Ligand and interfacial dynamics in a homodimeric hemoglobin

Prashant Kumar Gupta and Markus Meuwly

Ligand and interfacial dynamics in a homodimeric hemoglobin

Prashant Kumar Gupta[a)] and Markus Meuwly
Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland

The structural dynamics of dimeric hemoglobin (HbI) from *Scapharca inaequivalvis* in different ligand-binding states is studied from atomistic simulations on the μs time scale. The intermediates are between the fully ligand-bound (R) and ligand-free (T) states. Tertiary structural changes, such as rotation of the side chain of Phe97, breaking of the Lys96–heme salt bridge, and the Fe–Fe separation, are characterized and the water dynamics along the R-T transition is analyzed. All these properties for the intermediates are bracketed by those determined experimentally for the fully ligand-bound and ligand-free proteins, respectively. The dynamics of the two monomers is asymmetric on the 100 ns timescale. Several spontaneous rotations of the Phe97 side chain are observed which suggest a typical time scale of 50–100 ns for this process. Ligand migration pathways include regions between the B/G and C/G helices and, if observed, take place in the 100 ns time scale. © *2016 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License.* [http://dx.doi.org/10.1063/1.4940228]

I. INTRODUCTION

Multimeric proteins interacting with small ligands offer an opportunity to study cooperativity and its relevance to controlling chemical and biological processes.[1–4] The ligands inside the protein play important roles as biological messengers[5,6] and participate in chemical reactions.[7–12] Specifically, it is of interest to understand the dynamics of and at a protein-protein interface in view of the ligand dynamics within each of the subunits of a particular di- or multimeric protein. The eventual coupling between the two types of processes (within the monomers and at their interface) is potentially relevant for exercising control over ligand migration pathways and the binding kinetics in multimeric proteins. Because directly observing the real time spatial dynamics of ligands in proteins is experimentally very challenging,[13–16] atomistic simulations have become an attractive complementary technique for investigating such processes which is the approach pursued here.[17]

A system which is suitable to study such effects is the homodimeric hemoglobin HbI from *Scapharca inaequivalvis*. This protein has been previously characterized in structural,[18–21] thermodynamic,[22] kinetic,[2,23] and computational[24,25] terms. The accumulated information suggests that the functionally relevant dynamics is strongly influenced by structural changes at the interface and the number of water molecules between the two monomers. The most significant tertiary change between the deoxy (T, ligand unbound) and the fully ligand-bound (R) state is the orientation of the phenyl-sidechain of Phe97 in which the χ angle changes from 50° to 160° upon ligand binding. The conformational change is accompanied by a change in the degree of hydration of the interface in that 17 water molecules are present in the deoxy state and only 11

[a)]Present address: Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany.

water molecules are found for the ligand-bound protein. It has been proposed that the interfacial water molecules play a role in communication between the monomers.[24] However, phenyl rotation has never been directly observed experimentally.

The specific questions that arise for the homodimeric hemoglobin HbI concern the coupling between ligand migration pathways and their barriers on the one hand and structural changes and changes in hydration at the protein interface on the other hand. Specifically, it is of interest to understand how the heme-ligation-state in one monomer affects ligand migration in the second monomer and whether and how the dynamics at the interface transduces information. Atomistic simulations have repeatedly shown to be exquisitely useful to address such questions, both in qualitative and in quantitative terms. Characterization of the protein dynamics in atomistic detail cannot be easily obtained from experiment alone. For this, molecular dynamics (MD) simulations have provided useful insights and answered questions regarding protein dynamics and folding.[26–29]

For HbI, there are two major differences in the X-ray structure between the ligand-bound and ligand-free states: the orientation of the side-chain of Phe97 and the degree of hydration. Important amino acids are highlighted in Figure 1. Apart from this, the crystal structure of the deoxy state also shows water molecules in the distal site.[19] Further comparison of the two structures shows that the Lys96(monomer1)–propionate1(monomer2) and Lys96(monomer2)–propionate1(monomer1) salt bridges are replaced by Lys96(monomer1)–propionate2(monomer2) and Lys96(monomer2)–propionate2(monomer1) interactions, i.e., the Lys96 of one monomer switches interaction with the heme-propionates on the other monomer (see Figure 1). The breaking and formation of the salt-bridges acts as a gate for
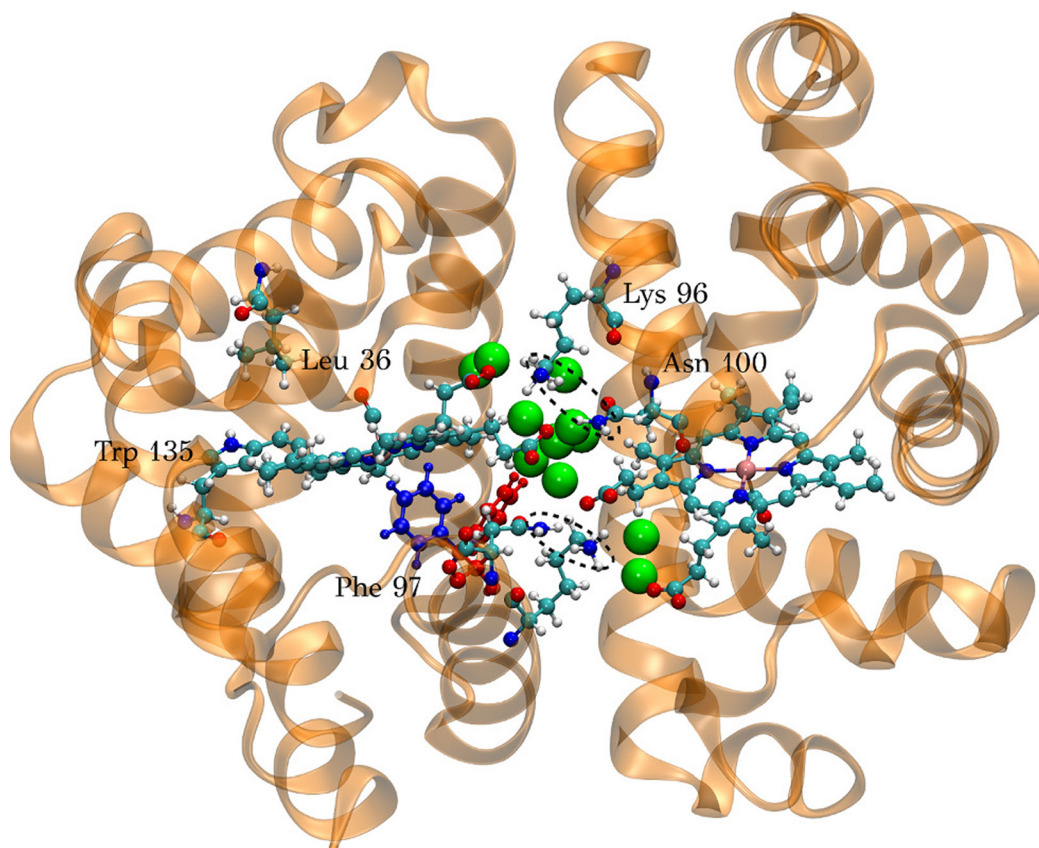

FIG. 1. Scapharca dimeric hemoglobin. The heme subunits, CO, Leu36, Trp135, and Phe97 residues are in CPK representation and the protein as orange cartoon. The conformation of Phe97 in the deoxy state is shown in blue and for the fully ligand-bound state in red with 11 interfacial water molecules (green spheres). The ligand-bound state exhibits the salt bridge between Lys96 of monomer1 and one of the heme-propionates of monomer2.

water molecules to move in and out of the interface and potentially affects additional structural changes at the interface.

The present work investigates the ligand and water dynamics of ligand-bound, ligand-free, and relevant intermediate (ligand-bound, ligand-unbound, and ligand-free monomers in different combinations) homodimeric hemoglobins HbI. The intermediates include partially ligated and unligated (but ligand-containing) monomers which are all physiologically relevant along the pathway between the ligand-free and ligand-bound states. Here, the particular focus is on the coupling of ligand and protein degrees of freedom and on the interplay between protein motion and water dynamics. First, results on the link between protein, ligand, and water dynamics are discussed. Then, ligand migration is considered in more detail.

II. MATERIALS AND METHODS

For the MD simulations, crystal structures 2GRH[30] and 2GRF[30] from the PDB database are used for oxygenated and deoxygenated HbI structures, respectively. Both structures are M37V mutants of wild–type *Scapharca Inequivalis* dimeric hemoglobin.[30] Experimentally, the mutation was introduced to enlarge the distal pocket which was found to lower geminate rebinding of oxygen in solution. At the same time, cooperativity in ligand binding is maintained and proceeds along structural transitions as in wild–type HbI. Simulations were started from the 1b2b X-ray structure which is nearly symmetrical for the two monomers; their root mean square deviation (RMSD) is 0.4 and 0.8 Å for backbone atoms and all heavy atoms, respectively. In order to assess the influence of initial symmetry in the dimer, an explicitly symmetrized C_2 structure for 1b2b was also set up, fully solvated and treated as described further below. The structure of the protein is shown in Figure 1.

