Ligand Uptake Modulation by Internal Water Molecules and Hydrophobic Cavities in Hemoglobins

Juan P. Bustamante,[†] Stefania Abbruzzetti,[‡] Agnese Marcelli,[§] Diego Gauto,[†] Leonardo Boechi,[∥] Alessandra Bonamore,[⊥] Alberto Boffi,[⊥] Stefano Bruno,[#] Alessandro Feis,[∇] Paolo Foggi,[‡,○,◆] Dario A. Estrin,*[,†] and Cristiano Viappiani*[,‡]

[†]Departamento de Química Inorgánica, Analítica y Química Física, INQUIMAE-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

[‡]Department of Physics and Earth Sciences "Macedonio Melloni", University of Parma, and IBF-CNR, Parma, Italy

[§]LENS, European Laboratory for Non-linear Spectroscopy, Florence, Italy

[∥]Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

[⊥]Istituto Pasteur, Fondazione Cenci Bolognetti, Department of Biochemical Sciences, University of Rome "La Sapienza", Rome, Italy

[#]Department of Biochemistry and Molecular Biology, University of Parma, Parma, Italy

[∇]Department of Chemistry "Ugo Schiff", University of Florence, Florence, Italy

[○]Department of Chemistry, University of Perugia, Perugia, Italy

[◆]INO-CNR, Florence, Italy

S *Supporting Information*

ABSTRACT: Internal water molecules play an active role in ligand uptake regulation, since displacement of retained water molecules from protein surfaces or cavities by incoming ligands can promote favorable or disfavorable effects over the global binding process. Detection of these water molecules by X-ray crystallography is difficult given their positional disorder and low occupancy. In this work, we employ a combination of molecular dynamics simulations and ligand rebinding over a broad time range to shed light into the role of water molecules in ligand migration and binding. Computational studies on the unliganded structure of the thermostable truncated hemoglobin from *Thermobifida fusca* (*Tf*-trHbO) show that a water molecule is in the vicinity of the iron heme, stabilized by WG8 with the assistance of YCD1, exerting a steric hindrance for binding of an exogenous ligand. Mutation of WG8 to F results in a significantly lower stabilization of this water molecule and in subtle dynamical structural changes that favor ligand binding, as observed experimentally. Water is absent from the fully hydrophobic distal cavity of the triple mutant YB10F-YCD1F-WG8F (3F), due to the lack of residues capable of stabilizing it nearby the heme. In agreement with these effects on the barriers for ligand rebinding, over 97% of the photodissociated ligands are rebound within a few nanoseconds in the 3F mutant case. Our results demonstrate the specific involvement of water molecules in shaping the energetic barriers for ligand migration and binding.
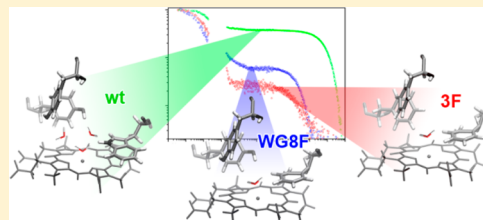
1. INTRODUCTION

It is widely accepted that water molecules play an active role in biomolecular recognition due to the partial desolvation of the ligand as well as of the region of the receptor involved in the ligand recognition process, a fact that often involves rearrangement of water molecules close to the active site.[1−5] Association processes between biomolecules immersed in an aqueous solvent require water reorganization in the contact surface. Reorganization in the solvation structure of the two binding partners leads to subtle changes in the water hydrogen-bonding network[6] that may have direct consequences over the whole binding process, especially when water molecules are embedded in a ligand-recognition cavity.[7,8] Solvent can be absent, transiently present in nonpolar cavities, or exchange between the bulk solvent and polar hydrated cavities.[9] In the latter case, the displacement of retained water molecules from protein surfaces or cavities by incoming ligands can promote favorable or disfavorable effects over the global binding process, depending on physicochemical properties of the biomolecular surfaces and cavities system. In this context, we have chosen a representative example, a small and well-characterized protein[10−13] in order to study the water influence on small ligand migration pathways from bulk solvent to protein matrix: the truncated hemoglobin of *Thermobifida fusca* (*Tf*-trHbO), a

small heme protein assigned to a distinct phylogenetic group within the globin superfamily.[14,15] The primary structure of trHbs is normally 20−40 residues shorter than mammalian Hbs, with shortened or missing α helices and modified loops.[16] The trHb family can be divided into three groups, termed I, II, and III (indicated by N, O, and P suffixes, respectively[15]). Although biochemical and physiological observations suggested scavenging activities toward NO,[17] sulfide,[18] or oxygen reactive species,[19,20] the functional role of these proteins is still unclear. The *Tf*-trHbO is the first identified thermostable group O trHb.[10] The active site is characterized by the invariant Fe−histidine covalent link on the proximal side, and by a highly polar distal environment in which WG8, YCD1, and YB10 provide three potential H-bond donors in the distal cavity to stabilize incoming ligands. WG8 and YCD1 were indeed found to be involved in the stabilization of exogenous ligands, namely, sulfide[18] and fluoride[21] in the ferric state, and CO in the ferrous state.[11]

Laser flash photolysis studies on *Tf*-trHbO showed that a rapid geminate recombination occurs with a time constant of 2 ns representing almost 60% of the overall reaction. This phase is followed by a small amplitude geminate recombination occurring around 100 ns, and a bimolecular rebinding phase extending to the millisecond time range.[12] Similarly, kinetic investigations on related trHbs showed the presence of an efficient and fast geminate recombination in the picosecond−nanosecond time scale.[13,22−25]

Experimental and theoretical investigations have shown the presence of water molecules in the distal pocket, in close contact with YB10, for the liganded structure of *Tf*-trHbO.[11] The presence of solvent in the binding pocket, although not coordinated to the active site, may impose kinetic barriers to ligand binding, due to the need to exchange water molecules for the ligand, as was previously reported by Olson and co-workers for myoglobin[26] and Ouellet et al. in the trHbN of *Mycobacterium tuberculosis*.[27]

Recent computational studies revealed that the native fold of *wild type* (wt) *Tf*-trHbO exhibits a highly polar primary docking site defined mainly by three polar amino acids WG8, YCD1, and YB10,[10] which is connected through a branched pathway with the solvent.[12] CO rebinding kinetics to wt *Tf*-trHbO was collected from the picoseconds to the milliseconds time scale and described using a model which took into account ligand migration through the dynamic system of tunnels.[12,28] On the basis of the so-far accumulated information, *Tf*-trHbO appears to be a well-characterized model to study through a combination of experimental and theoretical approaches that allow a detailed understanding of the water molecules' influence on the binding process.

Starting from the identified reaction scheme for CO rebinding kinetics in wt *Tf*-trHbO,[12] here, we focus on the consequences for kinetics, that mutation of distal site amino acids brings about. A thorough characterization of CO rebinding after laser photolysis, extending from 1 ps to 0.1 s, exposes the effects on free energy barriers for ligand binding and migration processes. To understand the structural and dynamical basis for the observed changes, we have employed two computational approaches based on unliganded protein molecular dynamics (MD) simulations: (i) analysis of the ligand migration free energy barriers along the tunnel that connects the solvent with protein matrix and (ii) probabilistic characterization of specific space regions, located inside the distal pocket, and harboring water molecules that potentially affect the global binding process. In the first place, the connection between the bulk solvent and the distal pocket is characterized by the presence of transient or permanent gates through which the dynamics of the protein allows water molecules to enter the pocket. In the second place, we have focused on the water molecules present in the distal pocket. In view of their specific interactions with distal site amino acids, the effects of mutation of WG8 to F were thoroughly analyzed. Furthermore, a triple mutant (WG8F, YCD1F, YB10F or merely 3F) where all polar interactions in the distal pocket are removed, was also considered in order to obtain a complete description of different polar cavities that enhance the water influence understanding.

2. EXPERIMENTAL PROCEDURES

2.1. Computational Modeling. *2.1.1. Setup of the Systems and Simulation Parameters.* The starting structure corresponds to the *Tf*-trHbO crystal structure (PDB entry 2BMM) as determined by Bonamore et al.[10] Amino acid protonation states were assumed to correspond to physiological pH; all solvent exposed His were protonated at the N-δ delta atom, as well as HisF8, which is coordinated to the heme iron. The system was immersed in a pre-equilibrated octahedral box of 10 Å in radius with 4912 TIP3P water molecules using the tLEaP module of the AMBER12 package.[29] All used residue parameters correspond to the parm99 Amber force field[30] except for the heme, which corresponds to those developed[31] and widely used in several heme-protein studies.[32−38] The charges and parameters for Fe(II) heme were determined by a standard procedure: partial charges were computed using the restricted electrostatic potential (RESP) recipe and DFT electronic structure calculations with the PBE functional and 6-31 G** basis sets. The calculation has been performed in the high-spin (HS) state. Equilibrium distances and angles, as well as force constants, were computed using the same methods and basis set used for computed charges. All simulations were performed using periodic boundary conditions with a 9 Å cutoff and particle mesh Ewald (PME) summation method for treating the electrostatic interactions. The hydrogen bond lengths were kept at their equilibrium distance by using the SHAKE algorithm, while temperature and pressure were kept constant with a Langevin thermostat and barostat, respectively, as implemented in the AMBER12 program.[29] The equilibration protocol consisted of (i) slowly heating the whole system from 0 to 300 K for 20 ps at constant volume, with harmonic restraints of 80 kcal per mol A2 for all C_α atoms and (ii) pressure equilibration of the entire system simulated for 1 ns at 300 K with the same restrained atoms. After these two steps, an unconstrained 100 ns molecular dynamics (MD) long simulation at constant temperature (300 K) was performed.

