Engineered chimeras reveal the structural basis of hexacoordination in globins: A case study of neuroglobin and myoglobin

Running Title: Determinants for globin hexacoordination

Engineered chimeras reveal the structural basis of hexacoordination in globins

Ignacio Boron[& 1, 2], Luciana Capece[& 1, 3], Francesca Pennacchietti[4], Diana E. Wetzler[1, 2], Stefano Bruno[5], Stefania Abbruzzetti[4], Lucía Chisari[1], .F. Javier Luque[6], Cristiano Viappiani[4], Marcelo A. Marti[1, 7], Dario A. Estrin[3, 7, *] and Alejandro D. Nadra[1, 2, *]

[1]Departamento de Química Biologica Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, Buenos Aires, C1428EHA, Argentina

[2] IQUIBICEN-CONICET, Argentina

[3] Departamento de Química Inorgánica,Analítica y Química Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, Buenos Aires, C1428EHA, Argentina

[4] Dipartimento di Fisica e Scienze della Terra, Universitá degli Studi di Parma, Italy

[5] Dipartimento di Farmacia, Universitá degli Studi di Parma, Italy

[6]Departament de Fisicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia, Universitat de Barcelona, Avda. Prat de la Riba 171, 08921 Santa Coloma de Gramenet, Spain.

[7] INQUIMAE-CONICET, Argentina

[&]These two authors contributed equally.

*To whom correspondence should be sent: Dario A. Estrin, e-mail: dario@qi.fcen.uba.ar, Alejandro D. Nadra, e-mail: anadra@qi.fcen.uba.ar

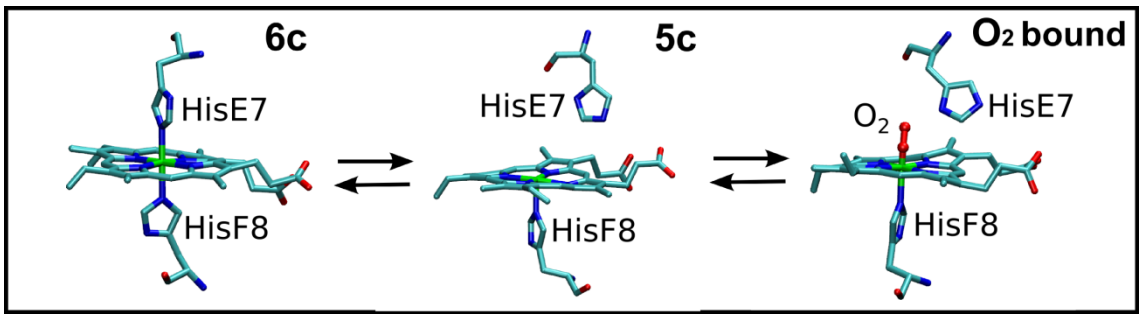
Abstract

Myoglobin (Mb) and neuroglobin (Ngb) are representative members of pentacoordinated and bis-histidyl hexacoordinated globins. In spite of their low sequence identity, they show surprisingly similar three-dimensional folds. The ability of Ngb to form a hexacoordinated bis-histidyl complex with the distal HisE7 has a strong impact on ligand affinity. The factors governing such different behavior have not been completely understood yet, even though they are extremely relevant to establish structure-function relationships within the globin superfamily. In this work we generated chimeric proteins by swapping a previously identified regulatory segment –the CD region- and evaluated comparatively the structural and functional properties of the resulting proteins by molecular dynamics simulations, spectroscopic and kinetic investigations. Our results show that chimeric proteins display heme coordination properties displaced towards those expected for the corresponding CD region. In particular, in the absence of exogenous ligands, chimeric Mb is found as a partially hexacoordinated bis-histidyl species, whereas chimeric Ngb shows a lower equilibrium constant for forming a hexacoordinated bis-histidyl species. While these results confirm the regulatory role of the CD region for bis-histidyl hexacoordination, they also suggest that additional sources contribute to fine tune the equilibrium.

Introduction

The globin superfamily is formed by a widely distributed group of heme proteins, which are responsible for a variety of biological functions in all kingdoms of life. [1] Many of these functions involve the coordination of small ligands to the prosthetic heme group, a process that is tightly controlled.[2-4] The heme resides deeply buried inside the protein, which generally folds in a globular 3-over-3 sandwich motif (formed by helices labeled A, B, E, F, G and H). The iron is coordinated to the four equatorial porphyrin nitrogen atoms and axially coordinated by the absolutely conserved proximal histidine F8 (HisF8), leading to the characteristic pentacoordinated (5c) state (Scheme 1). The vacant coordination at the distal site permits the binding of the diatomic ligand and allows the protein to fulfill its function. In several globins, however, the distal site is blocked by the coordination of an endogenous residue, usually the distal histidine (HisE7), thus resulting in a bis-histidyl hexacoordinated (6c) species. This process introduces an additional mechanism to control the function of globins, as has been noticed for neuroglobin (Ngb),[5], cytoglobin [6], and non-symbiotic plant hemoglobins. [7, 8]

It is important to recall that while globins are generally classified as penta or hexacoordinated, binding of exogenous ligands is only possible to the reactive 5c form. For hexacoordinated hemoglobins this is made possible thanks to the precisely regulated equilibrium between 6c and 5c states (Scheme 1). Fine tuning of the 6c⇆5c equilibrium provides a mechanism to modulate ligand affinity, since the more displaced the equilibrium to the 5c state, the higher the ligand affinity. An example of this regulation is Ngb, where Cys oxidation and disulfide bridge formation displaces the 6c⇆5c equilibrium towards the 5c state, resulting in a 10-fold increase in O_2 affinity. [9-11] In this scenario, understanding the molecular basis that underlies the regulation of the 5c⇆6c equilibrium is crucial for deciphering a fundamental regulatory mechanism of the physiological function of globins.



Scheme 1. The 6c⇆5c equilibrium and its role in the ligand binding process. Heme coordination states: bis-histidyl hexacoordination (left), ligand free pentacoordination (center), and ligand-bound hexacoordination (right). The heme group, the distal histidine E7, the proximal histidine F8 and the O_2 bound to the heme are shown.

Myoglobin (Mb) and Ngb can be considered the prototypical cases representative of each coordination state. Mb is a 5c globin, displaying overall association rate constants (k_{on}) for CO, O_2, and NO binding of 0.5, 17, and 22 μM^{-1} s^{-1}, respectively [12-14], which are characteristic of a vacant distal site. When the O_2 ligand is bound, it acquires a negative charge density through π back donation from the metal, which leads to a strong hydrogen bond (HB) with the distal HisE7 and hence to moderate oxygen dissociation rate (k_{off} = 12-22 s^{-1}) [14] and affinity (K_{O2} = 1.1 μM^{-1}). [13, 14] In the absence of exogenous bound ligands, Ngb is found as a bis-histidyl, 6c species, with the distal HisE7 occupying the sixth coordination site, as revealed by X-ray crystallography and spectroscopy. [15, 16] The competition between exogenous ligands and the endogenous HisE7 results in complex patterns of ligand binding kinetics. [17-20] Once oxygen is bound, formation of a hydrogen bond (HB) with HisE7 results in a small dissociation rate (k_{off}=0.5-1.5 s^{-1}) [17, 19] and an oxygen affinity similar to that of Mb.[21]

Although the sequence of many hundreds of globins is already known [2], and many of them have been characterized spectroscopically and structurally, leading to a precise definition of their coordination state [22], the factors that modulate the 6c⇆5c equilibrium are still unknown.

