Ligand migration through hemeprotein cavities: Insights from laser flash photolysis and molecular dynamics simulations

Abbruzzetti, Stefania; Spyrakis, Francesca; Bidon-Chanal, Axel; Luque, F. Javier; Viappiani, Cristiano.

Ligand migration through hemeprotein cavities: Insights from laser flash photolysis and molecular dynamics simulations. PHYSICAL CHEMISTRY CHEMICAL PHYSICS. 2013, 15(26) pp: 10686-10701.

DOI: 10.1039/c3cp51149

The publisher's version is available at:

http://pubs.rsc.org/en/Content/ArticleLanding/2013/CP/c3cp51149a#!divAbstract

Ligand migration through hemeprotein cavities: insights from laser flash photolysis and molecular dynamics simulations

Stefania Abbruzzetti, [1] Francesca Spyrakis, [2] Axel Bidon-Chanal, [3] F. Javier Luque, [3] Cristiano Viappiani [1,*]

[1] Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Parma, Parma, Italy

[2] Dipartimento di Scienze della Vita, Università degli Studi di Modena e Reggio Emilia, Modena, Italy

[3] Departament de Fisicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia, Universitat de Barcelona, Campus de l'Alimentació, Santa Coloma de Gramenet, Spain

KEYWORDS hemeprotein, ligand migration, reaction kinetics, molecular simulations

Corresponding Author

* Cristiano Viappiani, Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Parma, viale delle Scienze 7A, 43124, Parma, Italy, tel +39 0521 905208, fax +39 0521 905223, email: cristiano.viappiani@fis.unipr.it

Abstract

The presence of cavities and tunnels in the interior of proteins, in conjunction with the structural plasticity arising from the coupling to the thermal fluctuations of the protein scaffold, has profound consequences on the pathways followed by ligands moving through the protein matrix. In this perspective we discuss how quantitative analysis of experimental rebinding kinetics from laser flash photolysis, trapping of unstable conformational states by embedding proteins within the nanopores of silica gels, and molecular simulations can synergistically converge to gain insight into the migration mechanism of ligands. We show how the evaluation of the free energy landscape for ligand diffusion based on the outcome of computational techniques can assist the definition of sound reaction schemes, leading to a comprehensive understanding of the broad range of chemical events and time scales that encompass the transport of small ligands in hemeproteins.

1. Introduction

Recognition and binding of specific molecules is a process central to protein function, comprising as diverse events as transport of small diatomic ligands [1] and recognition of nucleotide sequences [2] or proteins. [3] The binding of ligands to their targets is governed by a number of factors, including shape complementarity and polar interactions at the binding site, reorganization of water molecules, conformational changes, enthalpy-entropy compensation and diffusion through kinetic pathways. [4-7] [8-12] Protein dynamics [13-17] and protein-solvent interactions [18, 19] also play a fundamental role in modulating ligand binding, from initial association to the target, to migration through the protein matrix, to adoption of the bound structure in the binding pocket. [20, 21]

The physicochemical rules that dictate the thermodynamics and kinetics of ligand binding also determine the binding of small gaseous ligands to hemeproteins. However, the ubiquitous relevance of the heme-based chemistry to living beings, the abundant availability of purified material, and the convenient spectroscopic markers, sensitive to ligation and oxidation state of the heme, have made binding reactions of diatomic ligands with hemeproteins the workbench for understanding the complex relationships between structure, dynamics, and function. Investigations on myoglobin (Mb) have had a profound impact on our understanding of the interrelation between structural plasticity and function, to the point that Mb was suggested to be "*the hydrogen atom of biology*". [22] Many questions are still open, and even in the apparently simple case of small ligand binding, the presence of allosteric effects, often accompanied by cooperative behavior, challenges our understanding of the fine details of ligand binding mechanisms. [1, 23-25] Covering the full body of work on ligand binding to hemeproteins exceeds the scope of this review. We instead prefer to limit our discussion to recently proposed methodologies and ideas, specifically dealing with ligand migration within the protein matrix, in view of new perspectives which were open thanks to an unprecedented understanding of the topological arrangement of internal cavities and tunnels in hemeproteins.

2. A brief historical perspective about ligand binding kinetics

Over the many years of hemoglobin studies, [26] an incredibly large number of experimental works have characterized the rebinding kinetics under a variety of experimental conditions (temperature, cosolvents, pressure, ...) for Mb [27-41], hemoglobin (Hb) [38, 42-51] and, more

recently, for other heme proteins as, for instance, neuroglobin (Ngb), [52-58] to mention just a very limited and arbitrarily chosen subset of the available literature.

The vast majority of the studies dealing with the kinetics of ligand binding to hemeproteins (mostly hemoglobins) relied on the fact that the chemical bond between the Fe atom and the diatomic ligand (CO, O_2, and NO) is photolabile. [59, 60] This was exploited to generate a measurable concentration of unliganded proteins (and free ligands) with a short pulse of light, and then monitor rebinding through the concomitant absorbance changes of the heme cofactor. [38, 61-68] This relaxation method was derived from flash photolysis, [69] for the development of which R.G.W. Norrish and G. Porter, together with M. Eigen, were awarded the Nobel prize in Chemistry in 1967. After dissociation from the heme iron, ligands can either rebind from the protein inner cavities with the so-called geminate recombination, [27, 70] or escape into the solvent through permanent or transiently available exit channels. Geminate recombination reduces the fraction of photodissociated ligands which subsequently recombine from the solvent [65] through a slower, bimolecular phase. [67]

When ligand recombination is investigated at cryogenic temperatures, the observed geminate rebinding kinetics indicate the existence of a frozen distribution of functionally distinct conformational substates. [27, 71] At higher temperatures, averaging of the conformational substates cancels the kinetic hole burning phenomena [72] observed at lower temperatures [73, 74] and new kinetic phases are observed, [27, 33, 40, 71, 75, 76] which apparently lead to an inverse temperature effect of the observed kinetics. Different explanations were offered for the observed anomalous temperature dependence. These included the existence of barrier-raising conformational relaxations [28, 34, 37, 77] and the presence of routes leading to slowly rebinding, temporary docking sites. [78, 79] These sites were identified in Mb as the packing defects referred to as Xenon cavities. [80-89] Therefore, the discrete kinetic phases of ligand rebinding upon flash photolysis received a new structural interpretation as due to the existence of kinetic traps for the photodissociated ligands, [90-101] likely regulated by protein dynamics. [102] The spectroscopic evidence of tertiary structural relaxation was reported in the transient absorbance signals following photolysis, [33, 35, 37, 39, 103, 104] and recently confirmed also from time-resolved X-ray crystallography data on Mb mutants. [88, 105] More complex quaternary structural relaxations were also detected as changes in the spectral shape in time-resolved absorption experiments. [24, 30, 38, 42-44, 46-49, 106, 107] The relevance of solvent slaved protein motions was thoroughly investigated and discussed, [14, 15, 19, 108-112] and the specific role of solvent in modulating protein reactivity was identified. [113, 114]

The presence of cavities and tunnels, in conjunction with their coupling to the protein dynamics, has been postulated to have functional consequences for hemeproteins, by facilitating and/or gating ligand diffusion to the reaction site. [115-118] Cavities and tunnels are considered instrumental to sustain multisubstrate reactions, as demonstrated for the well known NO dioxygenase activity elicited by Mb,[119, 120] and the truncated Hb N from *M. tuberculosis.*[121, 122] Therefore, understanding the topological arrangement of cavities inside proteins and their dynamical plasticity is crucial for elucidating the mechanism that underlies substrate delivery to the heme cavity, chemical processing, and product removal from the reaction site.

Time-resolved X-ray Laue diffraction can provide real time visualization of structural rearrangements and ligand movements through internal cavities with atomic resolution.[123] However, to date, experiments with only a few hemeproteins have been successful,[87, 89, 105, 124, 125] due to the very demanding requirements in terms of crystal quality and resistance to high intensity radiation-induced damage. For a limited number of favorable cases, it proved feasible to follow in real time ligand migration through the protein matrix after photodissociation, disclosing structural relaxations that trigger transient connections between otherwise separated cavities.[87-89] Furthermore, by monitoring the time evolution of electron density, it was possible to build ligand population at different positions in the protein matrix, thus retrieving kinetic information and reaction pathways.[88, 89, 105, 125]

The inherent experimental difficulties of time-resolved crystallographic techniques prevent their wide applicability, thus making it necessary to identify methodologies that are capable of affording kinetic data on ligand movements and linking this information to the topology of the migration pathways. In this context, a few spectroscopic methods can be utilized to detect gas migration through temporary docking sites. Among these, Temperature Derivative FTIR (TD-FTIR) spectroscopy has been used at cryogenic temperatures as a powerful tool to track migration through temporary sites for CO and NO complexes with hemeproteins. [53, 95, 126-129] Other approaches exploit ligand rebinding kinetics after laser flash photolysis to sense ligand migration at near-physiological temperatures, [57, 68, 130, 131] an advantage over TD-FTIR. In turn, the absorbance changes in the visible are not endowed with the remarkable sensitivity of the CO stretching absorption bands to the environment, a fact that allows to track the photodissociated ligand along its migration. [129] The geminate rebinding phase (i.e. the fraction of the rebinding kinetics independent of ligand concentration and associated with rebinding from within the protein matrix) is very informative about the multiple steps that

connect reaction site, internal cavities and solvent phase. The reason is that at room temperature the available thermal energy is high enough to allow sampling of remote docking sites by the ligand. Hence, rebinding kinetics can be turned into a means to follow the migration processes. However, while being a sensitive tool to assess reaction kinetics, it does not provide any structural information on the topology of migration pathways.

The extensive use of mutagenesis [91, 92, 132-135] and crystallography [135-138] allowed to map ligand egression pathways in Mb, which led to the concept that for this protein the main exchange pathway is through the so called His-gate. More recent computational studies have suggested that, when thermal fluctuation and protein dynamics are taken into account, additional exchange pathways other than the His gate, become active for Mb [139] and other related globins. [140] The latter finding is also supported by limited experimental evidences. [97, 141]

The structural features of inner cavities and tunnels can be retrieved from protein structures by using pocket detection algorithms, which mainly rely on i) the geometrical analysis of inner cavities, ii) the interaction energy with a suitable probe, and iii) the propensity of conserved residues in the binding site. Though the vast majority of cavity detection algorithms have been developed to treat static structures, few works have recently attempted to account for the structural plasticity of proteins in the identification of binding pockets and tunnels from conformational ensembles of the protein.[142-144] Alternative approaches for finding ligand migration pathways are Locally Enhanced Sampling (LES), which utilizes multiple non-interacting copies of the ligand that are propagated using a mean field approximation for the ligand-protein interactions,[145] the calculation of the interaction energy of a probe with the protein snapshots taken from a Molecular Dynamics (MD) simulation, as in the Implicit Ligand Sampling technique (ILS),[146] or the use of a perturbing probe in conjunction with Monte Carlo sampling and protein structure prediction algorithms, as implemented in PELE.[147]

In the following we discuss how experimental data can be coupled with the information derived from atomistic MD simulations to provide a structural interpretation to the reaction intermediates formed in ligand rebinding pathways, to disclose mechanisms for ligand migration, and eventually to gain insight into structure-function relationships in hemeproteins. To this end, section 3 discusses the use of strategies to enhance the sensitivity of experimental approaches for detecting reaction intermediates in rebinding kinetic assays.

Section 4 presents the quantitative treatment that allows the derivation of mechanistic schemes that relate the microscopic kinetic steps with the macroscopic rates for ligand binding. Section 5 revisits the interpretation and refinement of kinetic schemes based on the dynamical analysis of the topological features of inner cavities. Extension of the temporal resolution of rebinding kinetics to very short time scales is examined in section 6. Finally, section 6 discusses the current status and perspectives of molecular simulations techniques to provide structural information valuable to facilitate the quantitative analysis of ligand migration through hemeproteins. As a last remark, instead of discussing the theme providing a list of different cases, this review concentrates on a limited number of proteins to discuss more deeply these concepts.

