorter than Mb and exhibits a 2-over-2 sandwich fold involving only BC/E and G/H helices, in contrast to the classical 3-over-3 fold. The 2/2 Hbs have been proposed to act as small gas molecule sensors, oxygen carriers and pseudoenzimes.

Another of these lineages includes protoglobin and globin-coupled sensor (GCS) proteins, whose globin domain is bigger than Mb (~190 amino acids; [25-28]). Protoglobins are the first single-domain GCS-related globins found in *Archaea*. They can bind O_2, CO and NO reversibly *in vitro*, but so far their function is unknown [25,26,28] .

In the following sections we firstly describe the main features of Mb, which can be considered a prototypical example found in many textbooks, and three different globins with specific characteristics that make them to be representative examples of the structural and functional diversity of the globin family (Fig. 2):

- *Truncated hemoglobins*: Members of the 2/2 Hb subgroup, the second lineage of the globin family. In some cases, they have been related to the defense of the bacteria against nitrosative stress.

- *Protoglobin*: The 3D structure has been recently solved [29], though its function remains unknown. This protein hosts an unusually distorted heme and exhibits a very low dissociation constant for O_2.
- *Neuroglobin*: Member of the subgroup of globins with relevant internal hexacoordinated phenomena.

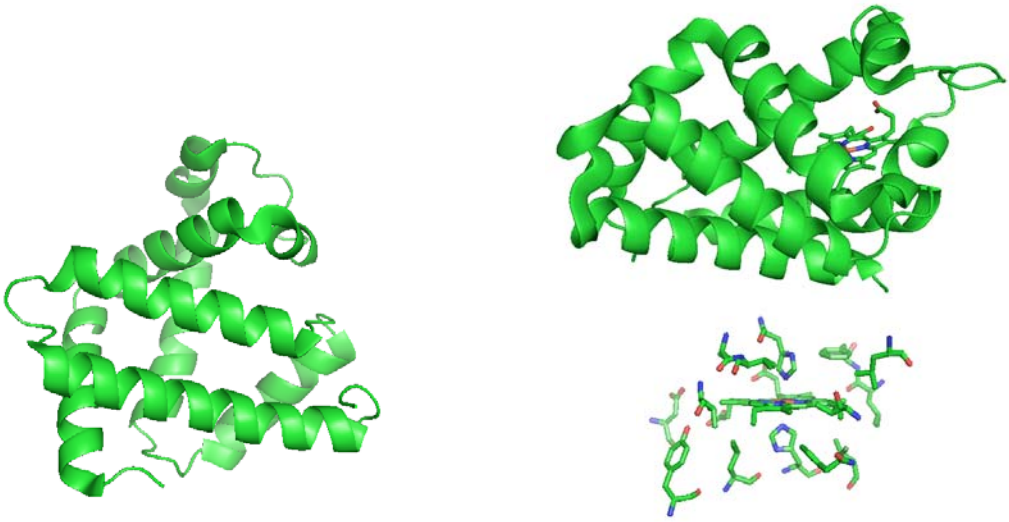
Brief survey of mammalian *myoglobin* (Mb)

Mb is not only a member of the first lineage of globins, but one of the most studied proteins. For this reason it is often referred to as the hydrogen atom of biology [30]. Its 3D structure was solved more that 50 years ago by J. Kendrew and coworkers [2], finding that deserved the Nobel Prize in Chemistry in 1962. Despite being one of the most studied proteins, there is still intense research effort due to its complexity, biological relevance and ongoing debate of its physiological function [31].

Mb is a cytoplasmic globin consisting of 154 amino acids expressed in cardiac mytocytes and oxidative skeletal muscle fibers. Like Hb, Mb reversibly binds O_2. However, Mb has a characteristic Michaelis-Menten hyperbolic O_2-saturation curve, while a sigmoid-shaped curve is seen in tetrameric Hb, thus reflecting the well known allosterism effect in this latter protein. From a biochemical point of view, Mb acts as O_2 storage protein in muscle, which is especially evident in marine mammals and birds that undergo extended periods of apnea, or in humans and other species living at high altitude. It has also been proposed as a buffer of intracellular O_2 pressure in a number of species [31], maintaining O_2 concentration relatively constant despite the occurrence of induced changes in O_2 level.

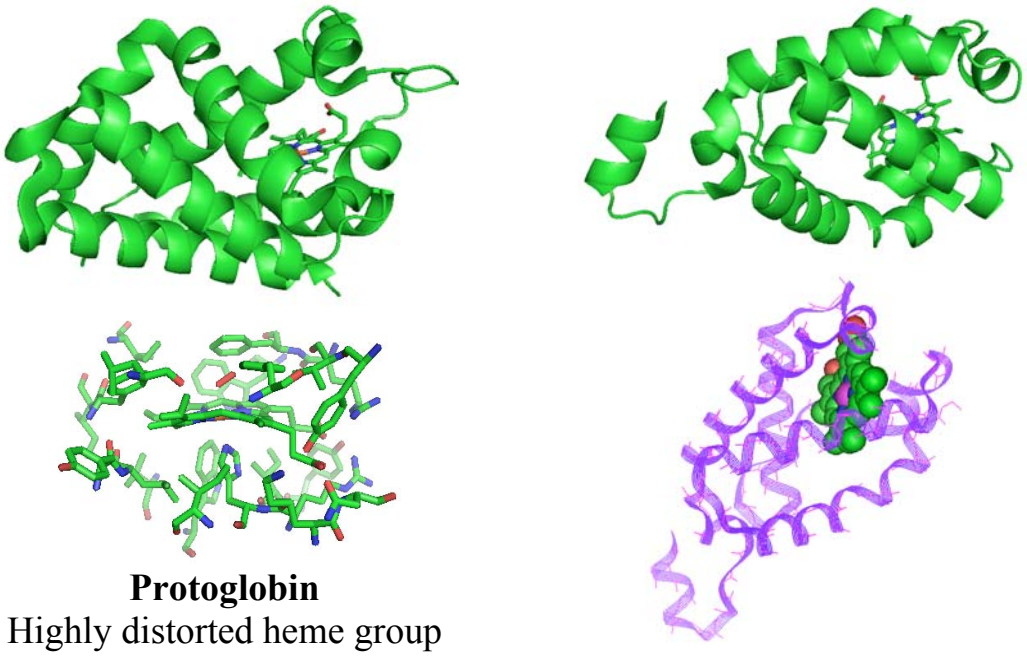
There is more controversy about the role of Mb in assisting O_2 diffusion in the cell. Desaturated Mb close to the cell membrane could bind O_2 and diffuse to the mitochondria, thus representing an alternative way of simple O_2 diffusion [31]. However, contrary to what it could be expected from its role, knockout experiments in mice with no Mb in skeletal muscles showed survival of the organism without severe biological consequences. This raises the question of whether other proteins could compensate for Mb absence. Beyond O_2 biochemistry, Mb has also been related to inactivation of NO and scavenging reactive O_2 species [31].