A first insight into the reasons that determine the intrinsic coordination state can be gained from the comparison of Mb and Ngb. At the sequence level the most important difference is found in the CD region, which is poorly conserved and presents changes both in length and sequence variability. Thus, the CD regions of Mbs tend to be shorter and comprise more charged residues while those of Ngbs are slightly longer and contain more aromatic residues. From a structural point of view, the two proteins are topologically very similar, the main differences being located in the CD region (see Figure 1). Interestingly, the presence of two cysteine residues in Ngb's CD region, capable of establishing an internal disulfide bridge, also affects the 6c⇆5c equilibrium. [9, 11] It can then be hypothesized that the nature of the CD loop is a key factor directly implicated in the coordination preferences of globins. Previous molecular dynamics (MD) studies suggested that the CD region plays a predominant role in the transition between 5c and 6c states, especially in Ngb where the CD region is highly flexible. [11, 23]

The aim of this study is to examine whether the CD region is the main structural determinant of the 6c⇆5c equilibrium in globins using both experimental and theoretical strategies. To this end, we have engineered two chimeric globins related to Mb (ChMb) and Ngb (ChNgb) that were designed by swapping of their CD regions. The detailed structural and dynamical properties associated to the 6c⇆5c equilibrium of the two chimeric globins have been examined by means of MD simulations. Experimentally, we have characterized the resulting chimeras in relation to their 6c⇆5c equilibrium and kinetic properties in the ferrous state, as well as their CO binding properties. We demonstrate that the flexibility of the CD region and the structural change associated to the 6c⇆5c process are significantly interchanged in the chimeric proteins, and that ChMb is able to form a bis-histidyl hexacoordinated species in the absence of exogenous ligands, whereas the inclusion of the Mb CD region in ChNgb results in a more

labile Fe-HisE7 bond. Altogether, our results confirm the role of the CD region in the modulation of the 6c⇆5c equilibrium in globins.

Results

Engineering Mb and Ngb CD region chimeras. Taking as a starting point the sequence and structural comparison of Mb and Ngb, we decided to build two new chimeric proteins (ChMb and ChNgb) by respectively swapping their CD region (Figure 1). Thus, ChMb displays Mb sequence except for residues 41 to 59 (Uniprot ID P02185), which are replaced by Ngb residues 39 to 58 (Uniprot ID Q9NPG2). In turn, ChNgb displays Ngb sequence except for residues 39 to 58, which are replaced by Mb residues 41 to 59.

MD simulations of ChMb and ChNgb in the 5c and 6c states. We first performed MD simulations to examine the preservation of the intrinsic flexibility properties of the CD region in the wild type (wt) proteins and the chimeras considering both 5c and 6c coordination states. The first 30 ns of the 100 ns trajectories were excluded from the analysis in order to allow the systems to equilibrate before collecting the data. Only in the case of ChMb-6c the first 50 ns were not considered in the analysis, as up to this time the system was still showing large fluctuations, and the trajectory was extended up to 130 ns (Figure S1).

The CD region in Mb is enriched in charged residues as compared with Ngb, which shows instead a higher number of apolar and aromatic residues. Upon insertion of the CD stretch into the other protein, each CD region retains, to a large extent, the intra-loop contacts as well as most interactions with conserved residues in the framework of the host protein (Table S1).

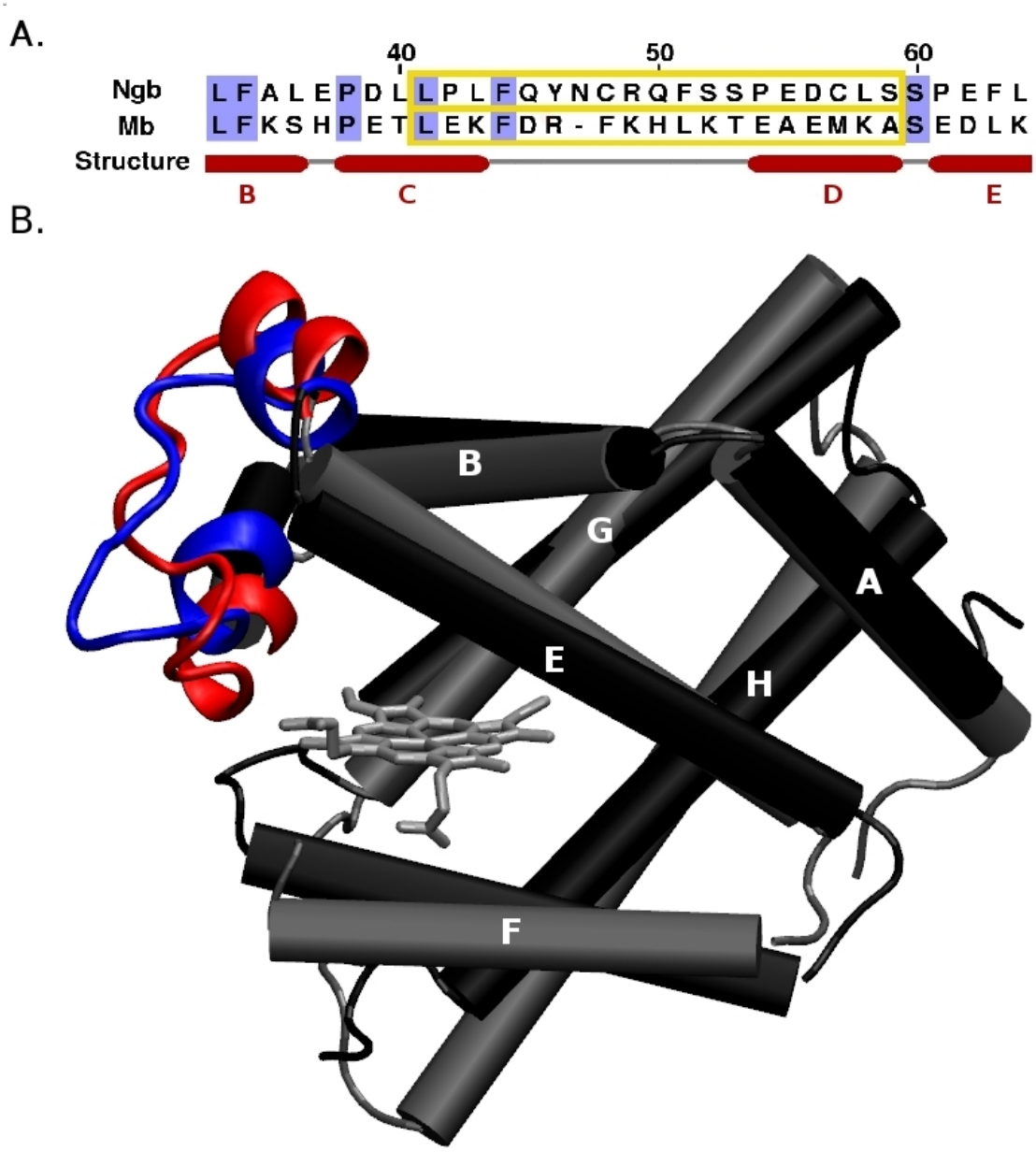


Figure 1. Comparison between Mb and Ngb. **A.** Structure-based sequence alignment of the CD region. The residues exchanged in the chimeric proteins are marked in the yellow box and conserved residues are highlighted in blue. Numbers above the sequence correspond to Mb (Uniprot ID P02185). Below the sequences, regions with defined α-helices in Mb crystal structure are indicated. **B.** Structural alignment of Mb (red/black) and Ngb (blue/gray) ligand free structures (PDB entries 1VXD and 1OJ6). The exchanged region is highlighted in color in both proteins. Only Ngb´s heme is depicted in gray sticks for the sake of clarity.

A first glimpse of the effect of the CD region exchange on the dynamics of the chimeras compared to the parent proteins can be obtained from the root-mean square fluctuations (RMSF) plots (Figure S2). In ChNgb the fluctuations of the CD region are significantly reduced in both 5c and 6c states when compared to the wt Ngb and are similar to those observed in Mb. The reduction is remarkably higher for the 6c state. The opposite trend is observed for the 5c state of ChMb, as the fluctuations of the CD loop are slightly larger than those found in Mb. In the case of Mb in the 6c state, however, no significant differences are observed between the fluctuations of the CD region of Mb and ChMb. Thus, even though the amplitude of the conformational deformations of the CD region is sensitive to the protein core, the results suggest that the intrinsic flexibility of the CD loop is largely preserved in the chimeras.