3. Trapping reaction intermediates with silica gels

As detailed in the introduction, ligand migration through the internal pathways, connecting reaction site and the solvent, can be tracked also by monitoring rebinding kinetics. However, ligands sampling internal protein cavities are elusive species. Hence, methodologies capable of enhancing their population at room temperature are needed, in order to better sense them. The trick to detect these reaction intermediates is to trap them: silica gels proved to be very good at doing that.[148-150] Perhaps the most dramatic example, demonstrating the effectiveness of silica gels in stabilizing thermodynamically unstable species, is the selective stabilization of the quaternary states *R* and *T* of human Hemoglobin A (HbA). [151-153] By encapsulating liganded *R* state or unliganded *T* HbA in wet silica gels, the quaternary states can be stabilized on the time scale of minutes to hours depending on factors such as temperature and allosteric effectors. [151, 153] The gel matrix does not prevent the *R*-to-*T* quaternary transition, rather it enormously slows it down by a factor of 10^6-10^8. [107, 153-156] This allowed the determination of oxygen affinity for *T* state Hb from equilibrium binding curves, [151, 152, 157] but also the study of reaction rates for CO binding to R and T states. [108, 154, 158-160] Distinct tertiary states within a fixed quaternary conformation were highlighted by equilibrium [157, 161] and kinetic investigations. [154, 162]

Increased viscosity, reduced hydration, altered water dynamics, and confinement effects frequently increase geminate rebinding, an effect which can be enhanced using cosolvents such as glycerol. [57, 102, 108, 109, 131, 162-166] From a mechanistic point of view, the increased

friction is expected to affect the availability of internal passages and cavities by slowing down molecular motions.

As a representative example, the effects of gel encapsulation and medium viscosity are illustrated in Figure 1A for the kinetics of the CO rebinding to AHb1 (class 1 non symbiotic hemoglobin from *A. thaliana*), a hemeprotein that shows bis-histidyl hexacoordination in the absence of exogenous ligands and is possibly acting as a NO-dioxygenase.[167] CO rebinding kinetics to AHb1 in solution (black curve) is only marginally affected by gel encapsulation of the protein (red curve). [68, 168] However, increasing the concentration of glycerol in the gel bathing solution (green, 80% by volume; blue, 100%) increases the geminate recombination, exposing multiple kinetic phases, which suggest complex reaction patterns. [131] In these experiments, the bulk viscosity dramatically increases from the low values of water at room temperature (~ 1 cP) to the very high value of glycerol (~ 10^3 cP), with an additional contribution coming from the gel.[169] All the features relevant for the protein functionality are preserved, including ligand exchange between the protein matrix and the solvent, and the bis-histidyl hexacoordination.

A qualitative appreciation of the multiple kinetic phases present in the rebinding kinetics can be obtained by retrieving model-independent lifetime distributions using a Maximum Entropy Method (MEM). [170, 171] It is assumed that the rebinding kinetics arises from a series of exponential decay functions with lifetimes characterized by a distribution of values. The MEM analysis affords the probability distribution of the lifetimes, without making any assumptions about the shape and the number of peaks in the distribution. Thus, for AHb1 in solution (Figure 1B) a dominant band corresponding to the bimolecular rebinding phase (note that its center is dependent on CO concentration) is found at ~1 ms, while minor bands at ~ 20 ns and ~ 100 ns are due to geminate rebinding.[172] In the gel, the minor band at ~ 100 ns is enhanced, and the bimolecular phase is broadened by the reaction of the heme with the distal His.[68] At increasing glycerol concentration, the amplitude of the bimolecular rebinding is decreased and the lifetime distribution of the geminate rebinding becomes more structured, with as many as four distinct reaction intermediates, most likely arising from rebinding of ligands which were temporarily docked at remote sites in the protein.[131]

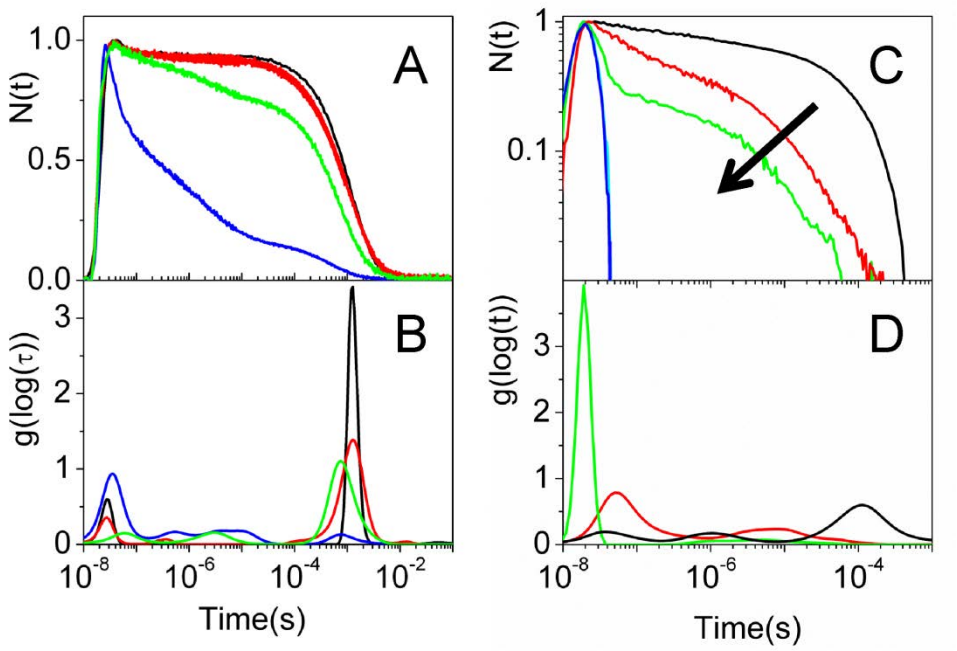


Figure 1. A. CO rebinding kinetics to AHb1 (reported as the fraction of unliganded molecules as a function of time after the nanosecond photolysis) in solution (black), immobilized in a silica gel bathed in buffer (red), in 80% glycerol/buffer mixture (green), and in anhydrous glycerol (blue). T = 10 °C, CO = 1 atm. [131] **B.** Lifetime distributions (reported as probability density for rebinding with time constant τ)[171] associated with the rebinding kinetics in A. [131] **C.** CO rebinding kinetics for trehalose-coated MbCO. The amount of residual water decreases in the direction of the arrow. [19] The blue trace was collected after drying the glass overnight in a vacuum and closely matches the instrumental response function (cyan). **D.** Lifetime distributions associated with the rebinding kinetics in **C.**

Silica gels have a remarkable advantage over other matrices often used to immobilize proteins at room temperature. Among them, a special place is held by trehalose, a sugar that forms glasses characterized by an enormous viscosity (>10^{15} cP) without affecting protein structure. When MbCO is embedded in a trehalose glass, photolyzed ligands cannot escape the protein matrix, and all CO molecules are rebound geminately.[173] This fact can be utilized to monitor the effect of hydration on internal migration at constant viscosity. [19, 93, 99] Thus, Figure 1 C shows that rebinding kinetics involves a multiphasic process due to ligand migration through temporary docking sites from which CO is rebound at later times, as highlighted by the MEM lifetime distributions reported in Figure 1D. The migration to farther sites is progressively inhibited as the residual water content is decreased. For

extremely dry samples (cyan curve), the rebinding occurs with a time profile that matches the Impulse Response Function of the detection setup (blue curve). The conclusion is that, in the absence of essential hydration water, protein fluctuations are hindered to such an extent that all fundamental migration processes are arrested. Experimental evidences like these have rich implications for ligand migration, but in order to fully exploit this potential a further step in data analysis is needed.

4. Building a reaction scheme: from macro- to microscopic rates

In order to make the observed kinetics (such as those reported in Section 3) meaningful in terms of mechanistic events, the qualitative analysis obtained by (model independent) MEM lifetime distributions needs to be superseded by a more quantitative (but model dependent) approach, which takes into account additional pieces of information. In the following we summarize the steps which allow to build a microscopic model from the data reported above. At the same time, this approach is a rather general procedure which is also valid for other similar systems. The transition between these models can take advantage of Singular Value Decomposition (SVD) applied to time-resolved differential absorption spectra, [46, 174] which helps retrieving the time profiles for ligand rebinding. Difference absorption spectra as a function of the delay from the laser pulse are calculated by subtracting from each of them the spectrum of the CO-bound form at equilibrium. A data matrix (**D**) is built, consisting of difference absorbance measured as a function of two variables: the wavelength of the probe beam and the time delay from the laser pulse. [68] The SVD of **D** can be written as $\mathbf{D} = \mathbf{USV}^T$, where the columns of **U** are a set of linearly independent, orthonormal basis spectra, the columns of **V** describe the time-dependent amplitudes of these basis spectra, and the matrix **S** is a diagonal matrix of non-negative singular values which describe the magnitudes of the contributions of each of the outer products of the i-th column vectors $U_iV_i^T$ to the data matrix **D**. [174] Rebinding can thus be followed through the time dependent amplitude V_i of the spectral component U_i, corresponding to the difference spectrum between CO and unliganded species, which normally shows the largest singular value.

For AHb1, a structural feature that has to be considered is the kinetic competition between the distal His and CO for binding to the heme upon photolysis of the CO adduct, as noted in the spectral changes associated with formation of the bis-histidyl complex.[68, 172] SVD analysis of the time-resolved difference absorption spectra after laser photolysis [46, 174] not

only retrieves the time profiles for ligand rebinding (first spectral component), but also the transient formation of the bis-histidyl complex (second spectral component, Figure 2A and B). Microscopic kinetic schemes can then be used to reproduce the observed rebinding kinetics and its response to physicochemical conditions (reactant concentration, viscosity, temperature). The differential equations derived from the kinetic scheme are solved numerically and the rate constants are treated as parameters to be optimized during the fitting procedure.[47, 130] It is important to note that the measured absorbance changes reflect only concentrations of the two relevant species (liganded vs unliganded), but the kinetics encompasses all coupled reactions.

The microscopic model (Scheme 1) built on the basis of kinetic and structural information (derived from optical and vibrational spectroscopy [172, 175] and from structural modelling) allows us to describe the CO rebinding kinetics to AHb1 in solution (Figure 2C) and in gels soaked in glycerol (Figure 2D).[131, 172] A clear advantage of this description is the possibility of obtaining, although indirectly, the time course of the concentration of the reaction intermediates, as noted in Figure 2C and D. The model outlined in this section accounts for a static picture of the connectivity between temporary docking sites, the distal cavity, and the solvent, but this description neglects one fundamental property of proteins, i.e. their dynamics. The next section deals with this issue and proposes methodological approaches to include dynamics into the kinetic modelling.

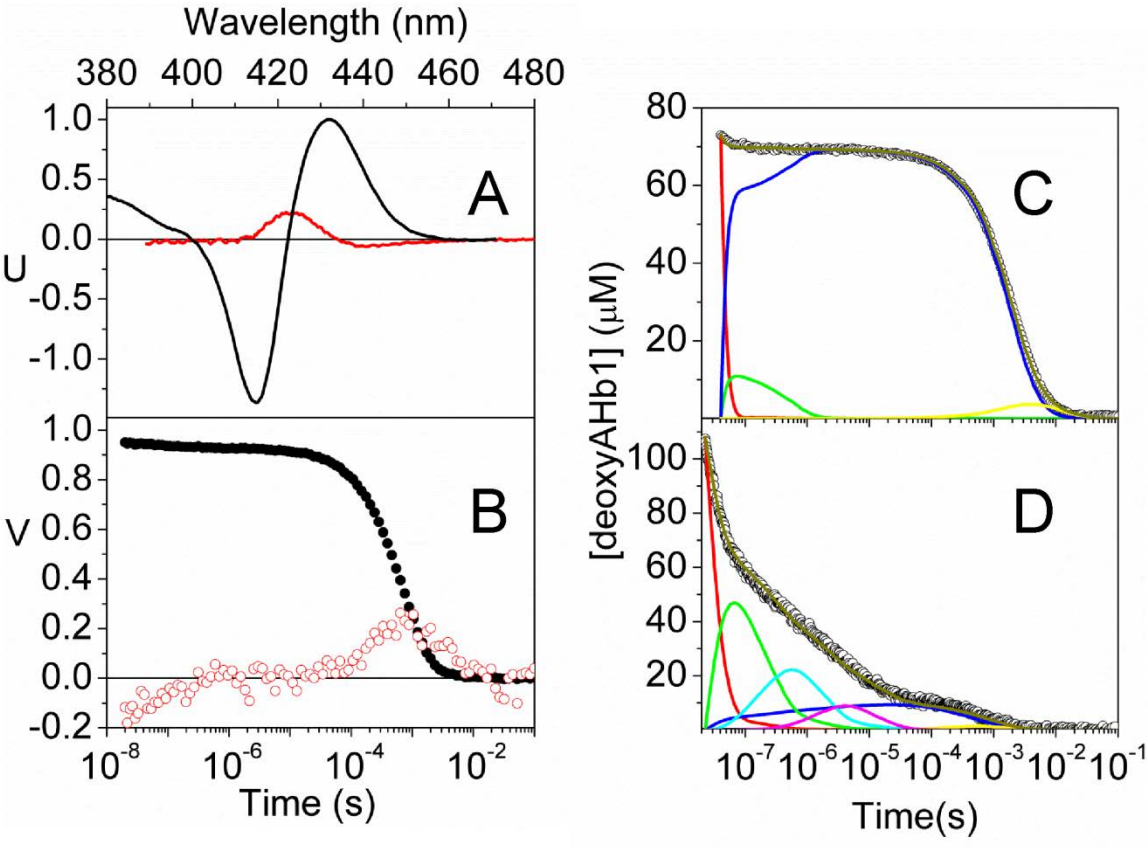
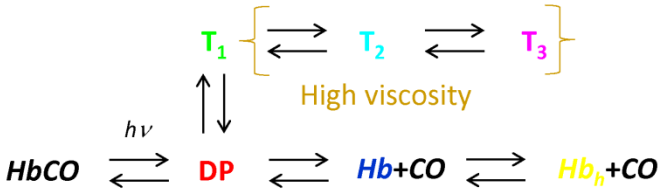