In silico mutant proteins, i.e., single mutant (WG8F) and triple mutant (YB10F-YCD1F-WG8F), were built starting from the same crystal structure as described above and mutated then using the tLEaP module of the AMBER12 package.[29] These mutant structures were equilibrated and simulated using the same protocol as that used for the wt form.

2.1.2. Analysis of the Ligand Migration Free Energy in Wild Type, Single and Triple Mutants of Tf-trHbO. The free energy for the CO migration process inside the protein tunnel/cavity system was computed by the Implicit Ligand Sampling (ILS) approach that uses computed MD simulation in the absence of the ligand and incorporates it afterward.[39] This method was thoroughly tested for heme proteins.[40] ILS

calculations were performed in a rectangular grid (0.5 Å resolution) that includes the whole simulation box (i.e., protein and the solvent); the probe used was a CO molecule. Calculations were performed on 5000 frames taken from the last 90 ns of simulation time. The values for grid size, resolution, and frame numbers were thoroughly tested in a previous work.[40] Analysis of the ILS data was performed using an ad hoc Fortran-90 program available upon request.[40] Besides this, the POcket Volume MEasurer program (POVME)[41] was used to estimate cavity system volume of each protein form.

2.1.3. Definition, Identification, and Characterization of Water Sites (WSs). WSs correspond to specific regions, adjacent to the protein region of interest, harboring a water molecule with a probability value higher than that of a water molecule surrounded by the bulk environment. As shown in our previous works,[42−44] these regions can be readily identified by computing the probability of finding a water molecule inside the correspondingly defined region during an explicit solvent MD simulation. The region volume used to identify the WS is arbitrarily set to 1 Å^3, and the WS center coordinates correspond to the average position of all the water oxygen atoms that visit the WS along the simulation. A water molecule is considered as occupying that WS as soon as the distance between the position of its oxygen atom and the WS center value is less than 0.6 Å. Once identified, for all putative WSs, we compute the following parameters:

(i) The water finding probability (WFP), corresponding to the probability of finding a water molecule in the region defined by the WS (using the arbitrary volume value of 1 Å^3) and normalized with respect to the bulk solvent probability to harboring a water molecule in a sphere of the same volume at the corresponding temperature and pressure values; thus, WFP is actually used as a cutoff value to decide which putative WSs are considered for further characterization. Hence, only WSs with WFPs values greater than 2 are retained.

(ii) The potential energy associated with the interaction of water molecules inside the WS with the protein and the rest of the solvent was computed, as well as the sum of Lennard-Jones 12-6 dispersion−repulsion term and a Coulombic electrostatic contribution along the MD simulation.

The computed contributions between the water located inside the WS and either the protein (E_p) or the other solvent molecules (E_w) were calculated taking into account the interaction to a cutoff distance of 8 Å, which has already been shown to yield reasonably converged results.[42] For each WS, the mean interaction energies $<E_x>$ were computed along the last 90 ns of the simulation. Total mean interaction energies $<E_t>$ of a water molecule inside the WS were then computed, as well as the standard error for all averages, being less than 0.01 kcal/mol.

2.2. Protein Expression and Purification. The acidic surface variant of *Tf*-trHbO was expressed and purified as described previously.[10,18] This engineered protein has shown a high recombinant expression level in soluble form and was obtained by mutating the surface-exposed residues F107 and R91 to E that remain exposed to the solvent, thus leaving the overall protein structure unchanged, without affecting thermo-stability or ligand binding properties. Therefore, ASV (in the following, referred to as wt *Tf*-trHbO) was taken as an engineered scaffold of the wt protein for subsequent site-
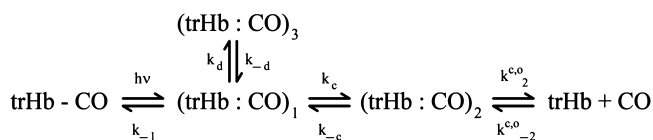
directed mutagenesis studies on the relevant residues of the distal heme pocket. In particular, our study included one single and triple mutants in which the polar distal amino acids [YB10(54), YCD1(67), and WG8(119)] were replaced with F residues. Two distal mutants of ASV were studied, namely, WG8F and YB10F-YCD1F-WG8F or merely 3F. The CO-adduct of the protein was prepared by adding a small amount (<10 mM) of sodium dithionite in the ferric protein solution at pH 7.2 under a CO atmosphere.

2.3. Femtosecond Transient Absorption Spectroscopy (TAS). The experimental setup was described previously.[45,46] Briefly, it is based on an amplified Ti:sapphire laser system delivering pulses with a time duration of ~100 fs. The output was frequency doubled for exciting the sample at 400 nm (pump energy = 0.5 μJ/pulse), while the time evolution of the excited protein was monitored by a second spectrally broad UV−visible pulse, the white continuum probe pulse, generated by focusing the fundamental beam on a calcium fluoride plate. The probe pulse is delayed with respect to the pump by means of a suitable optical line that allows scanning a time interval up to 2 ns after excitation. The repetition rate of the laser system was set at 100 Hz, and the sample was kept under continuous stirring by means of a small magnet inside the cuvette (path length = 2 mm). All the measurements were carried out by setting the relative pump−probe polarization at the magic angle (54, 7°). The detection system consists of two linear CCD arrays (Hamamatsu S8377-256Q), coupled to a spectrograph (Jobin Yvon CP 140-1824) and controlled by a homemade front-end circuit. The signals were fed into a simultaneous analog-to-digital conversion board (Adlink DAQ2010), and data were acquired by means of a LabVIEW written computer program. At each delay time, the transient absorption spectrum, from 410 nm up to 620 nm, was taken performing a pump−probe sequence of 900 shots. By repeating the sequence as a function of the pump−probe delay, we were able to obtain the dynamical evolution of the transient absorbance $\Delta A(\lambda, t)$. Kinetics extracted at different wavelengths were fitted with a multiexponential response function, convoluted with a Gaussian instrumental function (fwhm = 160 fs). Furthermore, global analysis[47] of kinetics recorded in the whole probed spectral range was applied. A sequential model was used to extract the spectral features of interest, associated with each transient.
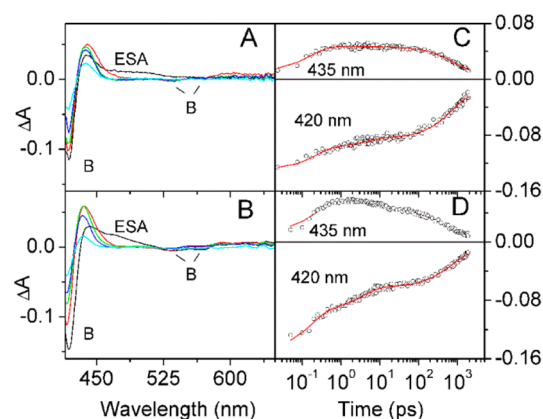
2.4. Nanosecond Flash Photolysis (LFP). The laser setup was described previously.[48] Photolysis was achieved by a frequency doubled (532 nm, 12 ns) nanosecond Nd:YAG laser (Spectron), and absorbance changes were monitored at 435 nm. Typically, 100 traces at a 0.5 Hz repetition rate were averaged to yield a single transient trace. Time-resolved spectra were acquired as described.[48] The sample holder is accurately temperature-controlled with a Peltier element (Flash100, Quantum Northwest, Inc.), allowing a temperature stability better than 0.1 °C. The concentration of the protein was ~30 μM.

The minimal model sketched in Scheme 1 was used to describe the rebinding kinetics. Numerical solutions to the set of coupled differential equations corresponding to Scheme 1 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).

Scheme 1. Extended Minimal Reaction Scheme for the Observed CO Rebinding Kinetics[a]

$$(trHb : CO)_3$$

$$trHb\text{-}CO \underset{k_{-1}}{\overset{h\nu}{\rightleftharpoons}} (trHb : CO)_1 \underset{k_{-c}}{\overset{k_c}{\rightleftharpoons}} (trHb : CO)_2 \underset{k^{c,o}_{-2}}{\overset{k^{c,o}_2}{\rightleftharpoons}} trHb + CO$$

with $k_d \big\uparrow\big\downarrow k_{-d}$ between $(trHb:CO)_3$ and $(trHb:CO)_1$.

[a] $(trHb: CO)_1$ and $(trHb: CO)_2$ indicate, respectively, the primary and secondary docking sites for the photodissociated CO inside the distal pocket along the exit (entry) pathway to (from) the solvent, while $(trHb: CO)_3$ represents a reaction intermediate with CO in a temporary docking site accessible from the distal site. For the wt protein, two static conformations are differing in the rate constants $k^{c,o}_{-1}$, $k^{c,o}_2$, and $k^{c,o}_{-2}$.