In a previous work, essential dynamics (ED) analysis of both Mb and Ngb [23] showed that in Ngb the main essential mode (EM), which accounts for 35% and 58% of the backbone flexibility in the 5c and 6c states, respectively, is primarily centered in the CD region. In the case of Mb, the first EM has a much lower contribution (between 20 and 25%) and has a more global character, as it includes contributions of B and H helices and the EF loop (especially in the 5c state). On the basis of these results, we decided to monitor the dynamics of chimeric protein using the same methodology. In the case of ChNgb-5c, the first EM accounts for 30% of the total structural variance, which implies a slight decrease compared to Ngb-5c (35%). Moreover, there is a reduction in the motion of the CD region as compared to Ngb. (Figure S3). In ChNgb-6c, the first EM accounts for 31% of the structural variance, which is notably lower than the 58% contribution obtained for Ngb. Furthermore, while in ChNgb it corresponds to a rigid movement of the CD loop and the D helix, in Ngb the movement includes a major conformational change of the CD loop, and importantly, of the N-terminal side of the E helix (residues 55-59) (Figure S3). The results for ChMb-5c (Figure S4) show a clear increase in the CD region contribution to the first EM. In this case, this mode accounts for 35% of the total

fluctuation (similarly to what is observed in Ngb-5c). In the case of ChMb-6c the CD region does not contribute significantly to the first EM.

To further analyze the 6c⇆5c transition we performed an ED analysis of the combined 5c and 6c trajectories for the chimeric proteins, as has been done in previous works for the wt proteins (see Computational Methods). Figure 2 shows the projection of the transition mode along the protein structure. In Ngb, the transition involves mainly the CD region and the N-terminal side of the E helix. On the other hand, in Mb the transition corresponds to a more global movement, involving the motion of protein regions that are highly rigid in the equilibrium MD simulations. The opposite behavior is found in the chimeric proteins. Thus, while in ChMb the transition involves the much more flexible CD region, in ChNgb the movement of the CD stretch is reduced, as it involves a smaller number of residues in the CD loop, and includes contributions from other regions of the protein.

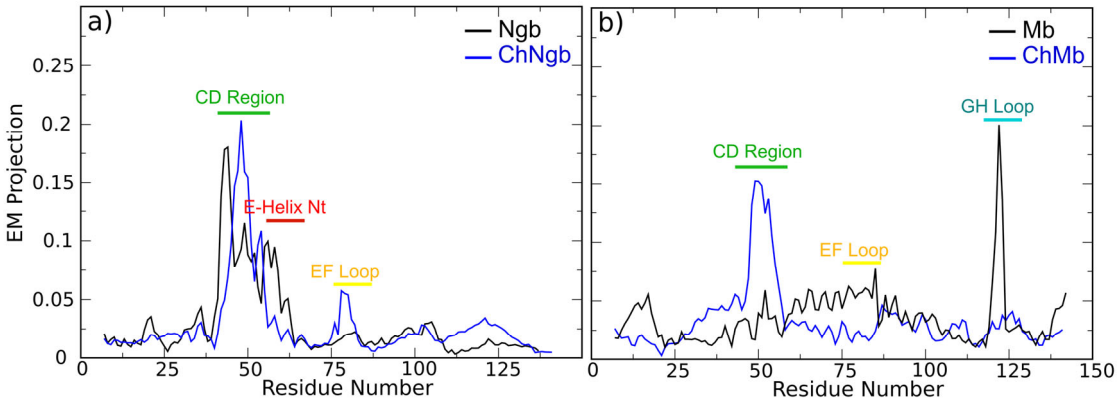


Figure 2. Projection of the transition EM along the protein structure for the wt and chimeric proteins. Results for Ngb and ChNgb are shown in panel a (black and blue lines, respectively). Results for Mb and ChMb are shown in panel b (black and blue lines, respectively).

In summary, the data obtained from the MD simulations strongly suggests that the flexibility pattern of the CD region is largely transferable between proteins. These data also suggest that the 6c⇆5c equilibrium is largely sensitive to the nature of the CD loop, and that the chimeras should exhibit different spectroscopic and kinetic signatures compared to the wt proteins.

Spectroscopic analysis of ChMb and ChNgb. Purified ChMb and ChNgb display typical absorption spectra of ferric hemoglobins. As expected, addition of sodium dithionite under Ar atmosphere results in heme Fe reduction. A summary of the peak positions in the absorption spectra for all protein states is given in Table 1. Whereas in ChNgb the resulting ferrous ligand free spectrum is almost indistinguishable from that of Ngb (see Figure S5), it is worth noting that the spectrum of ferrous ChMb appears to comprise features of both Mb and Ngb proteins, thus suggesting the presence of a mixture of 5c and bis-histidyl, 6c states (Figure 3A). Component analysis, using Ngb and Mb spectra as references, suggests the presence of ca. 25% of bis-histidyl 6c state for ChMb.

Table 1. UV-Vis spectra peak positions recorded for wt proteins and their chimeras. The three reported values correspond respectively to the Soret and alpha/beta peak positions.

Protein	Fe^{2+} (nm)	Fe^{3+} (nm)	Fe^{2+}-CO (nm)
ChMb	424 (435 shoulder), 528, 556	413, 535, 565 (shoulder)	422, 540, 570
Mb	435, 523 (shoulder), 556	-	423, 542, 580
ChNgb	424, 529, 558	413, 535, 562 (shoulder)	416, 536, 564
Ngb	424, 529, 558	413, 534, 560 (shoulder)	416, 538, 563

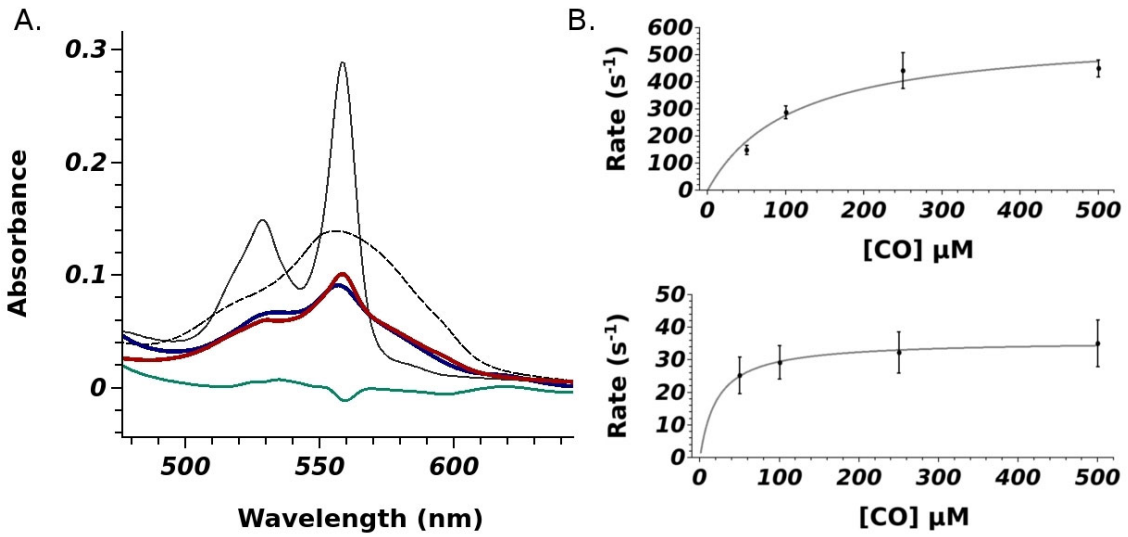


Figure 3 A. Absorption spectra of ferrous, ligand free ChMb (blue solid line), Ngb (black solid line), and Mb (black dashed line). The red solid line reports the best fit to the spectrum of ChMb obtained as a linear combination of the Ngb spectrum (taken as a reference for the bis-histidyl 6c state) and the Mb spectrum (taken as a reference for the 5c state). Best fit to ChMb spectrum is obtained with a mixture of 5c and 6c states in an approximate ratio of 75% to 25%, respectively (residuals are reported as green solid line). B. Apparent rates $k_{obs,1}$ and $k_{obs,2}$ for CO binding to ferrous ChMb as a function of CO concentration in stopped flow experiments. The time courses are measured at different CO concentrations ranging from 50 to 500 µM (after mixing). Continuous line corresponds to the fit to Equation 1.