Figure 2. A. Comparison of the first (black line, S_1 = 26.9) and second (red line, S_2 = 0.4, ×10) spectral components obtained from the SVD analysis on the time resolved spectra measured for AHb1 in solution. CO = 1 atm, T = 20 °C. **B.** Time evolution of the amplitudes of the first (black circles, plotted as the fraction of unliganded molecules as a function of time after the nanosecond photolysis) and second (red open circles) spectral components reported in **A.** The first component tracks the progress curve for ligand rebinding and is identical to the one (not shown) obtained from the absorbance changes measured at 436 nm (data from ref.[68]). The second component reflects formation and decay of the bis-histidyl complex. **C.** Fitting with the kinetic model detailed in ref.[172] of the CO rebinding kinetics to AHb1 in solution (CO = 1 atm, T = 10 °C, data from ref.[172]). **D.** Fitting with the kinetic model detailed in ref. [131] of the CO rebinding kinetics to AHb1 gels soaked in anhydrous glycerol (CO = 1 atm, T = 10 °C, data from ref.[131]). Color code of species in **C** and **D** as in Scheme 1.



Scheme 1. Minimal reaction scheme for the observed kinetics with sequential migration between internal hydrophobic cavities.[131] After photodissociation of the CO complex of AHb1 (*HbCO*), the photodissociated ligand can migrate from the distal pocket DP to a series of secondary sites (T_1, T_2, T_3) or exit to the solvent. The deoxy pentacoordinated species (*Hb*) is in equilibrium with the deoxy, bis-histidyl hexacoordinated species (Hb_h).

5. Telling the difference between still and moving structures

Location of internal cavities and their mutual connection can be achieved through the analysis of the three dimensional structure of proteins. The topological information retrieved from static structures provides a first evaluation of the availability and location of hydrophobic cavities, potentially capable of hosting diatomic ligands. However, the impact of factors such as heme coordination (i.e., bis-histidyl hexacoodination, pentacoordination, ligand binding at distal cavity) and protein dynamics on the number, shape and distribution of internal cavities cannot be overlooked, as illustrated by the comparison of the 3D structures of AHb1 in the bis-histidyl hexacoodinated, pentacoordinated, and ligand-bound states (Figure 3).

The structure of the bis-histidyl hexacoodinated AHb1 displays a set of internal hydrophobic cavities (Figure 3A), which are primarily shaped by hydrophobic residues (i.e., the cavity above the heme is formed by the side chains of residues Leu35, Phe36, Ile39, Phe50, Ala70, Val73, Phe74, Leu121 and Tyr145). These cavities may function as transient docking sites for the photodissociated ligand. Still, a direct connection between those cavities is not evident in the modelled structure nor there is a pathway for ligand egression to bulk solvent. It is thus difficult to justify the very low geminate recombination observed experimentally,[172] which suggests that relevant structural changes must occur upon change to pentacoordination or upon ligand binding. In contrast, the analysis of the ligand-bound hexacoordinated structure (Figure 3C) shows the existence of a hydrophobic tunnel in the protein matrix, which is mainly defined by residues Met24, Phe74, Cys77, Cys78, Ser80, Ala81, Leu84, Val90,

Trp141, Ala144, His147 and Leu148, thus leading from the heme to the end of helix E. This finding highlights a major change in the shape, extent and connectivity between internal docking sites compared to the bis-histidyl hexacoordinated structure.[175] Inspection of the topological features in the pentacoordinated state (Figure 3B) already shows the transition from the separate cavities found for the bis-histidyl hexacoordinated protein to the tunnel in the ligand-bound species, as noted in the formation of the docking site T_0, thus revealing the dramatic influence of the distal His coordination to the heme, and the concomitant rearrangement of helix E, on the nature and distribution of ligand docking sites. Hydrophobic and aromatic residues appear to play an important role in shaping cavities and tunnels.[176] Finally, Figure 3B also shows that a second pathway lines the distal HisE7 and extends toward the protein surface, though the completion of the accessibility to bulk solvent requires the opening of this residue,[175, 177] thus mimicking the HisE7 gating mechanism proposed for Mb. Overall, in this representation the distal pocket is linked, through an articulate tunnel, with the solvent.[175, 177]

These analyses challenge the kinetic model outlined above in Scheme 1, as they clearly show the simultaneous presence of two distinct pathways between the solvent and the binding site. Thus, while being capable of modelling the rebinding kinetics under all experimental conditions, the kinetic model outlined in Scheme 1 is inaccurate from a topological point of view. A revision of the reaction scheme which complies with both the kinetic data in solution and in gels and with the topological suggestions from MD simulations is reported in Scheme 2.

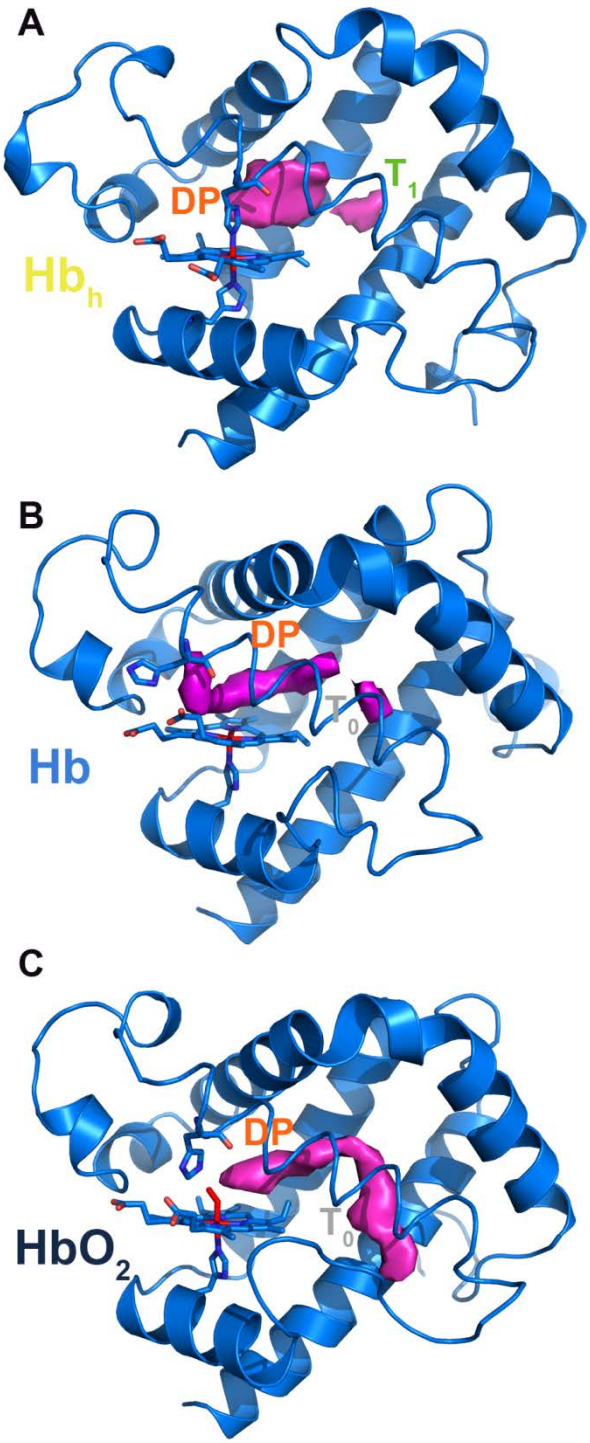
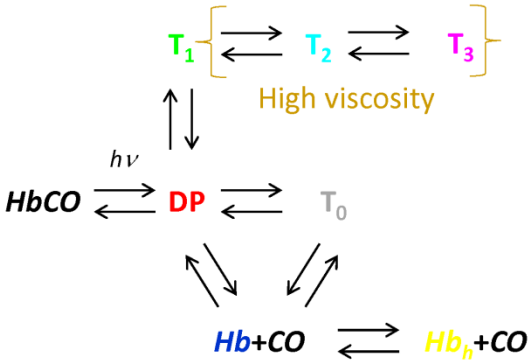


Figure 3. Topology of internal cavities and tunnels inside AHb1 as determined by the analysis of MD simulations with Fpocket.[178] A. Bis-histidyl hexacoordinated AHb1. B. Pentacoordinated AHb1 C. Ligand-bound hexacoordinated AHb1.



Scheme 2. Revised minimal reaction scheme for the observed kinetics with sequential migration between internal hydrophobic cavities. After photodissociation of the CO complex of AHb1 (*HbCO*), the photodissociated ligand in the distal pocket (DP) can: i) sequentially access secondary site T_1 (and under high viscosity conditions two other secondary sites T_2 and T_3 appear to be also accessible) ii) exit to the solvent (*Hb*) through a His-gate like mechanism; iii) migrate to docking site T_0 and then exit to the solvent. The deoxy pentacoordinated species (*Hb*) is in equilibrium with the deoxy, bis-histidyl hexacoordinated species (Hb_h).

The revised kinetic model leads to a description which only introduces minor improvements in comparison to Scheme 1, as judged from the representative analysis reported in Figure 4A. Nevertheless, Scheme 2 offers a description which much better fits within the topology of the ligand exchange pathways. Importantly, Scheme 2 results in time profiles (Figure 4A) for reaction intermediates that are quite different from the ones derived from Scheme 1 (Figure 2D). Thus, the additional information derived from modelling can fine tune our understanding of the way ligands migrate through the protein matrix, something that escapes the capabilities of numerical analysis of rebinding kinetics alone.