Time-resolved difference spectra were analyzed by Singular Value Decomposition (SVD),[48] using MATLAB (The Mathworks, Inc., Natick, MA). Most of the higher-order components of the SVD contain no real spectral information and correspond to noise with a random time dependence. A first criterion for the selection of usable components is the magnitude of the singular values, the higher values being the meaningful ones. The selected components can be further screened by evaluating the autocorrelations of the corresponding columns of U and V and rejecting the component if either autocorrelation falls below 0.8.[49] An additional procedure to increase the S/N ratio of the spectra and remove fluctuations of the baseline was applied, as described by Eaton and co-workers.[50]

3. RESULTS

3.1. CO Rebinding Kinetics. Taking advantage of the methodology developed for the wt protein,[12] which proved capable of characterizing the ligand rebinding kinetics over a temporal dynamics spanning more than 10 orders of magnitude, we have studied CO rebinding kinetics to WG8F and 3F *Tf*-trHbO by merging data obtained by transient absorption spectroscopy (TAS; ps to 2 ns) and laser flash photolysis (LFP; 20 ns to 100 ms).

TAS spectra of the CO complexes of the two mutant proteins (Figure 1A,B) appear very similar. Upon excitation in the Soret band, a strong bleach signal (B) at 420 nm and a very broad excited-state absorption (ESA) band appear in less than ~200 fs, our instrumental response function. In the red region of the probed spectral range, bleaching signals of the Q bands can be recognized with a very weak and broad ESA band at around 580 nm. As has been observed for wt *Tf*-trHbO,[12] the sharp positive band at ~435 nm in the transient spectra arises from the ground-state absorption of the pentacoordinated species (5c) and not from an excited state; for this reason, it is commonly referred to as the antibleaching (AB) band.[51] Figure 1C,D shows the kinetic profiles extracted at 420 nm (B maximum) and at 435 nm (AB maximum) for the WG8F and 3F mutants, along with the fits using multiexponential decay functions. The amplitudes and the time constants of the exponential decays are reported in Table 1.

The decay associated spectra (DAS) obtained by globally analyzing all the kinetic profiles are shown in the Supporting Information (Figure S1). The time constants well reproduce those obtained from the single-wavelength kinetic fits, reported in Table 1.

The spectral shapes and the time constants of the two faster transients observed for the CO complexes of WG8F and 3F *Tf*-trHbO nicely match the transients observed after excitation of

Figure 1. Transient absorption spectra of the CO complexes of WG8F (A) and 3F *Tf*-trHbO mutants (B) excited at 400 nm with femtosecond laser pulses at 20 °C. The spectra are shown at selected delay times: 0.2 ps (black), 1 ps (red), 10 ps (green), 100 ps (blue), and 1 ns (cyan). Kinetic profiles at single wavelength for WG8F (C) and 3F *Tf*-trHbO mutants (D). Solid red lines are the results of the fitting to a multiexponential decay. Parameters obtained from the fittings are reported in Table 1.

5c-wt *Tf*-trHbO, thus showing that these two transients arise from excited-state relaxation of the heme and do not reflect geminate recombination.[12]

Importantly, the spectral component associated with the slower time constant (400 ps for WG8F and 700−800 ps for 3F *Tf*-trHbO) resembles the properly scaled steady-state absorption difference spectrum (5c-trHb − CO-*Tf*-trHbO) for the thermally equilibrated ground-state protein. The intensity of these transient signals decreases in time, indicating the occurrence of geminate recombination.

For both mutants, the scaling factor applied to the static difference spectrum necessary for matching transient data gives an estimate of the molar fraction of photodissociated 5c-form, i.e., the initial photoproduct. Under our experimental conditions, we found values of 0.12 for WG8F and 0.16 for 3F *Tf*-trHbO. At the end of the probed time window (~2 ns), a residual absorbance change remains for both WG8F and 3F mutants, accounting for about 20% of the initial unliganded concentration.

The CO rebinding kinetics following nanosecond laser photolysis was recorded as a function of CO concentration and temperature and is reported as fraction of unliganded hemes versus time after photolysis in the Supporting Information (Figure S2). For both mutants, the rebinding curve is dominated by a large geminate recombination, which is unaffected by CO concentration and increases upon lowering temperature. The amplitudes and the rates of geminate rebinding for both mutants are larger than the one elicited by the wt protein. A bimolecular phase, with rate dependent on CO concentration, then follows on the millisecond time scale.

Time-resolved absorption spectra measured after nanosecond laser photolysis do not sense relevant conformational relaxations following CO photodissociation (see Figure S3, Supporting Information). The SVD analysis of time-resolved differential absorption spectra collected after nanosecond photolysis yields only one significant spectral component, superimposable to the spectral difference between CO-*Tf*-trHbO and 5c-*Tf*-trHbO. This demonstrates that single wavelength kinetics in LFP tracks the ligand rebinding kinetics.

Table 1. Amplitudes and Lifetimes Obtained by Fitting the Time Course of Absorbance Changes at the B (420 nm) and AB (435 nm) Maxima in the TAS[a] Data for WG8F and 3F *Tf*-trHbO at 20 °C[b]

	WG8F		5c-wt	3F	
	420 nm	435 nm	455 nm	420 nm	435 nm
A_1	−0.08	−0.06	−0.05	−0.06	−0.03
τ_1 (ps)	0.2 ± 0.1	0.2 ± 0.1	0.3 ± 0.1	0.2 ± 0.1	0.2 ± 0.5
A_2	−0.034	0.013	0.042	0.015	0.001
τ_2 (ps)	4 ± 1	11 ± 3	5.5 ± 0.7	4 ± 1	10 ± 3
A_3	−0.047	0.036		−0.064	0.036
τ_3 (ps)	600 ± 200	470 ± 70		700 ± 200	800 ± 100
A	−0.016	0.010		−0.021	0.011

[a]Transient absorption spectroscopy. [b]Parameters obtained from the fitting of the signal measured at 455 nm for the unliganded (5-c) wt *Tf*-trHbO are also reported.

Following the merging procedure of the kinetics acquired with TAS and LFP detailed in our previous work,[12] we built progress curves for CO binding to WG8F and 3F *Tf*-trHbO extending from a few picoseconds to several hundred milliseconds (Figure 2). For comparison, in the same plot,
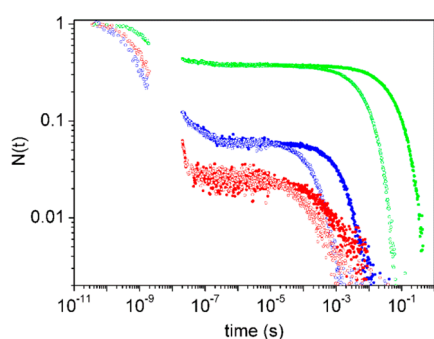
Figure 2. Merged rebinding curves measured with TAS and LFP experiments for wt (green), WG8F (blue), and 3F (red) *Tf*-trHbO mutants. Data are displayed as a fraction of the unliganded protein ($N(t) = \Delta A(t)/\Delta A(t_0)$), where t_0 is the time at which geminate rebinding begins after excited-state relaxation) as a function of the delay time after excitation. The LFP kinetic traces in the first 20 ns are affected by the instrumental function and are accordingly omitted in the plots. CO rebinding kinetics were measured in solution equilibrated with 1 (open circles) and 0.1 (filled circles) atm CO. T = 20 °C; λ = 436 nm. The range between 2 and 20 ns is not accessible to either technique used in this work.

we have also reported the progress curve for wt *Tf*-trHbO. For both mutants, the geminate rebinding is faster than that for wt *Tf*-trHbO and has larger amplitude. In particular, photodissociated CO molecules rebind geminately to WG8F with the highest rate constant. The amplitude of the geminate rebinding phase, estimated from the residual absorbance at ∼1 μs, is 97% in 3F and 94% in WG8F, in contrast with the value of 63% observed for wt *Tf*-trHbO.

As observed for wt *Tf*-trHbO, unimolecular processes show kinetic features that extend well beyond the picosecond time scale. This fact suggests that photodissociated ligands may migrate to transient docking sites located farther from the distal pocket. From a comparison between the signals on the nanosecond time scale, migration to secondary docking sites appears to be more favored in the WG8F and 3F mutants than in the wt protein.
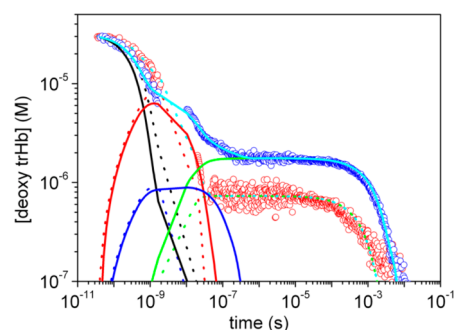
The complete time courses of the rebinding kinetics were analyzed using the microscopic model reported in Scheme 1, which was derived for the wt protein and is supported by MD simulations explained in the following sections.[12] Rebinding curves were modeled by optimizing the rate constants in Scheme 1. The fitted parameters are reported in Table 2. Figure

Table 2. Microscopic Rate Constants Determined from the Global Fit of the ps−ms Entire Time Course (at 1 and 0.1 atm CO) of CO Rebinding Kinetics to wt *Tf*-trHbO, WG8F *Tf*-trHbO, and 3F *Tf*-trHbO at 20 °C

	wt	WG8F	3F
k_{-1} (10^8 s^{-1})	3.0 ± 0.1	20.0 ± 0.6	11.5 ± 0.3
k_c (10^8 s^{-1})	1.9 ± 0.1	6.6 ± 0.3	6.5 ± 0.3
k_{-c} (10^6 s^{-1})	2.0 ± 0.2	84 ± 8	200 ± 20
k_d (10^7 s^{-1})	3 ± 1	8 ± 2	8 ± 2
k_{-d} (10^6 s^{-1})	6 ± 1	9 ± 1	230 ± 40
k_2^c (10^7 s^{-1})	3.9 ± 0.2	2.0 ± 0.1	0.9 ± 0.1
k_{-2}^c (10^6 M^{-1} s^{-1})	3.0 ± 0.3	5.0 ± 0.5	6.2 ± 0.6
k_2^o (10^7 s^{-1})	9 ± 4		
k_{-2}^o (10^7 M^{-1} s^{-1})	7 ± 3		

3 shows selected fitting curves (at T = 20 °C and CO = 0.1 atm) for WG8F and 3F *Tf*-trHbO and also reports the time courses of the reaction intermediates, determined through the fitting procedure. From the temperature dependence of the microscopic rate constants, we determined the free energy barriers associated with each reaction step, reported in Table 3.