Stopped flow measurements of ChMb and ChNgb CO association rates. To get further insight into the 6c⇆5c equilibrium, CO binding kinetics were measured as a function of CO concentration using a stopped flow technique [17]. Kinetic traces for CO binding to Mb and ChMb were measured through the absorption changes at the CO adduct peak position (λ=419 nm; see Figure S6). While CO binding to Mb could be properly described by a single exponential decay, CO binding to ChMb occurs through a biexponential relaxation (Figure S7). Thus, two observed rate constants are retrieved ($k_{obs,1}$ and $k_{obs,2}$) at each CO concentration. These

rates are plotted as a function of CO concentration in Figure 3B and reach limiting values at high ligand concentrations. On the basis of the absorption spectra (Figure 3A), which suggest that a mixture of 5c and bis-histidyl 6c species is present at equilibrium for the unliganded protein, we assign the two processes at high CO concentrations to binding to bis-histidyl 6c-like conformations, which are limited by HisE7 dissociation rates, $k_{off,1}$ = 36 s^{-1} and $k_{off,2}$ = 580 s^{-1} (Table 2).

Table 2. Ligand binding kinetic parameters. Kinetic microscopic constants for wt and chimeric proteins obtained from the fitting of laser flash photolysis (FP) and stopped flow (SF) traces at different CO concentrations.

Protein	$k_{on,CO}$ ($\mu M^{-1}s^{-1}$)	$k_{on,H}$ (s^{-1})	$k_{off,H}$ (s^{-1})	
	FP	FP	FP	SF
Ngb	105.0 ± 15.0	410 ± 5	0.42 ± 0.02	0.37± 0.1
ChNgb	26.0 ± 3.0 95.0 ± 15.0	190 ± 5 520 ± 10	1.1 ± 0.2 13 ± 3	0.71 ± 0.02 6.6 ± 0.4
Mb	0.7 ± 0.1	-	-	-
ChMb	2.6 ± 0.6 14.0 ± 3.0	- -	- -	580± 70 35.9 ± 0.8

The CO association to Ngb occurs with a kinetics well described by a single exponential decay, whose rate constant (k_{off} = 0.4 s^{-1}), at high CO concentration, is interpreted as the His dissociation rate [17]. Unlike the wt protein, ChNgb binds CO with a biphasic kinetics whose apparent rates at high CO concentration reach limiting values of 0.7 and 6.6 s^{-1} (Table 2), and are interpreted as arising from His dissociation rates for two distinct conformations. Thus,

stopped flow kinetics show that both chimeric proteins display partial bis-histidyl 6c character and exhibit heterogenous His dissociation kinetics.

CO rebinding kinetics to ChMb and ChNgb after laser flash photolysis. The competitive binding to the heme of the endogenous (HisE7) and exogenous (CO) ligands was investigated by nanosecond laser flash photolysis. Figure 4 compares CO rebinding kinetics to Mb, ChMb, Ngb and ChNgb.

Rebinding to Mb occurs with a well known biphasic kinetics with a non exponential geminate recombination (ca. 5% of rebinding) extending to the short microseconds, followed by a slower bimolecular phase. Even though the overall shape of CO rebinding to ChMb is similar to that observed for Mb, significant differences are found. The amplitude of geminate rebinding to ChMb increases to nearly 10% (inset to Figure 4) and the bimolecular rebinding appears faster than for Mb, although somewhat broader, a fact which suggests the presence of concurrent reactions occurring with different rates. Both geminate amplitude and bimolecular rebinding rate are intermediate between the corresponding figures for Mb and Ngb.

For Ngb the rebinding kinetics is characterized by three distinct phases: a geminate recombination of significantly larger amplitude than that observed for Mb (nearly 15% at 30 °C; Figure 4), a rather fast bimolecular recombination to the 5c species, and a much slower rebinding to the bis-histidyl hexacoordinated species, with a rate limited by distal His dissociation. [9, 18] The progress curve for CO rebinding to ChNgb follows a similar general trend as the wt protein, but with remarkable differences. Thus, the bimolecular reaction between CO and the 5c species is slower and much broader than for Ngb, leading to the formation of bis-histidyl hexacoordinated species to a higher extent. On the other hand, the His dissociation is a faster and broader process, suggesting a lower stability for the bis-histidyl hexacoordinated species in comparison to the wt protein.

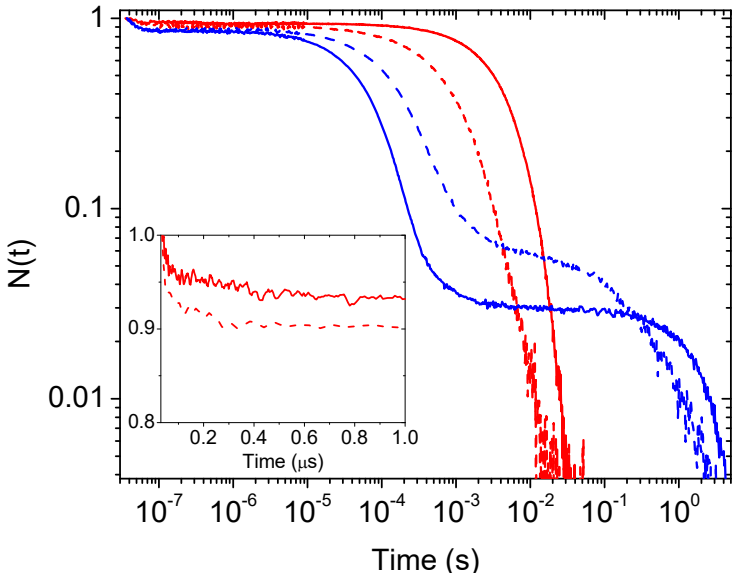


Figure 4. Comparison between the CO rebinding kinetic to chimeric (dashed lines) and wt (solid lines) proteins: Mb (red traces) and Ngb (blue traces), at T = 30°C and 0.1 atm of CO. **Inset.** Expanded view of the first 1 μs of geminate recombination showing the slight increase from Mb (solid) to ChMb (dashed).

The above findings for the chimeric proteins provide an intuitive demonstration that the tendency towards hexacoordination by the distal His is modulated by the CD corner. To turn this qualitative observation into a quantitative estimate of rate constants for the microscopic processes, we have analyzed the CO rebinding kinetics with the previously proposed kinetic schemes [20, 24, 25] or suitably modified in order to account for the observed kinetic heterogeneity, paying attention to the differences in bimolecular rebinding, distal His binding and dissociation kinetics (additional information on geminate recombination can be found in the Supporting Information).

Analysis of CO rebinding kinetics to Mb requires the use of a kinetic model detailed in Scheme 2, where the non-exponential geminate recombination is attributed to migration to an internal cavity and is followed by a bimolecular rebinding step. While this model describes well the CO

rebinding to Mb (see Table S2 and Figure S8), it fails to reproduce the progress curve for ChMb, mostly due to a broader bimolecular phase (Figure S9). Inclusion of the competitive binding reaction of the distal His in the kinetic model improves the agreement with the experimental curve (Figure S9). However, the retrieved rate for His dissociation ($\approx 10^3$ s^{-1}) is not fully consistent with the same parameter determined from stopped flow assays (580 s^{-1}; see Table 2). On the other hand, the bimolecular rebinding phase is well described if we assume the existence of two different conformations with distinct values of k_2 and k_{-2} (Figure 5A; see also Table S3). The total concentration of unliganded species is calculated as the weighted linear combination of the two conformations, where the weights (30% and 70%, respectively) estimated from biexponential fit of the bimolecular phase are independent of temperature and CO concentration.

$$\begin{array}{ccccc} & & T & & \\ & & k_c \uparrow\downarrow k_{-c} & & \\ MbCO & \underset{k_{-1}}{\overset{h\upsilon}{\rightleftharpoons}} & DS & \underset{k_{-2},k'_{-2}}{\overset{k_2,k'_2}{\rightleftharpoons}} & Mb+CO \end{array}$$

Scheme 2. Minimal reaction scheme for the observed CO rebinding kinetics to Mb and ChMb. After photolysis CO is in the primary docking site DS from which it can access an internal cavity T or exit to the solvent (Mb). While for Mb a single set of rates k_2 and k_{-2} is used to describe the bimolecular rebinding, for ChMb two conformations are assumed to exist in equilibrium, characterized by different rates k_2, k_{-2}, and k'_2, k'_{-2}, respectively.