Besides the described roles and the increasing number of studies done so far, there are still open questions about Mb that require further research. For instance, what are the factors that regulate its expression in response to hypoxic situations? The discovery of other tissue globins such as neuroglobin (presented in the last section) and cytoglobin raises the question about the complexity of the underneath physiological model governing skeletal muscle biology.



3-on-3 globin fold
Myoglobin

Hexacoordination
Neuroglobin:



Protoglobin
Highly distorted heme group

Truncated hemoglobin N:
2-2 Sandwich fold

Figure 2. Globin diversity represented by *myoglobin (upper-left), neuroglobin (upper-right), protoglobin (lower left) and truncated hemoglobin N (lower right).*

1. 2/2 Hbs (truncated) proteins

The 2/2 Hb family can be further divided into 3 groups: I, II, III (also known as N, O, P). Group I shows a typical tunnel system that connects the solvent with the distal site. Groups II and III share a common TrpG8 and

generally do not exhibit well delineated tunnels for ligand migration. A network of hydrogen bonds stabilizing the exogenous ligand in the heme distal site is found in all three groups.

Truncated hemoglobin N of *Mycobacterium tuberculosis* (Mt-trHbN) is the most studied member of the 2/2 globins. For that reason, we will focus on the description of the mechanism of ligand migration from the solvent to the heme pocket in this protein, and later we compare these results with other Hbs belonging to groups O and P.

The Mt-trHbN case

Mycobacterium tuberculosis is responsible for tuberculosis in humans [32]. During the first stages of the infection the bacteria is attacked by macrophages that generate large amounts of NO [33]. It has been shown that certain Hbs that are present in some microorganisms are related to its ability to detoxify NO [34,35]. These defense mechanism would be related to O_2-bound globins that could convert NO to nitrate anion following the reaction Fe(II)-O_2 + NO → Fe(III) + NO_3^-. In *M. tuberculosis* two 2/2 Hbs are capable of performing such detoxifying reaction: *truncated Hb N* and *truncated Hb O*. One of the most interesting characteristics of these globins is the presence of an apolar tunel system that connects the solvent with the active site [36], which is postulated to be involved in migration of ligands (O_2, NO). In particular, in Mt-trNHb there are two perpendicular tunnels: the so-called Short Tunnel G8 (STG8), which is around 8 Å long and is delineated by residues in helices G and B, and the Long Tunnel (LT), which is around 20 Å long and is mainly defined by helices B and E (Fig. 3).

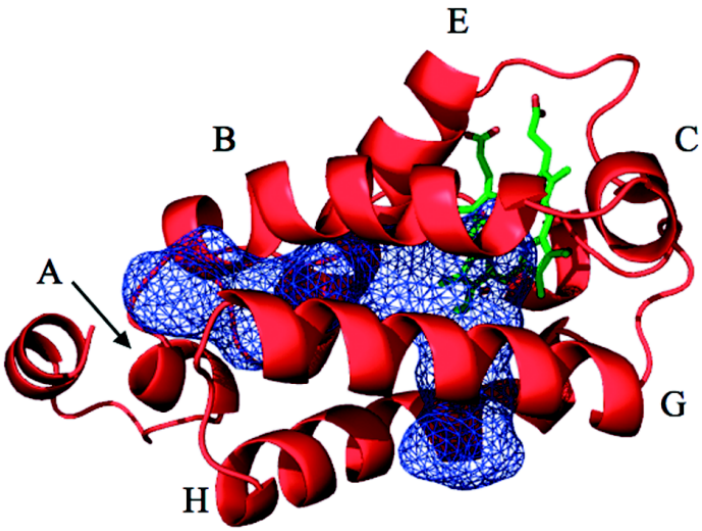


Figure 3. Representation of the two orthogonal branches of the apolar tunnel found in the protein matrix of Mt-trHbN.

In the crystal structure of the oxygenated globin (PDB code 1IDR [37]) PheE15, which is placed in the middle of the LT, shows two conformations defined by a rotation of ~63º along the C_α-$C_ß$ bond. In one conformation the benzene ring is parallel to the tunnel axis, while in the other it is roughly perpendicular. This suggests that the residue could act as a gate for ligand migration being these two conformations, which will be denoted open and closed states, respectively [37].

Molecular dynamics (MD) simulations of the oxygenated form of Mt-trHbN [10] showed several transitions between these two conformational states, in agreement with the conformational flexibility seen in the X-ray structure. In order to explore the functional implications of these findings, *Multiple Steered Molecular Dynamics* simulations coupled with Jarzinski's equality were used to obtain free energy profiles for NO migration through the tunnel for both open and closed conformations [10]. When PheE15 is in the open conformation, a small barrier (around 2 kcal/mol) has to be surpassed in order to access the heme cavity. On the other hand, for the closed state access to the active site is accompanied by a steep increase in the free energy, leading to a barrier of around 5 kcal/mol (Fig. 4). Similar studies performed for the STG8 show a higher barrier of around 7 kcal/mol. This means that NO entry for the oxygenated Mt-trHbN should mainly occur through the LT.

For the NO detoxifying reaction to take place, the protein must be firstly loaded with O_2. MD simulations run for the deoxygenated form of Mt-trHbN showed that PhE15 is only found in the closed state. This is consistent with the free energy barrier for the open/closed torsional transition obtained by

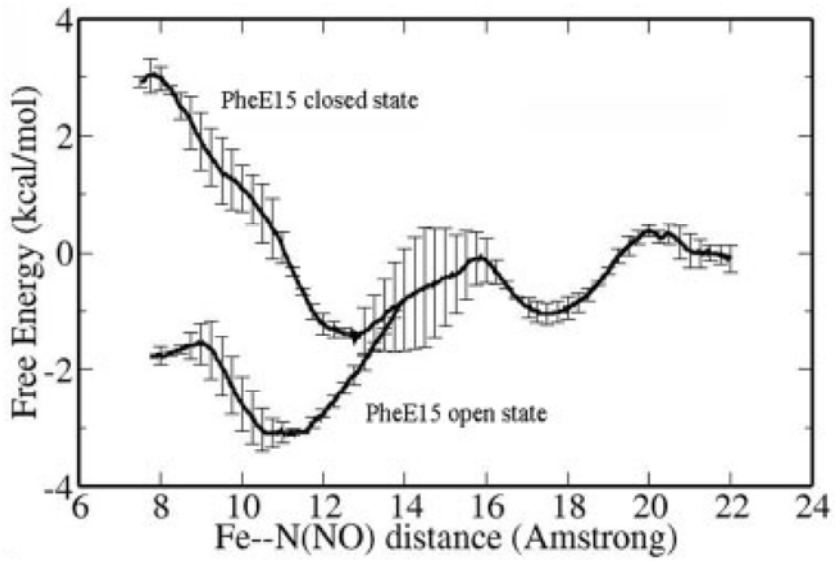


Figure 4. Free energy profile for the migration of NO through the LT in both open and closed states of the PheE15 gate.