HbI contains dimeric in which each monomer contains one heme group. Each monomer can be in a ligand-bound (b) or ligand-unbound (u, unbound; e, empty) state individually. The spectroscopic nomenclature labels the photodissociated, ligand-unbound state with an asterisk. Hence, 1u2u would correspond to 1(CO*)2(CO*). Experimentally, structures with both ligands bound (1b2b) or both ligands fully removed (1e2e) are known. Between those, several intermediate occupation and ligation states are possible: two mixed states with bound and unbound ligands (1b2u and 1u2b), both CO ligands unbound (1u2u), two mixed states with one ligand-bound and one ligand-free monomer (1b2e and 1e2b) and finally two states in which one monomer is ligand-free and the other one is ligand-unbound (1e2u and 1u2e). The intermediate states are prepared starting from the (1b2b, 2GRH) state by removing the bond between the heme-iron atom and the CO-carbon atom and leaving the CO either in the distal pocket (u state) or removing it entirely from the structure (e state). In both cases, the heme-unit is treated with a force field that correctly captures the domed, five-coordinated structure.[31,32] No structural information about the mixed states (binding intermediates) is available from previous experiments.

All the above systems were prepared using CHARMM,[33] and MD simulations were carried out using NAMD[34] with the CHARMM22 force field.[35] Each system was neutralized by adding counterions (16 Na^+ and 18 Cl^- ions) at physiological concentration (0.154 mol/l) and solvated in a cubic box of TIP3P water.[36] The fully solvated systems contains 46 216 atoms. The simulations consisted of relaxing the starting structures by conjugate gradient minimization, heating each system to 300 K for 1.2 ns, followed by 1 ns of equilibration at constant temperature and pressure (*NPT* ensemble) and finally several 100 ns of production in the *NPT* ensemble. The SHAKE algorithm[37] was used to constrain all rapid motions of hydrogen atoms. The nonbonded interactions were truncated at 16 Å, using a switching function starting from 12 Å. The nonbonded list was calculated for pairs within 16 Å. After running MD simulations for 200 ns of each systems, i.e., (1b2b, 1b2u, 1u2b, 1u2u, 1u2e, 1e2u, and 1e2e) in NAMD various structural and dynamics properties were analyzed.

To calculate the number of water molecules at the dimeric interface, a spherical region of radius 10.4 Å around the center of mass of the two heme subunits was considered.[24] Umbrella sampling was performed for estimating the energy barriers for CO migrating to internal Xe cavities and used a force constant of 5 kcal/mol/$Å^2$, with a typical step-size of 0.3 Å and the Fe-

CO distance as the reaction coordinate.[38] The step size is chosen such as to have between 25 and 35 overlapping windows per reaction coordinate. The simulation time for each window is 51 ps, 1 ps of relaxation and 50 ps for production. Overall, a total of 45 ns of simulation data were analyzed. The data were analyzed with the weighted histogram analysis method (WHAM).[39,40]

III. RESULTS

As a first validation, equilibrium simulations were carried out for the fully bound (1b2b) and unbound (deoxy-HbI, 1e2e) systems starting from their respective X-ray structures.[30] The structural RMSDs for the two dimers relative to their X-ray structures 2GRH (CO-bound) and 2GRF (deoxy) from a 25 ns simulation are 1.65 Å and 1.20 Å, respectively. This compares and agrees with RMSDs of 1.5 Å–2.0 Å for the ligand-bound and ligand-free monomers in a mixed dimer as investigated for 6 ns previously.[24] Although the average structures are quite well described, some of the distinguishing features of the ligand-bound and deoxy structures are only qualitatively reproduced, see Figure 10. The averaged (over 25 ns) Fe–Fe distance between the two heme-subunits for the 1b2b state is 18.5 Å (compared to 18.2 Å from experiment and 19.7 Å from previous simulations, averaged over 1 ns (Ref. 24)) whereas for the 1e2e (deoxy) state this decreases to 18.1 Å (compared to 16.6 Å from experiment and 17.2 Å from previous simulations, averaged over 1 ns (Ref. 24)). Similarly, the number of water molecules between the 1b2b and 1e2e states differs by 2 compared to a difference of 6 and 8 from experiment and previous simulations.

For the mixed dimers, the simulations were started from the 1b2b structure. Hence, there is always an initial relaxation from the starting structure towards the particular state considered which was found to take usually 25 ns. After that time, the RMSD levels out at 1.5 Å for all backbone atoms and 2.0 Å for all heavy atoms. Over the subsequent 175 ns the RMSD is stable. The stability of the intermediates found in the present work is consistent with the same finding of the mixed 1e2b (1d2o in the nomenclature of Ref. 24) and 1b2e (1o2d) dimers although the time scales of the two simulations differ by 2 orders of magnitude.

To directly assess the dimer-symmetry on the nanosecond time scale, one simulation was started from an explicitly symmetrized structure with initial RMSD = 0 between the two monomers. The dimer was again solvated in TIP3P water and heated, equilibrated, followed by 15 ns of free dynamics. Already after 1 ns of heating simulation the RMSD between the two monomers has increased to 1 Å. During the subsequent equilibration and free dynamics, the RMSD fluctuates around this value. Hence, on the nanosecond time scale, the exact C_2 structure is lost. This is the behaviour expected in NMR experiments. However, in a crystal, symmetry may either be preserved better or a slight asymmetry may be reinforced through crystal contacts as was, for example, found for the wild type (WT) homodimer for which substantial differences in ligand migration were observed for the two monomers.[25] The RMSDs between the two monomers in the X-ray structures of the M37V mutants are 0.8 and 0.4 Å (2GRH), 0.77 and 0.28 Å (2GRF) and for the WT they are 0.93 and 0.55 Å (3G46) for all heavy atoms and all backbone atoms, respectively. Hence, even in the crystals, the two monomers are not exactly symmetric, neither for the M37V mutant nor for the WT protein. This is also supported by the previous simulations which find an RMSD of 1.2 Å for main-chain atoms and 1.8 Å for heavy atoms between the two monomers.[24]

As explicit symmetry is lost on the nanosecond time scale, preparation of the various intermediates (see above) from one common structure (1b2b) with 2 ns of heating and equilibration is expected to lead to largely independent systems.

A. Interfacial water dynamics

Time series which describe the equilibrium dynamics of the 1b2u intermediate (i.e., close to the initial 1b2b) are reported in Figure 2. As the system is prepared from the fully ligand-bound structure, at early times the dynamics is that of the 1b2b state. This can be, for example, seen in the location of the CO ligand (Figure 2(a)). After less than 10 ns, the unbound CO
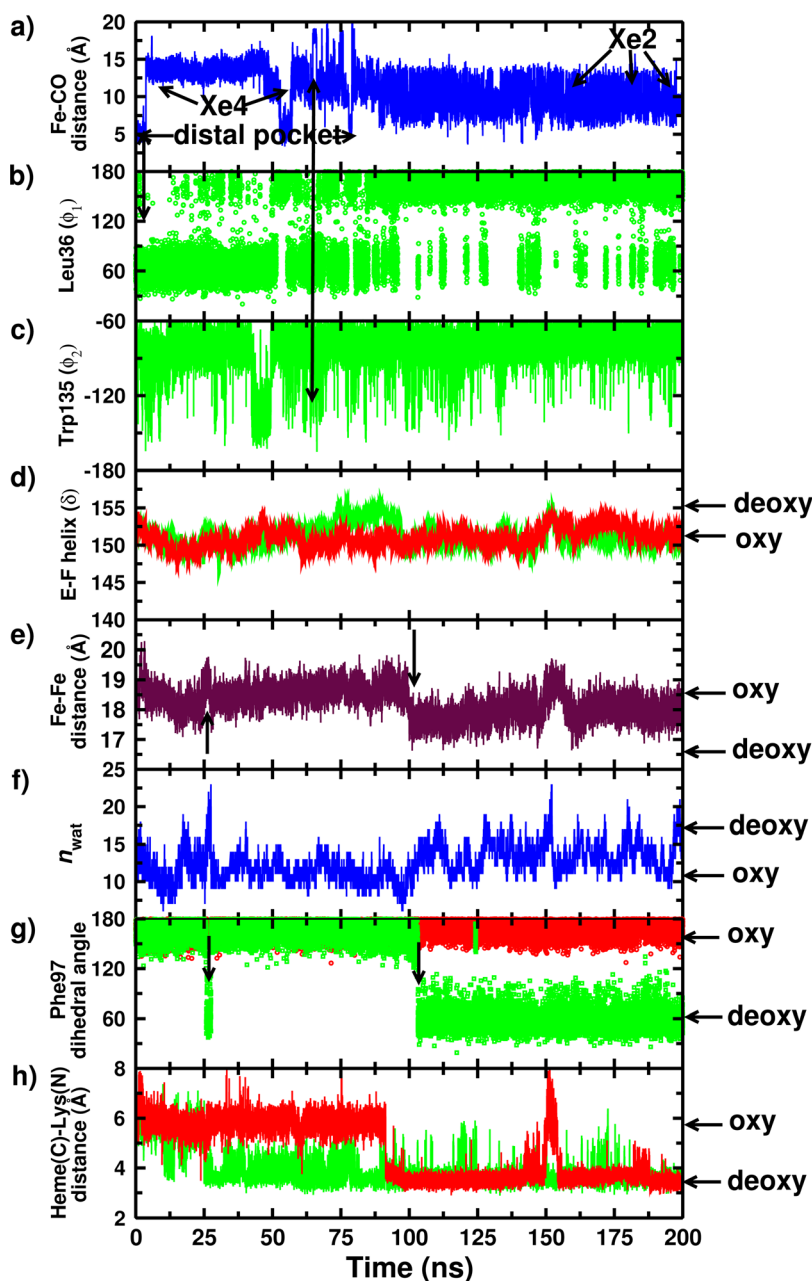
FIG. 2. A comprehensive comparison of different structural determinants for the 1b2u intermediate from a 200 ns simulation. (a) The CO position relative to the heme-Fe, (b) the Leu36 side chain dihedral angle for 2u, (c) the Trp135 side chain dihedral angle for 2u, (d) the angle between helices E and F (running average over 25 ps) in monomer 1b (red) and 2u (green), (e) the Fe–Fe distance, (f) n_{wat} at the interface, (g) Phe97 side chain dihedral angle χ for 1b (red) and 2u (green); (h) salt bridges between Lys96 and the heme propionate for 1b (red) and 2u (green).