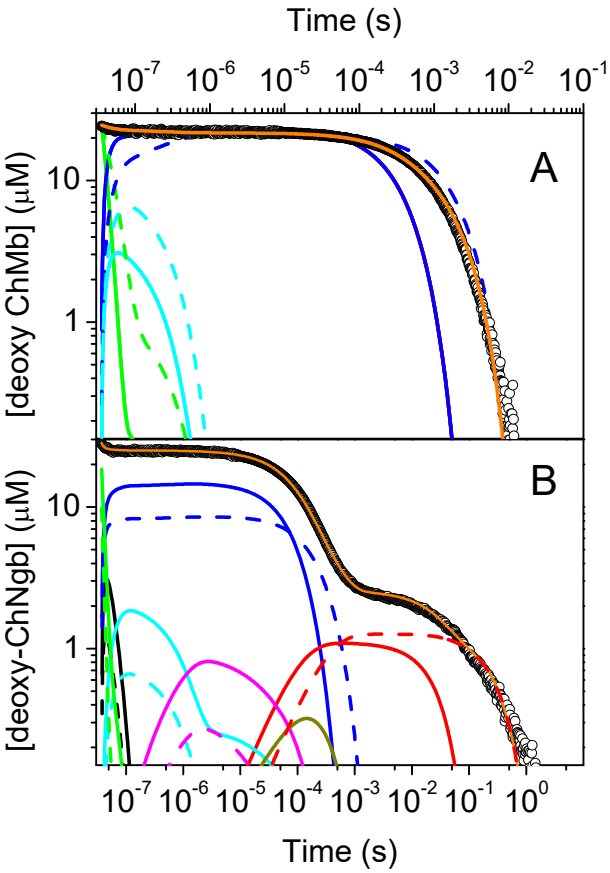
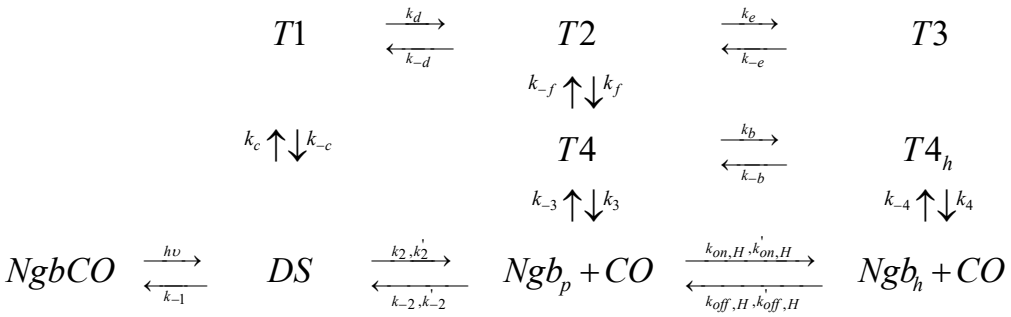


Figure 5 A. Representative analysis of the CO rebinding kinetics to ChMb, equilibrated with 0.1atm CO (black circles) at 20°C, with the model described in Scheme 2. The fitted curve (orange solid line) is superimposed to the experimental data and is calculated as the weighted average of the overall time course for rebinding to the two conformers. Curves representing rebinding to the two conformers are distinguished by the line style, dashed or solid. In the figure are also reported the time courses of reaction intermediates in Scheme 2: green, DS; cyan, T; blue, Mb_p. **B.** Analysis of the CO rebinding kinetics to ChNgb equilibrated with 0.1 atm CO (black circles) at 30°C. The fit (orange solid line) is superimposed to the experimental data. In the figure is also reported the time course of other relevant intermediates: blue, NGb_p; red, Ngb_h; green, DS; black, T1, cyan, T2; olive, T3; magenta, T4. The two dynamic species described in the text and identified by the different rates in Scheme 3 are distinguished by the line style, dashed or solid.

The analysis of CO rebinding to ChMb exposes the presence of two conformers, one of which shows a nearly thirty-fold increase in the rebinding rate k_{-2} in comparison to Mb, a fact that is taken as an indication that the entry pathway has properties similar to those of Ngb. Interestingly, recent theoretical work has shown that for Ngb a relevant escape pathway is located in proximity of the CD corner [20] and thus the observed change in k_{-2} may reflect the modulation introduced by replacing the Mb CD corner for the Ngb one. From the microscopic rates it is possible to estimate $k_{on,CO}$, which for Mb is 0.7×10^{6} $M^{-1}s^{-1}$, a value perfectly in line with the literature [12-14]. The value of $k_{on,CO}$ shows a remarkable increase for ChMb, for which two binding constants were determined, $k_{on,CO,1} = 2.6\times10^{6}$ $M^{-1}s^{-1}$ and $k_{on,CO,2} = 1.4\times10^{7}$ $M^{-1}s^{-1}$. While still smaller than the corresponding parameter for Ngb, the observed increase demonstrates that the modifications introduced in ChMb push the binding properties towards those of Ngb.

CO rebinding to Ngb occurs through a complex reaction pattern, determined by the presence of an articulate system of tunnels and cavities connecting the binding site with the solvent. Experiments under a variety of physicochemical conditions have led to the elaboration of a rather complex reaction scheme. [18-20, 25] While Scheme 3 reports the full reaction scheme, we concentrate on the main features of the reactions that are relevant to the identification of the effects of swapping the CD region. In particular, the most remarkable effects are expected on the rate constants for the equilibria DS $\rightleftarrows$ NGb_p + CO and NGb_p $\rightleftarrows$ NGb_h.



Scheme 3. Minimal reaction scheme for the CO rebinding to Ngb and ChNgb. After photolysis, CO is in the primary docking site DP from which it can access a series of internal cavities T1-T4, or exit directly to the solvent. The long lived reaction intermediates T4 and $T4_h$ offer alternative exit pathways. Given its very long lifetime, cavity T4 is allowed to relax to a $T4_h$ configuration, from which exit of CO to the solvent leads to formation of the Ngb_h species. For ChNgb two conformations exist, leading to the observed heterogeneity in CO binding. Their population is not dependent on ligation. In the model, the two species are characterized by different values of k_2, k_{-2}, $k_{on,H}$, and $k_{off,H}$.

The model can well describe the rebinding kinetics to Ngb (see Figure S9). Microscopic rates retrieved from the fitting are reported in Table S4 along with the activation parameters. Using the calculated microscopic rate constants, the equilibrium His binding constant is given by $K_H = k_{on,H}/k_{off,H} = 10^3$ at 20 °C, in general agreement with previous determinations [7, 26]

In the case of ChNgb CO rebinding to the pentacoordinated protein and to the bis-histidyl hexacoordinated species are biphasic processes, similarly to what was found for ChMb. Thus, the use of a reaction scheme as for Ngb results in unsatisfactory fits in the time range corresponding to bimolecular CO rebinding and His binding and dissociation (Figure S12). We thus assumed that two conformations exist in equilibrium and differ in rate constants k_2, k_{-2}, $k_{on,H}$ and $k_{off,H}$ as detailed in Scheme 3. The total concentration of unliganded species is calculated as the weighted linear combination of the two conformations, where the weights were estimated

from the amplitude of the decays of the bis-histidyl hexacoordinated species (Table S5). Figure 5B shows the result of the fitting of a representative rebinding kinetics, demonstrating the good agreement between the model and the measured kinetics.

The bimolecular rebinding observed for ChNgb is biphasic and slower in comparison to that for Ngb. The process is mostly ruled by the return rates k_{-2} and k'_{-2} (Table S6), both of which are still higher than the corresponding rate for Mb but not as different from those observed for ChMb. As for ChMb, it is important to evaluate the effect of the corner CD on $k_{on,CO}$. While for Ngb $k_{on,CO} = 1.05\times10^{8}$ $M^{-1}s^{-1}$, for the chimeric protein two binding rate constants are obtained: $k_{on,CO,1} = 2.6\times10^{7}$ $M^{-1}s^{-1}$ and $k_{on,CO,2} = 9.5\times10^{7}$ $M^{-1}s^{-1}$. These values are lower than the rate constant determined for Ngb, as expected for the modification performed on the protein, and demonstrate that the binding properties of ChNgb are tuned towards those of Mb.