Umbrella Sampling: while the open↔closed transition involves a barrier of 3 kcal/mol in the oxygenated protein, the barrier increases up to 6 kcal/mol in the deoxygenated protein. Oxygen entry is then supposed to be achieved through the STG8, as supported by the free energy profiles for ligand migration through this tunnel for the oxygenated and deoxygenated forms. While 7 kcal/mol are needed for NO to gain access to the active site in the oxygenated form, only 4 kcal/mol must be surpassed by O_2 in the deoxygenated form (Fig. 5).

This behavior suggests that once the Mt-trHbN-O_2 complex is formed, some residues sense the ligand in the distal cavity and favors the conformational change in PheE15, which triggers the aforementioned opening events. This allows NO to enter through the LT and reach the distal cavity for the detoxifying reaction to take place.

A plausible hypothesis to explain the sensing properties of Mt-trHbN relies on residues GlnE11 and TyrB10. In the deoxygenated protein those residues interact by hydrogen bonding between the amide group in GlnE11 and the hydroxyl group in TyrB10. MD simulations show that sometimes GlnE11 is acting as hydrogen-bond donor (and TyrB10 as acceptor), whereas in other snapshots the side chain carbonyl group of GlnE11 is hydrogen-bond acceptor (and TyrB10 is the donor). In all cases GlnE11 is primarily found in an extended *all-trans* conformation. This fluctuating hydrogen-bond network is drastically altered in the oxygenated protein. Thus, TyrB10 forms a hydrogen-bond with the heme-bound O_2 and forces GlnE11 to adopt a folded conformation in order to maintain a hydrogen-bond with TyrB10. In this conformational state, the side chain of GlnE11 is much closer to PheE15, and thermal fluctuations of the side chains would facilitate the opening of the gate.

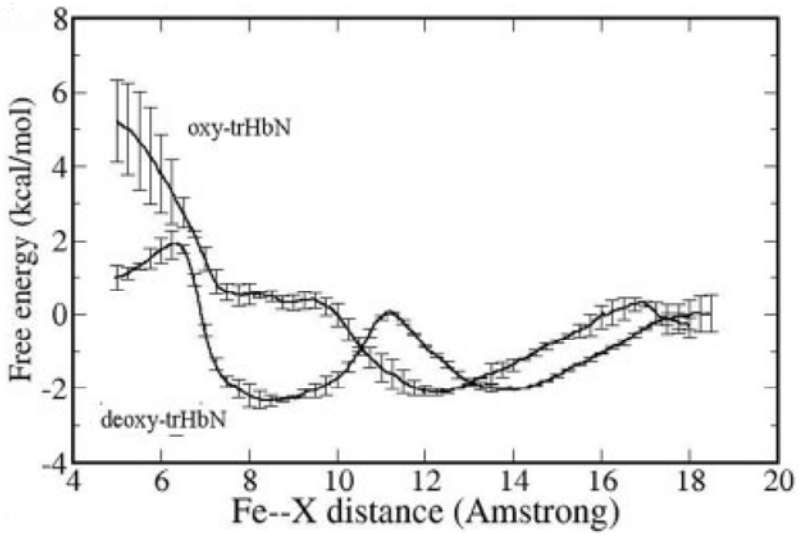


Figure 5. Free energy profile for the migration of NO through the STG8 in both deoxygenated and oxygenated states of the protein.

Support to this hypothesis also comes from MD simulations run for the TyrB10→Phe mutant [16]. Thus, the analysis of the trajectory reveals an increase in the distance between GlnE11 and PheE15, which implies a reduction in the mechanical pressure exerted by the former residue on the gate. This is due to the fact that in this mutant GlnE11 adopts an extended conformation, which enables the terminal amido group to form a hydrogen-bond directly with the heme-bound O_2. For the GlnE11→Ala mutant the protein is predicted to be also inactive, as the lack of the contacts between AlaE11 and PheE15 would make the gate to populate mainly the closed state in the mutant.

These results suggest that multiligand chemistry in Mt-HbN has evolved in such a way that there is a distinct access pathway to the active site for the molecules involved in the reaction: NO and O_2. Oxygen would first enter through the STG8 and bind to the heme. Upon oxygen binding, the hydrogen bond network formed by TyrB10 and GlnE11 is the key feature that regulates the mechanism that triggers the opening of PheE15 gate of the LT for NO entrance.

Other trHbs

Several efforts have been made in the last years to examine other trHbs by solving X-ray structures, or determining kinetic constants and spectroscopic data, as well as by using MD simulations. At this point, our group has been working on the ligand migration properties in *M. tuberculosis* trHb (Mt-trHbO) and *B. subtilis* trHb (Bs-trHbO).

For Mt-trHbO, even though a Leu residue occupies the E15 position (thus avoiding the PheE15 gate), the LT is blocked near the heme group. The only accessible tunnel for ligand migration is the Short Tunnel E7 (STE7), which is oriented toward the propionate groups (Fig. 6). This tunnel is topologically

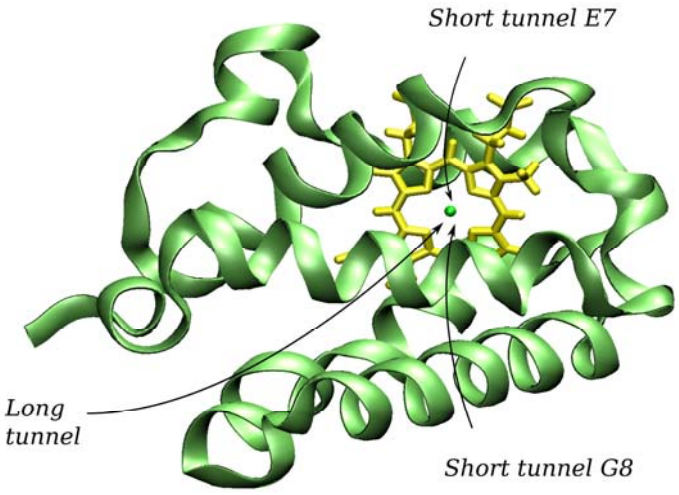


Figure 6. Representation of the backbone in Mt-trHbO and schematic view of the three main tunnels found in trHbs.

related to the E7 "gate" found in Mb. However, although STE7 showed a lower barrier for ligand migration than the LT one [11], it is even higher compared to that observed for Mt-trHbN, a fact that probably explains its low association rate.

Mutation of TrpG8 to smaller residues consistently lowers the O_2 entry barrier and increases the association rate, in agreement with the available experimental data [11]. These findings support the important role of TrpG8 in regulating the ligand migration through the LT (Fig. 7). In particular, the mutant TrpG8→Ala opens the STG8 (mentioned above for Mt-trHbN). Therefore, TrpG8 not only blocks LT, but also STG8. Noteworthy, sequence alignment shows that there is a Trp residue at position G8 along the whole O and P groups of this sub-family.