Figure 3. Results of global analysis of the complete course of CO binding kinetics to single mutant WG8F (blue circles) and triple mutant 3F *Tf*-trHbO (red circles) at T = 20 °C and 0.1 atm CO. The fits (cyan lines) are superimposed to the experimental data (circles). In the figure, we have also reported the time course of the other relevant species in Scheme 1: (trHb: CO)$_1$ (black), (trHb: CO)$_2$ (red), (trHb: CO)$_3$ (blue), (trHb) (green). Solid lines, single mutant WG8F; dotted lines, 3F mutant.

Table 3. Activation[a] Enthalpies ($\Delta H^\#$, kcal mol^{-1}), Entropies ($\Delta S^\#$, cal mol^{-1} K^{-1}), and Free Energies at 20 °C ($\Delta G^\#$, kcal mol^{-1}) Determined from the Global Fit of the Overall ps−ms Time Course (at 1 and 0.1 atm CO) of CO Rebinding Kinetics to wt, WG8F, and 3F *Tf*-trHbO

	wt			WG8F			3F		
	$\Delta S^\#$	$\Delta H^\#$	$\Delta G^\#$	$\Delta S^\#$	$\Delta H^\#$	$\Delta G^\#$	$\Delta S^\#$	$\Delta H^\#$	$\Delta G^\#$
k_{-1}	−19.7 ± 0.1		5.8 ± 0.1	−15.9 ± 0.1		4.7 ± 0.1	−16.9 ± 0.1		4.9 ± 0.1
k_c	−20.6 ± 0.1		6.0 ± 0.1	−18.16 ± 0.02		5.3 ± 0.1	−18.1 ± 0.1		5.3 ± 0.1
k_{-c}	−29.6 ± 0.1		8.7 ± 0.1	−22.19 ± 0.05		6.5 ± 0.1	−20.6 ± 0.1		6.3 ± 0.1
k_d	−24.3 ± 0.1		7.1 ± 0.1	−22.2 ± 0.1		6.5 ± 0.2	−22.3 ± 0.1		6.5 ± 0.1
k_{-d}	−26.7 ± 0.4	0.1 ± 0.1	7.9 ± 0.4	14 ± 9	12 ± 2	8 ± 5	−18.8 ± 0.2	0.4 ± 0.1	5.9 ± 0.1
k^c_2	−11 ± 3	4 ± 1	7 ± 3	−6 ± 1	5.5 ± 0.3	7.3 ± 0.5	−0.5 ± 5	8 ± 1	8 ± 3
k^c_{-2}	33 ± 3	18 ± 1	8 ± 3	21 ± 5	14 ± 1	8 ± 3	19 ± 12	13 ± 4	8 ± 7
k^o_2	6 ± 5	8 ± 1	6 ± 4						
k^o_{-2}	47 ± 13	20 ± 3	6 ± 10						

[a]Activation enthalpies $\Delta H^\#$ and entropies $\Delta S^\#$ were estimated from the linear Eyring plots for each rate constant k_i in the temperature range of 10−30 °C, according to the equation: $\ln(hk_i/k_BT) = \Delta S^\#/R - \Delta H^\#/RT$, where R is the gas constant, h is Planck's constant, and k_B is Boltzmann constant.
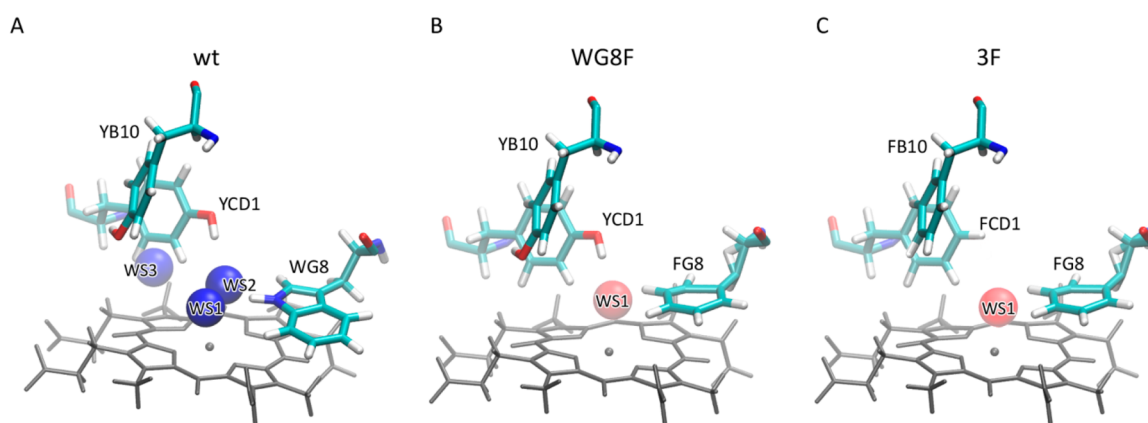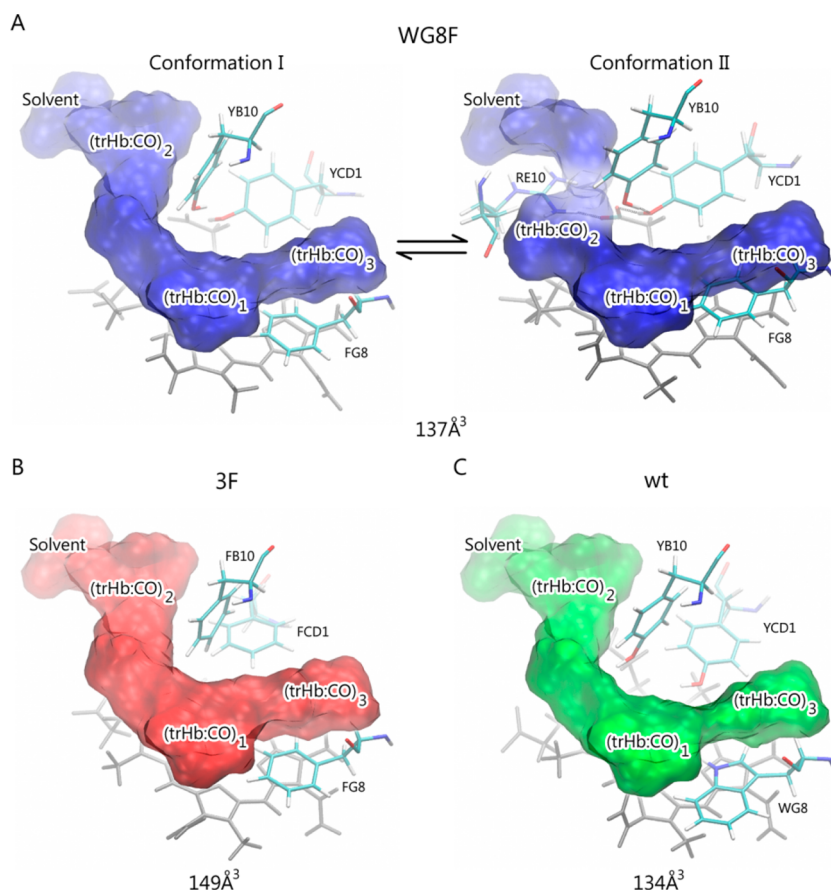

Figure 4. Schematic representation of WSs surrounding the heme group. WSs are depicted as balls from high probability to low probability on a blue to red color scale. Panels A, B, and C show results for wt, WG8F, and 3F *Tf*-trHbO, respectively.