Finally, the equilibrium constants for the binding of the distal histidine in the two conformations are $K_H = k_{on,H}/k_{off,H} = 40$ and $K_H' = k'_{on,H}/k'_{off,H} = 170$. Both equilibrium constants are lower than the value measured for the wt protein ($K_H \approx 1000$), demonstrating the decrease in affinity of the HisE7 for the heme Fe in the chimeric protein.

Discussion

In spite of their functional relevance, the identification of the regulatory mechanisms implicated in the equilibrium between bis-histdyl hexacoordinated and pentacoordinated globins is still a challenging question. [7, 11, 23, 27]. An interesting class of proteins showing bis-histidyl hexacoordination is formed by the non symbiotic plant hemoglobins (nsHbs), where incomplete hexacoordination is know to occur for class 1 nsHbs, whereas tighter binding is observed for class 2 nsHbs, resulting in full hexacoordination. [8] The structural rearrangements originated from the transition between 5c and 6c states have been identified in the case of class 1 plant nsHbs. Briefly, distal His dissociation from the heme iron is accompanied by rotation and translation of the E helix, formation of a D helix, and development of several new trans-helical

contacts in the EF turn [28]. Most of the structural changes between the 5c and bis-histidyl 6c species are associated with the E helix and its termini, and a structured D helix and stabilization of the EF turn were proposed as factors to support the 5c E-helix position. [27] This structure was found to be consistent with that shown by leg-hemoglobins, which are incapable of stabilizing the 6c configuration [29]

In contrast to the class 1 nsHbs, the experimental X-ray evidence available for Ngb indicate that the structural transition triggered upon ligand binding is quite small, being mainly limited to sliding away of the heme from the distal His to produce a binding site, with little rearrangement of the remainder of the structure. [30] This likely reflects the influence of the crystal packing in limiting the amplitude of the structural rearrangement, as we have shown in previous studies that the transition between 5c and 6c states leads to a marked tilting of the E helix and relocation of the D helix, in conjunction with structural rearrangements in C and D helices and the connecting loops. [11, 23] The largest fluctuations are found in the interhelical loops, particularly the CD region. [11] The ED analysis supported the fundamental role of the CD loop in controlling the 5c⇆6c equilibrium, as the first EM (accounting for more than 50% of the whole transition motion) corresponds to motion of the CD loop and a concerted movement of the E helix with respect to the F-helix.

The results presented in this work show that the CD region indeed plays a very important role in modulating the coordination equilibrium and the ligand binding kinetics, as each chimeric protein modifies the ligand binding and the distal His coordination properties towards those which are typical of the corresponding protein the CD region was swapped from. Thus, a peculiar aspect of the chimeric proteins is that they both exhibit heterogenous CO rebinding kinetics. The two bimolecular rebinding phases for ChNgb are characterized by quite different $k_{on,CO}$ values. The faster process has a $k_{on,CO}$ which is similar to the one measured for the wt protein, whereas the slower process has a 4-fold smaller $k_{on,CO}$. Although this value is still much larger than the corresponding value for Mb, the change in $k_{on,CO}$ is in the direction expected for

the perturbation introduced in the structure. Similarly, the bimolecular phase for CO rebinding to ChMb is also heterogeneous. Both species show $k_{on,CO}$ values larger than Mb, although still smaller than the one characteristic for Ngb.

Heterogeneity is also observed for HisE7 binding and dissociation in ChNgb. At present, the reason for such heterogeneity is unclear, though it might emerge from the presence of distinct conformations of the CD region, affecting the junction with E helix, the presence of different His tautomers [31] or even heme orientations. [32] Both configurations are characterized by rates that lead to a decrease in the hexacoordination equilibrium constant in comparison to that of Ngb. Thus, the CO rebinding kinetics afforded equilibrium binding constants K_H= 40 and K_H'= 170, much lower than the corresponding value for the wt protein (≈1000). On the other hand, the increased tendency to form 6c bis-histidyl species in ChMb is evident from the analysis of the absorption spectrum. Deconvolution of ChMb equilibrium spectra affords a 6c fraction of 25%, indicating a significant coordination equilibrium shift upon swapping the CD region.

Taken together, the above evidences suggest that swapping the CD region affects the reactivity of the proteins by tuning the tendency for hexacoordination, but also clearly demonstrate that simply swapping the CD regions is not enough to completely turn Mb into Ngb and vice versa. This result points out that other regions (as EF loop) are co-responsible for needed rearrangements and that the CD "subdomain" may introduce/inhibit interactions with the "framework" of the parental protein that are absent in the wild type.

Conclusions

Motivated by the evidence gained from MD simulations which suggested a key role of the CD loop in modulating globin hexacoordination, we have built chimeric proteins where the CD regions of Ngb and Mb have been swapped. *In silico* analysis of these chimeric proteins indicates that the peculiar features of the structure and the dynamics of the parent proteins are

generally transferred to the chimeras. Stopped flow and laser flash photolysis experiments on the recombinant chimeric proteins confirm this trend and demonstrate that CO rebinding rates and distal His binding and dissociation rates are affected in a way to make chimeric proteins similar to those from which the CD region was swapped from. Taken together, our results confirm the key role played by the CD region in the modulation of the 6c⇆5c equilibrium in globins.

Experimental Section

Protein cloning expression and purification. The pGX2T plasmid DNA containing the sequence of wild-type and chimera globins were transformed into *E. coli* strain BL21 (DE3) cells. The culture were grown at 37°C in TB medium (2.4% yeast extract, 1.2% bactotryptone, 0.4% glycerol, 72 mmol/L potassium phosphate buffer, pH 7.5) containing 100 μg/mL ampicillin and 8 μg/L hemin . The expression was induced as the optical density reached 0.8 at 600 nm by the addition of IPTG to a final concentration of 0.1 mM, and expression was continued overnight at 28°C in an orbital shaker. The cells were harvested at 4°C by centrifuging at 4000 g for 30min and washed with cold PBS and resuspended in lysis buffer (PBS buffer, 1% Tween 20, 1mM DTT) with protease inhibitors (0.2 mM PMSF, 0.4 μM Benzamidine, 3 μM Pepstatin A, 1.5 mM EDTA). Collected cells were then exposed to a freeze-thaw step and were sonicated until completely lysed. Lysate and cell debris were separated by centrifuging at 10.000 g for 45 min at 4°C. GST fusion globins were purified from the supernatant using a GSTrap system. Crude globin sample was bounded to the GSTtrap column prequilibrated with binding buffer (PBS buffer, 0.1% Tween 20, 1mM DTT). Column was washed with the binding buffer until no unbound material is detected in the effluent. Protein of interest was eluted with elution buffer (50 mM Tris-HCl, pH 8.0 , 10 mM reduced glutathione, 0.1% Tween 20, 1 mM DTT) followed at 280nm. The eluted goblin solution was concentrated using Amicon centrifugal filters and

stored in 0.1 M Phosphate buffer pH7, 1.5 mM EDTA, 1 mM DTT. A sample was taken for analyses on SDS-PAGE.

UV-Vis absorbance spectra. Electronic absorption spectra of purified proteins in 100 mM Phosphate buffer (pH 7) were recorded with a UV/ visible spectrophotometer (Jasco, V-650) at 20°C.

Kinetic stopped flow measurements of CO binding. Rapid mixing experiments were conducted with a thermostated stopped flow apparatus (Applied Photophysics, Salisbury, UK). Kinetics of CO binding to determine the k_{off} of distal His were measured on the deoxy state of chimeric and wild type globins. Solutions containing 10 μM protein in a 100 mM sodium phosphate, 1 mM DTT at pH 7.0 were degassed in a helium atmosphere and reduced with an equimolar concentration of sodium dithionite.

In order to obtain the individual kinetic constants we fit the k_{obs} as a function of CO concentration[7]

(1)

where $k_{off,H}$ corresponds to the limiting value of k_{obs} at high CO concentrations . The value of $k_{on,CO}$ was taken from flash photolysis experiments.