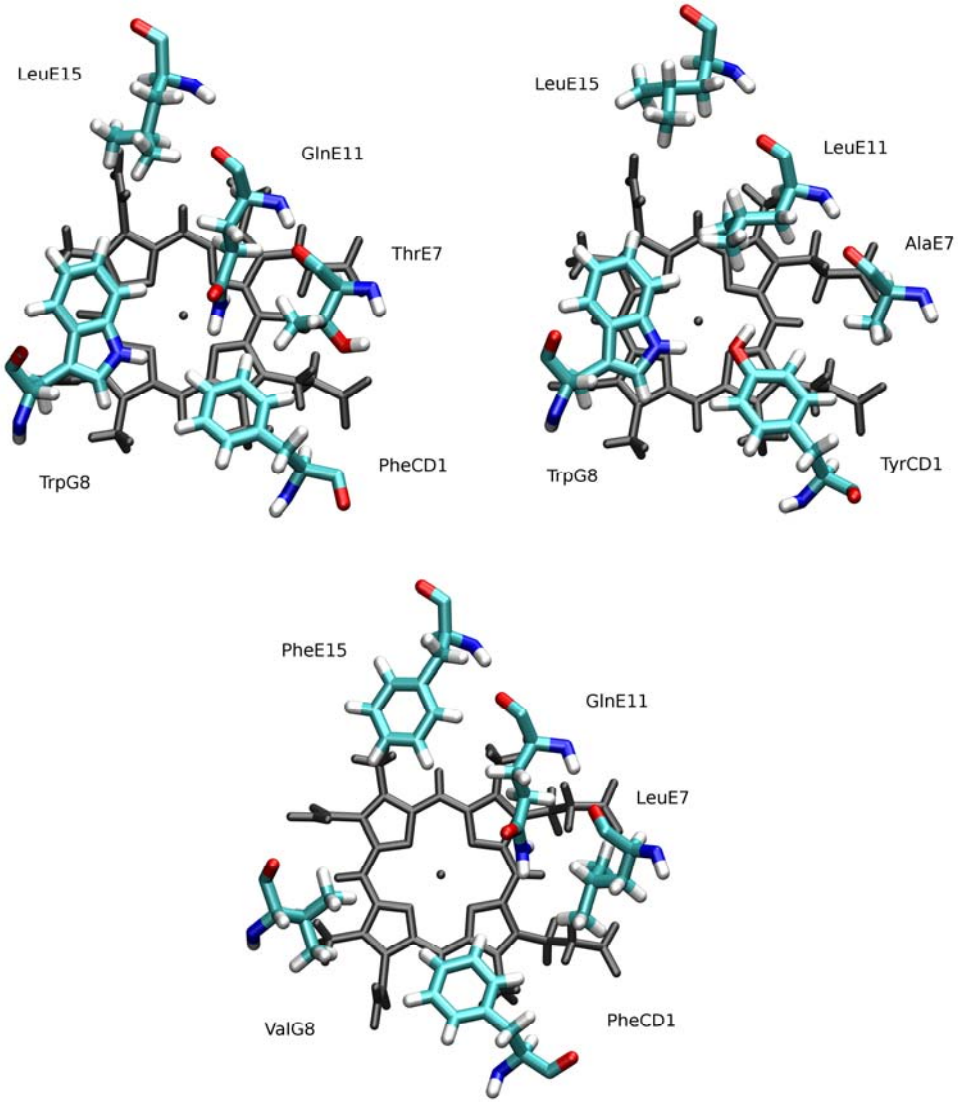


Figure 7. Representation of the main residues at the distal site in (*top, left*) Bs-trHbO, (*top, right*) Mt-trHbO, and (*bottom*) Mt-trHbN.

The case of Bs-trHbO is even more interesting. Even though inspection of the X-ray structure does not reveal a well-defined tunnel, this protein displays a high association rate [38]. The simulation results showed that GlnE11 adopts a different conformation compared to the position in the crystal structure, opening the LT (Fig. 7). Although Bs-trHbO presents TrpG8 (as in the case of Mt-trHbO), the presence of GlnE11 seems to suffice for opening of the LT. Thus, steered MD simulations run for the Bs-trHbO GlnE11→Leu mutant (designed to mimic Mt-trHbO) showed a large barrier for the migration through the LT, thus indicating that both TrpG8 and LeuE11 are responsible of blocking the LT. Although the presence of TrpG8 was expected to be enough to control ligand entry in these proteins, the presence of GlnE11 –as compared to Leu in Mt-trHbO– clearly contributes to opening of the LT and facilitates ligand entry, in agreement with the experimental values determined for the association rate constants.

Overall, this brief discussion suffices to highlight the close relationship between subtle changes in the nature of certain residues in the interior of the proteins, and the migration properties of gaseous ligands through the protein matrix of structurally related trHbs. Thus, the presence of certain residues at specific positions has a critical role in regulating ligand affinity and reactivity by controlling the barrier for migration of ligands toward the heme active site.

2. Protoglobin of *Methanosarcina acetivorans*

This section is focused in selected structural characteristics of *M. acetivorans* protoglobin (MaPgb), and particularly on the unusually distorted heme found in the X-ray structure of this protein [29], and its possible functional implications.

MaPgb contains around 190 residues and a total of 9 helices, including a pre-A segment named Z. The heme group is highly distorted (Fig. 8) and fully buried in the protein matrix. The propionates are thus inaccessible to the solvent due to the presence of extended CE and FG loops, and to a long (20 residue) N-terminal segment with no secondary structure. In general, globins host an almost planar heme and solvent exposed propionates. Therefore, the large distortion found for MaPgb suggests that heme distortion is due to some sort of tension exerted by the surrounding residues.

Figure 8. Heme distortion of (*right*) *MaPgb* compared to (*left*) the standard heme of Mb.

Access of diatomic ligands to the distal site seems feasible through a V-shaped tunnel (Fig. 9), which is topologically different from that found in other globins. Thus, two apolar tunnels connect the solvent with the heme and are delimited by the B/G (tunnel 1) and B/E (tunnel 2) helices.

As mentioned in the Introduction, the O_2 dissociation rate can be controlled by distal effects. For instance Mt-trHbN and *Ascaris* hemoglobin exhibit a very low k_{off} for O_2 due to the presence of multiple hydrogen-bond interactions [23]. In contrast to those proteins, MaPgb has no residues capable of establishing permanent hydrogen bond interactions with the ligand. The only residue that could act as hydrogen-bond donor is TyrB10, but extended MD simulations showed that TyrB10 is mainly involved in hydrogen bonding to LeuE4. Therefore, the low dissociation rate determined experimentally must be related to other mechanisms, like heme distortions.