ligand in monomer 2 starts to migrate, see Figure 2(a), as the Fe-CO bond in this monomer was removed at $t = 0$. The localization of the unbound CO ligand in the major pockets as a function of simulation time is also indicated. Unambiguous assignment of the CO ligand to a particular pocket requires more than a radial coordinate. Displaying such additional coordinates was, however, omitted for clarity. Instead, assignments were made based on information as that presented in Figure 3. After about 100 ns, several changes in relevant protein degrees of freedom occur, see Figures 2(e)–2(h). Closer inspection of the time series reveals that first the
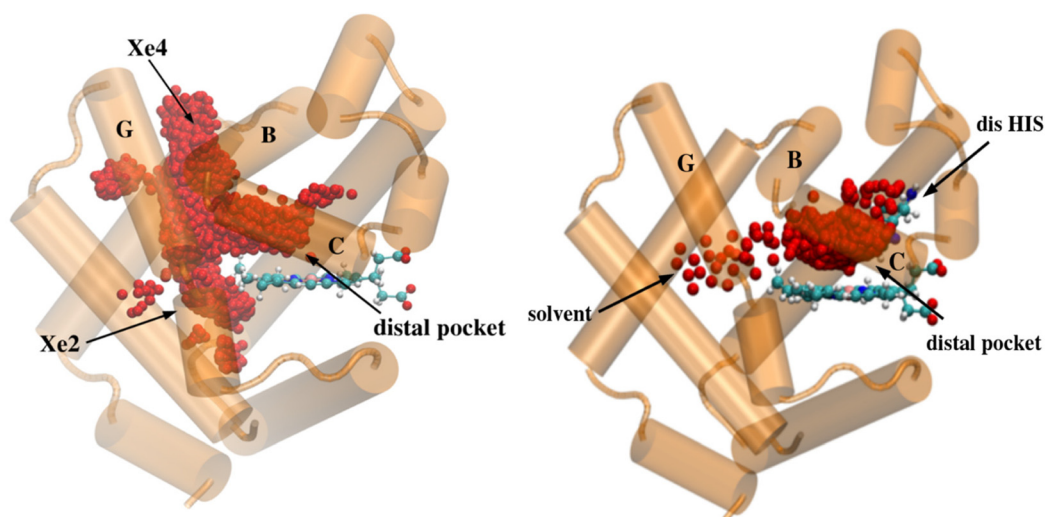
FIG. 3. Ligand migration pathway in monomer 2 of the 1b2u (left) and 1u2u (right) intermediates. Photodissociated CO (red sphere) starts from the distal pocket and diffuses into Xe2 and Xe4 (left panel), and Photodissociated CO mainly remains in distal pocket and then migrates into the solvent through helix G and C (right panel). The heme group is represented in CPK. The temporal evolution of the CO center of mass coordinates is reported in Figures 2(a) and 4(f). "dis HIS" labels the distal Histidine.

heme-propionate–Lys96 salt bridge breaks (at 92 ns) after which (at 100 ns) the Fe-Fe distance decreases by ≈ 1 Å approaching the value characteristic for the deoxy state. Over the next 10 ns, the number of water molecules changes from the value characteristic for the ligand-bound to that of the ligand-free state and the conformational transition of the Phe97 side chain occurs at 105 ns. Hence, several degrees of freedom are intimately linked on the nanosecond time scale. It is possible that water influx occurs in two stages, first following the breaking of the salt bridge and then following the conformational transition in Phe97. However, such a causal interpretation—Phe97 following salt bridge breaking—will require the observation of more such correlated events.

The analysis so far suggests that a conformational repositioning requires 10–15 ns and that breaking of the salt bridge triggers a change in the Fe-Fe distance and water occupation at the interface with reorientation of Phe97 occurring at later stages. This is further corroborated by a short period during which the salt bridge is reformed (Figure 2(h) at 150 ns) which results in almost immediate increase of the Fe-Fe distance and a burst of water molecules but no apparent change in the Phe97 conformation. On the other hand, at around 30 ns a reorientation of Phe97 leads to immediate increase in n_{wat} and the Fe–Fe separation which points towards direct involvement of Phe97 in the interfacial dynamics. Hence both, breaking the heme-propionate–Lys96 salt bridge and reorientation of the Phe97 side chain, lead to water influx.

The data in Figure 2 illustrate how conformational changes in the Phe97 dihedral angle lead to several concomitant changes at the dimerization interface such as the inclusion of water molecules at the interface. Repositioning of the Phe97 side chain is one of the tertiary structural changes between the ligand-bound and the deoxy-structure known from crystallography.[41] The Fe-Fe distance decreases sharply from 18.5 Å to 17.5 Å (see Figure 2(e)) which is half the value found in the X-ray structure which are 18.2 Å and 16.6 Å in the fully ligand-bound and deoxy states, respectively.[30]

Distribution functions for the number of water molecules n_{wat}, the Fe-Fe distance, and the dihedral angle of the side chain of Phe97 suggest that all intermediates considered (1b2u, 1u2b, 1e2u, 1u2e, 1u2u, 1b2b, and 1e2e) are structurally somewhere between ligand-bound and deoxy-HbI. However, the individual simulation time of 200 ns is still too short to observe full relaxation. For example, within the time scale of the present simulations, we have not observed any change in the rotation of one monomer relative to the other which has been found to be

3.3° based on static R and T X-ray structures.[42] The absence of more global changes is consistent with the experimental finding[30] that this rotation is observed on time scales of $80\,\mu$s in a recent time-resolved X-ray crystallography study.[43]

B. Ligand migration

1. Free dynamics

Monitoring ligand migration provides information about the architecture of proteins. First, the pathway of CO diffusion through the protein matrix after breaking the Fe-CO bond is followed. As shown in Figure 3 (left panel), the free CO molecule extensively samples monomer 2 in the 1b2u intermediate. The temporal evolution for this is illustrated in Figure 2(a) which highlights the repositioning of the ligand on the 200 ns time scale. Specifically, the ligand samples the distal pocket and pockets Xe2 and Xe4. In this particular case, no ligand escape from the protein is found.

Contrary to that, simulations for other intermediates exhibit ligand escape as shown in Figure 4. Trp135, which is one of the residues forming the Xe2 cavity, is usually involved when transitions into or out of this pocket occur, as can be seen in panels (a)–(c) and (e) of Figure 4. The change in the dihedral angle ϕ_2 of Trp135 from -60° to -160° creates a passageway for CO to move into or out of the Xe2 cavity. The 1u2u intermediate which has both CO molecules unbound but in the active site at the beginning of the simulation shows asymmetry in the ligand diffusion. In monomer 1 (1u) the CO ligand remains in the protein throughout the simulation whereas in monomer 2 (2u) it escapes after about 125 ns. In 1u, the ligand samples various internal states as indicated in Figure 4(e). Contrary to that, before escaping to the solvent, CO in 2u only samples the distal site extensively as is also reported in Figure 3. For this system, the initial RMSD for all heavy atoms is 0.8 Å and increases to 1.4 Å averaged over 25 ns. Hence, average structural differences of this magnitude lead to asymmetric dynamics in the two monomers.