Flash photolysis measurements of CO rebinding. Globins were diluted in 100 mM phosphate pH 7.0 to a final concentration of 40 μM and incubated overnight with 10 mM DTT. The samples were then reduced under anaerobic conditions with sodium dithionite at a final concentration of 10 mM and finally equilibrated at either 1 atm or 0.1 atm CO in a gas-tight cuvette. The laser-flash-photolysis setup has been described elsewhere. [20, 24] Photolysis was achieved with the second harmonic of a Q- switched nanosecond Nd:YAG laser (Spectron Laser). Absorbance changes were monitored using the monochromatic cw output of a 150 W Xe arc lamp coupled to a 0.25-m monochromator (AMKO gmbh). The transient absorbance traces were measured through a second 0.125-m monochromator (77250, LOT-Oriel) with a

5-stages photomultiplier (Applied Photophysics). The voltage signal was digitized by a digital oscilloscope (LeCroy Waverunner 104-Xi, 5 GS/s; 1 GHz). A custom dichroic filter (Omega optical) was positioned between the exit slit of the monochromator and the photomultiplier to remove residual stray light from the pump laser. A fast shutter (Vincent Associates, Uniblitz VS35 controlled by the driver VMM-T1) was positioned between the output of the first monochromator and the sample holder. The synchronisation of the laser and the shutter was controlled by a digital delay generator (Berkeley Nucleonics). The sample holder was accurately temperature controlled with a Peltier element, allowing temperature stability of at least 0.1°C.

Derivation of kinetic rate constants from flash photolysis kinetics. The minimal model necessary to account for the observed kinetics is sketched in Schemes 2 and 3. The differential equations associated with reaction Schemes were solved numerically, and the rate constants were optimized to obtain a best fit to the experimental data. Numerical solutions to the set of coupled differential equations were determined by using the function ODE15s within Matlab 7.0 (The MathWorks, Inc.). Fitting of the numerical solution to experimental data (and optimization of microscopic rate constants) was obtained with a Matlab version of the optimization package Minuit (CERN). In order to improve the retrieval of microscopic rate constants, data at the same temperature and different CO concentrations were simultaneously fitted, thus generating a set of rate constants at each temperature. This global analysis was repeated at several different temperatures between 10 °C and 40 °C. [33]

Computational Methods

Setup of simulated systems. Starting structures of wt 6c Ngb were built combining the murine and human X-ray structures (PDB entries 1Q1F [30] and 1OJ6 [16] respectively) as described previously. [11, 23] ChMb displays Mb sequence except for residues 41 to 59 that are replaced by Ngb residues 39 to 58. ChNgb displays Ngb sequence except for residues 39 to 58 that are replaced by Mb residues 41 to 59. The criteria to switch CD protein segments was to maximize

the flexible region to be exchanged while preserving flanking residues identical in both proteins. The 5c structure was obtained from the 6c one by slowly dissociating the distal histidine from the heme and then releasing the Fe-HisE7 bond constraint. In all these structures cysteines CD7 and D5 are in the reduced state (i.e., without the S-S bond). The initial 5c structure of Mb was taken from the X-ray structure of sperm whale Mb (PDB code 1VXD [34]). The 6c structure was obtained by slowly approaching Nε2(HisE7) to the Fe, and afterwards adding the Fe-HisE7 bond to the force field.

Classical molecular dynamics: Simulation parameters. MD simulations were performed using the Amber99SB force field [35] for all residues but the heme, whose parameters were developed and thoroughly tested by our group in previous works.[4] Standard protonation state at physiological pH was assigned to ionizable residues (Asp, Glu, Lys and Arg) while special attention was paid to the protonation of histidines, which were assigned on the basis of the hydrogen-bond pattern with neighboring residues. For distal (HisE7) and proximal (HisF8) histidines, protonation was chosen to be in the Nδ position, since this is the protonation state that allows coordination to the heme iron. Starting structures were finally immersed in an octahedral box of TIP3P waters. [36] All simulations were performed with the PMEMD module of the AMBER11 package,[37] using the default simulations parameters, consisting in periodic boundary conditions and Ewald sums to treat long range electrostatic interactions [36], the SHAKE algorithm to keep bonds involving hydrogen atoms at their equilibrium length [36], a 2 fs time step for the integration of Newton's equations, and the Berendsen thermostat and barostat to control the system temperature and pressure respectively. [38] Equilibration consisted of an energy minimization of the initial structures, followed by a slow, 500ps long heating up to 300 K (in the NVT ensemble), and a final 500ps long density equilibration procedure (in the NPT ensemble). MD runs consisted of 100ns long trajectories (with the exception of ChMb-6c, which was run for 130ns. The first 30 ns were not considered in the analysis, except for ChMb-6c for which we discarded the first 50 ns. Frames were collected at 2 ps intervals, which were

subsequently used to analyze the production trajectories. This simulation protocol and parameters are the same as those successfully used in our previous work for the study of Mb and Ngb. [23]

Essential dynamic analysis. To examine the dynamical properties of proteins and their influence on the 5c⇆6c transition, essential dynamics (ED) analysis was used. [39] ED involves diagonalization of the covariance matrices of atomic positions along the trajectory, yielding the eigenvectors that define the essential motions of the protein. This analysis was performed only for the backbone atoms. Terminal residues were excluded in order to avoid masking of the essential motions of the protein core by the high flexibility of terminal regions. ED analysis of combined trajectories was also performed to gain insight into the 6c⇆5c structural transition. Finally, projection of the MD trajectories onto selected essential motions was performed to analyze the configurational space explored along the MD run. This type of analysis has been valuable to explore the structural and dynamical relationships in similar cases. [11, 23, 40]

Acknowledgements

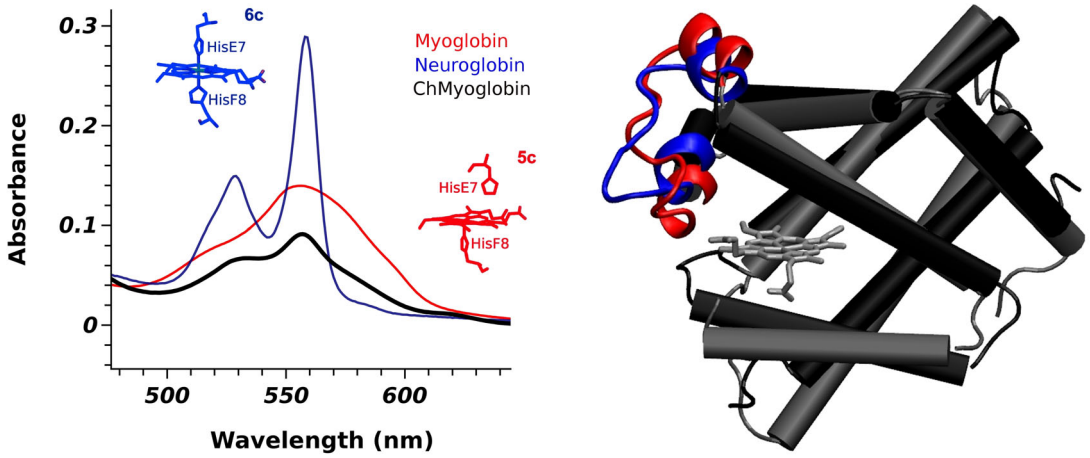
LC, DEW, DAE, MAM and ADN are staff members of CONICET. This study has been supported by PICT-2008-732, PICT-2010-0416, UBACYT 20020090300117, UBACYT 20020100100738, the Spanish Ministerio de Innovación y Ciencia (SAF2011-27642) and Generalitat de Catalunya [2009SGR249], and Ministero degli Affari Esteri, Direzione generale per la promozione del sistema Paese (Progetti di Grande Rilevanza, Italia-Argentina 2011–2013). FJL acknowledges the financial support from ICREA Academia.