3.2. Molecular Basis of Global Binding Process.

The strikingly peculiar kinetic profiles and the associated thermodynamical characterization of microscopic rate constants in mutant proteins, presented in the previous section, prompted us to investigate the dynamical behavior of internal cavities and water molecules inside wt, WG8F, and 3F *Tf*-trHbO, using MD simulations. Analysis of solvent molecules occupancy in the distal site environment, mainly determined by the dynamics of H-bonds with distal residues, reveals different occupancy degrees, or different probability values to find a water molecule in a particular position inside the ligand-recognition cavity. In this context, solvent molecules occupancy degrees are shown in Figure 4 for wt, WG8F, and 3F mutant forms. Particularly, Figure 4A shows that three water molecules are strongly stabilized by H-bond interactions in the wt protein, with a highly defined position (see water sites WS1, WS2, and WS3). Analysis of the simulations reveals that WG8 is the main responsible for water stabilization in WS1, with the assistance of YCD1. On the other hand, in the absence of WG8 (WG8F and 3F *Tf*-trHbO), WS1 becomes less stabilized since solvent molecules lack a very important H-bond donor. Only YCD1 partially retains a stabilizing role for the water molecules of WS1 in the WG8F mutant (Figure 4B). For the 3F mutant, all polar interactions in the distal pocket are lost (Figure 4C). Regarding the latter mutant, it is important to emphasize that the situation on the distal pocket could be quite different depending on the oxidation state of the iron heme. In 3F *Tf*-trHbO, we have previously described a water molecule coordinated to heme iron in the ferric state, and stabilized by interactions with additional solvent molecules.[52]

Because of the fact that the heme iron is in the ferrous state when the CO is bound, we assume that, when CO is photodissociated, the Fe remains in the same oxidation state. In the absence of the exogenous ligand, the position normally occupied by CO could be taken by those solvent molecules that are close to the distal pocket as was previously, and these molecules are markedly different since, for the Fe(II) state, no solvent molecules are coordinated to the metal. In this case, solvent molecules are only weakly stabilized by other water molecules that come temporarily in contact with them inside the distal cavity. Another issue that should be considered is that the distal residues' environment needs a small structural rearrangement after CO dissociation to achieve the equilibrium deoxy state. We assume that a picosecond−nanosecond time scale is completely enough to acquire the minimum structural rearrangements on the distal pocket and the subsequent water entry to identified WSs, based on the following: (i) only small side chain rearrangements of distal residues are required, and (ii) it is entirely feasible in this small time range that water molecules enter to WSs considering that they are in close proximity to thedistal pocket on the CO-bound heme state, as was previously reported by Droghetti et al.[11] From MD simulations, it is also possible to obtain thermodynamic information on the studied systems, as detailed above in the

Figure 5. Schematic representations of the heme distal residues and the tunnel and cavity system estimated with ILS for WG8F (A), 3F (B), and wt (c) *Tf*-trHbO. In (A), a conformational equilibrium between two conformations is shown, which accounts for the barrier encountered by the ligand along the exit pathways. Estimated volume of each tunnel is also shown.

Experimental Section. Water molecules in WS1 show stronger interactions for the wt protein in comparison to those observed for the investigated mutants. The electrostatic interaction energy is roughly 8.0 kcal/mol more negative for wt than for WG8F *Tf*-trHbO and 16.0 kcal/mol than for the 3F mutant. On the other hand, the Lennard-Jones interaction energy is less favorable in wt *Tf*-trHbO than in the two mutant proteins. This can be rationalized by considering that the WG8F and 3F mutants have less bulky residues and, therefore, impose a smaller steric hindrance in the distal cavity. However, this energy contribution (2.1 kcal/mol in wt, 0.5 kcal/mol in WG8F, and −1.6 kcal/mol in 3F *Tf*-trHbO) is not enough to overcome the favorable electrostatic stabilization due to the presence of polar groups capable of interacting and establishing H-bonds with waters molecules. Thus, the overall interaction energy is more favorable for wt (−14.8 kcal/mol) than for WG8F (−8.5 kcal/mol) or 3F *Tf*-trHbO (−2.6 kcal/mol).

Finally, we computed the water finding probability (WFP) that reflects the relative probability to find a water molecule inside the WS in comparison to the bulk solvent. In accordance with energetic and structural analysis, the WFP approach shows that the probability of finding a water molecule near the iron atom of the heme group in the distal pocket is larger for wt *Tf*-trHbO than for the mutant proteins, the WFP values being 19.2, 11.6, and 2.5 for wt, WG8F, and 3F *Tf*-trHbO, respectively.

Less-stabilized water molecules in the distal pocket site will be more easily displaced and leave room for incoming CO molecules. This finding perfectly fits with the picture emerging from CO rebinding curves. For the investigated mutants, the geminate phase is larger in both amplitude and rate than for the wt protein (see Figure 2). Accordingly, data by MD simulations show that water molecules inside the distal pocket are less retained in the mutants than in the wt form.

Rebinding curves shown in Figure 2 for the mutant proteins display different bimolecular recombination phases, which arise from two main sequential events, following the diffusion limited encounter of the protein and the diatomic ligand. The ligand needs to re-enter from the bulk solvent into the protein matrix, going through the on-pathway tunnel and the distal pocket. The ligand will eventually encounter a free energy barrier imposed by the presence of water molecules in the distal site. On the basis of our modeling, this barrier is expected to be different for wt and mutant proteins.

The first barrier, encountered by ligands when they diffuse through the tunnel that connects the distal pocket with the solvent, is expected to be strongly coupled with protein dynamics and the internal cavities.

The ILS computational method[40] was employed to analyze the ligand pathway that connects cavities in wt, WG8F, and 3F mutant proteins, and determine the dynamical accessibility of the distal pocket. POVME software[41] was used also to estimate the volume of each cavity system. In all cases, three main cavities were identified, $(trHb: CO)_1$, $(trHb: CO)_2$, and $(trHb: CO)_3$. Their overall average volume computed along 90 ns trajectories is roughly the same in wt ($134 \ \text{Å}^3$) and WG8F (137

G

dx.doi.org/10.1021/jp410724z | *J. Phys. Chem. B* XXXX, XXX, XXX−XXX

Å^3), whereas, for 3F Tf-trHbO, the volume is slightly larger, being around 149 Å^3. Figure 5 shows the dynamical pathways sampled by MD simulations and detected by ILS calculations for the studied proteins. A thorough analysis of the simulations including ILS results show that, for the WG8F mutant samples, there are two conformational states that we will call I and II, depicted in Figure 5A. During the time scale of our simulations, the population of the two conformations was evenly sampled. Conformation I presents a cavity system that remains very similar to the one observed for the wt protein (Figure 5C), but the connection between cavities seems to be slightly different. Conformation II implies a rearrangement of the distal site in which YCD1 is able to form a H-bond network together with RE10, YB10, and the heme group. In this conformation, the (trHb: CO)$_2$ cavity becomes smaller and roughly 3 Å closer to the distal pocket (trHb: CO)$_1$, due to partial and transient tunnel and cavity breakage. To quantify the contribution of mutated residues to the activation free energy barriers along the tunnel and cavity system shown in Figure 5, the corresponding free energy profiles for CO migration shown in Figure 6 were
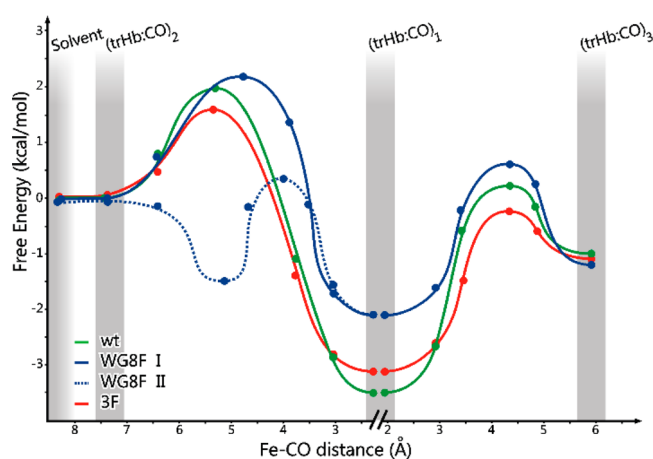
Figure 6. Free energy profiles along the connection between solvent and (trHb: CO)$_2$, (trHb: CO)$_1$, and (trHb: CO)$_3$ cavities. The results for wt, the two conformations for WG8F, and 3F Tf-trHbO are depicted in green, blue, and red lines, respectively. Circles represent calculated free energy values with the ILS method, and lines correspond to a fitting estimation on these calculated values. The x coordinate is defined by the Fe−CO distance through the pathway.

determined. The free energy was set to a value of 0 at 9 Å from the Fe atom, where CO ligand is fully solvated. The profiles do not present significant barrier for CO entry into the (trHb: CO)$_2$ cavity through the tunnel. A small barrier is encountered to reach the (trHb: CO)$_1$ cavity, with almost negligible changes between proteins when the free energy at ∼5 Å from the heme iron is considered. The free energy well at ∼2 Å corresponds to the primary docking site, which is a polar cavity in wt and a quite apolar cavity in 3F Tf-trHbO. An appreciable difference is observed regarding the depth of this free energy well, where the CO ligand is just above the iron heme center for bond formation. It is interesting to note that, for conformation II of 3F Tf-trHbO, a clear variation in the free energy profile toward (trHb: CO)$_2$ can be observed. As was explained before, this is a consequence of a rearrangement of the distal site in which (trHb: CO)$_2$ becomes smaller. Similar free energy barriers seem to separate the primary docking site and (trHb: CO)$_3$, WG8F Tf-trHbO showing the lowest one. The associated energetic

minimum in the (trHb: CO)$_1$ cavity of WG8F protein is the highest in energy, because WG8, the main polar residue responsible for ligand stabilization, is lost.