Figure 5 illustrates one CO displacement between the distal site and Xe4 together with repositionings of the relevant cavity-forming residues. It is found that Leu36 adopts a cavity-forming conformation (Figures 5(b) and 5(e)) during which passage of the CO ligand is possible. Hence, Leu36 is involved in controlling ligand migration between the distal site and Xe4. Contrary to that, residues Leu77, Ile114, and Ile118 remain almost invariant. Leu36 is generally quite flexible (see Figure 4) and adopts two well-defined conformations ($\phi_1 \approx 60^{\circ}$ and $\phi_1 \approx 150^{\circ}$) with rare excursions to $\phi_1 \approx -60^{\circ}$ whereas Trp135 covers a range between -60° and -150° in a more diffuse manner without adopting particularly well-defined orientations. The two alternate conformations for Leu36 were also reported in a 1.08 Å resolution room temperature static crystal structure of HbI-CO, as mentioned in Ref. 25.

2. Free energies

To understand the energy landscape of CO migration in different Xe cavities, the free energy for migration was determined from umbrella sampling simulations. This was done for the 1b2u and 1u2b intermediates and for migration between the distal pocket (B state) and Xe4 and for the B-to-Xe2 transition (only for 1u2b), as shown in Figure 6. These two occupation states were chosen in order to probe how symmetric the homodimer behaves during the dynamics. The migration barriers are of the order of a few kcal/mol. In order to establish how representative single umbrella sampling simulations along a predefined (but in this case well-defined) reaction coordinate are, migration between the B-state and Xe4 in 1b2u was repeated four times, starting from different initial structures taken from the equilibrium simulations (Figure 6(b)). Overall, the results are very similar in that the B-state is energetically unfavoured whereas the Xe4 pocket is a clear minimum. Contrary to that, in 1u2b, the configuration with CO in the B-state is energetically favoured compared to that in the Xe4 pocket. This further suggests that despite starting the simulations from one common, nearly symmetrical structure (1b2b) with a RMSD between the two monomers of 0.4 and 0.8 Å for backbone
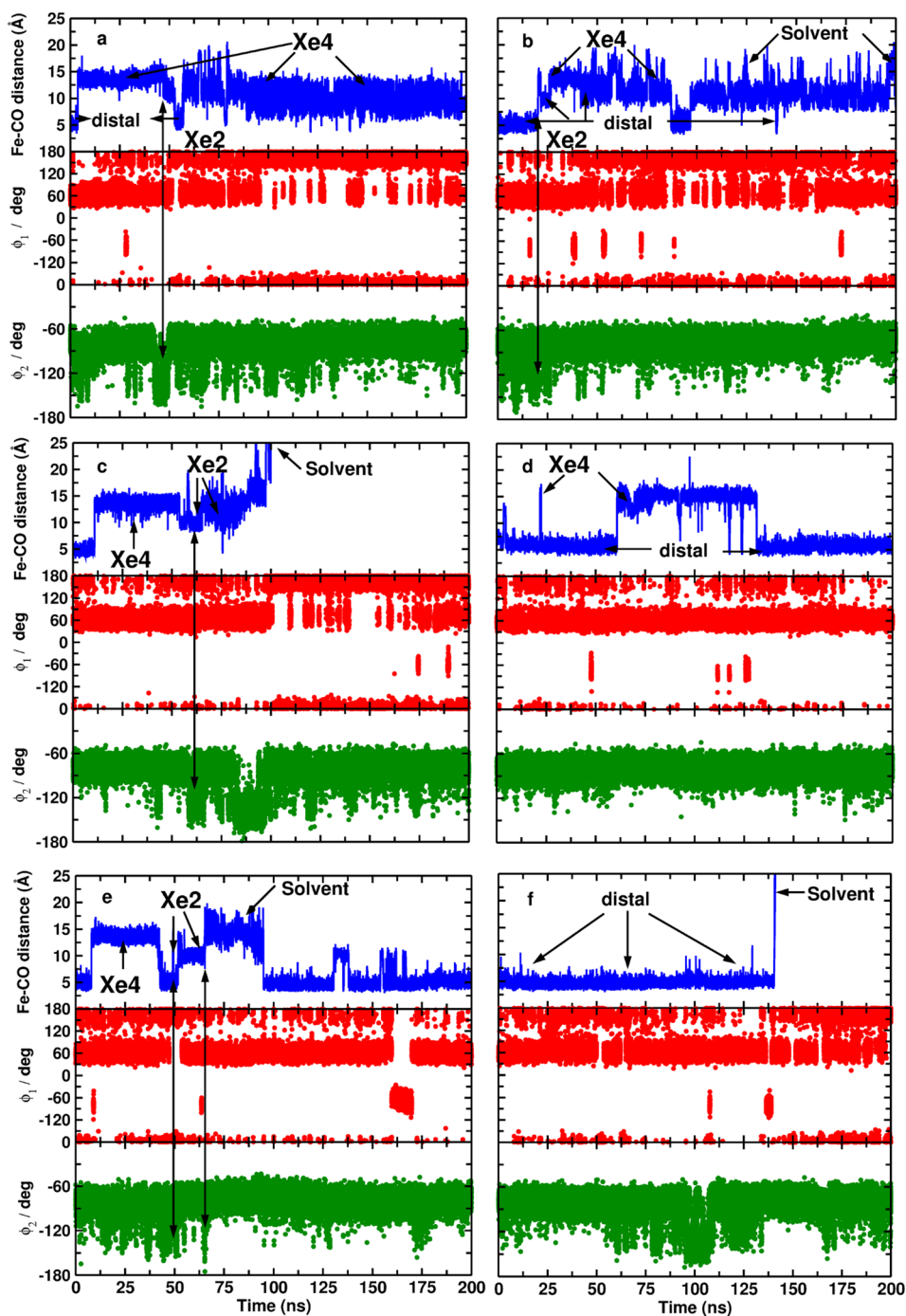
FIG. 4. Distance between Fe and CO (blue trace), the dihedral angle of the Leu36 side chain (red), and the dihedral angle of the Trp135 side chain (green) as a function of simulation time. The data shown here are for 5 different systems. (a) 1b2u, (b) 1u2b, (c) 1e2u, (d) 1u2e, (e) 1u2u monomer 1 and (f) 1u2u monomer 2. For the Fe-CO distance, the ligand localization pocket is indicated. In panel (b), the ligand attempts to escape to the solvent after 125 ns but only makes this transition towards the end of this simulation and in panel (e) the attempt after 100 ns does not lead to actual escape.

atoms and all atoms except hydrogens, the dynamics of the homodimer is asymmetric on the 100 ns time scale.

The observed asymmetry for ligand migration in the free energy simulations also agrees with the findings reported in Figures 4(e) and 4(f), where CO diffusion in the two monomers
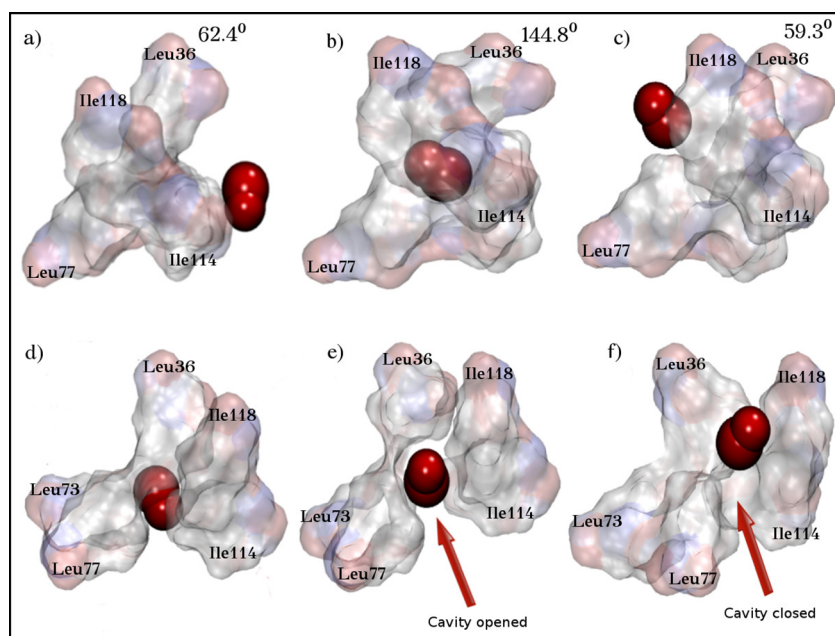
FIG. 5. CO migration in 1b2u from the distal site to the Xe4 pocket and simultaneous change in the side chain dihedral of Leu36. Surface representation of Leu36, Ile114, Ile118, Leu77, and Leu73 residues and changes of the dihedral angle of Leu36 is shown in (a)–(c). Panels (a)–(c) show a side view of the residues forming the gate between protein internal cavities and distal pocket, and (d)–(f) show a front view before, during and after the migration of CO from the distal pocket. Panel (b) emphasizes the change in Leu36 dihedral angle compared to the initial state as shown in (a) and (c). The CO molecule is shown in CPK representation and in red. All residues around the heme pocket are labeled.