Bikiel et al. [24] have recently examined the influence played by distinct deformations of the heme on the ligand affinity. A systematic classification of heme distortions, denoted as Normal-Coordinate Structural Decomposition (NSD), has been proposed by Jentzen and coworkers [39]. This technique identifies the most relevant out-of-plane (saddling, ruffling, doming, X-waving, Y-waving and propellering) and in-plane (meso-stretching, N-pyrrole stretching, pyrrole traslation (X,Y), breathing and pyrrole rotation) normal deformation modes that relate the structure of a distorted heme compared to a reference D_{4h} structure (Fig. 10).

Table 1 shows the difference in O_2 binding affinity for selected deformations of the heme compared to an ideal planar heme determined from

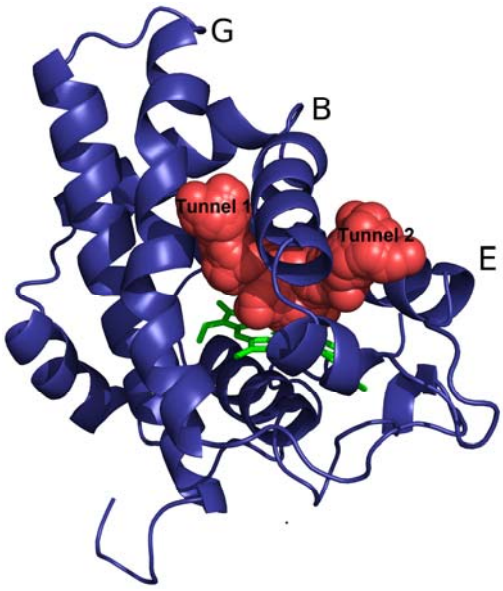


Figure 9. Representation of the V-shaped tunnel found in the X-ray structure of MaPgb.

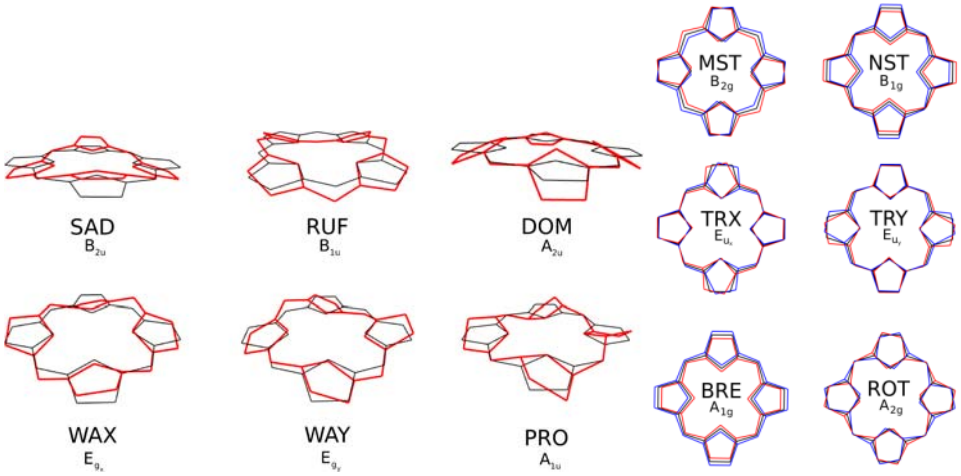


Figure 10. Representation of the normal (*left*) out-of-plane and (*right*) in-plane deformation modes of the heme.

Table 1. Difference in binding energy ($\Delta(\Delta E_{O2})$; kcal/mol) determined for selected distortions (Å) along out-of-plane and in-plane deformation modes relative to an ideal planar heme. The geometrical deviation from the planarity is measured by the root-mean square deviation (RMSD; Å).

Out-of-plane	**$\Delta(\Delta E_{O2})$**						
RMSD	**-1.2**	**-0.6**	**-0.3**	**0.0**	**0.3**	**0.6**	**1.2**
Saddling	1.9	0.5	0.1	0.0	0.1	0.5	2.4
Ruffling	2.9	0.8	0.4	0.0	-0.1	0.1	0.6
Doming	7.0	2.6	1.0	0.0	-0.3	0.0	2.1
X-waving/Y-waving	3.6	0.7	0.2	0.0	0.3	1.2	3.5
Propellering	2.2	0.5	0.1	0.0	0.1	0.5	2.2
In-plane	**$\Delta(\Delta E_{O2})$**						
RMSD	**-0.5**	**-0.25**	**-0.125**	**0.0**	**0.125**	**0.25**	**0.5**
Meso-stretching	2.1	0.8	0.2	0.0	-0.3	-0.2	0.1
N-pyrrole stretching	2.0	1.1	0.3	0.0	0.3	1.1	2.0
Pyrrole traslation X/Y	0.3	0.1	0.0	0.0	0.0	0.0	0.2
Breathing	11.3	5.6	3.1	0.0	-3.6	-6.7	-11.3
Pyrrole rotation	2.0	0.6	0.2	0.0	0.0	0.2	1.3

B3LYP/6G-31G** calculations [24]. All out-of-plane and in-plane distortions lead to an increase in the energy difference, which tends to reduce the ligand affinity. Fig. 11 shows the energy profiles arising from the ruffling (RUF) deformation as an example of this general trend: both positive and negative values of distortion decrease affinity. Only very small negative values are observed for slight distortions along the MST, DOM and RUF modes, but they do not alter the general trend. The only exception to this behavior is the in-plane BRE distortion, since a negative trend in Δ(ΔE) is clearly observed for positive values of BRE. This is represented in Fig. 11, where negative values for BRE distortion behave similarly to the rest of the normal modes, while positive values show negative Δ(ΔE), that is, compression of the heme (positive BRE) increases the ligand affinity.

The NSD analysis of the porphyrin ring in MaPgb shows that the main out-of-plane contribution to the heme deformation is ruffling, which accounts for a distortion of 1.42 Å (in Mb, this distortion only amounts to 0.02 Å). Regarding in-plane distortions, MaPgb results in a displacement around 10-fold larger compared to Mb, which is mainly due to breathing. Whereas the