References

[1] S. N. Vinogradov, D. Hoogewijs, X. Bailly, R. Arredondo-Peter, J. Gough, S. Dewilde, L. Moens and J. R. Vanfleteren, *BMC Evol Biol* 2006, *6*, 31.
[2] S. N. Vinogradov, M. Tinajero-Trejo, R. K. Poole and D. Hoogewijs, *Biochim Biophys Acta* 2013, *1834*, 1789-1800.
[3] S. N. Vinogradov, D. Hoogewijs, X. Bailly, R. Arredondo-Peter, M. Guertin, J. Gough, S. Dewilde, L. Moens and J. R. Vanfleteren, *Proc Natl Acad Sci U S A* 2005, *102*, 11385-11389.
[4] L. Capece, L. Boechi, L. L. Perissinotti, P. Arroyo-Manez, D. E. Bikiel, G. Smulevich, M. A. Marti and D. A. Estrin, *Biochim Biophys Acta* 2013, *1834*, 1722-1738.
[5] J. T. Trent, 3rd, R. A. Watts and M. S. Hargrove, *J Biol Chem* 2001, *276*, 30106-30110; T. Burmester, B. Weich, S. Reinhardt and T. Hankeln, *Nature* 2000, *407*, 520-523; M. Brunori and B. Vallone, *Cell Mol Life Sci* 2007, *64*, 1259-1268.
[6] T. Burmester, B. Ebner, B. Weich and T. Hankeln, *Mol Biol Evol* 2002, *19*, 416-421.
[7] S. Kakar, F. G. Hoffman, J. F. Storz, M. Fabian and M. S. Hargrove, *Biophys Chem* 2010, *152*, 1-14.
[8] B. J. Smagghe, J. A. Hoy, R. Percifield, S. Kundu, M. S. Hargrove, G. Sarath, J. L. Hilbert, R. A. Watts, E. S. Dennis, W. J. Peacock, S. Dewilde, L. Moens, G. C. Blouin, J. S. Olson and C. A. Appleby, *Biopolymers* 2009, *91*, 1083-1096.
[9] D. Hamdane, L. Kiger, S. Dewilde, B. N. Green, A. Pesce, J. Uzan, T. Burmester, T. Hankeln, M. Bolognesi, L. Moens and M. C. Marden, *J Biol Chem* 2003, *278*, 51713-51721.
[10] D. Hamdane, L. Kiger, S. Dewilde, B. N. Green, A. Pesce, J. Uzan, T. Burmester, T. Hankeln, M. Bolognesi, L. Moens and M. C. Marden, *Micron* 2004, *35*, 59-62.
[11] A. D. Nadra, M. A. Marti, A. Pesce, M. Bolognesi and D. A. Estrin, *Proteins* 2008, *71*, 695-705.
[12] J. S. Olson, A. J. Mathews, R. J. Rohlfs, B. A. Springer, K. D. Egeberg, S. G. Sligar, J. Tame, J. P. Renaud and K. Nagai, *Nature* 1988, *336*, 265-266.
[13] B. A. Springer, S. G. Sligar, J. S. Olson and G. N. Phillips, Jr., *Chemical Reviews* 1994, *94*, 699-714.
[14] E. Antonini and M. Brunori in *Hemoglobin and Myoglobin in Their Reactions with Ligands*, *Vol. 178* North-Holland Publishing Co., Amsterdam, 1971, p. 296.
[15] S. Dewilde, L. Kiger, T. Burmester, T. Hankeln, V. Baudin-Creuza, T. Aerts, M. C. Marden, R. Caubergs and L. Moens, *J Biol Chem* 2001, *276*, 38949-38955.
[16] A. Pesce, S. Dewilde, M. Nardini, L. Moens, P. Ascenzi, T. Hankeln, T. Burmester and M. Bolognesi, *Structure* 2003, *11*, 1087-1095.
[17] J. T. Trent, 3rd, A. N. Hvitved and M. S. Hargrove, *Biochemistry* 2001, *40*, 6155-6163.
[18] J. M. Kriegl, A. J. Bhattacharyya, K. Nienhaus, P. Deng, O. Minkow and G. U. Nienhaus, *Proc Natl Acad Sci U S A* 2002, *99*, 7992-7997.
[19] L. Kiger, J. Uzan, S. Dewilde, T. Burmester, T. Hankeln, L. Moens, D. Hamdane, V. Baudin-Creuza and M. Marden, *IUBMB Life* 2004, *56*, 709-719.
[20] D. Giordano, I. Boron, S. Abbruzzetti, W. Van Leuven, F. P. Nicoletti, F. Forti, S. Bruno, C. H. Cheng, L. Moens, G. di Prisco, A. D. Nadra, D. Estrin, G. Smulevich, S. Dewilde, C. Viappiani and C. Verde, *PLoS One* 2012, *7*, e44508.
[21] E. E. Scott, Q. H. Gibson and J. S. Olson, *J Biol Chem* 2001, *276*, 5177-5188.
[22] L. J. Smith, A. Kahraman and J. M. Thornton, *Proteins* 2010, *78*, 2349-2368; J. T. Lecomte, D. A. Vuletich and A. M. Lesk, *Curr Opin Struct Biol* 2005, *15*, 290-301.
[23] L. Capece, M. A. Marti, A. Bidon-Chanal, A. Nadra, F. J. Luque and D. A. Estrin, *Proteins* 2009, *75*, 885-894.
[24] S. Abbruzzetti, S. Bruno, S. Faggiano, E. Grandi, A. Mozzarelli and C. Viappiani, *Photochem Photobiol Sci* 2006, *5*, 1109-1120.
[25] S. Abbruzzetti, S. Faggiano, S. Bruno, F. Spyrakis, A. Mozzarelli, S. Dewilde, L. Moens and C. Viappiani, *Proc Natl Acad Sci U S A* 2009, *106*, 18984-18989.
[26] J. Uzan, S. Dewilde, T. Burmester, T. Hankeln, L. Moens, D. Hamdane, M. C. Marden and L. Kiger, *Biophys J* 2004, *87*, 1196-1204.
[27] J. A. Hoy, H. Robinson, J. T. Trent, 3rd, S. Kakar, B. J. Smagghe and M. S. Hargrove, *J Mol Biol* 2007, *371*, 168-179.
[28] F. Spyrakis, F. Lucas, A. Bidon-Chanal, C. Viappiani, V. Guallar and F. J. Luque, *Biochim Biophys Acta* 2013, *1834*, 1957-1967.
[29] M. S. Hargrove, J. K. Barry, E. A. Brucker, M. B. Berry, G. N. Phillips, Jr., J. S. Olson, R. Arredondo-Peter, J. M. Dean, R. V. Klucas and G. Sarath, *J Mol Biol* 1997, *266*, 1032-1042.

[30] B. Vallone, K. Nienhaus, M. Brunori and G. U. Nienhaus, *Proteins* 2004, *56*, 85-92.
[31] L. Boechi, M. Arrar, M. A. Marti, J. S. Olson, A. E. Roitberg and D. A. Estrin, *J Biol Chem* 2013, *288*, 6754-6762.
[32] W. Du, R. Syvitski, S. Dewilde, L. Moens and G. N. La Mar, *J Am Chem Soc* 2003, *125*, 8080-8081.
[33] S. Bruno, S. Faggiano, F. Spyrakis, A. Mozzarelli, S. Abbruzzetti, E. Grandi, C. Viappiani, A. Feis, S. Mackowiak, G. Smulevich, E. Cacciatori and P. Dominici, *J Am Chem Soc* 2007, *129*, 2880-2889.
[34] F. Yang and G. N. Phillips, Jr., *J Mol Biol* 1996, *256*, 762-774.
[35] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Structure, Function and Genetics* 2006, *65*, 712-725.
[36] A. R. Leach, *Molecular Modelling: Principles and Applications*, Pearson Education EMA, 2001, p.
[37] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham Iii, S. DeBolt, D. Ferguson, G. Seibel and P. Kollman, *Computer Physics Communications* 1995, *91*, 1-41.
[38] H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.* 1984, *81*, 3684-3690.
[39] A. Amadei, A. B. Linssen and H. J. Berendsen, *Proteins* 1993, *17*, 412-425.
[40] A. Bidon-Chanal, M. A. Marti, A. Crespo, M. Milani, M. Orozco, M. Bolognesi, F. J. Luque and D. A. Estrin, *Proteins* 2006, *64*, 457-464.

Table of Contents Art Work



Graphical Abstract

Chimeric proteins were engineered from myoglobin and neuroglobin by swapping their CD segments, previously suggested to modulate bis histidyl heme coordination. Spectroscopic and computational analysis results support the key role of the CD region in defining the heme coordination state in globins.