for 1u2u is described. After photodissociation, the ligand in monomer 1 samples the distal site, Xe2 and Xe4 but remains in the protein for the entire 200 ns of simulation time. Conversely, in monomer 2 the ligand samples only the distal site and escapes to the solvent. The barrier heights from umbrella sampling simulations support migration time scales in the 1–10 ns time scale as seen from the unbiased simulations in Figure 4. Consistent with both observations—asymmetric motion of CO through the protein matrix and motion of the photodissociated ligands on the nanosecond time scale—are previous time-resolved crystallography experiments on the WT protein which found substantially different photolysis levels in the two monomers already at 5 ns after photodissociation.[25]
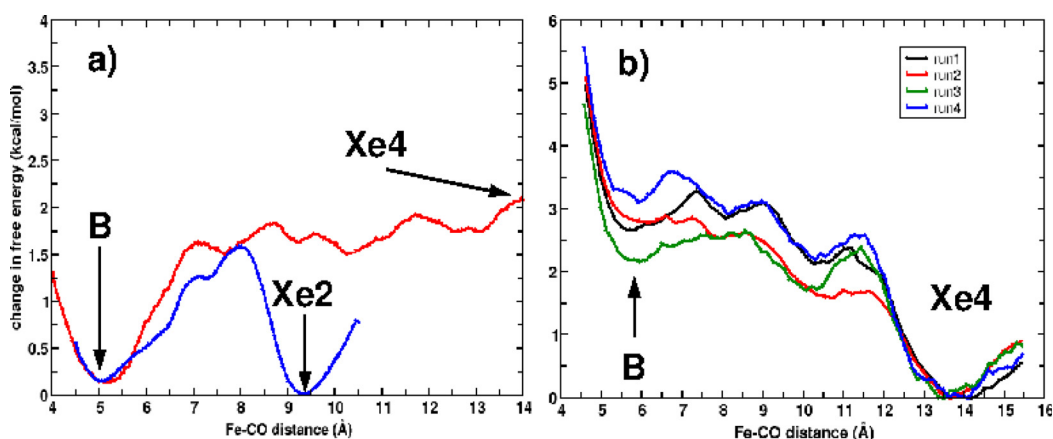

FIG. 6. Free energy profile for CO migration to the Xe2 and Xe4 pockets in 1u2b and 1b2u. (a) Free energy profile for migration of CO from distal pocket to Xe2 and Xe4 in 1u2b. (b) Free energy profile for migration of CO to Xe4 in 1b2u. Different colors in panel (b) represent different starting configurations for the umbrella sampling simulations.

3. Water diffusion

The presence of single water molecules in the distal pocket of the deoxy X-ray structure[30,41] (Figure 1) suggests that water molecules can enter the protein either around Phe97 or through the distal histidine (His69) gate at the dimeric interface. The distal pathway is indeed observed for 1b2u as shown in Figure 7(a) which is also supported by optical experiments.[21] Similar behavior for water diffusion to the distal pocket is also observed for 1u2b, see Figure 7(b). The present simulations find that water from the solvent diffuses into the protein through the BG pathway (Figure 8) and to the distal pocket in the presence of CO inside the protein which points towards the possibility of multiple ligand occupancy inside the protein.

C. Global structural changes

In order to correlate the migration/localization of CO within the protein and the dynamics/ fluctuation of the protein residues, the averaged residue-RMSD in different windows of the umbrella sampling simulations was determined. For this, 4 different windows were considered. The average flexibility of the residues for different locations of the CO ligand within the protein is shown in Figure 9. It is found that the RMSD of individual residues varies considerably depending on the location of the CO ligand inside the protein. Figure 9(a) shows that helices E and F have a low mobility when CO is in the distal pocket compared to other helices. However, as CO migrates away from the distal pocket, helices E and F become more dynamical. Interestingly, helix B to which Leu36 residue belongs and which acts as a gate for the distal pocket has an increased RMSD compared to the initial stage which is an indicative of the influence of CO movement on the individual residues (Figure 9(b)). Helices A, B, and G are involved in formation of the Xe4 pocket and their RMSD values after CO migration away from the distal site increase compared to the situation when CO is in the Xe4 pocket (Figures 9(c) and 9(d)). From this analysis, it is found that the protein motion is correlated with the
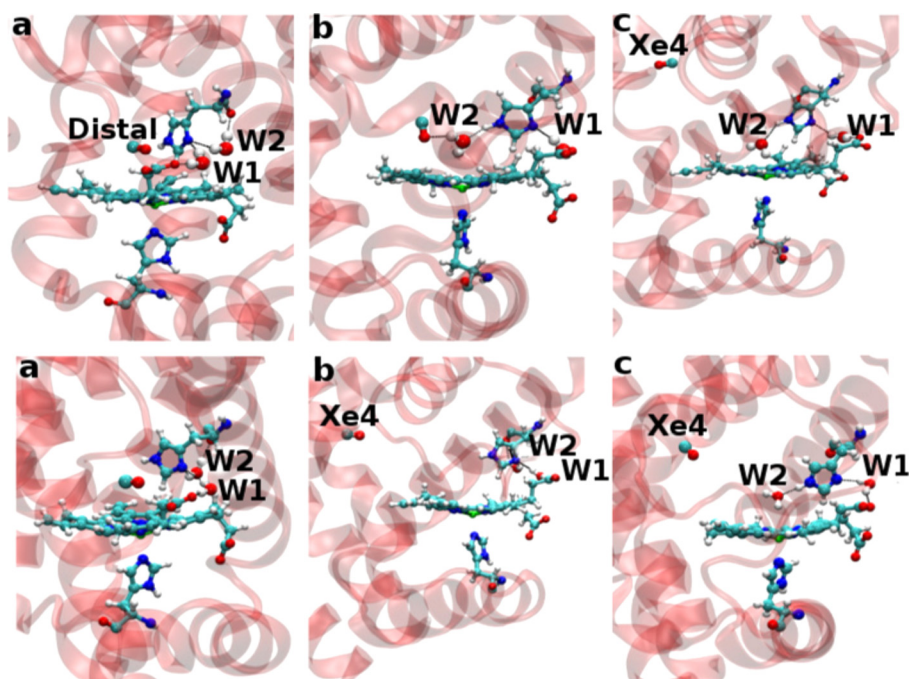

FIG. 7. Snapshots for CO migration to the Xe4 pocket. W1 and W2 are the crystallographic water molecule close to His69. W2 forms a hydrogen bond with His69 and W1 forms a hydrogen bond with His69 and the heme-propionate group. W1 molecule does not migrate during the whole course of the simulation. Upper row for 1b2u: (a) CO and W1 and W2 at their X-ray positions, (b) W2 and CO both in the distal pocket, and (c) water in the distal pocket and CO migrates to Xe4. Lower row for 1u2b: (a) CO and W1 and W2 at their X-ray positions, (b) W2 H-bonded to W1 and CO in Xe4, and (c) W2 in the distal pocket and CO in Xe4.
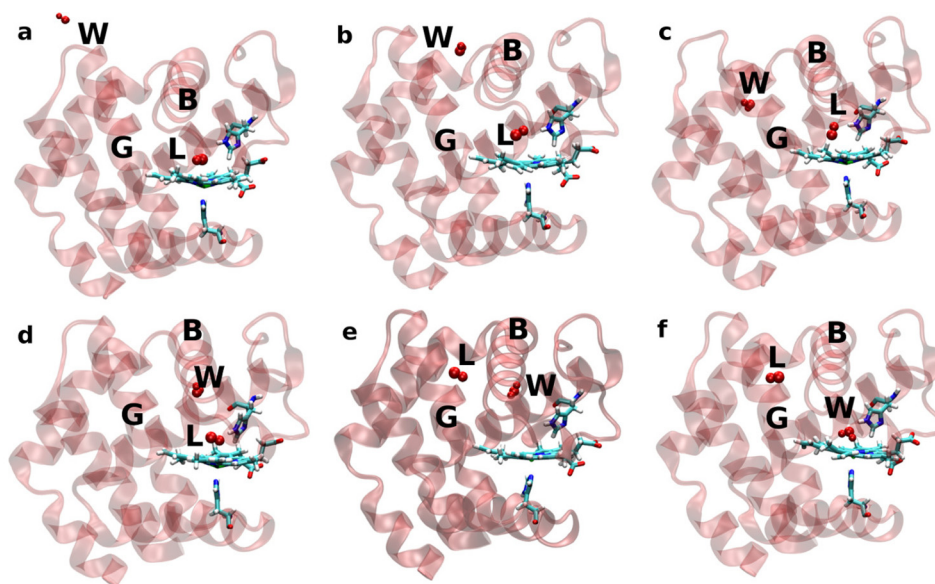
FIG. 8. Snapshots for the migration of a single water molecule (W) from the solvent to the internal cavities through the BG pathway (helices labelled) after the migration of CO (L) to the Xe4 pocket. (a) water outside the protein, (b) water close to the B and G helix, (c) water occupying the Xe4 pocket, (d) water below helix B, (e) migration of CO from the distal pocket, and (f) CO in the Xe4 pocket and water occupying the distal pocket.

positioning of the ligand inside the protein or, in other words, ligand and protein degrees of freedom are coupled.