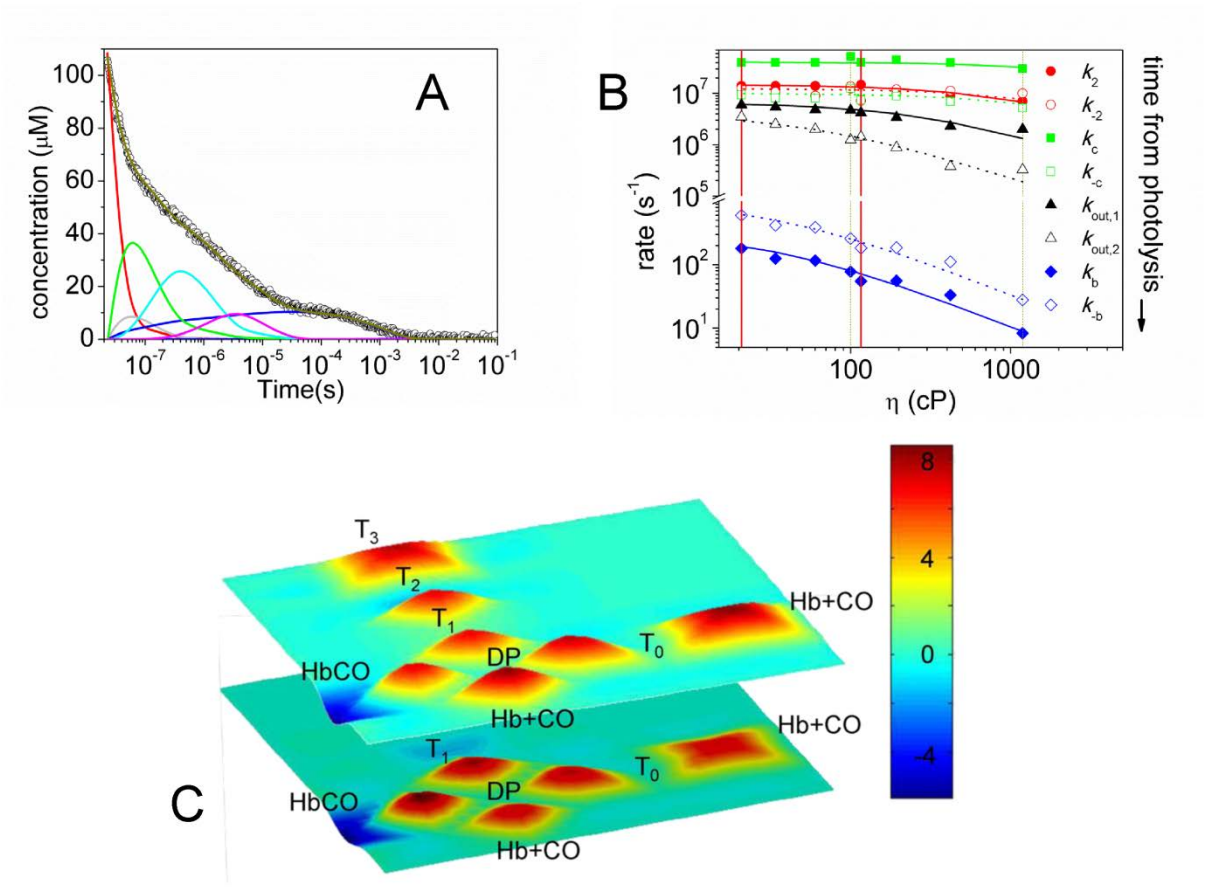


Figure 4 A. Representative fitting with the kinetic model detailed as Scheme 2 of the CO rebinding kinetics (shown as the concentration of unliganded protein as a function of time after laser photolysis) to AHb1 gels soaked in anhydrous glycerol (CO = 1 atm, T = 10 °C). Color code as in Scheme 2. B. Viscosity dependence of selected microscopic rates (filled symbols, forward rates; open symbols, reverse rates) determined using Scheme 2. The Kramers-like plots are built exploiting the temperature dependence (in the range 10°C – 40 °C) of the viscosity of the two glycerol-containing solutions (80% glycerol and 100% glycerol, ranges are marked by vertical lines). Lines are the best fit with a modified Kramers equation.[34] C. Schematic representation of the free energy surface (kcal/mol) for the ligand migration steps detailed in Scheme 2 (the top surface refers to high viscosity conditions (glycerol; ~10^3 cP), and the bottom surface to aqueous solution (~ 1 cP)).

Among other parameters, we note that the separation of distinct microscopic processes allows to expose the level of coupling of each rate constant with the solvent, as detailed in Figure 4B for selected rate constants. Processes which occur right after photolysis (on the nanosecond

time scale) do not sense the coupling between the protein fluctuations and the solvent, and show negligible dependence on bulk viscosity.[131] As time goes by, this coupling becomes evident,[14, 34, 108] but it is only for the slowest processes (like, e.g. distal His binding and dissociation) that the rates follow a Kramers-like behavior corresponding to full coupling between protein and solvent dynamics.

Finally, the temperature dependence of microscopic rate constants allows building free energy profiles for the modelled reactions.[57, 179, 180] Figure 4C shows a pictorial representation of the free energy surface determined along the migration pathways through cavities, exit to the solvent through the tunnel hosting cavity T_0, and directly from distal pocket through the HisE7 gate.

6. Looking at the shortest time scales of ligand binding kinetics

An obvious and largely underappreciated issue is the fact that the presence of inner pathways with barriers of rather different extent and conformational changes induced by deligation, often result in widely distributed rate constants, leading photodissociated ligands to return to their minimum-energy bound-state over extremely large time spans. The previous discussion has touched on the already enormous time span usually covered by nanosecond flash photolysis, which affords an adequate temporal resolution for CO rebinding to many hemeproteins. It is routine to monitor reaction kinetics from a few nanoseconds, where geminate rebinding often occurs, to the tens of seconds, as typically observed for CO binding to slowly reacting, bis-histidyl hexacoordinated species. [52, 53, 181-184] However, a number of cases have accumulated in the literature where the need for higher temporal resolution was dramatically evident, thus calling for extension of the dynamics on the short time scales side. [185-189] This may arise from higher reactivity towards the ligands (due to its chemical nature as is the case for NO and O_2) or from the existence of large barriers along the exit path from the distal cavity. A few recent examples are truncated Hbs [190-194] a heme-based GAF sensor domain [195], and nitrophorins. [196] Subnanosecond events may lead to modulation of the kinetics due to fast migration to internal cavities or to heterogeneous rate constants arising from coexistence of different conformations. [189, 191, 196] Thus, for those cases where reaction kinetics begins in the picoseconds, an additional challenge sets in.

The ability of describing the rebinding kinetics with a fully microscopic model requires that the complete time course of the progress curve be available. As of today, no single technique provides the necessary temporal dynamics, which needs to cover a tremendous extension in time, spanning over 12 order of magnitude. The approach that has been proposed to overcome the limited temporal dynamics of optical detection methods is the merging of separate time ranges acquired with different methodologies. [191, 196] With a full time course available, and having identified the structural determinants for ligand migration, it becomes possible to apply a quantitative analysis of the rebinding process over its entire time course. [191]

The case of type II truncated Hb from *Thermobifida fusca* (*Tf*-trHb) well describes how this approach merges the topological information available from MD with the experimental kinetic pattern. Panels A and B in Figure 5 show the 3D layout of the internal cavities detected in *Tf*-trHb with ILS (the method is discussed in Section 7). [191] The presence of two conformations (A and B) for the YCD1 residue makes the cavity T_1 more (B) or less (A) accessible to ligands in the distal cavity (DP). This results in a modulation of the rebinding kinetics occurring over the short nanosecond time scale. Unlike other Hbs, this protein has no direct connection between DP and the solvent, and the exchange of ligands between the binding site and the solvent occurs through a rather narrow hydrophobic tunnel. Besides the primary docking site (DP), ILS has identified two main minima in the energy contours, an on-pathway docking site (T_2) and an off-pathway docking site (T_1), the latter being transiently accessible through the rotation of YCD1 side chain. The use of the microscopic model sketched in Figure 5D allows us to describe the observed kinetics (Figure 5C), as demonstrated by the excellent agreement between experimental and modelled curve over the entire time span, and retrieve rate constants for every step. [191] It is foreseen that for those cases showing substantial sub-nanosecond geminate rebinding, the use of such integrated approach will lead to a dramatic improvement in the precision of the microscopic rate constants determined from the analysis of the rebinding curves.

In the cases discussed so far the electronic absorption of the prostetic group heme is exploited as a spectroscopic marker for sensing ligand migration. While providing high sensitivity, thanks to the dramatic changes in the absorption spectrum upon binding, this approach is detecting ligand migration only indirectly, and provides no information on the location of the gas. Probing directly the IR stretching absorption of CO, which is quite sensitive to the polarity of its environment, appears to be an interesting approach to pursue. Recent

femtosecond pump and probe experiments with multichannel detection in the mid infrared, have shown that for *Tf*-trHb and *B. subtilis* trHb it is possible to track photodissociated CO at remote docking sites through its transient absorption at ~2200 cm^{-1}, [192] just like in TDS-FTIR spectroscopy. [129] Ligand recombination is followed through the bleaching of the ~1900 cm^{-1} band. A similar approach could be in principle applied using step scan FTIR detection [197] and thus allow the extension of the kinetics to the micro-to-milliseconds range.

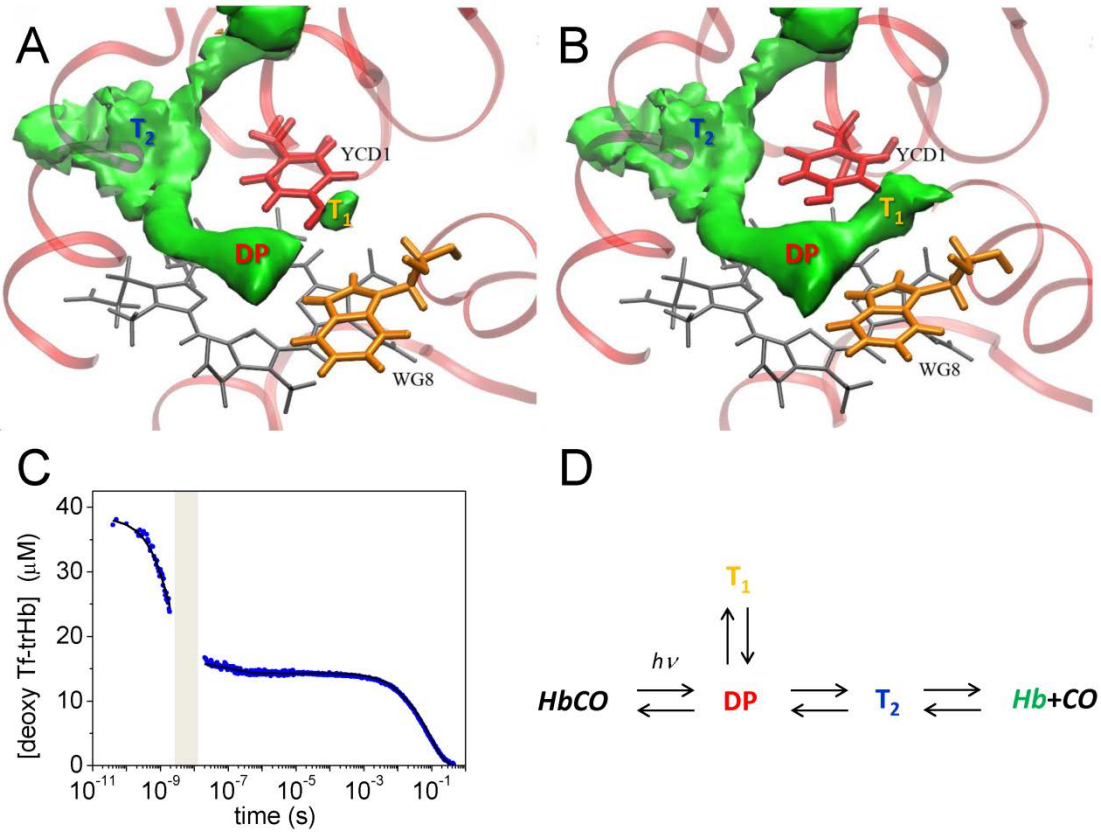


Figure 5. Top. Representations of the heme distal pocket, the tunnel and cavity system. Two spatial conformations (A and B) of YCD1 (in red) are depicted, the second one (B) making accessible the T_1 cavity (adapted from ref [191]). On the basis of the topological arrangement of inner cavities and tunnels, a minimal reaction scheme (D) can be proposed to describe the observed CO rebinding kinetics (shown as the concentration of unliganded protein as a function of time after laser photolysis) to *Tf*-trHb reported in C (blue dots, experimental data, black line, fitted curve; grey shaded area, time range inaccessible to either technique). DP and T_2 indicate respectively the primary and secondary docking site for the photodissociated CO inside the distal pocket along the exchange pathway with the solvent, while T_1 represents

a reaction intermediate with CO in a temporary docking site accessible from the distal site (see ref [191]for details).

7. One step beyond: closing the gap between simulations and experiments

The cases discussed in the previous paragraphs clearly demonstrate that the availability of topological information can assist in the definition of suitable reaction schemes to be used in the analysis of rebinding kinetics. Experimental and computational information thus becomes tightly interwoven to produce a quantitative description of the migration processes.

However, as remarked before, the major challenge still lies in moving from a qualitative description of observed phenomena by means of modeled dynamics to a quantitative comparison of microscopic rates and energetic profiles determined from theory and experiments, as recently suggested by several research groups actively developing theoretical approaches[198] or trying to merge directly experiments and modeling.[199] Clearly, a synergistic approach that combines the structural information about internal cavities with accurate estimates of the free energy landscape for ligand migration would be very useful to provide a microscopic interpretation of the intermediates and reaction rates derived from experimental kinetic studies. To this end, several computational methodologies have been developed to identify the ligand migration pathways through the protein matrix, to characterize the potential energy surface for ligand diffusion, and to derive eventually kinetic rate constants.