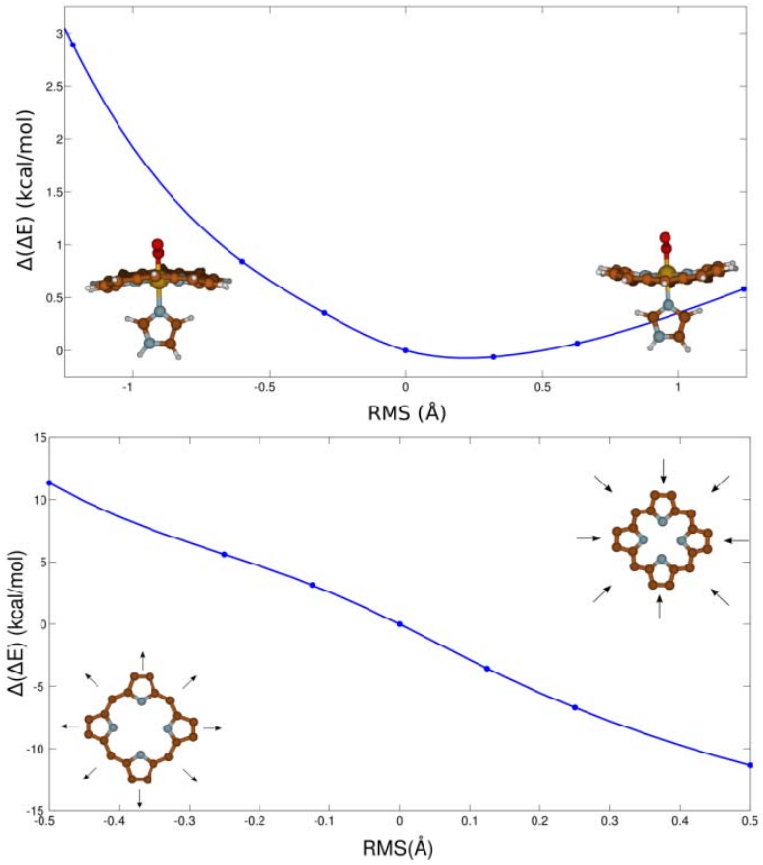


Figure 11. Representation of energy difference profiles due to (*top*) RUF and (*bottom*) BRE deformation modes.

out-of-plane distortions in MaPgb reduce the binding affinity by 1.5 kcal/mol, the positive breathing enhances ligand binding by 4.7 kcal/mol. The net effect is an overall increase in oxygen affinity for MaPgb due to heme distortion.

These findings suggest that evolution has led to the design of a tighter cavity for the heme in MaPgb, which induces an overall heme compression and in turn an increase in oxygen affinity. Although the function of MaPgb is still unkown, these results can shed light on an inherent modulation mechanism for ligand affinity in globins.

3. Neuroglobin

Neuroglobin (Ngb) is a member of the group formed by endogenous reversibly hexacoordinated globins. Many globins contain HisE7, normally responsible for O_2 stabilization by H bonding upon heme-O_2 complex is formed. However, in some globins, HisE7 is directly bound to the sixth coordination site of Fe in the heme, forming an hexacoordinated (6c) globin. This is the case of Ngb.

Ngb was firstly identified by Burmester *et al.* in 2000 [40] in man and mouse, and the X-ray structure of human and murine Ngbs have been solved [41, 42]. Later Ngb has also been identified in rat, pufferfish and zebrafish [43, 44], which suggests a widespread distribution among vertebrate species. Its function might be related to the protection of cells from stroke damage, amyloid toxicity and injury due to lack of oxygen, and neuroprotection [45]. Although the exact mechanisms by which Ngb protects cells are still unclear, it is suggested that it maintains the function of mitochondria and regulates the concentration of important chemicals in the cell. Increased risk of Alzheimer's disease has been related to low levels of Ngb [46]. Ngb mRNA and protein have been shown to be upregulated by hypoxia and post-anoxia re-oxygenation, which suggest a putative role as a reactive oxygen species scavenger [47].

Ngb is a monomeric 151-residue globin that shows less than 25% sequence identity with Mb. Though it retains the canonical 3-over-3 α-helical sandwich of mammalian Mb and Hb, notable structural deviations are found in the CD-D region and in the N-terminal half of the E helix. As a result of hexacoordination through HisE7, E helix is pulled toward the heme relative to Mb.

Pentacoordinated (5c) Ngb exhibits a very high affinity for O_2 but because of the endogenous hexacoordination, it results in a moderate O_2 affinity (P_{O2}=2 torr), similar to Mb. The reversible hexacoordination could serve as a way to fine tune the O_2 affinity. When 5c state is favoured affinity should be higher, though it also depends on the residues present in the distal cavity. This hypothesis was confirmed by experiments that drive the 5c-6c equilibrium toward the pentacoordinated state under oxidizing conditions that favor Cys disulfide bridge formation [48, 49].

MD simulations (MD), essential dynamics and steered molecular dynamics have been used to gain a molecular-level picture of the 5c↔6c transition and to estimate its thermodynamic properties [50]. Fig. 12 shows the free energy profile for the 6c to 5c transition in Cys reduced and oxidized states, called Cred and Cox, respectively. Reduction of disulfide bridge increases the barrier for the transition from 6.2 to 7.3 kcal/mol, with a 2.6 kcal/mol increase in the free energy of 5c-Cred state (relative to the 5c.Cox) state. In terms of rate constant, this implies a decrease by a factor of 6, and around 78 in the equilibrium constant.

These trends agree with the experimental results reported by Hamdane and coworkers [48, 49]. The overall conclusion is that Cys oxidation stabilizes the 5c state, favoring more avid-for-oxygen species, which supports a mechanism of oxygen release in case of hypoxic conditions, thus suggesting an O_2 storage function for Ngb. If oxygen concentration becomes critically low, the disulfide bridge in Ngb would be in the reduced state, which would in turn release oxygen to the cell. If oxygen concentration increases the disulfide bridge will form, triggering an increase in affinity for molecular oxygen.

Capece et al. [51] gained insight into the 5c↔6c equilibrium molecular determinants by means of MD simulations of Ngb and Mb at normal (1 bar) and high (3 kbar) pressure conditions in both coordination states. The overall conclusion of these studies is that the main differences between both proteins are located in the CD loop, whose structure is much more sensitive to both pressure and coordination state in Ngb than in Mb. This is reflected in the free energy profiles for the 5c↔6c transition in Ngb and Mb (Fig. 13). In Ngb the barrier increases from 7 kcal/mol at 1 bar to 13 kcal/mol at 3 kbar, with a destabilization of 4 kcal/mol of the 5c state. As shown experimentally, the

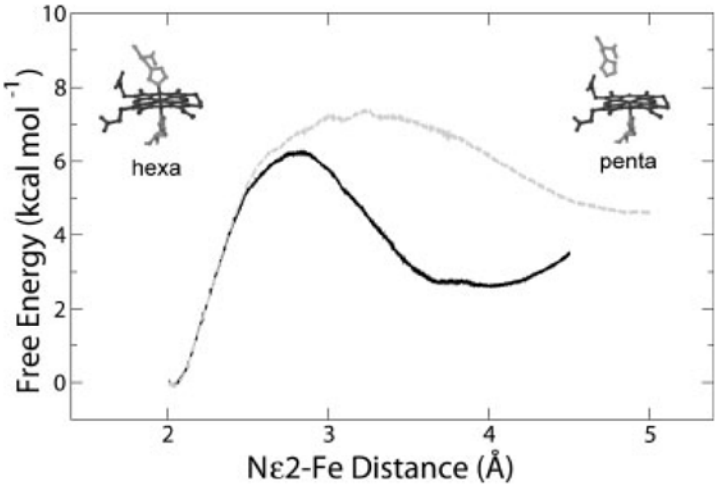


Figure 12. Free energy profile for the 5c↔6c transition with Cys residues in both reduced (gray) and oxidized (black) states.