The importance of fluctuations among conformational substates for ligand diffusion through the protein matrix has been suggested before based on experimental observations and simulations.[44,45] Explicit coupling between ligand and protein motion has been found in the previous
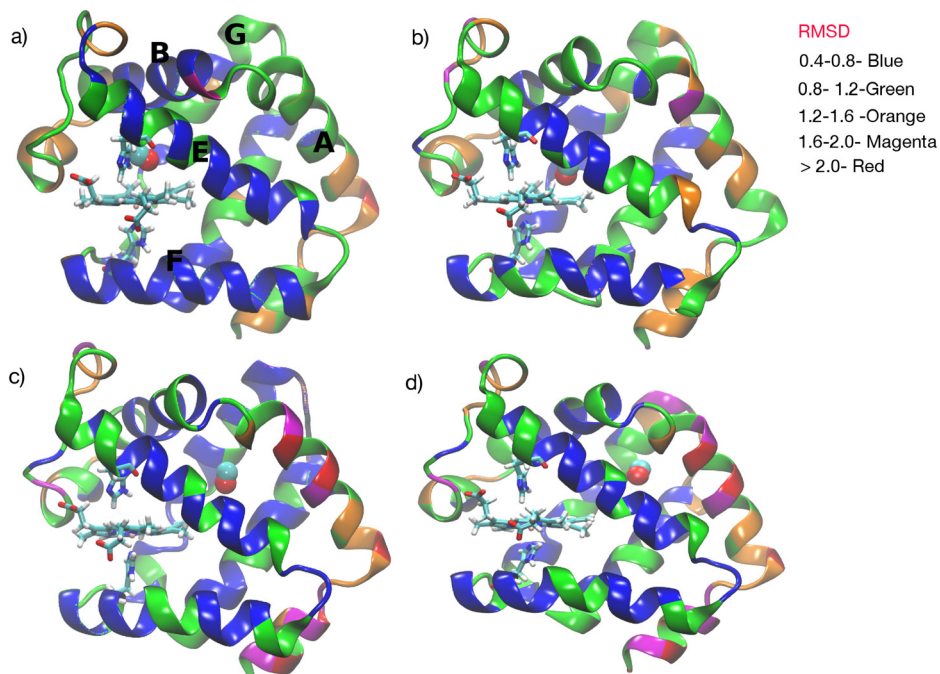

FIG. 9. RMSD value of individual residues of protein in different umbrella windows which depends on the location of CO inside the protein. Color codes: blue ($0.4\,\text{Å} \leq \text{RMSD} < 0.8\,\text{Å}$); green ($0.8\,\text{Å} \leq \text{RMSD} < 1.2\,\text{Å}$); orange ($1.2\,\text{Å} \leq \text{RMSD} < 1.6\,\text{Å}$); magenta ($1.6\,\text{Å} \leq \text{RMSD} < 2.0\,\text{Å}$); and red ($\text{RMSD} \geq 2.0\,\text{Å}$).

work on CO migration in myoglobin (Mb).[46] The mutual influence of protein and ligand degrees of freedom is further highlighted by the fact that for O_2 diffusion in truncated hemoglobin (trHbN) a Markov-state-model built on a combined analysis of protein structural changes together with the O_2 positions faithfully reproduces the underlying kinetics.[47] Truncated Hb exhibits Xenon-pockets similar to Mb and HbI, and analysis of the ligand degrees of freedom only in the Markov model is unable to capture the ligand dynamics in the protein as found in the molecular dynamics.[48] Such coupling is absent in implicit ligand sampling which neglects the influence of the ligand on the protein motion and vice versa and may affect ligand migration barriers derived from it.[49]

IV. DISCUSSION

The present simulations show a cascade of events for intermediates between the fully ligand-bound and ligand-free states. Of particular relevance are changes in the Fe-Fe distance, the number of water molecules at the dimerization interface, the conformational change in the dihedral angle of Phe97, and the salt bridge involving Lys96 and the heme-propionate. This data is summarized in Figure 10 for all systems studied with an accumulated simulation time in excess of 1 μs. The illustration attempts to provide an overview of the changes observed in the simulations by placing them in a logical, but necessarily non-unique or time-disordered fashion. Hence, the x-axis does not show a total simulation time but rather reports the 200 ns intervals which were run for each individual state. Also, conformationally averaged Fe–Fe separations and values for n_{wat} are reported in the panels.

At the far left both CO molecules are bound (1b2b) to the heme whereas at the far right both CO molecules have been removed from the system (1e2e). Starting at 1b2b, the first change is to break the Fe-CO bond in one monomer. The choice of placing 1b2u next to 1b2b
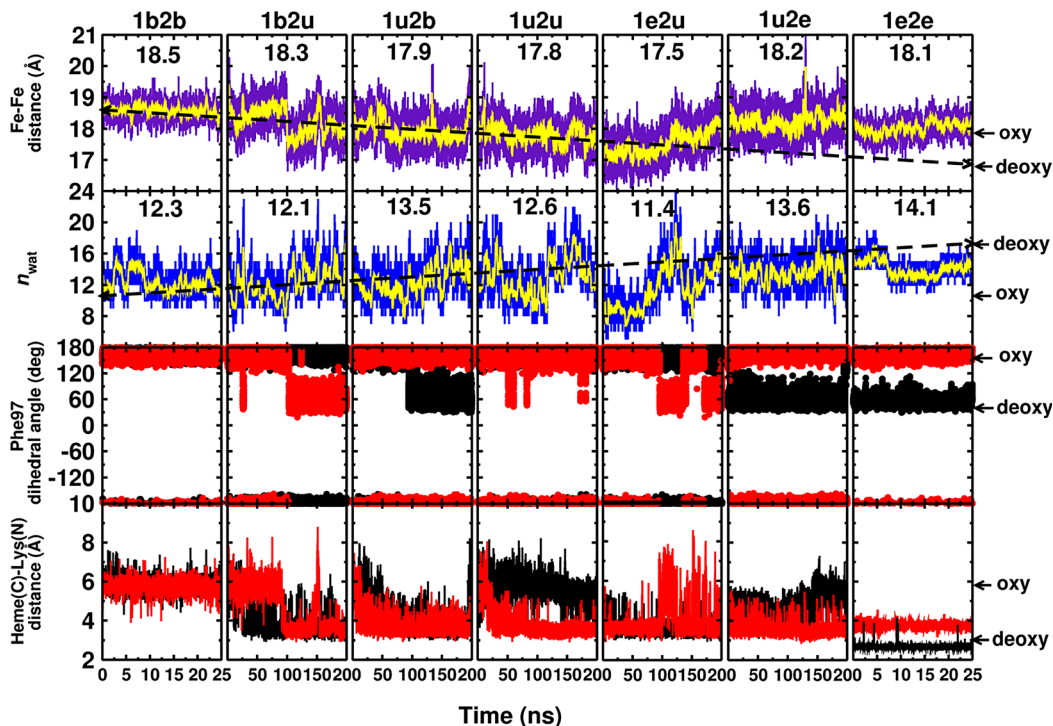

FIG. 10. Data combined for all 7 studied systems. First row: Fe–Fe distance as a function of simulation time. Second row: Number of water molecules at the dimeric interface. Third row: conformational change of the dihedral angle of Phe97 for each monomer unit (black and red for monomers 1 and 2). Fourth row: Lys96(monomer 1)–heme-propionate(monomer 2) (black) and Lys96(monomer 2)–heme-propionate(monomer 1) (red) distances. The experimentally observed values from the crystal structures are marked with an arrow and conformationally averaged Fe–Fe distances and values for n_{wat} are given in the top two rows. For additional explanations see text.

is arbitrary, though. Next, the bond to both CO ligands is broken, which yields 1u2u. The same remark concerning the ordering applies to the panels to the right of 1u2u.