4. DISCUSSION

The presence of water molecules in the distal pocket was previously evidenced in the liganded structure of Tf-trHbO.[10] The large entropic barrier to solvent entry has to be overcome by favorable internal electrostatic interactions with polar amino acids. According to MD simulation of the wt protein presented in this work, a water molecule enters into the active site via the typical "E7 gate" found in myoglobin and hemoglobin[53,54] and remains in intimate contact with YB10 near the E7 position. Although only one water molecule remains in the active site of the protein at any given time, solvent molecules are dynamically exchanged between bulk solvent and distal pocket.[11] A H-bonding network involving the heme 7-propionyl group, two water molecules, and YB10 was also evident in the X-ray crystal structure.[10] In two recent works, the influence of water molecules in the distal pocket in ligand rebinding on trHbs N and O from *Mycobacterium tuberculosis*[25,27] (Mt-trHbN and Mt-trHbO) has been previously studied. In both cases, this influence is analyzed by means of performing simulations with and without the presence of these key water molecules. The latter case gives insight only at short time scales, since, at longer time scales, water molecules are supposed to equilibrate and enter the distal cavity if this is thermodynamically favored. The first work reports results for wt and selected apolar mutants of Mt-trHbN. Simulations of the deoxy protein with and without a key water molecule show that its presence affects significantly the ligand rebinding. More recently, Jasaitis et al.[25] reported results of MD simulations of Mt-trHbO in the ferrous state with a free CO in the distal cavity. The authors have performed simulations with the water molecules present in the X-ray initial structure, as well as an additional simulation deleting these water molecules. The large differences observed by the authors in the two situations reinforce our conclusion of the key role of water molecules in ligand migration. Even if there is a difference in the simulation design, since, in our case, we focus on the water equilibrated state instead of performing simulations with and without structural water molecules, the main conclusion of both works reinforce the key role of water molecules in ligand rebinding. MD simulations on the unliganded structures of Tf-trHbO presented in this work provide clear evidence that the distal pocket hosts water molecules that form stable interactions with the side chain of WG8 and YCD1. Although none of these water molecules are coordinated to the heme Fe, their position would clearly clash with a bound diatomic ligand. Hence, CO ligand coordination by the heme requires that the water molecule at WS1 is displaced. This process is expected to impose a detectable free energy barrier. The clearly more favorable interaction energy between water and the distal pocket residues for wt Tf-trHbO (−14.8 kcal/mol vs −8.5 kcal/mol for WG8F and −2.6 kcal/mol for 3F Tf-trHbO) suggests that the barrier for binding should be higher for the wt protein than for mutants in which WG8 is replaced by a residue that cannot provide a H-bond to the water molecule.

A simple comparison between the picosecond geminate rebinding to wt, WG8F, and 3F Tf-trHbO (Figure 2 and Table 2) confirms that rebinding to wt Tf-trHbO occurs with the longest apparent time constant, i.e., with the largest energy barrier. The time constant decreases when WG8 is removed

H

dx.doi.org/10.1021/jp410724z | J. Phys. Chem. B XXXX, XXX, XXX−XXX

and then slightly increases when additional F residues are introduced at CD1 and B10 positions.

This qualitative trend is confirmed when quantitative analysis is performed with the microscopic model in Scheme 1. Table 2 shows that k_{-1} undergoes a 7-fold increase when WG8 is mutated to F, associated with an ~1 kcal/mol decrease in the corresponding free energy barrier (Table 3).

We attempted to detect the presence of water molecules in the distal pocket through their effects on the spectral shape in the Q absorption bands. Recent spectrokinetic studies on myoglobin demonstrated that it is possible to detect the presence of disordered water molecules in the distal pocket through small, yet appreciable, spectral shifts in the Q absorption bands.[7,55] However, due to a too low S/N ratio, SVD analysis of the TAS data in the Q-band region only afforded a main spectral component, with the characteristic shape of the CO-Tf-trHbO minus 5c-Tf-trHbO spectrum. An additional minor spectral component, which could be attributed to solvation effects in the distal pocket, was evident only for the wt protein and not for the mutated proteins. Unfortunately, this component did not meet the statistical validation criteria. Although this may be taken as a hint that water molecules may be detectable only for the wt protein, in agreement with MD results, the quality of the present data prevents a final assessment. Future improvements in the quality of the data may allow reaching a more conclusive analysis of these spectral effects.

It should be mentioned that, unlike in the nano- to microsecond time range, a further complication in the short picoseconds arises from the strong overlap of the ESA bands with the Q-band region, potentially sensing solvation effects through very small spectral changes. Given the too low signal-to-noise ratio in our data, these small contributions may be obscured by the overwhelmingly higher intensity of the ESA bands.

The geminate phase is further modulated by the presence of migration to secondary docking sites, from which ligands can either access the solvent and escape or return to the distal pocket and are rebound at later times. Whereas forward rates (k_c and k_d) toward cavities show a modest increase for WG8F and 3F mutants, the reverse rate constants (k_{-c} and k_{-d}) undergo a rather large (~100 fold) increase. This increase is qualitatively matched by the free energy barriers for these elementary processes (Table 3) as well as by the free energy profiles for ligand migration (Figure 6). A complete quantitative agreement is not straightforward to obtain in view of (i) the limited experimental sensitivity of the experimental data in some of the relevant time ranges, (ii) the limitations of the theoretical estimate of the energetic profiles, and (iii) the possible role in the modulation of the barriers played by the different levels of stabilization of water molecules.

A second source of modulation of the free energy barriers for ligand migration is the connection between cavities. For the WG8F mutant in conformation II (Figure 5A), the (trHb: CO)$_2$ cavity becomes 3 Å closer to the primary site (trHb: CO)$_1$ than in the wt protein, thus enhancing the communication between them, but resulting in a narrower passage toward the solvent. In contrast, the second conformation of WG8F (conformation I in Figure 5A) displays an open connection to the solvent, thus allowing exchange of the ligand with the exterior of the protein.

The free energy profiles estimated by ILS (Figure 6) do not show any significant barrier for CO entry from the bulk solvent into the tunnel of the three studied proteins. Further inside the protein, a small barrier is encountered to reach the (trHb: CO)$_1$ cavity, which is only marginally affected by the studied mutations. Instead, an appreciable difference is observed in the depth of the free energy well associated with the (trHb: CO)$_1$ cavity. The loss of WG8, the main polar residue responsible for ligand stabilization, results in a higher energetic minimum. ILS also demonstrates that a distinct energy profile is observed for the second conformation of WG8F Tf-trHbO, for which the rearrangement of the distal site results in a smaller size of the (trHb: CO)$_2$ cavity. Although the different energetic barriers along the migration pathway between the solvent and the distal cavity suggest that exit to, and rebinding from, the solvent may occur with two distinct rates, no such heterogeneity was evident in the bimolecular phase of the rebinding kinetics. This may be due to the fact that the two conformers are in equilibrium on a much faster time scale than the one probed by the bimolecular rebinding. Thus, the retrieved rate constants (Table 2) and free energies (Table 3) reflect average values for the two conformations.

Comparison of the ILS free energy barriers between the primary docking site and the (trHb: CO)$_3$ cavity suggests similar values for the three proteins, a fact that is also observed experimentally. The barriers for the reverse reaction, (trHb: CO)$_3$ → (trHb: CO)$_1$, are instead affected by mutations, with 3F Tf-trHbO being the lowest one, although the experimental values for 3F Tf-trHbO are admittedly affected by large errors.

According to the data in Figure 6, the barriers for the process (trHb: CO)$_1$ → (trHb: CO)$_3$ are generally smaller than those for (trHb: CO)$_1$ → (trHb: CO)$_2$, which would lead to preferential escape of photodissociated ligands to (trHb: CO)$_3$ than (trHb: CO)$_2$, thus contributing to the large geminate recombination observed experimentally. Nevertheless, the estimates of the barriers from experimental data do not provide a similar clear-cut distinction between the barriers associated with the two paths. Experiments tracking directly the infrared absorption of the migrating ligand may be able to assist in a better determination of these parameters.

An additional source of modulation in the rate constants k_{-1} can be tracked to the size of the cavity system. In the 3F Tf-trHbO mutant, the cavity shows the largest volume; hence, the conformational space that can sample the CO molecule is the largest, resulting in a higher entropy of the system. This may also account for a slight stabilization of the ligand in the cavity system, which would then explain the slower rebinding observed for this mutant in comparison with the WG8F mutant. Consistently, the entropic barrier estimated for k_{-1} is by 1 kcal/mol smaller for 3F Tf-trHbO than for WG8F Tf-trHbO.

Recent computational results have put forward the idea that phenylanine residues may play a relevant role in modulating CO migration.[56−58] Using a revised force field, the stacking interaction between CO and the phenylalanine side chain was found to lead to persistence of CO nearby the phenyl rings of F33, F43, and F46 in myoglobin.[59] Although it may not be the sole source of the slower rebinding rate k_{-1} to 3F Tf-trHbO, the presence of three phenyl rings in the distal cavity may reinforce the stabilization coming from the larger size of the cavity.