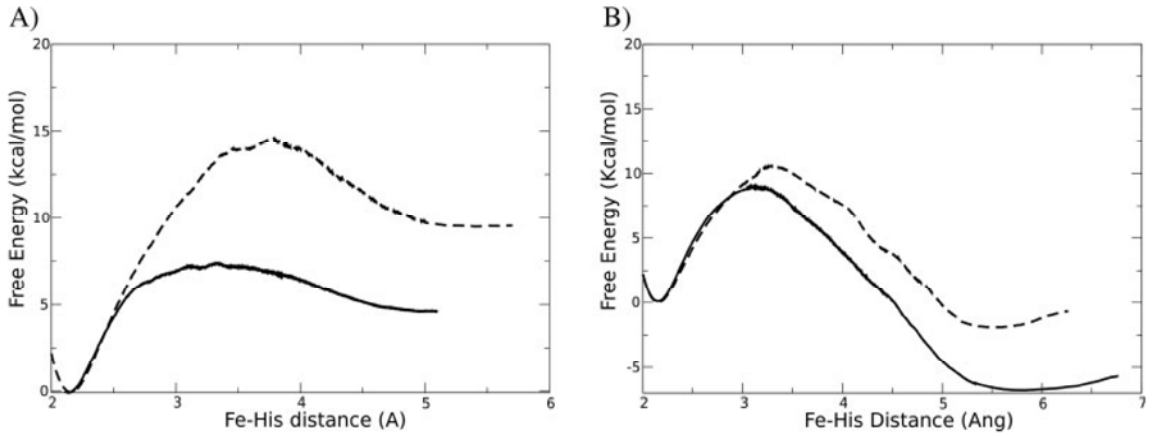


Figure 13. Free energy profile for the 5c↔6c transition at 1 bar (solid line) and 3 kbar (dashed line) in (A) Ngb and (B) Mb.

hexacoordination state is favored by an increase in pressure [52]. For Mb, the pentacoordinated state in more stable at 1 bar than the hexacoordinated state by 6 kcal/mol, whereas this difference is reduced to 1 kcal/mol at 3 kbar, with a reduction in the barrier for the transition.

Following the previous work by Nadra et al. [50], the disulfide bridge is found to have an important effect on the structure of the CD region in both coordination states, being the effect smaller than the variation with pressure. It seems that the oxidation state of the key Cys has an impact focused on the CD loop, while pressure has a more general influence on the overall structure of the protein.

The preceding results show that pressure alters the dynamics of globins, specifically reducing mobility and shifting the 5c↔6c equilibrium toward the hexacoordinated state. However the reasons for this shift are different in Mb and Ngb. In Ngb, pressure mainly affects the mobility in the CD region and increases the barrier for the 5c↔6c transition. In contrast, the coordination equilibrium in Mb involves a more global structural rearrangement and pressure destabilizes the 5c state. Although the overall trend for the 5c↔6c transition is the same in both cases –the higher the pressure, the more favored the hexacoordination–, the underlying differences highlight the existence of diverse response mechanisms upon changes in external conditions within the globin world.

Conclusion

The advances made in the last decade on the structural and functional variation found within the globin world reinforces the idea that the thoroughly studied Mb and Hb are just specialized cases in a broad evolutionary

superfamily that evolved specifically to the demands of circulatory systems and muscles. The three cases presented here (Ngb, MaPgb and MtHbN) are representative examples of the growing diversity discovered within globins. Even though they reflect the basic chemical properties of a heme group buried in a conserved *globin fold*, it is clear that a proper understanding of the structural and dynamical differences between globins is fundamental to gain insight into the functional role of novel globins.

Acknowledgements

This work was partially supported by the University of Buenos Aires, ANPCyT (PICT-25667), CONICET, the Spanish Ministerio de Innovación y Ciencia (SAF200805595 and PCI2006-A7-0688), the Generalitat de Catalunya (2009-SGR00298), and the EU FP7 program (project NOStress).

References

1. Smith, L. J., Kahraman, A., Thornton, J. M. *Proteins ,* 78, 2349.
2. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., Phillips, D. C. 1958, *Nature,* 181, 662.
3. Ghosh, A. 2007, *The smallest biomolecules: diatomics and their interactions with heme proteins*, Elsevier.
4. Paoli, M., Marles-Wright, J., Smith, A. 2002, *DNA Cell. Biol.* 21, 271.
5. Vinogradov, S. N., Moens, L. 2008, *J. Biol. Chem.,* 283, 8773.
6. Frauenfelder, H., McMahon, B. H., Fenimore, P. W. 2003, *Proc. Natl. Acad. Sci. USA,* 100, 8615.
7. Scott, E. E., Gibson, Q. H., Olson, J. S. 2001, *J. Biol. Chem.,* 276, 5177.
8. Wittenberg, J. B., Wittenberg, B. A. 2003, *J. Exp. Biol.,* 206, 2011.
9. Kapp, O. H., Moens, L., Vanfleteren, J., Trotman, C. N., Suzuki, T., Vinogradov, S. N. 1995, *Prot. Sci.,* 4, 2179.
10. Bidon-Chanal, A., Marti, M. A., Crespo, A., Milani, M., Orozco, M., Bolognesi, M., Luque, F. J., Estrin, D. A. 2006, *Proteins,* 64, 457.
11. Boechi, L., Marti, M. A., Milani, M., Bolognesi, M., Luque, F. J., Estrin, D. A. 2008, *Proteins,* 73, 372.
12. Boechi, L., Manez, P. A., Luque, F. J., Marti, M. A., Estrin, D. A. 2009, *Proteins,* 78, 962.
13. Nadra, A. D., Martí, M. A., Pesce, A., Bolognesi, M., Estrin, D. A. 2007, *Proteins,* 71, 695.
14. Brunori, M., Vallone, B. 2007, *Cell. Mol. Life Sci.,* 64, 1259.
15. Lama, A., Pawaria, S., Bidon-Chanal, A., Anand, A., Gelpi, J. L., Arya, S., Marti, M., Estrin, D. A., Luque, F. J., Dikshit, K. L. 2009, *J. Biol. Chem.,* 284, 14457.
16. Bidon-Chanal, A., Marti, M. A., Estrin, D. A., Luque, F. J. 2007, *J. Am. Chem. Soc.,* 129, 6782.