On the right hand side, the limiting values for each of the variables in the 1b2b and 1e2e states from the X-ray structures[41] are indicated. On average, the Fe–Fe distance, the number of water molecules n_{wat}, the Phe97 dihedral angle, and the length of the salt bridge for the various systems are bracketed by the limits given by experiment. Each system displays potentially rich dynamics between these limits, such as 1e2u. For example, this intermediate samples a "dry" state ($n_{wat} = 8$) with an exceptionally small number of water molecules at the interface at the beginning of the simulation, followed by an increase in the degree of hydration to $n_{wat} = 20$ during the following 100 ns. This change is accompanied by a gradual increase in the Fe–Fe separation. The observation of such a "dry" state has also been reported in earlier simulations of a 1b2e intermediate.[24]

The T state (1e2e) is not fully stabilized in the present simulations. This is reminiscent of the situation for tetrameric Hb for which the same was observed.[50] While for the salt bridge (bottom row in Figure 10) the deoxy structure is maintained, only one Phe97 is in the orientation found in the X-ray structure. The degree of hydration is more towards the deoxy state but the number of water molecules at the interface $n_{wat} = 14$ differs from the experimentally observed value which is $n_{wat} = 17$. The same applies to the Fe–Fe distance which is smaller for 1e2e than for 1b2b. This is in accord with the experimental finding of a decreasing distance upon removal of the ligand but is too long by 1 Å in absolute terms. Hence, the qualitative changes between the R- and T-states can be reproduced by the simulations but not necessarily all quantitative aspects.

In the present work, the focus was on on-pathway intermediates between fully ligand-bound and ligand-free HbI. A concise summary of the findings is as follows:

- *Asymmetry:* The dynamics of the homodimeric protein HbI is asymmetric *on the 100 ns time scale* in view of the ligand migration pathways and the energetic preference of the different states as suggested in Figures 4 and 6. Comparing the first migration time between 1u2b and 1b2u from our simulation shows asymmetry in the ligand migration time as was also found from time-resolved crystallography.[25] The difference in the occupation state of CO after photo-dissociation (more probable for Xe4 in 1b2u compared to Xe2 in 1u2b, not shown) in internal Xe cavities manifests an inherent asymmetry between the two monomers regardless of the global symmetry of the entire protein on the time scale of the present simulations. Similar observations of structural asymmetry were reported before.[18,24,25]

- *Conformational transitions:* The Phe97 rotation between the two known orientations has been explicitly observed and occurs in the 50–100 ns time scale (about 15 events during 1 μs). This transition is followed by water influx at the interface and migration of CO in the Xe cavities and suggests coupling of these degrees of freedom. Also, the Fe–Fe separation responds to changes in the orientation of the Phe97 side chain. Furthermore, the salt bridge between Heme and Lys96 influences the Fe–Fe distance. The two Leu36 orientations known from crystallography and relevant for ligand migration between the distal site and the Xe4 pocket interconvert on the time scale of the simulations.

- *Water-influx:* Breaking the heme-propionate–Lys96 salt bridge and reorientation of the Phe97 side chain lead to water influx. This occurs primarily through the distal pathway.

- *Ligand migration:* The present unbiased simulations find ligand escape from the protein on the ≈100 ns time scale, see Figure 4(c). This compares with a time scale of 10 ns from recent biased (locally enhanced sampling) simulations.[51] These simulations investigated O_2 diffusion in WT and a number of HbI mutants and found the following ligand-escape routes: between the B and G helices (5 cases, also observed here), between the E and F helices (4), between the B and E helices (2), and between the C and G helices (1, also observed here). However, in these locally enhanced sampling simulations, the histidine gate was not found to be an important ligand escape route,[51] at variance with previous spectroscopic work.[21]

- *Coupling:* The present simulations support coupling of ligand- and protein-motion as suggested by recent NMR experiments.[52] This is also consistent with findings for Mb and trHbN.[46,47]

In two of the intermediates (1b2u and 1u2u), CO migration to the solvent through Xe4 was found. Previous work on the wild type protein found that Xe4 is not on the major exit pathway. This was concluded indirectly from the observation that the geminate CO rebinding fraction did not increase upon blocking the Xe4 site.[25] However, the possibility of a minor pathway through the G and H helices and exit at the dimeric interface cannot be excluded. In the present simulations, CO attempts to exit through this pathway, but diffusion out of the protein is never observed. The histidine gate, which was found to allow water entrance in the present work, has been proposed to be the major ligand exit route from the previous spectroscopic work[21] but was not found to be operative in simulations of O_2 diffusion in 14 different HbI variants and in the present work.[51] Concerning the time scale of ligand escape, the present study finds such a process on the 100 ns time scale which is about an order of magnitude longer than escape times from biased simulations such as locally enhanced sampling.[51]

V. CONCLUSION

In summary, the present simulations highlight that tertiary structural changes are correlated with each other. A triggering event (e.g., rotation of Phe97 or breaking of Lys96–heme salt bridge) is followed by additional changes at the interface (e.g., water influx or repositioning of the Fe–Fe separation). Ligand migration—which is at the heart of the function of such proteins—in one monomer is affected by the occupation and ligation state in the other monomer. Such behaviour is a hallmark of allostery. Given that the energetics of the chemical step (affinity of CO towards the heme-Fe) is identical in both monomers, variations in the affinity must arise from the monomeric ligation and occupation state. A mild correlation between protein motion and ligand diffusion was found for the movement of Trp135 whereas other changes in the protein structure did not correlate in an obvious fashion with ligand migration. The energy differences between the relevant changes are small, i.e., on the order of a few kcal/mol. This has important consequences not only for the system investigated here but also for the allosteric systems at large. It is anticipated that "the allosteric effect" is probably not one singular event but an orchestrated and coordinated change involving several elementary steps.

ACKNOWLEDGMENTS

[1] J. Monod, J. Wyman, and J. P. Changeux, "On the natute of allosteric transitions: A plausible model," J. Mol. Biol. **12**, 88–118 (1965).

[2] W. A. Eaton, E. R. Henry, J. Hofrichter, and A. Mozzarelli, "Is cooperative oxygen binding by hemoglobin really understood?," Nat. Struct. Biol. **6**, 351–358 (1999).

[3] M. Brunori, "Half a century of hemoglobins allostery," Biophys. J. **109**, 1077–1079 (2015).

[4] H. S. Chung, S. Piana-Agostinetti, D. E. Shaw, and W. A. Eaton, "Structural origin of landscape roughness in protein folding from single-molecule {FRET} and all-atom molecular dynamics simulations," Biophys. J. **108**, 347a (2015).

[5] T. M. Dawson and S. H. Snyder, "Gases as biological messengers: nitric oxide and carbon monoxide in the brain," J. Neurosci. **14**, 5147–5159 (1994).

[6] R. J. P. Williams, "Nitric oxide in biology: Its role as a ligand," Chem. Soc. Rev. **25**, 77–83 (1996).

[7] M. W. Merx, A. Goedecke, U. Floegel, and J. Schrader, "Myoglobin function reassessed," FASEB. J. **19**, 1105–1107 (2005).

[8] D. J. Garry, A. Meeson, Z. Yan, and R. S. Williams, "Life without myoglobin," Cell. Mol. Life Sci. **57**, 896–898 (2000).

[9] E. De Marinis, L. Casella, C. Ciaccio, M. Coletta, P. Visca, and P. Ascenzi, "Catalytic peroxidation of nitrogen monoxide and peroxynitrite by globins," IUBMB Life **61**, 62–73 (2009).

[10] A. Hausladen, A. Gow, and J. S. Stamler, "Flavohemoglobin denitrosylase catalyzes the reaction of a nitroxyl equivalent with molecular oxygen," Proc. Natl. Acad. Sci., U.S.A. **98**, 10108–10112 (2001).

[11] S. Mishra and M. Meuwly, "Atomistic simulation of no dioxygenation in group i truncated hemoglobin," J. Am. Chem. Soc. **132**, 2968–2982 (2010).

[12] J. B. Wittenberg and B. A. Wittenberg, "Myoglobin function reassessed," J. Exp. Biol. **206**, 2011–2020 (2003).

[13] H. S. Cho, N. Dashdorj, F. Schotte, T. Graber, R. Henning, and P. Anfinrud, "Protein structural dynamics in solution unveiled via 100-ps time-resolved x-ray scattering," Proc. Natl. Acad. Sci., U.S.A. **107**, 7281–7286 (2010).