A key issue connected to the lack of any possible polar interactions in the distal pocket is that solvent molecules are not stabilized. Hence, due to the absence of the three mutated

residues, it is plausible to observe in mutant proteins a significant decrease of kinetic barriers between cavities and no thermodynamic barriers for solvent molecules stabilized at primary docking site.

5. CONCLUSIONS

The presence of water molecules coordinated to distal site amino acids modulates ligand binding, as evidenced by the barrier of the innermost step in the rebinding kinetics to *Tf*-trHbO. Mutation of the amino acids involved in the stabilization of water molecules in the distal site results in drastic reduction in this barrier. Ligand exchange between the binding site and the solvent is further tuned by a dynamic system of cavities with branched on-pathway and off-pathway temporary docking sites. These integrated studies exploit molecular dynamics simulations and ligand binding kinetics to provide insight into functional roles of solvent, otherwise difficult to obtain.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Figure S1, decay associated spectra obtained by the global analysis of the transient data of WG8F and 3F *Tf*-trHb; Figure S2, CO rebinding kinetics to WG8F and 3F *Tf*-trHb solutions at $T = 10\ °C$, $T = 20\ °C$, and $T = 30\ °C$, at 1 and 0.1 atm CO; Figure S3, comparison between the time courses of the amplitudes V_1 and the normalized absorbance change measured at 436 nm at the same CO pressure and temperature for WG8F *Tf*-trHb and 3F *Tf*-trHb). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*Phone: +541145763375. E-mail: dario@qi.fcen.uba.ar (D.A.E.).
*Phone: +390521905223. E-mail: cristiano.viappiani@fis.unipr.it (C.V.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ ABBREVIATIONS

trHb, truncated hemoglobin; *Tf*, *Thermobifida fusca*; ASV, acid surface variant of *Tf* containing two single site mutations F107E and R91E; TAS, femtosecond transient absorption; LFP, nanosecond laser flash photolysis; 5c-, pentacoordinated hemoglobin; PDB, Protein Data Bank; MD, molecular dynamics; ILS, Implicit Ligand Sampling; H-bond, hydrogen bond; wt, *wild type*

■ REFERENCES

(1) Li, Z.; Lazaridis, T. The Effect of Water Displacement on Binding Thermodynamics: Concanavalin A. *J. Phys. Chem. B* **2005**, *109*, 662−670.

(2) Li, Z.; Lazaridis, T. Thermodynamic Contributions of the Ordered Water Molecule in HIV-1 Protease. *J. Am. Chem. Soc.* **2003**, *125*, 6636−6637.

(3) Michel, J.; Tirado-Rives, J.; Jorgensen, W. Prediction of the Water Content in Protein Binding Sites. *J. Phys. Chem. B* **2009**, *113*, 13337−13346.

(4) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817−2831.

(5) Schmidtke, P.; Luque, J. F.; Murray, J. B.; Barril, X. Shielded Hydrogen Bonds as Structural Determinants of Binding Kinetics: Application in Drug Design. *J. Chem. Theory Comput.* **2011**, *133*, 18903−18910.

(6) Hummer, G. Molecular Binding: Under Water's Influence. *Nat. Chem.* **2010**, *2*, 906−907.

(7) Goldbeck, R. A; Bhaskaran, S.; Ortega, C.; Mendoza, J. L.; Olson, J. S.; Soman, J.; Kliger, D. S.; Esquerra, R. M. Water and Ligand Entry in Myoglobin: Assessing the Speed and Extent of Heme Pocket Hydration After CO Photodissociation. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1254−1259.

(8) Baron, R.; Setny, P.; McCammon, J. A. Water in Cavity-Ligand Recognition. *J. Am. Chem. Soc.* **2010**, *132*, 12091−12097.

(9) Baron, R.; Mccammon, J. A. Dynamics, Hydration, and Motional Averaging of a Loop-Gated Artificial Protein Cavity: The W191G Mutant of Cytochrome *c* Peroxidase in Water as Revealed by Molecular Dynamics Simulations. *Biochemistry* **2007**, *46*, 10629−10642.

(10) Bonamore, A.; Ilari, A.; Giangiacomo, L.; Bellelli, A.; Morea, V.; Boffi, A. A Novel Thermostable Hemoglobin from the Actinobacterium *Thermobifida fusca*. *FEBS J.* **2005**, *272*, 4189−4201.

(11) Droghetti, E.; Nicoletti, F. P.; Bonamore, A.; Boechi, L.; Arroyo Mañez, P.; Estrin, D. A.; Boffi, A.; Smulevich, G.; Feis, A. Heme Pocket Structural Properties of a Bacterial Truncated Hemoglobin from *Thermobifida fusca*. *Biochemistry* **2010**, *49*, 10394−10402.

(12) Marcelli, A.; Abbruzzetti, S.; Bustamante, J. P.; Feis, A.; Bonamore, A.; Boffi, A.; Gellini, C.; Salvi, P. R.; Estrin, D. A.; Bruno, S.; et al. Following Ligand Migration Pathways from Picoseconds to Milliseconds in Type II Truncated Hemoglobin from Thermobifida fusca. *PLoS One* **2012**, *7*, e39884.

(13) Lapini, A.; Di Donato, M.; Patrizi, B.; Marcelli, A.; Lima, M.; Righini, R.; Foggi, P.; Sciamanna, N.; Boffi, A. Carbon Monoxide Recombination Dynamics in Truncated Hemoglobins Studied with Visible-Pump MidIR-Probe Spectroscopy. *J. Phys. Chem. B* **2012**, *116*, 8753−8761.

(14) Wittenberg, J. B.; Bolognesi, M.; Wittenberg, B. A.; Guertin, M. Truncated Hemoglobins: A New Family of Hemoglobins Widely Distributed in Bacteria, Unicellular Eukaryotes, and Plants. *J. Biol. Chem.* **2002**, *277*, 871−874.

(15) Vuletich, D. A.; Lecomte, J. T. J. A Phylogenetic and Structural Analysis of Truncated Hemoglobins. *J. Mol. Evol.* **2006**, *62*, 196−210.

(16) Nardini, M.; Pesce, A.; Milani, M.; Bolognesi, M. Protein Fold and Structure in the Truncated (2/2) Globin Family. *Gene* **2007**, *398*, 2−11.

(17) Frey, A. D.; Kallio, P. T. Nitric Oxide Detoxification − A new Era for Bacterial Globins in Biotechnology. *Trends Biotechnol.* **2005**, *23*, 69−73.

(18) Nicoletti, F. P.; Comandini, A.; Bonamore, A.; Boechi, L.; Boubeta, F. M.; Feis, A.; Smulevich, G.; Boffi, A. Sulfide Binding Properties of Truncated Hemoglobins. *Biochemistry* **2010**, *49*, 2269−2278.

(19) Crespo, A.; Martí, M. A.; Kalko, S. G.; Morreale, A.; Orozco, M.; Gelpi, J. L.; Luque, F. J.; Estrin, D. A. Theoretical Study of the Truncated Hemoglobin HbN: Exploring the Molecular Basis of the NO Detoxification Mechanism. *J. Am. Chem. Soc.* **2005**, *127*, 4433−4444.

(20) Ouellet, H.; Ranguelova, K.; Labarre, M.; Wittenberg, J. B.; Wittenberg, B. A.; Magliozzo, R. S.; Guertin, M. Reaction of *Mycobacterium tuberculosis* Truncated Hemoglobin O with Hydrogen Peroxide: Evidence for Peroxidatic Activity and Formation of Protein-Based Radicals. *J. Biol. Chem.* **2007**, *282*, 7491−7503.

(21) Nicoletti, F. P.; Droghetti, E.; Boechi, L.; Bonamore, A.; Sciamanna, N.; Estrin, D. A.; Feis, A.; Boffi, A.; Smulevich, G. Fluoride as a Probe for H-bonding Interactions in the Active Site of Heme Proteins: The Case of *Thermobifida fusca* Hemoglobin. *J. Am. Chem. Soc.* **2011**, *133*, 20970−20980.

(22) Guallar, V.; Lu, C.; Borrelli, K.; Egawa, T.; Yeh, S. Ligand Migration in the Truncated Hemoglobin-II from *Mycobacterium tuberculosis*: The Role of G8 Tryptophan. *J. Biol. Chem.* **2009**, *284*, 3106−3116.

(23) Boechi, L.; Martí, M. A.; Milani, M.; Bolognesi, M.; Luque, F. J.; Estrin, D. A. Structural Determinants of Ligand Migration in *Mycobacterium tuberculosis* Truncated Hemoglobin O. *Proteins: Struct., Funct., Genet.* **2008**, *73*, 372−379.

(24) Feis, A.; Lapini, A.; Catacchio, B.; Brogioni, S.; Foggi, P.; Chiancone, E.; Boffi, A.; Smulevich, G. Unusually Strong H-Bonding to the Heme Ligand and Fast Geminate Recombination Dynamics of the Carbon Monoxide Complex of *Bacillus subtilis* Truncated Hemoglobin. *Biochemistry* **2008**, *47*, 902−910.

(25) Jasaitis, A.; Ouellet, H.; Lambry, J. C.; Martin, J. L.; Friedman, J. M.; Guertin, M.; Vos, M. H. Ultrafast Heme−Ligand Recombination in Truncated Hemoglobin HbO from *Mycobacterium tuberculosis*: A Ligand Cage. *Chem. Phys.* **2012**, *396*, 10−16.