The LES technique was proposed by Elber and Karplus as an approach for "enhanced sampling in ligand binding and related processes in macromolecular dynamics" and the range of applications was illustrated by exploring the ligand binding pathways in Mb.[145] In this study, the system was separated into two subsystems: the ligand and the protein. The ligand involved an ensemble of 60 non-interacting CO molecules, which moved through a protein matrix simulated by a single MD trajectory subjected to the effective potential obtained by averaging over the ensemble of ligands. The mean-field effect obtained in this way smoothes the energy landscape, and thus improves sampling through reduction of barrier heights. [200] However, it has been shown that although every local energy minimum in the potential surface of the real system is also a minimum on the LES potential energy landscape, there are minima on the LES potential energy surface that are not found on the potential surface of the real system. [201] Hence, the LES trajectories should be primarily valuable for identifying

migration pathways within the protein matrix, which can then serve as starting points for more accurate calculations of the reaction dynamics. The main advantage of this approach is the reduction in computational time for ligand sampling, as the calculation is approximately a factor of N (N being the number of ligand copies) faster than the N protein-ligand trajectories required for separate standard MD simulations.

The LES approach has found wide application in the identification of pathways for binding and escape to Hbs, as noted in few representative examples. The study of CO diffusion through Mb by Elber and Karplus highlighted that structural fluctuations are very significant for facilitating the ligand migration, as the analysis of the LES trajectories showed that barriers between cavities are reduced transiently by coupled motions of the protein backbone and side chains. [143] LES simulations have been used by Orlowski and Nowak to investigate oxygen transport in the miniHb from *C. lacteus* [202] and in human cytoglobin. [203] In this latter study five cavities accessible to ligands were reported and, apart from the direct exit from the heme distal cavity, at least three additional ligand migration routes were identified. The role of the most feasible route for ligand migration has been supported by recent studies, [204] which have also shown the large structural plasticity that affects both the internal volumen and the number of nature of the tunnels depending on the coordination state of cytoglobin. Golden and Olsen have utilized this technique to explore the ligand-binding pathways in three truncated Hbs (from *M. tuberculosis*, *P. caudatum*, and *C. eugametos*), demonstrating the existence of at least two paths common for the three proteins. [205] As a last example, Olsen and coworkers have recently used LES MD simulations to investigate ligand migration in the truncated Hb O from *M. tuberculosis* and its TrpG8Phe mutant. [206] The results showed that the three diffusion pathways found in the wild type protein are reduced to only two in the mutated protein, thus supporting the critical role of TrpG8 in regulating the ligand escape from the heme pocket.

The ILS technique [146] is a computationally inexpensive approach to describe the energetics of ligand migration through a post-processing of the MD trajectory sampled for the protein in the absence of the ligand. By defining an ensemble of representative snapshots, which are superimposed and enclosed in a 3D grid, ILS computes the interaction energy of the ligand, located at every grid point and considering all the possible rotations in the local reference frame, with the protein. Then, the free energy of placing the ligand at a given position is estimated from the Boltzmann-weighted average of interaction energies determined for the ensemble of snapshots. At the end, ILS provides a complete 3D map of the potential of mean

force of gas ligand placement at any position inside the solvated protein. The main limitation of this strategy is the neglect of the coupling between ligand and protein dynamics, which should result in overestimated energy barriers. [139, 207, 208] However, this assumption can be reasonable for weakly interacting ligands, since then the interaction can be treated as a perturbation to the equilibrium dynamics of the protein.

Cohen et al. used ILS to characterize the docking sites and the pathways between these sites for a series of small ligands (O2, NO, CO and Xe) inside Mb. [139] The calculations reproduced the experimentally measured locations of the Xe binding sites. They also showed similar cavities and pathways for all the ligands, but the method was also able to identify differences in the energy values corresponding to the interaction with docking sites and energy barriers between cavities. ILS was also used to compare the migration pathways for a series of 12 monomeric globins. [140] Despite the preservation of the 3D structural fold, it was shown that there is a large variability in the shape and topology of the migration pathways, which are found to correlate with the location of large hydrophobic residues. The technique has also been utilized to characterize the migration of ligands in other proteins, including the truncated Hb N from *M. tuberculosis*, [209, 210] a comparative analysis of fish and human Ngb, [211] and the effect of dimerization and ligand binding on the channels in protoglobin from *M. acetivorans*, [212] among others. [213, 214]

GRID-MD is a related computational approach proposed for the exploration of ligand migration pathways that relies on the use of a 3D energy grid for the interaction of suitable probes with an ensemble of protein conformations. [215] Once the potential energy has been calculated, the force generated by the protein on a particle at any grid point is determined (including a Boltzmann weighting of the grids built up for each protein conformation) and subsequently used in the context of Brownian dynamics to explore the motion of the probe along the protein matrix. The suitability of the technique was tested for three proteins, including the truncated Hb N from *M. tuberculosis*, confirming the ligand migration through the two major branches of the tunnel found in the X-ray structure [216] and atomistic MD simulations, [217, 218] but also identified shorter channels that might act as product-release routes. For oxygen-bound hexaccordinated form of AHb1, GRID-MD confirmed the formation of the tunnel leading from the heme cavity to the protein surface (shown in Figure 3C), which might be implicated in the NO detoxification function of this protein. [177]

The Protein Energy Landscape Exploration (PELE) [147] represents a different strategy for determining the energetics of ligand migration. PELE combines a Monte Carlo stochastic approach with energy minimization calculations to predict ligand migration pathways and energetics at a very reduced computational cost. The algorithm involves the consecutive iteration of three main moves: a ligand and protein (backbone) perturbation, a side-chain sampling, and a minimization of a region including, at least, all residues local to the atoms involved in the previous moves. These three steps lead to a new structure that is accepted (defining a new minimum) or rejected based on a Metropolis criterion. The collection of accepted steps forms a stochastic trajectory and an effective exploration of the protein energy landscape associated with ligand migration. PELE has been utilized to revisit the migration of carbon monoxide in truncated Hb O from *M. tuberculosis*, [190] and human Hb, taking into account both tense and relaxed states. [219]

A powerful method to model ligand migration through the protein matrix is Multiple Steered Molecular Dynamics (MSMD).[220-224] In this strategy different MD simulations are performed to push the ligand from the protein active site toward the solvent (or to pull it toward the active site from the solvent) by means of an external guiding potential. When applied to ligand migration, the aim of the potential is to guide the ligand along the tunnel/cavity system and to overcome the possible entry/exit barriers. For each steered molecular dynamics simulation (SMD) the irreversible work performed by the guiding potential is measured along the ligand migration path. The free energy is obtained by computing the exponential average of the work values, as described by Jarzynski's equality. [208, 225] Estrin and coworkers have used this technique to investigate the energetics of ligand migration and the effect of specific mutations on the migration pathways in a variety of proteins, such as the miniHb from *C. lacteus*, [226] the truncated Hb O from *B. subtillis* [227] and *M. tuberculosis*, [228] and nitrophorins 2 and 4 from *R. prolixus*, [229, 230] Other enhanced sampling techniques, such as umbrella sampling, metadynamics and temperature-accelerated MD, have been alternatively utilized to characterize the free energy surface for ligand migration in hemeproteins. [231-235]

The increase in computer power seen in the last years facilitates estimating kinetic constants by analyzing multiple independent MD simulations of photolyzed ligand, leading to a quantitative description of the free energy surface for ligand migration. [235-237] This approach provides a time-dependent distribution of the ligand, which can then be compared with the experimental relaxation kinetics induced by photolysis. Moreover, by combining the ligand distribution obtained from MD simulations with the reaction free energy profile for ligand

binding to the heme, one can derive the kinetics for the whole process involved in ligand-heme bond formation/disruption and ligand migration. This approach is exemplified by the study of CO migration in Mb by D'Abramo et al., [238] which combined equilibrium atomistic MD simulations with quantum mechanical (QM) calculations for CO binding in order to derive kinetic rate constants. The results yielded rate constants in good agreement with the experimental values, showing that CO migration is characterized by relaxations in the picosecond range for CO transitions among Mb cavities, and in the nanosecond range for transitions between the distal pocket and the Xe4 binding site. Similar studies have been recently reported for Ngb. [239] A distinct approach has been undertaken by Blumberger and coworkers, as they utilize a multiscale simulation that combines kinetic data from both equilibrium simulations and enhanced samplig techniques.[240, 241] This information is combined to construct a master equation that describes the movement of gas molecules within the protein. The results obtained for H_2 and O_2 transport in a [NiFe]-hydrogenase can be fitted to the phenomenological rate used to interpret experiments and the diffusion rates reproduce the experimental data, even though the simulations point out the existence of a diverse network of accessible pathways by which the gas molecules can reach the active site.

Although the large number of computational studies focused on hemeproteins prevents a comprehensive review that will surpass the scope of this work, and the references only include an aribtrary selection of representative studies, we believe that they will suffice to demonstrate the progressive impact of theoretical methods in deciphering the structural and energetic determinants of ligand migration and how a judicious choice of techniques can be valuable to reduce the gap with experimental models of ligand binding.

8. Conclusions

A detailed characterization of ligand migration in hemeproteins demands a multifaceted approach that synergistically takes into account the topological features and structural plasticity of inner cavities as well as thermodynamic and kinetic properties determined from experimental measurements. In the last years, significant progress has been made in both simulation methods and experimental techniques, such as the advent of time-resolved crystallography, which permits to follow the time evolution of ligand diffusion, and the use of graphical processors units in conjunction with optimized simulation codes, which permit to run multiple trajectories for longer simulation times. These advances will facilitate the

exploration of the variety of molecular processes that encompasses the kinetic data derived from rebinding studies of photolyzed ligands, including recombination to the heme, diffusion between pockets, migration through tunnels, exit to the solvent, transition between bis-histidyl hexacoordination and pentacoordination, and thermal relaxation of protein backbone and side chains. Further experimental advances like extension of the temporal range available to kinetics investigations and the use of vibrational spectroscopy will increase the quality of the experimental data. In addition, the advances in computer power and the development of enhanced sampling techniques will afford a more comprehensive understanding of the topological and energetic features of ligand migration pathways. In conjunction with refined descriptions of the ligand binding to the heme provided by QM methods, this information will be a fundamental complement to assist the interpretation of the mechanistic events observed in experimental rebinding assays. We can predict that these advances will facilitate the transition from qualitative to quantitative descriptions of the kinetic mechanisms, with implications for substrate delivery to the heme cavity, chemical reactivity, and product egression. This will contribute to a comprehensive understanding of the structure-function relationships in hemeproteins, finally providing a framework which may be of help for other protein families.

Acknowledgment

The authors acknowledge the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (PRIN 2008, 2008BFJ34R, Azioni Integrate Italia Spagna 2009, IT10L1M59M), Ministero degli Affari Esteri, Direzione generale per la promozione del sistema Paese (Progetti di Grande Rilevanza, Italia-Argentina 2011–2013), Spanish Ministerio de Economia y Competitividad (SAF2011-27642), Generalitat de Catalunya (SGR2009-298), XRQTC and the Barcelona Supercomputer Center.

REFERENCES

1. W. A. Eaton, E. R. Henry, J. Hofrichter, S. Bettati, C. Viappiani and A. Mozzarelli, *IUBMB Life*, 2007, **59**, 586-599.
2. G. D. Stormo and Y. Zhao, *Nat Rev Genet*, 2010, **11**, 751-760.
3. A. Bellahcene, V. Castronovo, K. U. E. Ogbureke, L. W. Fisher and N. S. Fedarko, *Nat Rev Cancer*, 2008, **8**, 212-226.
4. E. C. Hulme and M. A. Trevethick, *British Journal of Pharmacology*, 2010, **161**, 1219-1237.
5. L. Wang, B. J. Berne and R. A. Friesner, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 1326-1330.
6. D. L. Mobley and K. A. Dill, *Structure (London, England : 1993)*, 2009, **17**, 489-498.
7. J. T. Henry and S. Crosson, *Annual Review of Microbiology*, 2011, **65**, 261-286.
8. H. Gohlke and G. Klebe, *Angew. Chem. Int. Ed.*, 2002, **41**, 2644-2676.
9. C. Bissantz, B. Kuhn and M. Stahl, *J. Med. Chem.*, 2010, **53**, 5061-5084.
10. C. A. Hunter, *Angew. Chem. Int. Ed.*, 2004, **43**, 5310-5324.
11. D. H. Williams, E. Stephens, D. P. O'Brien and M. Zhou, *Angew. Chem. Int. Ed.*, 2004, **43**, 6596-6616.
12. G. Kuzu, O. Keskin, A. Gursoy and R. Nussinov, *Methods Mol. Biol.*, 2012, **819**, 59-74.
13. S. J. Teague, *Nat Rev Drug Discov*, 2003, **2**, 527-541.
14. H. Frauenfelder, G. Chen, J. Berendzen, P. W. Fenimore, H. Jansson, B. H. McMahon, I. R. Stroe, J. Swenson and R. D. Younge, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 5129–5134.
15. H. Frauenfelder, B. H. McMahon, R. H. Austin, K. Chu and J. T. Groves, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 2370–2374.
16. P. Csermely, R. Palotai and R. Nussinov, *TIBS*, 2010, **35**, 539-546.
17. F. Spyrakis, A. Bidon-Chanal, X. Barril and F. J. Luque, *Curr. Topics Med. Chem.*, 2011, **11**, 192-210.
18. H. Frauenfelder, P. W. Fenimore and B. H. McMahon, *Biophys. Chem.*, 2002, **98**, 35–48.
19. L. Cordone, G. Cottone, S. Giuffrida, G. Palazzo, G. Venturoli and C. Viappiani, *Biochim. Biophys. Acta - Proteins and Proteomics*, 2005, **1749**, 252– 281.
20. I. Buch, T. Giorgino and G. De Fabritiis, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 10184-10189.
21. R. O. Dror, A. C. Pan, D. H. Arlow, D. W. Borhani, P. Maragakis, Y. Shan, H. Xu and D. E. Shaw, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 13118-13123.
22. H. Frauenfelder, B. H. McMahon and P. W. Fenimore, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 8615–8617.
23. J. Monod, J. Wyman and J. P. Changeux, *J. Mol. Biol.*, 1965, **12**, 88-118.
24. W. A. Eaton, E. R. Henry, J. Hofrichter and A. Mozzarelli, *Nature Struct. Biol.*, 1999, **6**, 351-358.
25. J. P. Changeux and S. J. Edelstein, *Science*, 2005, **308**, 1424-1428.
26. Q. H. Gibson, in *Comprehensive Biochemistry*, eds. G. Semenza and A. J. Turner, 2004, vol. 43, pp. 101–197.
27. R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder and I. C. Gunsalus, *Biochemistry*, 1975, **14**, 5355-5373.