[14] R. Neutze, "Opportunities and challenges for time-resolved studies of protein structural dynamics at x-ray free-electron lasers," Philos. Trans. R. Soc. Lond. B **369**, 20130318 (2014).

[15] K. Moffat, "Time-resolved crystallography and protein design: signalling photoreceptors and optogenetics," Philos. Trans. R. Soc. B **369**, 20130568 (2014).

[16]F. Schotte, M. Lim, T. A. Jackson, A. V. Smirnov, J. Soman, J. S. Olson, G. N. Phillips, M. Wulff, and P. A. Anfinrud, "Watching a protein as it functions with 150-ps time-resolved x-ray crystallography," Science **300**, 1944–1947 (2003).

[17]Q. Cui and M. Karplus, "Allostery and cooperativity revisited," Prot. Sci. **17**, 1295–1307 (2008).

[18]S. Song, L. Rothberg, D. L. Rousseau, A. Boffi, and E. Chiancone, "Metastable CO binding sites in the photoproduct of a novel cooperative dimeric hemoglobin," Biophys. J. **65**, 1959–1962 (1993).

[19]W. E. Royer, W. A. Hendrickson, and E. Chiancone, "Structural transitions upon ligand binding in a cooperative dimeric hemoglobin," Science **249**, 518–521 (1990).

[20]J. E. Knapp, M. A. Bonham, Q. H. Gibson, J. C. Nichols, and W. E. Royer, "Residue F4 plays a key role in modulating oxygen affinity and cooperativity in Scapharca dimeric hemoglobin," Biochemistry **44**, 14419–14430 (2005).

[21]K. Nienhaus, J. E. Knapp, P. Palladino, W. E. Royer, and G. U. Nienhaus, "Ligand migration and binding in the dimeric hemoglobin of scapharca inaequivalvis," Biochemistry **46**, 14018–14031 (2007).

[22]E. Chiancone and A. Boffi, "Structural and thermodynamic aspects of cooperativity in the homodimeric hemoglobin from Scapharca inaequivalvis," Biophys. Chem. **86**, 173–178 (2000).

[23]J. Choi, S. Muniyappan, J. T. Wallis, W. E. Royer, and H. Ihee, "Protein conformational dynamics of homodimeric hemoglobin revealed by combined time-resolved spectroscopic probes," Chem. Phys. Chem. **11**, 109–114 (2010).

[24]Y. Zhou, H. Zhou, and M. Karplus, "Cooperativity in Scapharca dimeric hemoglobin: Simulation of binding intermediates and elucidation of the role of interfacial water," J. Mol. Biol. **326**, 593–606 (2003).

[25]J. E. Knapp, R. Pahl, J. Cohen, J. C. Nichols, K. Schulten, Q. H. Gibson, V. Srajer, and W. E. Royer, "Ligand migration and cavities within Scapharca Dimeric HbI: Studies by time-resolved crystallography, Xe binding, and computational analysis," Structure **17**, 1494–1504 (2009).

[26]M. Karplus and D. L. Weaver, "Protein-folding dynamics," Nature **260**, 404–406 (1976).

[27]M. Karplus and D. L. Weaver, "Protein folding dynamics: The diffusion-collision model and experimental data," Prot. Sci. **3**, 650–668 (1994).

[28]C. M. Dobson and M. Karplus, "The fundamentals of protein folding: bringing together theory and experiment," Curr. Opin. Struct. Biol. **9**, 92–101 (1999).

[29]D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers, "Atomic-level characterization of the structural dynamics of proteins," Science **330**, 341–346 (2010).

[30]J. E. Knapp, R. Pahl, V. Šrajer, and W. E. Royer, "Allosteric action in real time: Time-resolved crystallographic studies of a cooperative dimeric hemoglobin," Proc. Natl. Acad. Sci., U.S.A. **103**, 7649–7654 (2006).

[31]J. W. Petrich, J. C. Lambry, K. Kuczera, M. Karplus, C. Poyart, and J. L. Martin, "Ligand binding and protein relaxation in heme proteins: A room temperature analysis of nitric oxide geminate recombination," Biochemistry **30**, 3975–3987 (1991).

[32]M. Meuwly, O. M. Becker, R. Stote, and M. Karplus, "No rebinding to myoglobin: A reactive molecular dynamics study," Biophys. Chem. **98**, 183–207 (2002).

[33]B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, "CHARMM: The biomolecular simulation program," J. Comp. Chem. **30**, 1545–1614 (2009).

[34]J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," J. Comp. Chem. **26**, 1781–802 (2005).

[35]J. A. D. MacKerell, D. Bashford, M. Bellott, J. R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wirkiewicz-Kuczera, D. Yin, and M. Karplus, "All-atom empirical potential for molecular modeling and dynamics studies of proteins," J. Phys. Chem. B **102**, 3586–3616 (1998).

[36]W. L. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, and M. Klein, "Comparison of simple potential functions for simulating liquid water," J. Chem. Phys. **79**, 926 (1983).

[37]W. V. Van Gunsteren and H. J. C. Berendsen, "Algorithms for macromolecular dynamics and constraint dynamics," Mol. Phys. **34**, 1311–1327 (1977).

[38]G. Torrie and J. Valleau, "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling," J. Comp. Phys. **23**, 187–199 (1977).

[39]S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method," J. Comp. Chem. **13**, 1011–1021 (1992).

[40]M. Souaille and B. Roux, "Extension to the weighted histogram analysis method: Combining umbrella sampling with free energy calculations," Comput. Phys. Commun. **135**, 40–57 (2001).

[41]W. E. Royer, "High-resolution crystallographic analysis of a co-operative dimeric hemoglobin," J. Mol. Biol. **235**, 657–681 (1994).

[42]J. E. Knapp and W. E. Royer, "Ligand-linked structural transitions in crystals of a cooperative dimeric hemoglobin," Biochemistry **42**, 4640–4647 (2003).

[43]K. H. Kim, S. Muniyappan, K. Y. Oang, J. G. Kim, S. Nozawa, T. Sato, S.-Y. Koshihara, R. Henning, I. Kosheleva, H. Ki, Y. Kim, T. W. Kim, J. Kim, S.-I. Adachi, and H. Ihee, "Direct observation of cooperative protein structural dynamics of homodimeric hemoglobin from 100 ps to 10 ms with pump-probe X-ray solution scattering," J. Am. Chem. Soc. **134**, 7001–7008 (2012).

[44]P. J. Steinbach, A. Ansari, J. Berendzen, D. Braunstein, K. C. Hu, B. Cowen, D. Ehrenstein, H. Frauenfelder, J. Johnson, D. C. Lamb, S. Luck, J. R. Mourant, G. U. Nienhaus, P. Ormos, R. Philipp, A. Xie, and R. Young, "Ligand-binding to Heme-proteins—Connection between dynamics and function," Biochemistry **30**, 3988–4001 (1991).

[45]J. Z. Ruscio, D. Kumar, M. Shukla, M. G. Prisant, T. M. Murali, and A. V. Onufriev, "Atomic level computational identification of ligand migration pathways between solvent and binding site in myoglobin," Proc. Natl. Acad. Sci., U.S.A. **105**, 9204–9209 (2008).

[46]N. Plattner and M. Meuwly, "Quantifying the importance of protein conformation on ligand migration in m yoglobin," Biophys. J. **102**, 333–341 (2012).

[47]P.-A. Cazade, G. Berezovska, and M. Meuwly, "Coupled protein-ligand dynamics in truncated hemoglobin N from atomistic simulations and transition networks," Biochem. Biophys. Acta **1850**, 996–1005 (2015).

[48]P. A. Cazade, W. Zheng, D. Prada-Gracia, G. Berezovska, F. Rao, C. Clementi, and M. Meuwly, J. Chem. Phys. **142**, 025103 (2015).

[49]J. Cohen, A. Arkhipov, R. Braun, and K. Schulten, "Imaging the migration pathways for O_2, CO, NO, and Xe inside myoglobin," Biophys. J. **91**, 1844–1857 (2006).

[50]J. S. Hub, M. B. Kubitzki, and B. L. de Groot, "Spontaneous quaternary and tertiary T-R transitions of human hemoglobin in molecular dynamics simulation," PLoS Comput. Biol. **6**, e1000774 (2010).

[51]K. Trujillo, T. Papagiannopoulos, and K. W. Olsen, "Effects of mutations on the molecular dynamics of oxygen escape from the dimeric hemoglobin of scapharca inaequivalvis," F1000Research, **4**, 65 (2015).

[52]J. M. Laine, M. Amat, B. R. Morgan, W. E. Royer, and F. Massi, "Insight into the allosteric mechanism of scapharca dimeric hemoglobin," Biochemistry **53**, 7199–7210 (2014).