(26) Eich, R. F.; Li, T.; Lemon, D. D.; Doherty, D. H.; Curry, S. R.; Aitken, J. F.; Mathews, A. J.; Johnson, K. A.; Smith, R. D.; Phillips, G. N.; et al. Mechanism of NO-Induced Oxidation of Myoglobin and Hemoglobin. *Biochemistry* **1996**, *35*, 6976−6983.

(27) Ouellet, Y. H.; Daigle, R.; Lagüe, P.; Dantsker, D.; Milani, M.; Bolognesi, M.; Friedman, J. M.; Guertin, M. Ligand Binding to Truncated Hemoglobin N from *Mycobacterium tuberculosis* is Strongly Modulated by the Interplay Between the Distal Heme Pocket Residues and Internal Water. *J. Biol. Chem.* **2008**, *283*, 27270−27278.

(28) Abbruzzetti, S.; Spyrakis, F.; Bidon-chanal, A.; Luque, F. J.; Viappiani, C. Ligand Migration Through Hemeprotein Cavities: Insights from Laser Flash Photolysis and Molecular Dynamics Simulations. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10686−10701.

(29) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput. Phys. Commun.* **1995**, *91*, 1−41.

(30) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.

(31) Marti, M. A.; Capece, L.; Bidon-Chanal, A.; Crespo, A.; Guallar, V.; Luque, F. J.; Estrin, D. A. Nitric Oxide Reactivity with Globins as Investigated Through Computer Simulation. *Methods Enzymol.* **2008**, *437*, 477−498.

(32) Marti, M. A.; Crespo, A.; Capece, L.; Boechi, L.; Bikiel, D. E.; Scherlis, D. A.; Estrin, D. A. Dioxygen Affinity in Heme Proteins Investigated by Computer Simulation. *J. Inorg. Biochem.* **2006**, *100*, 761−770.

(33) 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.; et al. Modeling Heme Proteins Using Atomistic Simulations. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5611−5628.

(34) Forti, F.; Boechi, L.; Bikiel, D.; Martí, M. A.; Nardini, M.; Bolognesi, M.; Viappiani, C.; Estrin, D. A.; Luque, F. J. Ligand Migration in *Methanosarcina acetivorans* Protoglobin: Effects of Ligand Binding and Dimeric Assembly. *J. Phys. Chem. B* **2011**, *115*, 13771−13780.

(35) Capece, L.; Lewis-ballester, A.; Marti, M. A.; Estrin, D. A.; Yeh, S. R. Molecular Basis for the Substrate Stereoselectivity in Tryptophan Dioxygenase. *Biochemistry* **2011**, *50*, 10910−10918.

(36) Arroyo Mañez, P.; Lu, C.; Boechi, L.; Martí, M. A.; Shepherd, M.; Wilson, J. L.; Poole, R. K.; Luque, F. J.; Yeh, S. R.; Estrin, D. A. Role of the Distal Hydrogen-Bonding Network in Regulating Oxygen Affinity in the Truncated Hemoglobin III from *Campylobacter jejuni*. *Biochemistry* **2011**, *50*, 3946−3956.

(37) Giordano, D.; Boechi, L.; Samuni, U.; Vergara, A.; Martí, M. A.; Estrin, D. A.; Friedman, J. M.; Mazzarella, L.; Prisco, G.; Grassi, L. The Hemoglobins of the Sub-Antarctic Fish *Cottoperca gobio*, a Phyletically Basal Species − Oxygen-Binding Equilibria, Kinetics and Molecular Dynamics. *FEBS J.* **2009**, *276*, 2266−2277.

(38) Perissinotti, L. L.; Marti, M. A.; Doctorovich, F.; Luque, F. J.; Estrin, D. A. A Microscopic Study of the Deoxyhemoglobin-Catalyzed Generation of Nitric Oxide from Nitrite Anion. *Biochemistry* **2008**, *47*, 9793−9802.

(39) Cohen, J.; Olsen, K. W.; Schulten, K. Finding Gas Migration Pathways in Proteins Using Implicit Ligand Sampling. *Methods Enzymol.* **2008**, *437*, 439−457.

(40) Forti, F.; Boechi, L.; Estrin, D. A.; Marti, M. A. Comparing and Combining Implicit Ligand Sampling with Multiple Steered Molecular Dynamics to Study Ligand Migration Processes in Heme Proteins. *J. Comput. Chem.* **2011**, *32*, 2219−2231.

(41) Durrant, J. D.; de Oliveira, C. A. F.; McCammon, J. A. POVME: An Algorithm for Measuring Binding-Pocket Volumes. *J. Mol. Graphics Modell.* **2011**, *29*, 773−776.

(42) Di Lella, S.; Martí, M. A.; Alvarez, R. M. S.; Estrin, D. A.; Ricci, J. C. D. Characterization of the Galectin-1 Carbohydrate Recognition Domain in Terms of Solvent Occupancy. *J. Phys. Chem. B* **2007**, *111*, 7360−7366.

(43) Gauto, D. F.; Di Lella, S.; Guardia, C. M. A.; Estrin, D. A.; Martí, M. A. Carbohydrate-Binding Proteins: Dissecting Ligand Structures through Solvent Environment Occupancy. *J. Phys. Chem. B* **2009**, *113*, 8717−8724.

(44) Gauto, D. F.; Di Lella, S.; Estrin, D. A.; Martí, M. A. Structural Basis for Ligand Recognition in a Mushroom Lectin: Solvent Structure as Specificity Predictor. *Carbohydr. Res.* **2011**, *346*, 939−948.

(45) Marcelli, A.; Foggi, P.; Moroni, L.; Gellini, C.; Salvi, P. R. Excited-State Absorption and Ultrafast Relaxation Dynamics of Porphyrin, Diprotonated Porphyrin, and Tetraoxaporphyrin Dication. *J. Phys. Chem. A* **2008**, *112*, 1864−1872.

(46) Foggi, P.; Neuwahl, F. V. R.; Moroni, L.; Salvi, P. R. S1 → Sn and S2 → Sn Absorption of Azulene: Femtosecond Transient Spectra and Excited State Calculations. *J. Phys. Chem. A* **2003**, *107*, 1689−1696.

(47) Van Wilderen, L. J. G. W.; Lincoln, C. N.; van Thor, J. J. Modelling Multi-Pulse Population Dynamics from Ultrafast Spectroscopy. *PLoS One* **2011**, *6*, e17373.

(48) Abbruzzetti, S.; Bruno, S.; Faggiano, S.; Grandi, E.; Mozzarelli, A.; Viappiani, C. Time-Resolved Methods in Biophysics. 2. Monitoring Haem Proteins at Work with Nanosecond Laser Flash Photolysis. *Photochem. Photobiol. Sci.* **2006**, *5*, 1109−1120.

(49) Henry, E. R.; Hofrichter, J. Singular Value Decomposition: Application to Analysis of Experimental Data. *Methods Enzymol.* **1992**, *210*, 129−192.

(50) Jones, C. M.; Ansari, A.; Henry, E. R.; Christoph, G. W.; Hofrichter, J.; Eaton, W. A. Speed of Intersubunit Communication in Proteins. *Biochemistry* **1992**, *31*, 6692−6702.

(51) Ye, X.; Demidov, A.; Champion, P. M. Measurements of the Photodissociation Quantum Yields of MbNO and MbO$_2$ and the Vibrational Relaxation of the Six-Coordinate Heme Species. *J. Am. Chem. Soc.* **2002**, *124*, 5914−5924.

(52) Nicoletti, F. P.; Droghetti, E.; Howes, B. D.; Bustamante, J. P.; Bonamore, A.; Sciamanna, N.; Estrin, D. A.; Feis, A.; Boffi, A.; Smulevich, G. H-bonding Networks of the Distal Residues and Water Molecules in the Active Site of *Thermobifida fusca* Hemoglobin. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 1901−1909.

(53) Perutz, M. F.; Mathews, F. S. An X-ray Study of Azide Methaemoglobin. *J. Mol. Biol.* **1966**, *21*, 199−202.

(54) Scott, E. E.; Gibson, Q. H.; Olson, J. S. Mapping the Pathways for O_2 Entry Into and Exit from Myoglobin. *J. Biol. Chem.* **2011**, *276*, 5177−5188.

(55) 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. Optical Detection of Disordered Water within a Protein Cavity. *J. Am. Chem. Soc.* **2009**, *131*, 12265−12272.

(56) Borrelli, K. W.; Vitalis, A.; Alcantara, R.; Guallar, V. PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. *J. Chem. Theory Comput.* **2005**, 1304−1311.

(57) Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K. Imaging the Migration Pathways for O_2, CO, NO, and Xe inside Myoglobin. *Biophys. J.* **2006**, *91*, 1844−1857.

(58) Spyrakis, F.; Luque, F. J.; Viappiani, C. Structural Analysis in Nonsymbiotic Hemoglobins: What Can We Learn from Inner Cavities? *Plant Sci.* **2011**, *181*, 8−13.

(59) Lucas, M. F.; Guallar, V. An Atomistic View on Human Hemoglobin Carbon Monoxide Migration Processes. *Biophys. J.* **2012**, *102*, 887−896.