28. D. Beece, L. Eisenstein, H. Frauenfelder, D. Good, M. C. Marden, L. Reinisch, A. H. Reynolds, L. B. Sorensen and K. T. Yue, *Biochemistry*, 1980, **19**, 5147-5157.
29. W. Doster, D. Beece, S. F. Bowne, E. E. DiIorio, L. Eisenstein, H. Frauenfelder, L. Reinisch, E. Shyamsunder, K. H. Winterhalter and K. T. Yue, *Biochemistry*, 1982, **21**, 4831-4839.
30. E. R. Henry, J. H. Sommer, J. Hofrichter and W. A. Eaton, *J. Mol. Biol.*, 1983, **166**, 443-451.
31. A. Ansari, J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder and R. D. Young, *Biochemistry*, 1985, **82**, 5000-5004.
32. V. Srajer, L. Reinisch and P. M. Champion, *J. Am. Chem. Soc.*, 1988, **110**, 6656-6670.
33. P. J. Steinbach, A. Ansari, J. Berenden, D. Braunstein, K. Chu, B. R. Cowen, D. Ehrenstein, H. Frauenfelder, J. B. Johnson, D. C. Lamb, S. Luck, J. R. Mourant, G. U. Nienhaus, P. Ormos, R. Philipp, A. Xie and R. D. Young, *Biochemistry*, 1991, **30**, 3988-4001.
34. A. Ansari, C.M.Jones, E. R. Henry, J. Hofrichter and W. Eaton, *Science*, 1992, **256**, 1796-1798.
35. W. D. Tian, J. T. Sage, V. Srajer and P. M. Champion, *Phys. Rev. Lett.*, 1992, **68**, 408-411.
36. F. Post, W. Doster, G. Karvounis and M. Settles, *Biophys. J.*, 1993, **64**, 1833-1842.
37. A. Ansari, C.M.Jones, E. R. Henry, J. Hofrichter and W. Eaton, *Biochemistry*, 1994, **33**, 5128-5145.
38. E. Chen, R. A. Goldberg and D. S. Kliger, *Ann. Rev. Biophys. Biomol. Struct.*, 1997, **26**, 327-355.
39. R. M. Esquerra, R. A. Goldbeck, D. B. Kim-Shapiro and D. S. Kliger, *Biochemistry*, 1998, **37**, 17527-17536.
40. T. Kleinert, W. Doster, H. Leyser, W. Petry, V. Schwarz and M. Settles, *Biochemistry*, 1998, **37**, 717-733.
41. B. H. McMahon, B. P. Stojkovic, P. J. Hay, R. L. Martin and A. E. Garcia, *J. Chem. Phys.*, 2000, **113**, 6831-6850.
42. J. Hofrichter, J. H. Sommer, E. R. henry and W. A. Eaton, *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 2235-2239.
43. L. P. Murray, J. Hofrichter, E. R. Henry and W. A. Eaton, *Biophys. Chem.*, 1988, **29**, 63-76.
44. L. P. Murray, J. Hofrichter, E. R. Henry, M. Ikeda-Saito, K. Kitagishi, T. Yonetani and W. A. Eaton, *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 2151-2155.
45. E. W. Findsen, J. M. Friedman and M. R. Ondrias, *Biochemistry*, 1988, **27**, 8719-8724.
46. C. M. Jones, A. Ansari, E. R. Henry, G. W. Christoph, J. Hofrichter and W. A. Eaton, *Biochemistry*, 1992, **31**, 6692-6702.
47. E. R. Henry, C. M. Jones, J. Hofrichter and W. A. Eaton, *Biochemistry*, 1997, **36**, 6511-6528.
48. S. C. Bjorling, R. A. Goldbeck, S. J. Paquette, S. J. Milder and D. S. Kliger, *Biochemistry*, 1996, **35**, 8619-8627.
49. R. A. Goldbeck, S. J. Paquette, S. C. Björling and D. S. Kliger, *Biochemistry*, 1996, **35**, 8628-8639.
50. R. M. Esquerra, R. A. Goldbeck, S. H. Reaney, A. M. Batchelder, Y. Wen, J. W. Lewis and D. S. Kliger, *Biophys. J.*, 2000, **78**, 3227-3239.
51. R. A. Goldbeck, S. J. Paquette and D. S. Kliger, *Biophys. J.*, 2001, **81**, 2919–2934.

52. 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.
53. J. M. Kriegl, A. J. Bhattacharyya, K. Nienhaus, P. Deng, O. Minkow and G. U. Nienhaus, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 7992-7997.
54. 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.
55. J. Uzan, S. Dewilde, T. Burmester, T. Hankeln, L. Moens, D. Hamdane, M. C. Marden and L. Kiger, *Biophys. J.*, 2004, **87**, 1196–1204.
56. D. Hamdane, L. Kiger, G. H. B. Hoa, S. Dewilde, J. Uzan, T. Burmester, T. Hankeln, L. Moens and M. C. Marden, *J. Biol. Chem.*, 2005, **280**, 36809–36814.
57. S. Abbruzzetti, S. Faggiano, S. Bruno, F. Spyrakis, A. Mozzarelli, S. Dewilde, L. Moens and C. Viappiani, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 18984–18989.
58. L. Astudillo, S. Bernad, V. Derrien, P. Sebban and J. Miksovska, *Biophys. J.*, 2010, **99**, L16–L18.
59. J. Haldane and J. Lorrain-Smith, *J. Physiol.*, 1896, **20**, 497–520.
60. Q. H. Gibson and S. Ainsworth, *Nature*, 1957, **180**, 1416-1417.
61. Q. H. Gibson, *J. Physiol.*, 1956, **134**, 112–122.
62. Q. H. Gibson, *Biochem. J.*, 1959, **71**, 293-303.
63. R. W. Noble, M. Brunori, J. Wyman and E. Antonini, *Biochemistry*, 1967, **6**, 1216-1222.
64. M. Brunori, G. M. Giacometti, E. Antonini and J. Wyman, *Proc. Natl. Acad. Sci. USA*, 1973, **70**, 3141-3144.
65. W. A. Saffran and Q. H. Gibson, *J. Biol. Chem.*, 1977, **252**, 7955-7958.
66. L. J. Parkhurst, *Ann. Rev. Phys. Chem.*, 1979, **30**, 503-546.
67. Q. H. Gibson, *Bioch. Soc. Trans.*, 1990, **18**, 1-6.
68. S. Abbruzzetti, S. Bruno, S. Faggiano, E. Grandi, A. Mozzarelli and C. Viappiani, *Photochem. Photobiol. Sci.*, 2006, **5**, 1109-1120.
69. R. G. W. Norrish and G. Porter, *Nature*, 1949, **164**, 658.
70. D. A. Duddell, R. J. Norris, N. J. Muttucumaru and J. T. Richards, *Photochem.Photobiol.*, 1980, **31**, 479-484.
71. R. H. Austin, K. Beeson, L. Eisenstein, H. Frauenfelder, I. C. Gunsalus and V. P. Marshall, *Science*, 1973, **181**, 541-543.
72. J. Huang, A. Ridsdale, J. Wang and J. M. Friedman, *Biochemistry*, 1997, **36**, 14353-14365.
73. B. F. Campbell, M. R. Chance and J. M. Friedman, *Science*, 1987, **238**, 373-376.
74. N. Agmon, *Biochemistry*, 1988, **27**, 3507-3511.
75. W. Doster, T. Kleinert, F. Post and M. Settles, in *Protein-Solvent Interactions*, ed. R. B. Gregory, Marcel Dekker, 1993, pp. 375-.
76. N. Agmon, W. Doster and F. Post, *Biophys. J.*, 1994, **66**, 1612-1622.
77. N. Agmon and J. Hopfield, *J. Chem. Phys.*, 1983, **79**, 2042-2053.
78. L. Powers, B. Chance, M. Chance, B. Campbell, J. Friedman, S. Khalid, C. Kumar, A. Naqui, K. S. Reddy and Y. Zhou, *Biochemistry*, 1987, **26**, 4785-4796.
79. V. Srajer, L. Reinisch and P. M. Champion, *Biochemistry*, 1991, **30**, 4886-4895.
80. B. P. Schoenborn, *Nature*, 1965, **208**, 760-762.
81. R. F. J. Tilton, I. D. J. Kuntz and G. A. Petsko, *Biochemistry*, 1984, **23**, 2849-2857.
82. I. Schlichting, J. Berendzen, G. N. Phillips and R. M. Sweet, *Nature*, 1994, **371**, 808-812.
83. H. Hartmann, S.Zinser, P.Komninos, R. T. Schneider, G. U. Nienhaus and F. Parak, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 7013-7016.

84. J. Vojtechovsky, K. Chu, J. Berendzen, R. M. Sweet and I. Schlichting, *Biophys. J.*, 1999, **77**, 2153-2174.
85. A. Ostermann, R. Washipky, F. G. Parak and G. U. Nienhaus, *Nature*, 2000, **404**, 205-208.
86. M. Brunori, B. Vallone, F. Cutruzzolà, C. Travaglini-Allocatelli, J. Berendzen, K. Chu, R. M. Sweeti and I. Schlichting, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 2058–2063.
87. V. Srajer, Z. Ren, T. Y. Teng, M. Schmidt, T. Ursby, D. Bourgeois, C. Pradervand, W. Schildkamp, M. Wulff and K. Moffat, *Biochemistry*, 2001, **40**, 13802-13815.
88. D. Bourgeois, B. Vallone, F. Schotte, A. Arcovito, A. E. Miele, G. Sciara, M. Wulff, P. Anfinrud and M. Brunori, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 8704–8709.
89. F. Schotte, M. Lim, T. A. Jackson, A. V. Smirnov, J. Soman, J. S. Olson, G. N. P. Jr., M. Wulff and P. A. Anfinrud, *Science*, 2003, **300**, 1944-1947.
90. M. Brunori, F. Cutruzzolà, C. Savino, C. Travaglini-Allocatelli, B. Vallone and Q. H. Gibson, *Biophys. J.*, 1999, **76**, 1259–1269.
91. E. E. Scott and Q. H. Gibson, *Biochemistry*, 1997, **36**, 11909-11917.
92. E. E. Scott, Q. H. Gibson and J. S. Olson, *J. Biol. Chem.*, 2001, **276**, 5177–5188.
93. F. Librizzi, C. Viappiani, S. Abbruzzetti and L. Cordone, *J. Chem. Phys.*, 2002, **116**, 1193-1200.
94. D. Dantsker, U. Samuni, A. J. Friedman, M. Yang, A. Ray and J. M. Friedman, *J. Mol. Biol.*, 2002, **315**, 239-251.
95. K. Nienhaus, P. Deng, J. M. Kriegl and G. U. Nienhaus, *Biochemistry*, 2003, **42**, 9647-9658.
96. K. Nienhaus, P. Deng, J. M. Kriegl and G. U. Nienhaus, *Biochemistry*, 2003, **42**, 9633-9646.
97. C. Tetreau, Y. Blouquit, E. Novikov, E. Quiniou and D. Lavalette, *Biophys. J.*, 2004, **86**, 435–447.
98. S. Sottini, S. Abbruzzetti, C. Viappiani, L. Ronda and A. Mozzarelli, *J. Phys. Chem. B*, 2005, **109**, 19523 - 19528.
99. S. Abbruzzetti, S. Giuffrida, S. Sottini, C. Viappiani and L. Cordone, *Cell Biochem. Biophys.*, 2005, **43**, 431-438.
100. D. Dantsker, C. Roche, U. Samuni, G. Blouin, J. S. Olson and J. M. Friedman, *J. Biol. Chem.*, 2005, **280**, 38740–38755.
101. D. Lavalette, C. Tetreau and L. Mouawad, *Bioch. Soc. Trans.*, 2006, **34**, 975-978.
102. D. Dantsker, U. Samuni, J. M. Friedman and N. Agmon, *Biochim. Biophys. Acta*, 2005, **1749**, 234 – 251.
103. M. Lim, T. A. Jackson and P. A. Anfinrud, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 5801-5804.
104. M. Levantino, A. Cupane, L. Zimanyi and P. Ormos, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 14402–14407.
105. D. Bourgeois, B. Vallone, A. Arcovito, G. Sciara, F. Schotte, P. A. Anfinrud and M. Brunori, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 4924–4929.
106. W. A. Eaton, E. R. Henry and J. Hofrichter, *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 4472–4475.
107. E. R. Henry, S. Bettati, J. Hofrichter and W. A. Eaton, *Biophys. Chem.*, 2002, **98**, 149-164.
108. U. Samuni, C. J. Roche, D. Dantsker and J. M. Friedman, *J. Am. Chem. Soc.*, 2007, **129**, 12756-12764.
109. U. Samuni, D. Dantsker, C. J. Roche and J. M. Friedman, *gene*, 2007, **398**, 234-248.
110. G. Cottone, G. Ciccotti and L. Cordone, *J. Chem. Phys.*, 2002, **117**, 9862-9866.