17. Goldbeck, R. A., Pillsbury, M. L., Jensen, R. A., Mendoza, J. L., Nguyen, R. L., Olson, J. S., Soman, J., Kliger, D. S., Esquerra, R. M. 2009, *J. Am. Chem. Soc.,* 131, 12265.
18. Goldbeck, R. A., Bhaskaran, S., Ortega, C., Mendoza, J. L., Olson, J. S., Soman, J., Kliger, D. S., and Esquerra, R. M. 2006, *Proc. Natl. Acad. Sci. USA,* 103, 1254
19. Ouellet, Y. H., Daigle, R., Laguee, P., Dantsker, D., Milani, M., Bolognesi, M., Friedman, J. M., Guertin, M. 2008, *J. Biol. Chem.,* 283, 27270.
20. Laverman, L. E., Ford, P. C. 2001, *J. Am. Chem. Soc.,* 123, 11614.
21. Milani, M., Pesce, A., Nardini, M., Ouellet, H., Ouellet, Y., Dewilde, S., Bocedi, A., Ascenzi, P., Guertin, M., Moens, L., Friedman, J. M., Wittenberg, J. B., Bolognesi, M. 2005, *J. Inorg. Biochem.,* 99, 97.
22. Bikiel, D. E., Boechi, L., Capece, L., Crespo, A., De Biase, P. M., Di Lella, S., González Lebrero, M. C., Martí, M. A., Nadra, A. D., Perissinotti, L. L., Scherlis, D. A., Estrin, D. A. 2006, *Phys. Chem. Chem. Phys.,* 8, 5611.
23. Marti, M. A., Crespo, A., Capece, L., Boechi, L., Bikiel, D. E., Scherlis, D. A., Estrin, D. A. 2006, *J. Inorg. Biochem.,* 100, 761.
24. Bikiel, D. E., Forti, F., Boechi, L., Nardini, M., Luque, F. J., Martí, M. A., Estrin, D. A. 2010, *J. Phys. Chem. B,* 114, 8536.
25. Vinogradov, S. N., Hoogewijs, D., Bailly, X., Mizuguchi, K., Dewilde, S., Moens, L., Vanfleteren, J. R. 2007, *Gene,* 398, 132.
26. Vinogradov, S., Hoogewijs, D., Bailly, X., Arredondo-Peter, R., Gough, J., Dewilde, S., Moens, L., Vanfleteren, J. 2006, *BMC Evol. Biol.,* 6, 31.
27. Freitas, T. A. K., Hou, S., Dioum, E. M., Saito, J. A., Newhouse, J., Gonzalez, G., Gilles-Gonzalez, M.-A., Alam, M. 2004, *Proc. Natl. Acad. Sci. USA,* 101, 6675.
28. Hou, S., Freitas, T., Larsen, R. W., Piatibratov, M., Sivozhelezov, V., Yamamoto, A., Meleshkevitch, E. A., Zimmer, M., Ordal, G. W., Alam, M. 2001, *Proc. Natl. Acad. Sci. USA,* 98, 9353.
29. Nardini, M., Pesce, A., Thijs, L., Saito, J. A., Dewilde, S., Alam, M., Ascenzi, P., Coletta, M., Ciaccio, C., Moens, L., Bolognesi, M. 2008, *EMBO Rep.,* 9, 157.
30. Tomita, A., Sato, T., Ichiyanagi, K., Nozawa, S., Ichikawa, H., Chollet, M., Kawai, F., Park, S.-Y., Tsuduki, T., Yamato, T., Koshihara, S.-Y., Adachi, S.-I. 2009, *Proc. Natl. Acad. Sci. USA,* 106, 2612.
31. Ordway, G. A., Garry, D. J. 2004, *J. Exp. Biol.,* 207, 3441.
32. Bloom, B. 1994, *Tuberculosis: pathogenesis, protection and control*, ASM Press, Washington, DC.
33. MacMicking, J. D., North, R. J., LaCourse, R., Mudgett, J. S., Shah, S. K., Nathan, C. F. 1997, *Proc. Natl. Acad. Sci. USA,* 94, 5243.
34. Ouellet, H., Ouellet, Y., Richard, C., Labarre, M., Wittenberg, B., Wittenberg, J., Guertin, M. 2002, *Proc. Natl. Acad. Sci. USA,* 99, 5902.
35. Couture, M., Yeh, S.-R., Wittenberg, B. A., Wittenberg, J. B., Ouellet, Y., Rousseau, D. L., Guertin, M. 1999, *Proc. Natl. Acad. Sci. USA,* 96, 11223.
36. Pesce, A., Nardini, M., Milani, M., Bolognesi, M. 2007, *IUBMB Life*, 59, 535.
37. Milani, M., Pesce, A., Ouellet, Y., Ascenzi, P., Guertin, M., Bolognesi, M. 2001, *EMBO J.,* 20, 3902.

38. Giangiacomo, L., Ilari, A., Boffi, A., Morea, V., Chiancone, E. 2005, *J. Biol. Chem.,* 280, 9192.
39. Jentzen, W., Song, X. Z., Shelnutt, J. A. 1997, *J. Phys. Chem. B,* 101, 1684.
40. Burmester, T., Weich, B., Reinhardt, S., Hankeln, T. 2000, *Nature,* 407, 520.
41. Pesce, A., Dewilde, S., Nardini, M., Moens, L., Ascenzi, P., Hankeln, T., Burmester, T., Bolognesi, M. 2003, *Structure,* 11, 1087.
42. Vallone, B., Nienhaus, K., Brunori, M., Nienhaus, G. U. 2004, *Proteins,* 56, 85.
43. Awenius, C., Hankeln, T., Burmester, T. 2001, *Biochem. Biophys. Res. Commun.,* 287, 418.
44. Zhang, C., Wang, C., Deng, M., Li, L., Wang, H., Fan, M., Xu, W., Meng, F., Qian, L., He, F. 2002, *Biochem. Biophys. Res. Commun.,* 290, 1411.
45. Greenberg, D. A., Jin, K., Khan, A. A. 2008, *Curr. Opin. Pharmacol.,* 8, 20.
46. Raychaudhuri, S., Skommer, J., Henty, K., Birch, N., Brittain, T. 2009, *Apoptosis,* 15, 401.
47. Nayak, G. H., Milton, S. L., Prentice, H. M. 2007, *FASEB J.,* 21, A924.
48. Hamdane, D., Kiger, L., Dewilde, S., Green, B. N., Pesce, A., Uzan, J., Burmester, T., Hankeln, T., Bolognesi, M., Moens, L., Marden, M. C. 2004, *Micron,* 35, 59.
49. Hamdane, D., Kiger, L., Dewilde, S., Green, B. N., Pesce, A., Uzan, J., Burmester, T., Hankeln, T., Bolognesi, M., Moens, L., and Marden, M. C. 2003, *J Biol. Chem.,* 278, 51713
50. Nadra, A. D., Martí, M. A., Pesce, A., Bolognesi, M., Estrin, D. A. 2008, *Proteins,* 71, 695.
51. Capece, L., Marti, M. A., Bidon-Chanal, A., Nadra, A., Luque, F. J., Estrin, D. A. 2008, *Proteins,* 75, 885.
52. Hamdane, D., Kiger, L., Hoa, G. H. B., Dewilde, S., Uzan, J., Burmester, T., Hankeln, T., Moens, L., Marden, M. C., 2005, *J. Biol. Chem.,* 280, 36809.