111. H. Frauenfelder, P. W. Fenimore, G. Chen and B. H. McMahon, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 15469–15472.
112. P. W. Fenimore, H. Frauenfelder, B. H. McMahon and F. G. Parak, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 16047–16051.
113. R. A. Goldbeck, S. B. C. Ortega, J. L. Mendoza, J. S. Olson, J. Soman, D. S. Kliger and R. M. Esquerra, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 1254–1259.
114. R. A. Goldbeck, M. L. Pillsbury, R. A. Jensen, J. L. Mendoza, R. L. Nguyen, J. S. Olson, J. Soman, D. S. Kliger and R. M. Esquerra, *J. Am. Chem. Soc.*, 2009, **131**, 12265–12272.
115. M. Brunori and Q. H. Gibson, *EMBO Reports*, 2001, **2**, 674-679.
116. M. Brunori, A. Giuffre`, K. Nienhaus, G. U. Nienhaus, F. M. Scandurra and B. Vallone, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 8483–8488.
117. M. Milani, A. Pesce, M. Nardini, H. Ouellet, Y. Ouellet, S. Dewilde, A. Bocedi, P. Ascenzi, M. Guertin, L. Moens, J. M. Friedman, J. B. Wittenberg and M. Bolognesi, *J. Inorg. Biochem.*, 2005, **99**, 97–109.
118. W. J. Song, G. Gucinski, M. H. Sazinsky and S. J. Lippard, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 14795–14800.
119. M. Brunori, *TIBS*, 2001, **26**, 209–210.
120. M. Brunori, *TIBS*, 2001, **26**, 21–23.
121. R. Pathania, N. K. Navani, A. M. Gardner, P. R. Gardner and K. L. Dikshit, *Mol. Microbiol.*, 2002, **45**, 1303-1314.
122. M. Couture, S. Yeh, B. A. Wittenberg, J. B. Wittenberg, Y. Ouellet, D. L. Rousseau and M. Guertin, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 11223-11228.
123. M. Schmidt, T. Graber, R. Henning and V. Srajer, *Acta Crystallog. Sect. A*, 2010, **A66**, 198–206.
124. J. Key, V. Srajer, R. Pahl and K. Moffat, *Biochemistry*, 2007, **44**, 4627-4635.
125. J. E. Knapp, R. Pahl, J. Cohen, J. C. Nichols, K. Schulten, Q. H. Gibson, V. Šrajer and W. E. Royer, *Structure*, 2009, **17**, 1494-1504.
126. B. Vallone, K. Nienhaus, K. Matthes, M. Brunori and G. U. Nienhaus, *Proteins*, 2004, **56**, 85.
127. G. U. Nienhaus and K. Nienhaus, *J. Biol. Phys.*, 2002, **28**, 163–172.
128. K. Nienhaus, P. Dominici, A. Astegno, S. Abbruzzetti, C. Viappiani and G. U. Nienhaus, *Biochemistry*, 2010, **49**, 7448–7458.
129. K. Nienhaus and G. U. Nienhaus, *Biochim. Biophys. Acta - Proteins and Proteomics*, 2011, **1814**, 1030-1041.
130. S. Sottini, S. Abbruzzetti, F. Spyrakis, S. Bettati, L. Ronda, A. Mozzarelli and C. Viappiani, *J. Am. Chem. Soc.*, 2005, **127**, 17427-17432.
131. S. Abbruzzetti, E. Grandi, S. Bruno, S. Faggiano, F. Spyrakis, A. Mozzarelli, P. Dominici and C. Viappiani, *J. Phys. Chem. B*, 2007, **111**, 12582-12590.
132. J. S. Olson and G. N. J. Phillips, *J. Biol. Chem.*, 1996, **271**, 17593-17596.
133. F. Draghi, A. E. Miele, C. Travaglini-Allocatelli, B. Vallone, M. Brunori, Q. H. Gibson and J. S. Olson, *J. Biol. Chem.*, 2002, **277**, 7509–7519.
134. J. S. Olson, J. Soman and G. N. P. Jr, *IUBMB Life*, 2007, **59**, 552 – 562.
135. M. D. Salter, G. C. Blouin, J. Soman, E. W. Singleton, S. Dewilde, L. Moens, A. Pesce, M. Nardini, M. Bolognesi and J. S. Olson, *J. Biol. Chem.*, 2012, **287**, 33163-33178.
136. F. Yang and G. N. Phillips, Jr., *J. Mol. Biol.*, 1996, **256**, 762–774.
137. M. Bolognesi, E. Cannillo, P. Ascenzi, G. M. Giacometti, A. Merli and M. Brunori, *J. Mol. Biol.*, 1982, **158**, 305–315.

138. D. Ringe, G. A. Petsko, D. Kerr and P. R. Ortiz de Montellano, *Biochemistry*, 1984, **23**, 2-4.
139. J. Cohen, A. Arkhipov, R. Braun and K. Schulten, *Biophys. J.*, 2006, **91**, 1844–1857.
140. J. Cohen and K. Schulten, *Biophys. J.*, 2007, **93**, 3591–3600.
141. Y. Nishihara, M. Sakakura, Y. Kimura and M. Terazima, *J. Am. Chem. Soc.*, 2004, **126**, 11877-11888.
142. E. Chovancová, A. Pavelka, P. Beneš, O. Strnad, J. Brezovský, B. Kozlíková, A. Gora, V. Šustr, M. Klvaňa, P. Medek, L. Biedermannová, J. Sochor and J. Damborský, *PLoS Comput. Biol.*, 2012, **8**, e1002708.
143. P. Schmidtke, A. Bidon-Chanal, F. J. Luque and X. Barril, *Bioinformatics*, 2011, **27**, 3276-3285.
144. T.-L. Lin and G. Song, *Proteins*, 2011, **79**, 2475-2490.
145. R. Elber and M. Karplus, *J. Am. Chem. Soc.*, 1990, **112**, 9161-9175.
146. J. Cohen, K. W. Olsen and K. Schulten, *Methods in Enzymology - Globins and other NO-reactive Proteins in Microbes, Plants and Invertebrates*, 2008, **437**, 439–457.
147. K. W. Borrelli, A. Vitalis, R. Alcantara and V. Guallar, *J. Chem. Theory Comput.*, 2005, **1**, 1304-1311.
148. L. Ellerby, C. R. Nishida, F. Nishida, S. A. Yamanaka, B. Dunn, J. S. Valentine and J. I. Zink, *Science*, 1992, **255**, 1113-1115.
149. S. Bettati, B. Pioselli, B. Campanini, C. Viappiani and A. Mozzarelli, in *Encyclopedia of Nanoscience and Nanotechnology*, ed. H. S. Nalwa, American Scientific Publishers, 2004, vol. 9, pp. 81-103.
150. S. Bruno, L. Ronda, S. Abbruzzetti, C. Viappiani, S. Bettati, S. Maji and A. Mozzarelli, in *Encyclopedia of Nanoscience and Nanotechnology.*, ed. H. S. Nalwa, American Scientific Publishers, 2011, pp. 481-517.
151. N. Shibayama and S. Saigo, *J. Mol. Biol.*, 1995, **251**, 203-209.
152. S. Bettati and A. Mozzarelli, *J. Biol. Chem.*, 1997, **272**, 32050-32055.
153. N. Shibayama and S. Saigo, *J. Am. Chem. Soc.*, 1999, **121**, 444-445.
154. C. Viappiani, S. Bettati, S. Bruno, L. Ronda, S. Abbruzzetti, A. Mozzarelli and A. W. Eaton, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 14414–14419.
155. G. Schirò and A. Cupane, *Biochemistry*, 2007, **46**, 11568-11576.
156. L. Ronda, S. Abbruzzetti, S. Bruno, S. Bettati, A. Mozzarelli and C. Viappiani, *J. Phys. Chem. B*, 2008, **112**, 12790–12794.
157. S. Bruno, M. Bonaccio, S. Bettati, C. Rivetti, C. Viappiani, S. Abbruzzetti and A. Mozzarelli, *Protein Sci.*, 2001, **10**, 2401-2407.
158. S. Abbruzzetti, C. Viappiani, S. Bruno, S. Bettati, M. Bonaccio and A. Mozzarelli, *J. Nanosci. Nanotech.*, 2001, **1**, 407-415.
159. U. Samuni, D. Dantsker, L. J. Juszczak, S. Bettati, L. Ronda, A. Mozzarelli and J. M. Friedman, *Biochemistry*, 2004, **43**, 13674-13682.
160. E. S. Peterson, R. Shinder, I. Khan, L. Juczszak, J. Wang, B. Manjula, S. A. Acharya, C. Bonaventura and J. M. Friedman, *Biochemistry*, 2004, **43**, 4832-4843.
161. N. Shibayama and S. Saigo, *FEBS Letters*, 2001, **429**, 50-53.
162. I. Khan, C. F. Shannon, D. Dantsker, A. J. Friedman, J. Perez-Gonzales-de-Apodaca and J. M. Friedman, *Biochemistry*, 2000, **39**, 16099-16109.
163. L. J. Juszczak and J. M. Friedman, *J. Biol. Chem.*, 1999, **274**, 30357-30360.
164. U. Samuni, D. Dantsker, A. Ray, J. B. Wittenberg, B. A. Wittenberg, S. Dewilde, L. Moens, Y. Ouellet, M. Guertin and J. M. Friedman, *J. Biol. Chem.*, 2003, **278**, 27241–27250.

165. D. Dantsker, U. Samuni, Y. Ouellet, B. A. Wittenberg, J. B. Wittenberg, M. Milani, M. Bolognesi, M. Guertin and J. M. Friedman, *J. Biol. Chem.*, 2004, **279**, 38844-38853.
166. U. Samuni, C. J. Roche, D. Dantsker, L. J. Juszczak and J. M. Friedman, *Biochemistry*, 2006, **45**, 2820-2835.
167. M. Perazzolli, P. Dominici, M. C. R. Puertas, E. Zago, J. Zeier, M. Sonoda, C. Lamb and M. Delledonne, *Plant Cell*, 2004, **16**, 2785-2794.
168. S. Bruno, S. Faggiano, F. Spyrakis, A. Mozzarelli, E. Cacciatori, P. Dominici, E. Grandi, S. Abbruzzetti and C. Viappiani, *Gene*, 2007, **398**, 224–233.
169. A. M. Massari, I. K. Finkelstein and M. D. Fayer, *J. Am. Chem. Soc.*, 2006, **128**, 3990-3997.
170. P. J. Steinbach, *J. Chem. Inf. Comput. Sci.*, 2002, **42**, 1476-1478.
171. P. J. Steinbach, R. Ionescu and C. R. Matthews, *Biophys. J.*, 2002, **82**, 2244-2255.
172. 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.
173. S. J. Hagen, J. Hofrichter and W. A. Eaton, *Science*, 1995, **269**, 959-962.
174. E. R. Henry and J. Hofrichter, in *Numerical computer methods*, eds. L. Brand and M. L. Johnson, Academic Press, Inc., San Diego, 1992, vol. 210, pp. 129-192.
175. S. Faggiano, S. Abbruzzetti, F. Spyrakis, E. Grandi, C. Viappiani, S. Bruno, A. Mozzarelli, P. Cozzini, A. Astegno, P. Dominici, S. Brogioni, A. Feis, G. Smulevich, O. Carrillo, P. Schmidtke, A. Bidon-Chanal and F. J. Luque, *J. Phys. Chem. B*, 2009, **113**, 16028–16038.
176. F. Spyrakis, F. J. Luque and C. Viappiani, *Plant Science*, 2011, **181**, 8-13.
177. F. Spyrakis, S. Faggiano, S. Abbruzzetti, P. Dominici, E. Cacciatori, A. Astegno, E. Droghetti, A. Feis, G. Smulevich, S. Bruno, A. Mozzarelli, P. Cozzini, C. Viappiani, A. Bidon-Chanal and F. J. Luque, *J. Phys. Chem. B*, 2011, **115**, 4138–4146.
178. V. LeGuilloux, P. Schmidtke and P. Tuffery, *BMC Bioinformatics*, 2009, **10**, 168.
179. S. Abbruzzetti, L. Tilleman, S. Bruno, C. Viappiani, F. Desmet, S. V. Doorslaer, M. Coletta, C. Ciaccio, P. Ascenzi, M. Nardini, M. Bolognesi, L. Moens and S. Dewilde, *PLoS ONE*, 2012, **7**, e33614.
180. S. Abbruzzetti, C. He, H. Ogata, S. Bruno, C. Viappiani and M. Knipp, *J. Am. Chem. Soc.*, 2012, **134**, 9986–9998.
181. M. S. Hargrove, *Biophys. J.*, 2000, **79**, 2733–2738.
182. A. Fago, C. Hundahl, S. Dewilde, K. Gilany, L. Moens and R. E. Weber, *J. Biol. Chem.*, 2004, **279**, 44417-44426.
183. T. Hankeln, B. Ebner, C. Fuchs, F. Gerlach, M. Haberkamp, T. L. Laufs, A. Roesner, M. Schmidt, B. Weich, S. Wystub, S. Saaler-Reinhardt, S. Reuss, M. Bolognesi, D. DeSanctis, M. C. Marden, L. Kiger, L. Moens, S. Dewilde, E. Nevo, A. Avivi, R. E. Weber, A. Fago and T. Burmester, *J. Inorg Biochem*, 2005, **99**, 110-119.
184. S. Kakar, F. G. Hoffman, J. F. Storz, M. Fabian and M. S. Hargrove, *Biophys. Chem.*, 2010, **152**, 1-14.
185. K. A. Jongeward, D. Magde, D. J. Taube, J. C. Marsters, T. G. Traylor and V. S. Sharma, *J. Am. Chem. Soc.*, 1988, **110**, 380-387.
186. J. W. Petrich, C. Poyart and J. L. Martin, *Biochemistry*, 1988, **27**, 4049-4060.
187. J. L. Martin and M. H. Vos, *Ann. Rev. Biophys. Biomol. Struct.*, 1992, 199–222.
188. M. H. Vos and J. L. Martin, *Biochim. Biophys. Acta*, 1999, **1411**, 1–20.
189. M. H. Vos, *Biochim. Biophys. Acta - Bioenegetics*, 2008, **1777**, 15–31.
190. V. Guallar, C. Lu, K. Borrelli, T. Egawa and S. R. Yeh, *J. Biol. Chem.*, 2009, **284**, 3106-3116.

191. A. Marcelli, S. Abbruzzetti, J. P. Bustamante, A. Feis, A. Bonamore, A. Boffi, C. Gellini, P. R. Salvi, D. A. Estrin, S. Bruno, C. Viappiani and P. Foggi, *PLoS ONE*, 2012, **7**, e39884.
192. A. Lapini, M. DiDonato, B. Patrizi, A. Marcelli, M. Lima, R. Righini, P. Foggi, N. Sciamanna and A. Boffi, *J. Phys. Chem. B*, 2012, **116**, 8753−8761.
193. H. Ouellet, L. Juszczak, D. Dantsker, U. Samuni, Y. H. Ouellet, P. Y. Savard, J. B. Wittenberg, B. A. Wittenberg, J. M. Friedman and M. Guertin, *Biochemistry*, 2003, **42**, 5764-5774.
194. A. Jasaitis, H. Ouellet, J. C. Lambry, J. L. Martin, J. M. Friedman, M. Guertin and M. H. Vos, *Chem. Phys.*, 2012, **396**, 10-16.
195. M. H. Vos, L. Bouzhir-Sima, J. C. Lambry, H. Luo, J. J. Eaton-Rye, A. Ioanoviciu, P. R. OrtizdeMontellano and U. Liebl, *Biochemistry*, 2012, **51**, 159-166.
196. A. Benabbas, X. Ye, M. Kubo, Z. Zhang, E. M. Maes, W. R. Montfort and P. M. Champion, *J. Am. Chem. Soc.*, 2010, **132**, 2811–2820.
197. I. Radu, M. Schleeger, C. Bolwien and J. Heberle, *Photochemical & Photobiological Sciences*, 2009, **8**, 1517-1528.
198. R. Elber, *Curr. Opin. Struct. Biol.*, 2010, **20**, 162–167.
199. S. V. Lepeshkevich, S. A. Biziuk, A. M. Lemeza and B. M. Dzhagarov, *Biochim. Biophys. Acta - Proteins and Proteomics*, 2011, **1814**, 1279-1288.
200. A. Roitberg and R. Elber, *J. Chem. Phys.*, 1991, **95**, 9277-9287.
201. C. M. Stultz and M. Karplus, *J. Chem. Phys.*, 1998, **109**, 8809-8815.
202. S. Orlowski and W. Nowak, *Theor. Chem. Acc.*, 2007, **117**, 253-258.
203. S. Orlowski and W. Nowak, *J. Mol. Model.*, 2007, **13**, 715–723.
204. M. Gabba, S. Abbruzzetti, F. S. S. Bruno, A. Mozzarelli, C. Viappiani, P. Cozzini, M. Nardini, F. Germani and M. B. L. M. S. Dewilde, *PLoS ONE*, 2013, **8**, e49770.
205. S. D. Golden, K. W. Olsen and R. K. Poole, in *Methods in Enzymology*, Academic Press, 2008, vol. Volume 437, pp. 459-475.
206. M. S. Heroux, A. D. Mohan and K. W. Olsen, *IUBMB Life*, 2011, **63**, 214-220.
207. M. W. Lee and M. Meuwly, *J. Phys. Chem. B*, 2012, **116**, 4154-4162.
208. F. Forti, L. Boechi, D. A. Estrin and M. A. Marti, *J. Comput. Chem.*, 2011, **32**, 2219–2231.
209. R. Daigle, J.-A. Rousseau, M. Guertin and P. Lagüe, *Biophys. J.*, 2009, **97**, 2967-2977.
210. A. Lama, S. Pawaria, A. Bidon-Chanal, A. Anand, J. L. Gelpi, S. Arya, M. Martì, D. A. Estrin, F. J. Luque and K. L. Dikshit, *J. Biol. Chem.*, 2009, **284**, 14457-14468.
211. D. Giordano, I. Boron, S. Abbruzzetti, W. VanLeuven, F. Nicoletti, F. Forti, S. Bruno, C. H. C. Cheng, L. Moens, G. diPrisco, A. Nadra, D. Estrin, G. Smulevich, S. Dewilde, C. Viappiani and C. Verde, *PLoS ONE*, 2012, **7**, e44508.
212. F. Forti, L. Boechi, D. Bikiel, M. A. Martí, M. Nardini, M. Bolognesi, C. Viappiani, D. Estrin and F. J. Luque, *J. Phys. Chem. B*, 2011, **115**, 13771–13780.
213. J. Saam, I. Ivanov, M. Walther, H.-G. Holzhuetter and H. Kuhn, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 13319-13324.
214. C. K. Regmi, Y. R. Bhandari, B. S. Gerstman and P. P. Chapagain, *J. Phys. Chem. B*, 2013, **117**, 2247-2253.
215. O. Carrillo and M. Orozco, *Proteins*, 2008, **70**, 892-899.
216. M. Milani, A. Pesce, Y. Ouellet, P. Ascenzi, M. Guertin and M. Bolognesi, *EMBO J.*, 2001, **20**, 3902-3909.
217. A. Crespo, M. A. Marti, S. G. Kalko, A. Morreale, M. Orozco, J. L. Gelpi, F. J. Luque and D. Estrin, *J. Am. Chem. Soc.*, 2005, **127**, 4433-4444.

218. A. Bidon-Chanal, M. A. Martí, A. Crespo, M. Milani, M. Orozco, M. Bolognesi, F. J. Luque and D. A. Estrin, *Proteins*, 2006, **64**, 457-464.
219. M. F. Lucas and V. Guallar, *Biophys. J.*, 2012, **102**, 887-896.
220. M. Orzechowski and P. Cieplak, *Journal of Physics: Condensed Matter*, 2005, **17**, S1627.
221. S. Park, F. Khalili-Araghi, E. Tajkhorshid and K. Schulten, *J. Chem. Phys.*, 2003, **119**, 3559-3566.
222. B. Isralewitz, M. Gao and K. Schulten, *Curr. Opin. Struct. Biol.*, 2001, **11**, 224-230.
223. H. Lu and K. Schulten, *Proteins: Struct. Funct. Genet.*, 1999, **35**, 453-463.
224. H. Grubmüller, B. Heymann and P. Tavan, *Science*, 1996, **271**, 997-999.
225. C. Jarzynski, *Phys. Rev. Lett.*, 1997, **78**, 2690-2693.
226. M. A. Martí, D. E. Bikiel, A. Crespo, M. Nardini, M. Bolognesi and D. A. Estrin, *Proteins: Struct. Funct. Bioinf.*, 2006, **62**, 641-648.
227. L. Boechi, P. A. Manez, F. J. Luque, M. A. Marti and D. A. Estrin, *Proteins: Struct. Funct. Bioinf.*, 2010, **78**, 962-970.
228. L. Boechi, M. A. Martı, M. Milani, M. Bolognesi, F. J. Luque and D. A. Estrin, *Proteins*, 2008, **73**, 372–379.
229. M. A. Marti, M. C. G. Lebrero, A. E. Roitberg and D. A. Estrin, *J. Am. Chem. Soc.*, 2008, **130**, 1611-1618.
230. J. M. Swails, Y. Meng, F. A. Walker, M. A. Marti, D. A. Estrin and A. E. Roitberg, *J. Phys. Chem. B*, 2009, **113**, 1192-1201.
231. M. Ceccarelli, R. Anedda, M. Casu and P. Ruggerone, *Proteins: Struct. Funct. Bioinf.*, 2008, **71**, 1231-1236.
232. A. Bocahut, S. Bernad, P. Sebban and S. Sacquin-Mora, *J. Phys. Chem. B*, 2009, **113**, 16257–16267.
233. N. Plattner and M. Meuwly, *Biophys. J.*, 2012, **102**, 333-341.
234. M. Lapelosa and C. F. Abrams, *J. Chem. Theory Comput.*, 2013, **9**, 1265-1271.
235. S. Mishra and M. Meuwly, *Biophys. J.*, 2009, **96**, 2105-2118.
236. S. Mishra and M. Meuwly, *Biophys. J.*, 2010, **99**, 3969-3978.
237. M. A. Scorciapino, A. Robertazzi, M. Casu, P. Ruggerone and M. Ceccarelli, *J. Am. Chem. Soc.*, 2009, **131**, 11825-11832.
238. M. D'Abramo, A. DiNola and A. Amadei, *J. Phys. Chem. B*, 2009, **113**, 16346–16353.
239. M. Anselmi, A. DiNola and A. Amadei, *J. Phys. Chem. B*, 2011, **115**, 2436–2446.
240. P.-h. Wang, R. B. Best and J. Blumberger, *J. Am. Chem. Soc.*, 2011, **133**, 3548-3556.
241. P.-h. Wang, R. B. Best and J. Blumberger, *Phys. Chem. Chem. Phys.*, 2011, **13**, 7708-7719.